An odyssey in modern quantum many-body physics

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029677">http://nrs.harvard.edu/urn-3:HUL.InstRepos:42029677</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
An odyssey in modern quantum many-body physics

A dissertation presented
by
Liujun Zou
to
The Department of Physics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Physics

Harvard University
Cambridge, Massachusetts
April 2019
An odyssey in modern quantum many-body physics

Abstract

This thesis is a collection of research work in quantum many-body physics that I have done during the past few years. The topics include: 1. a theory of anomalous magnetotransport from mass anisotropy of the fermions on a Fermi surface; 2. attempts to understand and build a model for magic-angle bilayer graphene; 3. phenomena of dimensional decoupling at various continuous Mott transitions; 4. possible field-induced neutral Fermi surface and QCD$_3$-Chern-Simons quantum criticalities in Kitaev materials; 5. examples showing that the commonly-employed methods for detecting topological orders may generate spurious results; 6. a classification of three-dimensional symmetry enriched $U(1)$ quantum spin liquids, and the applications of the ideas therein to study topological crystalline insulators.
## Contents

0 **Introduction**  
  0.1 Quantum phases and phase transitions .......................... 2  
  0.2 Outline of this thesis ........................................... 7

1 **Theory of anomalous magnetotransport from mass anisotropy**  
  1.1 Model of an anisotropic Fermi pocket ............................ 13  
  1.2 "Familiar" regimes ................................................. 15  
  1.3 "Fake" high field regime: $\omega_{h}\tau \ll 1 \ll \omega_{l}\tau$ ........... 17  
  1.4 Discussion .......................................................... 18

2 **Origin of Mott insulating behavior and superconductivity in twisted bilayer graphene**  
  2.1 Introduction .......................................................... 21  
  2.2 Electronic Structure of Twisted Bilayer Graphene: General Considerations ............. 24  
  2.3 Low Energy Theory: Two Band Projection .......................... 28  
  2.4 Mirror and chirality .................................................. 34  
  2.5 Inter-valley coherent order: phenomenological motivation ............ 36  
  2.6 Simple theory of the IVC ordered state ............................ 38  
  2.7 Inter-valley coherent Mott insulators: generalities and a concrete example ........... 41  
  2.8 Other possible Mott insulating states ............................. 46
# 2.9 Proposed future experiments ................................. 48
# 2.10 Conclusion .................................................. 49

## 3 Faithful tight-binding models and fragile topology of magic-angle bi-layer graphene in twisted bilayer graphene 51
### 3.1 Introduction ................................................. 51
### 3.2 Symmetry and topology of TBG .................................. 54
### 3.3 Tight-binding models .......................................... 56
### 3.4 Fragile topology .............................................. 61
### 3.5 Discussion ................................................... 63

## 4 Dimensional decoupling at continuous quantum critical Mott transitions 65
### 4.1 Introduction ................................................... 65
### 4.2 Preliminaries ................................................ 70
### 4.3 Warm-up: Dimensional decoupling in a quantum phase transition between a Fermi liquid and an orthogonal metal ................................. 73
### 4.4 Dimensional decoupling in continuous Mott transitions .............................................. 78
### 4.5 Interlayer conductivity ......................................... 86
### 4.6 Crossovers out of criticalities .................................. 89
### 4.7 Discussion ................................................... 94

## 5 Field-induced neutral Fermi surfaces and QCD$$_3$$-Chern-Simons quantum criticalities in Kitaev materials 97
### 5.1 Introduction ................................................... 97
### 5.2 Symmetries of the materials and models ......................... 103
### 5.3 Field-induced neutral Fermi surface ............................... 105
### 5.4 QCD$$_3$$-Chern-Simons quantum criticalities ......................... 116
### 5.5 Discussion ................................................... 125

## 6 Spurious long-range entanglement and replica correlation length 129
6.1 Introduction .................................................. 129
6.2 Example: 2D cluster state on a triangular lattice ......................... 134
6.3 A sufficient condition for a nonzero sub-leading term of entanglement entropy under
the cylinder extrapolation method .................................. 139
6.4 Generic behavior of $\gamma$: replica correlation length .................. 147
6.5 Replica correlation functions ................................... 149
6.6 Discussion .................................................... 163

7 Symmetry enriched U(1) quantum spin liquids 166
7.1 Introduction ................................................... 166
7.2 $U(1)$ quantum spin liquids enriched by time reversal and $SO(3)$ spin rotational sym-
metries .......................................................... 170
7.3 Anomalous quantum spin liquids with $SO(3) \times T$ symmetry .......... 178
7.4 Fractional topological paramagnets ................................ 184
7.5 Combining $U(1)$ QSLs and bosonic SPTs protected by $SO(3) \times T$ ........ 190
7.6 A general framework to classify symmetry enriched $U(1)$ quantum spin liquids .. 197
7.7 $U(1)$ quantum spin liquids enriched by $Z_2 \times T$ symmetry .......... 204
7.8 $U(1)$ quantum spin liquids enriched by some other symmetries .......... 216
7.9 Discussion .................................................... 219

8 Bulk characterization of topological crystalline insulators 222
8.1 Introduction ................................................... 222
8.2 Properties of bulk monopoles from a symmetric surface state ............ 226
8.3 Examples ...................................................... 229
8.4 Relation to symmetry enriched $U(1)$ quantum spin liquids .............. 237
8.5 Discussion .................................................... 241

9 Conclusion 244

Appendix A Appendices of chapter 1 246
A.1 Review of the conventional regimes ................................ 246
A.2 Corrections to the conductivities at the order of $(\omega l\tau)^{-1}$ and $\omega h\tau$ .......................... 250

APPENDIX B APPENDICES OF CHAPTER 2 ............................. 253
B.1 Lattice and symmetries ........................................... 253
B.2 Valley Symmetry and Wannier Obstruction ...................... 255
B.3 Mirror-eigenvalue obstruction and chirality-obstruction .......... 257

APPENDIX C APPENDICES OF CHAPTER 3 ............................ 263
C.1 Symmetry representations ....................................... 263
C.2 Details of the Bloch Hamiltonian ................................ 266
C.3 Deformation to explicit atomic limits for the six-band model .... 274
C.4 Five-band model .................................................. 274
C.5 “Wilson loop” analysis ........................................... 278

APPENDIX D APPENDICES OF CHAPTER 4 ............................ 283
D.1 An example that does not display dimensional decoupling: non-Fermi liquid near an Ising-nematic critical point ........................................ 283
D.2 RG equation of perturbation $\delta L_1$ ............................... 285
D.3 Electron spectral function at criticality of the chemical potential tuned Mott transition ......................... 286
D.4 Effects on the interlayer pairing of (4.20) .......................... 288
D.5 Derivation of the formula for interlayer electric conductivity .......... 290
D.6 Calculations of the imaginary parts of the spinon self-energies in the IFL and MFL regimes, and of the in-plane conductivity in the MFL regime ..................... 291
D.7 Validity of the perturbative calculations of the interlayer electric conductivity ........ 295

APPENDIX E APPENDICES OF CHAPTER 5 ............................. 297
E.1 Symmetry actions of the Kitaev model and the neutral Fermi surface state ............... 297
E.2 Kitaev model in terms of $SU(2)$ partons .......................... 300
E.3 Derivation of duality of critical theories ........................... 303
E.4 Parton mean field of the fermionic $U(2)$ critical theory ............ 308
Appendix F Appendices of chapter 6

F.1 Intertwined chains of cluster state .................................................. 319
F.2 Time reversal and lattice symmetries ................................................. 321
F.3 Cylinder extrapolation method on the 2D cluster state at nonzero temperature . 327

Appendix G Appendices of chapter 7

G.1 Some remarks on time reversal and charge-conjugation symmetry .......... 329
G.2 An SPT: $eCm_2$ ................................................................................. 331
G.3 Anomalous spin liquids as surface states of some $4+1$-d systems ............ 334
G.4 Classification of some SPTs ............................................................... 337
G.5 Relations between the classification of $SO(3)\times T$ symmetric $U(1)$ quantum spin liquids and the classification of some relevant SPTs ................................................................. 342
G.6 Bosonic SPTs with $SO(3) \times T$ symmetry ......................................... 344
G.7 Constraints on the Hall conductances due to time reversal and spin rotational symmetries ................................................................. 345
G.8 Non-edgeability of some $Z_2$ topological orders in the presence of nontrivial particles 347
G.9 Projective representations: the electric (standard), the magnetic (twisted) and the dyonic (mixed) ones ......................................................... 354
G.10 Examine the anomalies of $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids with $\theta = 0$ 361
G.11 $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids with $\theta = \pi$ and $Z_2$ not acting as a charge conjugation ................................................................. 398
G.12 $U(1)$ quantum spin liquids with some other symmetries ..................... 400

Appendix H Appendices of chapter 8

H.1 Time reversal symmetric generalizations of the mirror symmetric TCI .......... 404
H.2 Surface topological orders of various TCIs ............................................. 406

References 437
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The simple model for the Fermi surface pocket used in these notes. It consists of several circular arcs. The light segments in green have radius $k_l$ and Fermi velocity $v_l$, while the heavy segments in red have radius $k_h$ and Fermi velocity $v_h$.</td>
</tr>
<tr>
<td>1.2</td>
<td>Magnetic field dependence of the Hall coefficient (left) and longitudinal resistivity (right) for the model with fixed $m_l = m_e$ and various $m_h$. In the calculations, $k_h = 0.2k_l$, the pocket is fixed to have an area of 1.9% of the Brillouin zone, and the relaxation time is $\tau = 0.2ps$. Only when the mass enhancement is minimal is the Hall coefficient negative at all fields. For large mass enhancement a sign change in the Hall coefficient occurs at a high value of the magnetic field, for instance at $\approx 170, T$ for the case $m_c = 8.3m_e$ shown dashed in blue. When the mass enhancement is substantial, there is also a regime of large $B$-linear magnetoresistance, as described in section 1.3.</td>
</tr>
<tr>
<td>1.3</td>
<td>The dependence of the Hall coefficient on approach to a QCP at which $m_h$ diverges, with $m_l = m_e$ held fixed. Other parameters are as in Figure 1.2. Data are shown at various values of the magnetic field, showing that the sign change always precedes the QCP.</td>
</tr>
</tbody>
</table>


1.4 A schematic illustration of the effects of finite temperature on the Hall coefficient under the scenario explored in this paper. A small electron pocket persists until a critical doping $p_c$, with the Hall coefficient changing sign well below $p_c$ due to mass renormalization and curvature effects. At zero temperature, there is a sharp jump of the Hall coefficient at $p_c$, where the electron pocket transforms into the large Fermi surface. At finite temperature, this jump is rounded, leaving a maximum in the Hall coefficient at a doping $p_{\text{max}} < p_c$.

2.1 Effective symmetries and constraints on band structures. (a) The effective symmetries of the twisted bilayer graphene system can be inferred by inspecting the point-group symmetry of a hexagon center in the real space, taken to be the rotation axis of the layers. (b) Schematic band structure along a high-symmetry path in the moiré Brillouin zone. (c-f) Effect of symmetry breaking. (c) Breaking the $C_2$ rotation will gap out the Dirac points. (d) An external perpendicular electric field breaks the mirror $M_y$ symmetry, which only modifies the energetics but cannot open a band gap at charge neutrality. (e) When $C_3$ rotation is broken, but the combined symmetry of two-fold rotation $C_2$ and time-reversal $\mathcal{T}$ is preserved, the Dirac points remain protected, although unpinned from $K_M$ and $K'_M$. (f) When valley conservation $U_v(1)$ symmetry is broken, one can no longer label the bands using their valley index. The gaplessness at charge neutrality is no-longer symmetry-required, although, depending on detailed energeticss, there can still be remnant Dirac points. In contrast, at quarter filling relevant for the observed Mott physics, there are necessarily Dirac points present in this case. The other symmetry breaking patterns listed above also do not open band gaps at quarter filling.

2.2 A $C_3$ symmetry breaking order which preserves other symmetries.
3.1 Real-space orbitals. (a) Lattices and conventions. The shaded region indicates the relative positions of the sites assigned to the same unit cell. (b) The constructed quasi-orbital $t_{p^z}^{B}$ in the real space. $\xi = e^{i2\pi/6}$ and $\omega = \xi^2$. Going from top to bottom, the entries in the three-component vectors attached to the triangular sites denote the amplitude for the $p_z$, $p_+$ and $p_-$ orbitals; that attached to the kagome sites denote the amplitude of their associated $s$ orbital.

3.2 Band structures. (a,b) Bands from the ten-band Hamiltonian $\hat{H}(t_0, \delta)$. For both panels, we choose $t_0 \equiv 130 \text{ meV}$, and the wave-function parameters $a = 0.110$, $b = 0.033$, $c = 0.033$ and $d = 0.573$. We set $\delta = 0$ in (a) and 1 in (b). (c) Bands obtained from the continuum theory for twisted bilayer graphene with a twist angle of $\theta = 1.05^\circ$. The ten bands around charge neutrality are highlighted. (d-f) A zoom-in of the two bands at charge neutrality for the corresponding panels in (a-c). The three-dimensional plots in (e,f) are plotted over the first Brillouin zone centered at $\Gamma$, showing the presence of exactly two Dirac points pinned to $K$ and $K' = -K$. Note that (e) is generated from our tight-binding model, whereas (f) is generated from the continuum model.

3.3 Band structures from a six-band model. (a) Color code for the orbital characters. (b) The broad energetic features can be set up using only the intra-orbital dispersion. (c) Band structure from the full model, with parameters detailed in Appendix C.2.1.
3.4 Deformation to atomic limits. (a) The band gaps $\Delta$ below and above the two nearly flat bands stay open for all values of $\mu \in [-\mu_0, \mu_0]$ in Eq. (3.4). More than $4 \times 10^4$ momenta are sampled in the Brillouin zone in determining $\Delta$. We choose $\mu_0 = t_0 = 130$ meV. (b–d) Schematic band diagrams at various limits of the deformation. (b) When $\mu = -\mu_0$, the lowest four bands arise solely from the $(\eta, p_{\pm})$ orbitals and are therefore strictly atomic. Correspondingly, the upper six bands, altogether, are also in a strict atomic limit. (Dashed boxes indicate strictly atomic bands.) The same is true for the case of $\mu = \mu_0$ in (d), but with the role of the lowest and highest bands exchanged. Since the band gaps are maintained throughout the entire deformation, we can infer that all the (light and dark) purple blocks correspond to trivial, atomic bands. The nontrivial nearly flat bands at charge neutrality, therefore, must feature fragile topology.

4.1 Dimensional decoupling across the phase transition. On the side with $g > g_c$ we have a 3D Fermi liquid, and at the quantum critical point and on the other side we have a state that behaves as a stack of many decoupled layers at low energies.

4.2 The schematic phase diagram and crossover structure of an orthogonal metal transition. When $g > g_c$ we have a Fermi liquid (FL), while we have an orthogonal metal (OM) when $g < g_c$. At finite temperature, there is a quantum critical regime (QC) where the system is non-Fermi liquid like.

4.3 Phase diagram and crossover structure of a chemical potential tuned Mott transition. The critical point corresponds to $\mu = 0$ and $T = 0$. The $\mu < 0$ side corresponds to a spin liquid Mott insulator (MI), and the $\mu > 0$ side corresponds to a Fermi liquid metal (FL). The quantum critical regime is highly non-Fermi liquid like (QC nFL). In crossing over to the Fermi liquid, the system has to first go through an intermediate incoherent Fermi liquid regime (IFL).
4.4 Phase diagram and crossover structure of the bandwidth controlled Mott transition. The $g < g_c$ side corresponds to a spin liquid Mott insulator (MI), and the $g > g_c$ side corresponds to a Fermi liquid (FL) metal. The quantum critical (QC) regime is highly non-Fermi liquid like. In crossing over to the Fermi liquid, the system has to go through an intermediate marginal Fermi liquid regime (MFL). In crossing over to the Mott insulator, the system has to go through a marginal spinon liquid (MSL) regime, but this regime will not be discussed in this paper.

4.5 Spinon self-energy. The wavy line represents a gauge boson with momentum and frequency $(q, i\Omega)$, and the dashed line represents a spinon with momentum and frequency $(k - q, i(\omega - \Omega))$.

4.6 Schematic crossover behaviors of the resistivity with respect to temperature for chemical potential tuned Mott transition (left) and for bandwidth controlled Mott transition (right). The solid line represents the out-of-plane conductivity and the dashed line represents the in-plane conductivity.

5.1 Under a Zeeman field, a Kitaev material may go through four different phases (scenario A): a zigzag ordered state, a non-Abelian chiral quantum spin liquid with Ising topological order (ITO), a neutral Fermi surface (NFS) coupled to an emergent dynamical $U(1)$ gauge field, and a trivially polarized state. The phase transitions between these phases are described by the 1) QCD$_3$-Chern-Simons theory, 2) gauged pairing transition, and 3) gauged Lifshitz transition. Depending on the microscopic details (e.g., direction and strength of the Zeeman field), there can also be a direct continuous quantum phase transition between the ITO state and the trivial polarized state, and the critical theory is described by another QCD$_3$-Chern-Simons theory (scenario B). The QCD$_3$-Chern-Simons transitions require a high symmetry of the system to be stable, which can be satisfied if the magnetic field is in the $ac^*$ plane. All other phases and transitions do not rely on the high symmetries of the system.
5.2 The honeycomb lattice. The \( x \)-bond, \( y \)-bond and \( z \)-bond are along the direction of \( a_1 \), \( a_2 \) and \( a_3 \), respectively. The \( a \)- and \( b \)-axes are shown in the figure, and the \( c^* \)-axis is perpendicular to the paper and pointing outwards.

5.3 (a) Cylinder geometry used in DMRG simulations. Here we plot an example with \( L_y = 3 \) unit cells (6 sites). The sites labeled by the same letters are identified. (b), (c) The dependence of the correlation length \( \xi \) on the magnetic field strength \( h \) and bond dimension \( \chi \) on the (b) \( L_y = 2 \) cylinder and the (c) \( L_y = 3 \) cylinder.

5.4 (a), (b): Schematic illustration of Brillouin zone (BZ) and Fermi pockets. The dashed line represent the momenta points accessible on (a) the \( L_y = 2 \) cylinder and (b) the \( L_y = 3 \) cylinder. The spinons have anti-periodic boundary condition on the \( L_y = 2 \) cylinder, and periodic boundary condition on the \( L_y = 3 \) cylinder. (c), (d): Central charge fitting for the (c) \( L_y = 2 \) cylinder and (d) \( L_y = 3 \) cylinder.

5.5 The static structure factor of \( S^x \) on the \( L_y = 3 \) cylinder. The value of \( 2k_f \) decreases as the field strength increases, which is a signal that the NFS shrinks as the field increases.

5.6 Representative dispersions of the Hamiltonian (5.9). The parameters we use are \( J'_1/J_1 = 3.75 \), \( J_2/J_1 = -0.75 \), and \( h/J_1 = 1, 2, 3, 4 \) for the four dispersions, respectively. The definitions of the parameters \( J_1 \), \( J_2 \) and \( J'_1 \) are given in Appendix E.2. In all these dispersions, the chemical potential is tuned so that the gauge constraint (5.7) is satisfied at the level of expectation values: \( \langle f_i^\dagger f_i \rangle = 1 \).

5.7 Brillouin zone of the honeycomb lattice. The red dashed line represents the high symmetric line which connects the \( K \) and \( K' = -K \) point; The blue dashed line connects the \( M_3 \) and \( M_3 \) point.

6.1 Two methods of extracting TEE. Left: Kitaev-Preskill prescription divides the system into four parts and extract TEE by using (6.2). Right: DMRG calculations put the system on an infinite cylinder and divide the system into two parts, then calculate the entanglement entropy \( S(L) \) between the two parts for different circumferences \( L \) of the cylinder. Fitting the results into (6.1), TEE is identified as \(-S(L = 0)\).
6.2 A 7-spin interaction in the summation of Eq. (6.8), which is a product of the $\sigma^x$ operator on the center site and the $\sigma^z$ operators on the sites that surround it.

6.3 Triangular lattice with an entanglement cut parallel to one of the sides of a triangle. The upper region is $A$, and the lower is $B$.

6.4 The reduced 1D chain of the 2D cluster state. The red circles represent qubits in region A and the green circles represent qubits in region B.

6.5 Bravyi’s example. There is a qubit on each vertex of the zig-zag chain, and the dashed line represents an entanglement cut that divides the chain into two halves. If the chain is in the cluster state of Eq. (6.10), the entanglement entropy is equal to $\frac{1}{2}\#(\text{bond cuts}) - 1$ in the units of log 2.

6.6 (a) $\rho$ obtained by tracing out one physical qudits for every site. (b) $\rho^2$. Connected bonds are contracted. Wiggling lines represent complex conjugation. The horizontal virtual bonds are contracted due to the periodic boundary condition, which is not drawn. $\text{Tr}(\rho)$ and $\text{Tr}(\rho^2)$ are computed by contracting the upper vertical wiggling bonds with the lower straight ones. For this purpose, it is enough and more efficient to consider the transfer matrix designated by the dotted rectangle; see Eq. (6.21).

6.7 Transfer matrix and its symmetry. The left-most diagram of (b) represents the transfer matrix $T_3$ for $\rho^3$. The ensuing equalities are direct consequences of the symmetry lifted to the virtual level. This implies that the transfer matrix $T_\alpha$ has degeneracy $q^\alpha - 1$ for some integer $q > 1$.

6.8 Measuring replica correlation functions. Observables are inserted in the circles. Even if the global state $\rho_{AB}$ has a short correlation length, the positive semi-definite operator $\rho_A^2$, treated as a normalized state $\rho_A^2 / \text{Tr}(\rho_A^2)$, may have much longer correlation length. The latter length scale, which we call as the replica correlation length, can be simply measured in numerical calculations, and is the relevant length scale for the subleading term in the entanglement entropy.
6.9 Replica correlation function calculation. One prepares two copies of the state, and apply the swap operator on the shaded region; insert observables in the circles, and compute the overlap with the original, unswapped state. The overlap is generally exponentially small in the boundary length of the shaded region, but after normalization this reveals the replica correlation length.

6.10 Double-semion model defined in (6.58). (a) shows the lattice configuration and the location of the degrees of freedom. Each edge of the hexagon accommodates two spins, denoted by the blue dots on the edge. The dashed line is the entanglement cut. (b-d) pictorially show the three terms in the Hamiltonian for a plaquette, vertex and edge. The symbol near a spin denotes the operator acting on this spin, where $X$ means $\sigma_x$, $Z$ means $\sigma_z$, $\sqrt{Z} = \text{diag}(1,i)$, and $\text{Id}$ is the identity matrix. The total Hamiltonian is the summation over all plaquettes, vertices and edges, with appropriate sign factors defined in (6.58).

7.1 Charge-monopole lattice at $\theta = 0 \pmod{2\pi}$. 

7.2 Charge-monopole lattice at $\theta = \pi \pmod{2\pi}$.

7.3 Symmetry protected distinctions among symmetry enriched $U(1)$ quantum spin liquids. For example, with $SO(3) \times T$ symmetries, two phases, $E_b M_b$ and $E_{b1} M_{f1}$ in this example, cannot be connected without crossing a phase transition. When the symmetry is broken, they can be connected without crossing a phase transition.
7.4 There is a systematic ungauging procedure that takes a symmetry enriched $U(1)$ quantum spin liquid to its corresponding SPT. Consider the time reversal symmetric $U(1)$ quantum spin liquid $(E_{FT}M_f)_\theta$ (the upper left system), and we will try to get its corresponding SPT from the perspective of the electric charge. To do this, we first introduce an auxiliary trivial time reversal symmetric insulator made of fermions that are Kramers doublets, where these fermions are denoted by $c$ (the lower left system). Next we condense the bound state of $E$, the electric charge of the $U(1)$ spin liquid, and $c^\dagger$, the holes in the auxiliary trivial insulator. This bound state is a boson and a Kramers singlet, so this condensation will preserve the time reversal symmetry. The dynamical $U(1)$ gauge field in the $U(1)$ gauge theory will be confined, and the resulting state is precisely the Fu-Kane-Mele topological insulator, which, viewed from the perspective of the electric charge, is the corresponding SPT of $(E_{FT}M_f)_\theta$ (the right system).

7.5 That the $G$ symmetric $U(1)$ quantum spin liquid is anomaly-free is equivalent to that it has a corresponding SPT, which is in turn equivalent to that this SPT can have a consistent (but possibly anomalous) 2D surface state.

8.1 The notion of symmetry protected distinction of quantum phases of matter. As long as the relevant symmetries are preserved, the two phases cannot be connected without crossing a phase transition. However, if symmetry-breaking perturbations are allowed, the phase transition can be avoided.

C.1 Coupling terms in the six-band model. The full coupling terms consist of the indicated hopping together with their Hermitian conjugates. We always take the center site to be a triangular site in the “home” unit cell, and indicate the unit cell coordinates $la_1 + ma_2$ of connected sites by $(lm)$ with $\bar{l} \equiv -l$. For the kagome sites in (e,f), we further specify their sublattice indices. The strength of the terms in panels (a-f) are respectively denoted by $t_{p\pm p\pm}^+, t_{p\mp p\pm}^-, t_{p\pm p\pm}^+, t_{p\pm p\pm}^-, t_{\kappa p\pm}^+$, and $t_{\kappa p\pm}^-$. 

C.2 The leading perturbation to the ten-band model, which breaks the undesirable $\tilde{T}$-invariance. Each circle denotes a honeycomb site, with the orbital $p_\pm$ indicated.
C.3 Evolution of the band gaps above and below the two nearly flat bands from Eq. (C.21), which establishes an adiabatic deformation of the lowest four bands of our six-band model to a strict atomic insulator. More than $4 \times 10^4$ momenta are sampled in the BZ in determining $\Delta$.

C.4 Real-space orbitals for the five-band model. (a) On each of triangular sites (filled yellow circles), we consider the three $p$ orbitals $p_z$ and $p_\pm \equiv p_x \pm ip_y$; on each of the honeycomb sites (open circles), we consider an $s$ orbital. The centers of the nearest-neighbor bonds between the honeycomb sites form a kagome lattice. (b) The constructed quasi-orbital $\rho_s^{(1)}$ centered at a kagome site, indicated by a cross. Going from top to bottom, the entires in the three-component vectors attached to the triangular sites denote the amplitude for the $p_z$, $p_+$, and $p_-$ orbitals; that attached to the honeycomb sites denote the amplitude of their associated $s$ orbital. The other two quasi-orbitals labeled by $l = 2, 3$ can be obtained through symmetries.

C.5 Band structure from the five-band model. (a) The full spectrum. (b) A zoom-in for the two nearly flat bands. (c) The band gaps above and below the two nearly flat bands at charge neutrality never close as we deform the bands into strict atomic limits. More than $4 \times 10^4$ momenta are sampled in the BZ in determining $\Delta$.

C.6 Wilson loop spectra for the following set of bands in the ten-band model: (a) the two nearly flat bands at charge neutrality; (b) the lowest four bands; and (c) the six bands of (a) and (b) combined. Note that a nontrivial spectral flow, which forbids any atomic description, is found only in (a). This is consistent with the fragile nature of the band topology.

C.7 Wilson loop spectra for the following set of bands in $H^{(5)}_{\mu=0}$: (a) the two nearly flat bands at charge neutrality; (b) the lowest three bands; and (c) the five bands of (a) and (b) combined.

D.1 Feynman diagrams that contribute to the RG equation of $\delta L_1$ to the one-loop order.

E.1 The hopping terms of the parton mean-field ansatz.
F.1 Deforming the state while respecting the lattice reflection or inversion symmetry. From the original 1D chain which consists of red ($a_i$) and green ($b_i$) qudits, one can insert an auxiliary yellow ($c_i$) qudit. Then one can apply the swap operator that exchanges the red qudits and yellow qudits circled by the dashed ellipses. This swap operation can be implemented continuously without breaking the lattice reflection symmetry. The numbers in each qudits label the positions of the corresponding qudits before the swap operation. It is understood that we have a 1D chain of qudits, although only a few sites are shown here.

F.2 Deforming state while respecting symmetry. In addition to the red ($a_i$) and green ($b_i$) qudits, a pair of time reversal-invariant spin singlets are inserted in each site (the yellow qudits are $a'_i$ and blue qudits are $b'_i$). The swap unitary is applied to the qudits circled by the dashed ellipses, so that the entanglement across the cut solely arises from the inserted singlet. The swap can be implemented continuously during which the time-reversal symmetry is unbroken. It is understood we have a 1D chain of qudits, although only one site is shown here.

G.1 Layer construction of the 4+1-d system whose surface realize an anomalous spin liquid.

H.1 Gapping out the helical fermions by a $Z_4$ STO.
To all my mentors,
especially the two who have been with me from the very beginning: my parents, Ao Zou and Aihua Liu.
Acknowledgments

First and foremost, I would like to thank my thesis advisor T. Senthil for giving me the chance to work with him, for his enthusiastic support, wise guidance, and insightful advice. I would also like to thank the two other members of my thesis committee, Subir Sachdev and Philip Kim, for invaluable discussions and encouragement. I thank Ashvin Vishwanath for fruitful collaborations, and thank him and Daniel Jafferis for giving me the opportunity to be their teaching assistant. I thank the students, postdocs and other faculties at Harvard and MIT for providing an amazing dynamic environment for creative research, from which I have benefited tremendously. In particular, I thank Jeongwan Haah and Chong Wang for leading me into the field and helping me build up my basis. They are two of my earliest mentors at the beginning of my PhD, and they have always been my role models. I have had the pleasure to discuss physics with and learn from Debanjan Chowdhury, Adam Nahum, Lucile Savary, Brian Swingle, Max Metlitski, Xiao-Gang Wen, Liang Fu, Patrick Lee, Michael Pretko, Zhehao Dai, Yi-Zhuang You, Yin-Chen He, Samuel Lederer, Adrian Po, Zhen Bi, Xueda Wen, Yang Qi, Meng Cheng, Ya-Hui Zhang, Dan Mao, Wenjie Ji, Shang Liu, Ruihua Fan, Muqing Xu, Xue-Yang Song, Linda Ye and others. I am deeply grateful to all the friends I made during my PhD and thank them for making my life in Cambridge enjoyable. I would like also to thank Jacob Barandes, Lisa Cacciabaudo and Carol Davis for all the help. Without all of them, this thesis would not have been possible.

My PhD research has been supported in part by the Purcell Fellowship at Harvard University, a US Department of Energy grant de-sc0008739, and NSF grant DMR-13001648.
Quantum many-body physics is a subject that involves the studies of the rationalization, detection, realization, exploration, characterization and classification of quantum phases and phase transitions.

Traditionally, quantum phases of matter are often distinguished by different patterns of spontaneous symmetry breaking, and the low-energy excitations can be described in terms of long-lived well-defined quasi-particles. In the traditional wisdom, quantum phase transitions are often characterized by fluctuating order parameters of certain symmetries [1, 2].

Over the last decades, exotic quantum phases and phase transitions beyond the conventional paradigm are uncovered. The most well established examples may be the fractional quantum Hall (FQH) states, which are nontrivial phases of matter not associated with spontaneous symmetry-
breaking. The low-energy excitations in the FQH states are still long-lived quasi-particles, but these quasi-particles are fractionalized, i.e., they can carry fractional statistics and fractional charge [1].

Another class of long-sought exotic phases are quantum spin liquids (QSLs) [1]. A QSL is a spin system with a nontrivial structure of long-range entanglement in the ground state. QSLs can be broadly classified into two groups: gapped QSLs and gapless QSLs. A gapped QSL has an energy gap that separates its ground state(s) from its excited states, and its low-energy excitations also contain fractionalized long-lived quasi-particles, just as in the FQH states. In a gapless QSL, the dynamics is more complex because the excitations are gapless, and it may not be possible to view these excitations in terms of well-defined quasi-particles.

Motivated by these discoveries, further experimental and theoretical exploration in exotic quantum phases and phase transitions has become a frontier of modern quantum many-body physics. This thesis is a collection of my studies (in collaboration with many other great physicists) in this subject. To have a global viewpoint on this subject, below we will first have a quick review of the landscape of quantum phases and phase transitions known to date. We will only consider physics in thermal equilibrium or close to equilibrium, and we will mainly discuss systems with local interactions, although physics beyond such scenarios is very interesting and rich.

0.1 **Quantum phases and phase transitions**

It turns out to be easier if we discuss gapped phases, gapless phases and quantum phase transitions separately.

0.1.1 **Gapped phases**

Gapped phases are also known as topological phases. By now there is a relatively systematic understanding of the broad structure of gapped phases. They can be roughly organized into the following classes.

1. Featureless states. These are gapped states that preserve all global symmetries of the system and possess no long-range entanglement. More precisely, these states can be smoothly
connected to a trivial product state without breaking the symmetry or encountering a phase transition. A classic example of these states is the paramagnetic phase of a transverse-field Ising model [2].

2. Gapped states with spontaneous breaking of a discrete symmetry. These are gapped states that can be smoothly connected to a product state while preserving the unbroken symmetries without encountering a phase transition, but these states spontaneously break a global discrete symmetry. The difference between these states and featureless states is that the former has a ground state degeneracy due to spontaneous symmetry breaking, while the latter does not have protected ground state degeneracy. A classic example of these states is the ferromagnetic phase of a transverse-field Ising model [2].

3. Symmetry-protected topological (SPT) states, also known as symmetry-protected trivial states. These are gapped states that preserve all the symmetries and have symmetry-protected distinctions from a trivial product state. Namely, these states can be smoothly connected to a trivial product state without encountering a phase transition if and only if the relevant symmetries are preserved. Classic examples of these states include one-dimensional Haldane chain [3], and topological insulators and superconductors in various dimensions [4–6]. Notice we often also regard the featureless state as a trivial SPT. In passing, SPTs are also referred to as invertible states with symmetry protection.

4. Invertible states without symmetry protection. These are states that cannot be smoothly connected to a trivial product state without encountering a phase transition, no matter whether the symmetries are preserved or broken. Furthermore, by certain local operations all excitations in these states can be created, with no expense of creating other excitations. These states are, in a sense, rare. In one dimension, the only example is the Kitaev chain [7]. In two dimensions, the only two examples are the fermionic $p \pm ip$ superconductor [8] and the bosonic $E_8$ state [9], or multiple copies of them. There is no such example in three dimensions, while in four dimensions there is an example [10]. Together with SPTs, these states are referred to as invertible states (with or without symmetry protection).

5. Topologically ordered states. Just like invertible states, these are also states that cannot
be smoothly connected to a trivial product state without encountering a phase transition, no matter whether the symmetries are preserved or broken. The difference is that in these states there are necessarily fractional excitations, namely, excitations that cannot be created alone by local operations. Classic examples of these states are fractional quantum Hall states and gapped QSLs [1].

6. Gapped fractonic states. In all the above classes of states, excitations are well-defined quasi-particles (there are also quasi-string excitations in three-dimensional topological orders) that can move freely. Fractonic states have excitations with restricted mobility. A classic example of these states is Haah’s cubic code [11].

Notice the above classes should be viewed as “roots” of more complicated gapped phases, which can be (conceptually) obtained by stacking multiple states from different classes in these roots.

0.1.2 Gapless phases

Next, let us turn to gapless phases, whose structure is much less understood. An important difference between gapped and gapless phases is that the former is always stable against a weak perturbation due to the presence of a finite gap, while the latter requires some other reason to be a stable phase. There are a few different mechanisms that can protect the gaplessness of a system, and we can organize the classes of gapless phases according to the mechanism responsible for its protection.

1. Gapless states with spontaneously-broken continuous symmetry. Due to the Goldstone theorem, as long as the system only allows for local interactions, a spontaneous-broken continuous symmetry will imply a gapless mode, the Goldstone mode. Classic examples of this type of gapless phases include solids with gapless phonon, superfluid helium, etc.

2. Gapless states protected by global symmetries. In these states, as long as certain global symmetries are preserved, they are stable against weak perturbations. This phenomenon is closely related to the Lieb-Schultz-Mattis theorems and quantum anomalies. Classic exam-
amples of this type of gapless phases include Dirac and Weyl semimetals, Fermi liquid metals, certain $U(1)$ Dirac QSLs [12], etc.

3. Gapless states with an intrinsic entanglement pattern that makes it stable. In some cases, regardless of the global symmetries, the entanglement pattern of the system is sufficient to make all local weak perturbation to be irrelevant to this gapless state. Classic examples of this type of gapless phases include a three-dimensional $U(1)$ QSL with gapped matter fields [13–19], a $U(1)$ QSL with a neutral Fermi surface [20–29], composite Fermi liquids in quantum Hall systems [28,29], etc.

Part of the reason that gapless phases are much more difficult to capture is that the excitations in some of them cannot be viewed as any well-defined quasi-particles. That is, due to the strong interactions among excitations in such systems, any would-be quasi-particles will have a very short lifetime so that it is inappropriate to view them as quasi-particles. The gapless Goldstone modes can often be viewed as well-defined quasi-particles, so can the excitations in Dirac and Weyl semimetals and that in Fermi liquid metals, as well as the excitations in a three-dimensional $U(1)$ QSL with gapped matter fields. The dynamics of these systems are relatively easier to capture. In contrast, excitations in certain $U(1)$ Dirac QSL and $U(1)$ QSL with a neutral Fermi surface cannot be given a quasi-particle-based description, and the universal properties of these systems are still under intense study. Understanding gapless phases without quasi-particle excitations remains an outstanding challenge in quantum many-body physics.

0.1.3 Quantum phase transitions

Finally, we turn to quantum phase transitions. Before going into the details, let us first justify why it is interesting and even important to understand a quantum phase transition between two quantum phases, given that a phase transition only occurs at the phase boundary, which is a measure-zero region in the full parameter space. The first reason to understand quantum phase transitions is because many of them exhibit interesting universal properties, just like quantum phases. The second reason is, if the phase transition is continuous, understanding the phase transition usually requires understanding the dynamics of the relevant degrees of freedom of the two
nearby phases, which enables us to understand the two phases in a unified fashion. Furthermore, the properties of a quantum critical point control the physics at the intermediate length scales of the system in either phases close to the critical point.

Despite its importance and the intense studies on it for many decades, we are still far from having a systematic understanding of the possible types of quantum phase transitions and the universal critical properties of them. In principle, one can imagine a quantum phase transition between any two phases. The most conventional quantum phase transition is perhaps the ones between a featureless state and a (gapped or gapless) state with spontaneously-broken symmetries. The classic examples include the transition between the paramagnetic state and the ferromagnetic state described by the transverse-field Ising model, and the transition between a superfluid and a Mott insulator described by the Bose-Hubbard model [2]. The theories that describe these transitions can be formulated in terms of fluctuating local order parameters of the relevant symmetries, and application of the idea of renormalization-group (RG) to these theories uncovers the universal physics of the critical point. This framework of critical theories for a quantum phase transition will be referred to as the Landau-Ginzburg-Wilson framework.

Another type of simple quantum phase transitions is between two phases of non-interacting fermions, say, between a band insulator and a band metal, between trivial band insulator and a topological band insulator, etc. The critical theory of such phase transitions can be formulated in terms of these non-interacting fermions. Due to their non-interacting nature, such theories are easy to solve. Notice it is again the local degrees of freedom that enter the critical theories in this scenario.

However, it turns out to be very difficult to capture many other types of quantum phase transitions by formulating the critical theory purely in terms of local degrees of freedom. For instance, to the best of our knowledge, there is no example of a critical field theory in terms of local degrees of freedom, which describes a quantum phase transition between a topologically ordered state and a featureless state. In such cases, it turns out to be easier to formulate the critical field theory in terms of some fractionalized degrees of freedom, in spite of the fact that the notion of fractionalization is often ill-defined due to the absence of well-defined quasi-particles at those critical points. A classic example of such critical theories is the conjectured deconfined
quantum critical points between a magnetically ordered state and a valence bond solid state [30].

So up to date the usual strategy to construct a critical theory between two phases is to consider the possible dynamics of the (local or non-local) degrees of freedom in one of the two phases, and check whether changing the dynamics of those degrees of freedom will turn the system into another phase. This approach has produced fruitful results. However, if one is interested in generating novel examples of quantum phase transitions, another approach is to start from a critical field theory and ask which phases can be reached once relevant perturbations are added. The general philosophy behind this approach is to recognize that an infrared (IR) critical field theory can have different ultra-violet (UV) regulations, and depending the UV regulations, this IR field theory can describe different quantum phase transitions. An elegant example of work within this approach is Ref. [31].

At this point we have only discussed the construction of a critical theory, namely, a theory that can be in two different desired phases depending on the parameters of this theory. We have not covered the dynamics of a critical theory, which is a similar problem to the dynamics of some of the gapless phases. This problem remains an outstanding theoretical challenge due to the absence of any quasi-particle description of the system in many cases.

0.2 Outline of this thesis

During the past few years I have worked on a variety of topics in modern quantum many-body physics, and the rest of the thesis will be the details of these works. Below I will give an outline of this thesis.

Some of the topics that I worked on are directly relevant to current experiments. In chapter 1, I will present our work on a theory of anomalous magnetotransport due to anisotropy of the effective mass of the fermionic quasi-particles on a Fermi surface. Conventionally it is thought that the Hall coefficient is able to indicate the charge and the density of the electric carriers of a system, and experimentally people often use the result of the Hall coefficient to interpret the carrier type and carrier density. In our work, we give an alert that the Hall coefficient is directly related to the carrier type or carrier density only under special circumstances, which puts doubts on the interpretation of some recent experiments [32].
Another piece of work in this flavor is on twisted bilayer graphene, which is the focus of chapters 2 and 3. Realizing correlated insulators which become superconducting upon doping in twisted bilayer graphene is a great recent progress in experimental condensed matter physics, and it started tremendous excitement to study such moiré structures [33, 34]. In a series papers, we attempted to understand the phenomena observed in the experiments on twisted bilayer graphene, and to build up an effective model to capture the physics in such a system. Along the way, we uncovered an intriguing type of band topology in twisted bilayer graphene.

Besides the works that are directly relevant to experiments, some of my works are motivated by recent experimental and numerical findings, but the predictions of the theories we developed have not been experimentally tested. The flavors of these theories are, in a sense, exotic. One such work is to demonstrate the property of dimensional decoupling at some quantum critical continuous Mott transitions between a Fermi liquid metal and a $U(1)$ QSL with a neutral Fermi surface. This work was motivated by experiments on $\kappa$-(ET)$_2$Cu$_2$(CN)$_3$ [35–37], and its outcome can be potentially applied to detect the non-Fermi liquid physics in various layered quasi-two-dimensional materials by measuring interlayer electric transport. This work is presented in chapter 4.

Another piece of such work is to study the phase diagram of the Kitaev materials, and to study the critical theories between various pairs of phases in the phase diagram, as detailed in chapter 5. This work was motivated by the recent progress in Kitaev materials, especially $\alpha$-RuCl$_3$ [38]. We pointed out from a simplified model that the Kitaev materials may host a $U(1)$ QSL state with neutral Fermi surfaces. Furthermore, we constructed highly nontrivial critical theories to describe the quantum phase transitions out of the long-sought non-Abelian QSL, which could be realized by the Kitaev materials.

Besides these works, I also did some research that are more conceptual, and chapters 6, 7 and 8 are devoted to them. In chapter 6, we constructed explicit examples and proved that in such examples the popular way of detecting topological orders may lead to spurious results. In chapter 7, we classify three-dimensional symmetry-enriched $U(1)$ QSLs with gapped matter fields. This work is one of the first systematic studies of three-dimensional symmetry-enriched long-range entangled gapless phases. Along the way, we establish many results on SPTs and quantum anoma-
lies. The idea involved in this work is later applied to study three-dimensional topological crystalline insulators (TCIs) in chapter 8. We show that the topological nature of some TCIs is robust under strong interactions, and we demonstrate the crystalline equivalence principle through these examples.

After all these chapters, in chapter 9 I will reflect on what I should have done better during my PhD. These mainly consist of a set of research problems which I have not made any contribution to but are interesting to study in the future.
Recent high magnetic field experiments [39] have shed new light on the underdoped regime of several cuprate high temperature superconductors (particularly YBa$_2$Cu$_3$O$_{6+x}$ or YBCO), revealing a field-induced metallic state at low temperature that exhibits quantum oscillations, a hallmark of Fermi liquid behavior. A precise theoretical description of this phase (or phases) could provide a valuable framework by which to understand, among other phenomena, the pseudogap regime that prevails at higher temperature. However, even relatively basic questions about the phenomenology of the field-induced metallic regime, such as the number of phases present and the symmetries they break, remain unsettled.
In YBCO, one unsettled question concerns the extent of charge density wave (CDW) order within the metallic region of the phase diagram. Recent high field measurements [32] show a substantial variation of the Hall coefficient over doping levels \( p \) between 0.16 and 0.205, consistent with the scenario of a quantum critical point (QCP) near \( p = 0.19 \). Furthermore, the Hall coefficient is positive, opposite in sign to that observed at lower dopings.

How should we think about this sign change of the Hall coefficient? Since the negative Hall coefficient for \( p < 0.16 \) is associated with CDW order [40], the authors of Ref. [32] suggest that CDW order terminates at or below \( p = 0.16 \). In this understanding, the region between 0.16 < \( p < 0.19 \) contains, at high fields, a distinct metallic phase featuring small hole pockets. A candidate state with such hole pockets is a spin density wave metal—however to date there is no indication of long range spin density wave order at these doping levels in YBCO in zero magnetic field. Whether such order is induced by the magnetic field is not currently known, and is a good target for future experimental work. Another class of proposals [27, 41–44] posit an interesting metallic state which does not break any symmetries, but nevertheless has small hole pockets violating the standard Luttinger theorem. Such a state necessarily has fractionalized excitations in addition to the Fermi pockets, and hence is known as a Fractionalized Fermi Liquid [45]. Clear evidence in support of a Fractionalized Fermi Liquid in the cuprates would be a tremendously exciting development, and is again a fascinating target for future experimental work.

However, some caution is warranted in the interpretation of the results on the Hall coefficient. One factor is the elevated temperature (\( T \sim 50K \)) of the measurements, which makes them difficult to compare to measurements in the regime of quantum oscillations at lower doping and \( T \sim 4K \). Another factor is that the conventional interpretation of the Hall coefficient, as a measure of the number and sign of charge carriers, may fail near the putative QCP. A breakdown of Fermi liquid and/or Boltzmann transport theory would naturally invalidate this conventional interpretation. However, even if Fermi liquid and Boltzmann transport theory are valid, it is not clear that the conventional interpretation is necessarily correct, as we demonstrate in this paper. Our result, along with previous work [46, 47], raises the possibility that the small pocket that exists at lower doping persists all the way to \( p = 0.19 \) (presumably along with the CDW order), but nevertheless has a sign change of its measured Hall effect. It remains to be seen if this
is what actually happens in YBCO, or if another state, such as those discussed in the previous paragraph, is realized.

We study a Fermi liquid metal in a situation where the quasiparticle effective mass varies strongly around the Fermi surface. In the context of the underdoped cuprates, precisely such a highly anisotropic effective mass was proposed [48] to account for seemingly conflicting measurements of the effective mass in underdoped cuprates. Here we study the magnetotransport properties of such a Fermi liquid metal within the standard Boltzmann framework. To be concrete, we will treat a Fermi surface similar to the model proposed by Harrison and Sebastian [49], in which a diamond-shaped electron pocket and CDW order are both present. Following [48], we will assume that there are ‘heavy’ portions near the corner of the Fermi pocket, and ‘light’ portions near the zone diagonal. We have two principal results, which are not limited to the specific form of the model chosen. The first is that there is generically a change in sign of the Hall coefficient as the ratio of the heavy to the light quasiparticle masses is increased at fixed magnetic field. The second is that when this ratio is large there exists a parametrically broad regime of magnetic fields, distinct from the familiar weak- and strong-field limits, in which simple Drude-like formulas badly mischaracterize the system. In this regime, the Hall number bears no systematic relation to the number of carriers or their charge, and the magnetoresistance is linear in the field, with a coefficient independent of disorder strength and of the effective mass.

In Ref. [48] it was further proposed that this ratio of the heavy and light effective masses diverges as a putative quantum critical point around \( p = 0.19 \) is approached. Within this proposal it follows from our results that there will generically be a sign change of the Hall effect as the doping is increased towards 0.19 even if the CDW order persists, without any fundamental change in the Fermi surface topology.

Our results should be of broader interest in the theory of metals (apart from just the cuprates). Large variations of the quasiparticle effective mass may simply occur from band structure effects (such as proximity to a van Hove singularity [50,51]), but also from fluctuations which renormalize the band structure. An example is in metals proximate to a density wave instability. There, the soft density wave fluctuations will couple strongly to fermions at “hot spots” where the ordering wave vector nests the Fermi surface, leading to an enhanced effective mass near the hot
Figure 1.1: The simple model for the Fermi surface pocket used in these notes. It consists of several circular arcs. The light segments in green have radius $k_l$ and Fermi velocity $v_l$, while the heavy segments in red have radius $k_h$ and Fermi velocity $v_h$.

spots but little effect elsewhere. Other examples are heavy fermi liquids in rare earth alloys. The strong enhancement of quasiparticle effective mass that characterizes these metals likely occurs in some portions of the Fermi surface but not in others, thereby leading to strong variations of the effective mass around the Fermi surface [52]. Our results, for instance the regime of linear magnetoresistance, are pertinent to all such metals.

The remainder of this paper is organized as follows: we introduce the model Fermi surface in section 1.1; we then review the weak- and strong-field regimes in section 1.2, and the novel intermediate asymptotic regime in section 1.3; we close with a discussion of the implications of our results for the interpretation of experiments in the underdoped cuprates.

1.1 Model of an anisotropic Fermi pocket

Figure. 1.1 shows a simplified model of a 2D electron pocket with a diamond-like shape similar to that proposed by Harrison and Sebastian. It consists of circular arcs of radius $k_h$ and $k_l$ with
Fermi velocities \( v_h \) and \( v_l \) respectively\(^1\), the subscripts \( h \) and \( l \) refer to “heavy” and “light”. For a given magnetic field \( B \), the cyclotron orbits have angular velocities \( \omega_{h,l} \equiv eBv_{h,l}/k_{h,l} \), on the heavy and light segments, and the cyclotron period is

\[
\frac{2\pi}{\omega_c} = \frac{4\pi}{\omega_h} + \frac{2\pi}{\omega_l} = \frac{2\pi h}{eB} \left( \frac{2k_h}{v_h} + \frac{k_l}{v_k} \right)
\]

(1.1)

If we introduce the effective masses \( m_{h,l} \equiv eB/\omega_{h,l} \), then the cyclotron effective mass is \( m_c \equiv eB/\omega_c = 2m_h + m_l \).

While this model is artificial, the conclusions we glean from it rest on only a few essential assumptions. The first is that the Fermi surface is an electron pocket consisting primarily of segments with hole-like curvature, an assumption common to most proposals \([49, 53]\) to explain high-field magnetotransport in YBCO. The second assumption is that the segments of electron-like curvature connecting the hole-like segments are especially subject to mass enhancement as the CDW transition is approached. This relies on the CDW terminating in a QCP (or weakly first order transition), and was previously argued \([48]\) to account for seemingly conflicting measurements of the effective mass in underdoped cuprates. We will be interested in the behavior of the model on approach to a QCP at which \( m_h/m_l \) diverges\(^2\), and we will simplify the discussion by keeping \( m_l \) fixed to equal the electron mass \( m_e \).

We employ the relaxation time approximation to Boltzmann transport. For simplicity we work at zero temperature, with an isotropic relaxation time \( \tau = 0.2\text{ps} \) \([54]\). To compute the DC conductivity we use Chambers’ formula \([55]\), which is valid for arbitrary magnetic field:

\[
\sigma_{ij} = \frac{e^2}{2\pi^2\hbar} \int \frac{dk}{|v(k)|} v_j(k) \int_0^\infty dt v_i(q(t)) e^{-t/\tau}
\]

(1.2)

\(^1\)The piecewise definition here cannot arise from a smooth dispersion relation. This entails modest singularities in magnetic field dependence as \( B \to 0 \), which are not of concern here. For the numerical calculations, we round the jump in Fermi velocity between “light” and “heavy” segments.

\(^2\)\( k_h \) would also vanish at the transition, but this effect changes none of our qualitative conclusions, so we fix \( k_h = 0.2k_l \).

\(^3\)A finite value of \( \tau \) at zero temperature derives from disorder scattering. For a fixed density of point-like scatterers, and assuming the coupling of quasiparticles to disorder is not singularly renormalized, \( \tau \) will be proportional to the inverse density of states, and therefore tend to vanish as \( m_c \) diverges. We neglect this effect here, but it may be important for the interpretation of transport experiments on approach to various QCPs.
for the model with fixed $m_l = m_e$ and various $m_h$. In the calculations, $k_h = 0.2k_l$, the pocket is fixed to have an area of 1.9\% of the Brillouin zone, and the relaxation time is $\tau = 0.2\text{ps}$. Only when the mass enhancement is minimal is the Hall coefficient negative at all fields. For large mass enhancement a sign change in the Hall coefficient occurs at a high value of the magnetic field, for instance at $\approx 170\ T$ for the case $m_c = 8.3m_e$ shown dashed in blue. When the mass enhancement is substantial, there is also a regime of large $B$-linear magnetoresistance, as described in section 1.3.

Here the $\mathbf{k}$ integral is over the Fermi surface, $\mathbf{q}(t = 0) = \mathbf{k}$, and the time evolution of $\mathbf{q}$ is given by the Lorentz force law $\hbar \dot{\mathbf{q}} = -e\mathbf{v}(\mathbf{q}) \times \mathbf{B} = -eB\mathbf{v}(\mathbf{q}) \times \hat{z}$.

1.2 “FAMILIAR” REGIMES

1.2.1 WEAK FIELD: $\omega_l \tau, \omega_h \tau \ll 1$

In the weak field regime, a quasiparticle can only travel a small fraction of a Fermi surface segment before decaying, so (1.2) can be rewritten as a single integral over the Fermi surface. We will be concerned with the Hall conductivity, which in this limit is conveniently expressed using Ong’s “geometric” formula [56]:

$$\sigma_{xy} = \frac{2e^3}{\hbar^2} \cdot A_l,$$

where $A_l$ is the signed area swept out by the mean free path vector $\mathbf{l} \equiv \mathbf{v}_F \tau$ as the Fermi surface is traversed. For our model, the mean free path sweeps out one circle of radius $v_l \tau$, (from the
Figure 1.3: The dependence of the Hall coefficient on approach to a QCP at which $m_h$ diverges, with $m_l = m_e$ held fixed. Other parameters are as in Figure 1.2. Data are shown at various values of the magnetic field, showing that the sign change always precedes the QCP.

Far from the QCP, $v_h \approx v_l$ and the Hall conductivity is negative, as expected for an electron-like pocket. As the QCP is approached, $v_h$ is reduced, eventually yielding a positive Hall conductivity despite the fixed electron-like topology of the pocket.

1.2.2 Strong field: $\omega_l \tau, \omega_h \tau \gg 1$

In the strong field regime, a quasiparticle executes numerous cyclotron orbits before decaying, and an expansion of the conductivity in powers of $1/B$ can be obtained by Taylor expanding the exponential of Eq.(1.2). Standard manipulations relate the the conductivity to the area of the Fermi surface as

$$\sigma_{ij} = -\frac{eA}{2\pi^2 B} \epsilon_{ij} + O(1/B^2)$$

(1.5)
Here $\epsilon$ is the Levi-Civita symbol and $A$ is the signed area of the pocket(s), defined so that an electron (hole) pocket has positive (negative) area. The Hall conductivity goes like $1/B$ at large field, while the closure of the Fermi surface prohibits a $1/B$ term in the longitudinal conductivity, which goes as $1/B^2$. Accordingly, the Hall coefficient obtains its familiar classical expression

$$R_H = \frac{\rho_{yx}}{B} \approx \frac{1}{B\sigma_{xy}} \approx -\frac{1}{ne},$$

(1.6)

where $n = 2A/(2\pi)^2$ is the signed number density of carriers. Meanwhile, the longitudinal resistivity saturates: $\rho_{xx} \approx \sigma_{xx}/\sigma_{xy}^2 \sim B^0$

1.3 “Fake” High Field Regime: $\omega_h \tau \ll 1 \ll \omega_l \tau$

In a typical metal, the only asymptotic regimes of magnetotransport are the weak- and strong-field regimes described above, with a crossover between them when the cyclotron period is of order $\tau$. In the presence of large mass anisotropy, an additional asymptotic regime exists, in which a quasiparticle does not complete a full cyclotron orbit before decaying, but rapidly traverses parts of the Fermi surface of low effective mass. In our model, transport coefficients in this regime can be expressed as a double expansion in $(\omega_l \tau)^{-1}$ and in $\omega_h \tau$. The zeroth order term in this expansion involves only the geometric properties of the light segments, so that $k_l$ is the only model parameter that enters.

To obtain this zeroth order term, we set the exponential damping factor in Eq. (1.2) to unity when $q(t)$ lies in a light segment, and zero when it lies in a heavy segment. We rewrite the integration measure $dt = dq/|q| = h dq/(e|v(q)|B)$ and write the conductivity as a sum over light segments:

$$\sigma_{ij} = \frac{e}{2\pi^2 B} \sum_{\text{light segments}} \int_{k_0^\alpha}^{k_1^\alpha} \frac{dk}{|v(k)|} v_j(k) \cdot \int_k^{k_1^\alpha} dq/|v(q)| v_i(q),$$

(1.7)

where the light segment $\alpha$ begins at $k_0^\alpha$ and ends at $k_1^\alpha$. Evidently, both the longitudinal and Hall conductivity are proportional to $1/B$ at this level of approximation. For the model in ques-
The Hall coefficient and longitudinal resistivity are

\[
R_H = \frac{2\pi^2}{ek_f^2} \left( \frac{\pi - 2}{(\pi - 2)^2 + 4} \right) \left[ 1 + \mathcal{O} \left( \omega_h \tau, (\omega_l \tau)^{-1} \right) \right]
\]

\[
\rho_{xx} = \frac{4\pi^2 B}{ek_f^2} \left( \frac{1}{(\pi - 2)^2 + 4} \right) \left[ 1 + \mathcal{O} \left( \omega_h \tau, (\omega_l \tau)^{-1} \right) \right]
\]

(1.8)

The Hall coefficient remains positive, and is about 0.68 times its weak field value, while the longitudinal resistivity exhibits an unusual $B$-linear magnetoresistance. The latter phenomenon has been shown to arise near density-wave QCPs when mass enhancement effects are neglected [57, 58], due to the sharp curvature of the Fermi surface near the hot spots. $B$-linear magnetoresistance arises in our context in a regime of higher magnetic field, and from a mechanism in which Fermi surface curvature plays no direct role\(^4\). Unlike the curvature effect, it is also apparent for current flowing perpendicular to the two-dimensional plane treated in this work.

1.4 Discussion

A wealth of experimental evidence points to QCP near $p = 0.19$ in YBCO [59], one which is likely relevant for superconductivity [60] and for the strange metal regime that prevails at temperatures above $T_c$. At high magnetic fields and lower dopings, $0.08 < p < 0.16$, a metallic state with a small electron pocket [39] and CDW order is well established. The high field measurements of Ref. [32] cover the doping region between this metal and the QCP, and show strong doping dependence of the Hall number, as well as a change of its sign relative to that at $p < 0.16$.

The strong doping dependence of the Hall number is striking, but it does not obviously constrain theories of the QCP. A cusp-like singularity in the Hall number is predicted, within Boltzmann theory, in theories of a variety of order parameter transitions, including $d$-density wave [61], spin-density wave [62–64], and nematicity [65], among others. A transition between a Fermi liquid and fractionalized Fermi liquid (FL\(^*\)) state would be expected to feature a discontinuous jump in the Hall number, but this would be inevitably rounded by finite temperature.

The sign change of the Hall number has been interpreted by the authors of [32] to rule out

\(^4\)Furthermore, B-linear magnetoresistance due to curvature is absent if the Fermi velocity at the turning points is sufficiently small, as is the case in our study
Figure 1.4: A schematic illustration of the effects of finite temperature on the Hall coefficient under the scenario explored in this paper. A small electron pocket persists until a critical doping $p_c$, with the Hall coefficient changing sign well below $p_c$ due to mass renormalization and curvature effects. At zero temperature, there is a sharp jump of the Hall coefficient at $p_c$, where the electron pocket transforms into the large Fermi surface. At finite temperature, this jump is rounded, leaving a maximum in the Hall coefficient at a doping $p_{max} < p_c$. CDW order at dopings above $p = 0.16$. However, our calculations show that a sign change of the Hall number precedes the loss of CDW order under reasonable assumptions about Fermi surface anisotropy. Therefore, it is premature to rule out CDW order in the region $0.16 < p < 0.19$. High field measurements directly sensitive to charge order would clearly be useful to better understand the phenomenology in this doping regime.

Though our calculations rationalize the sign change without postulating an additional phase in the region $0.16 < p < 0.19$, the rapid rise of the Hall number with doping in this region requires a different explanation. It is natural to postulate that this is tied to thermal crossover physics around the QCP near $p_c = 0.19$, but we do not attempt an analysis of that QCP here. However it is plausible that $R_H$ has a sharp drop as the doping is increased through $p_c$ at very low $T$. At higher $T$, this will then lead to a rounded peak in $R_H$, which when combined with our calculations at lower $p$, can lead to the observed behavior. A schematic plot of the Hall coefficient
versus doping in this scenario is shown in Fig. 1.4.

The results of [32] underscore the centrality of a QCP near $p = 0.19$ to the phenomenology of YBCO, but evidently offer minimal phenomenological constraints on theories of that QCP. That said, such phenomenological constraints do exist in the literature. Some of these have been discussed in [48], where various general possibilities for the $T = 0$ evolution from the overdoped to the underdoped metal were described. The absence of an elastic peak near $(\pi, \pi)$ in neutron scattering renders a spin-density wave or d-density wave transition less likely, though neutron measurements in high magnetic field would be necessary to definitively rule out such proposals. Also, the presence of anti-nodal electron pockets just below a continuous density wave transition precludes the identification of $p = 0.19$ with the opening of the pseudogap, which is conventionally understood to be an anti-nodal phenomenon. At lower dopings, close to $p = 0.10$, the presence of both nodal and anti-nodal pockets would likely result in electronic specific heat in excess of that measured [66].

These problems are mitigated, though not solved entirely, by the scenario of an FL to FL$^*$ transition\footnote{Case B of Ref. [48]. The comments below also pertain to Cases C and D of that paper.}. Such a transition can be accompanied by a discontinuous jump in the size of the Fermi surface even when the transition is continuous, yielding an anti-nodal gap, and lower specific heat than the continuous onset of an order parameter. However, it is not known whether the FL to FL$^*$ transition can be continuous, and novel experiments will be necessary to establish fractionalization in this doping regime. Evidently, the results of [32] and the analysis pursued in this paper leave considerable uncertainty in the phenomenological constraints on a QCP in the near-optimally doped cuprates.

Moving beyond the cuprates, our calculations point out a previously unappreciated regime of magnetotransport, with a novel orbital mechanism for $B$-linear magnetoresistance. That phenomenon is observed in a wide variety of correlated metals [67–70], but often inadequately understood. Our results motivate a more detailed experimental account of the anisotropy of mass enhancement in correlated metals.
2

Origin of Mott insulating behavior and superconductivity in twisted bilayer graphene

2.1 Introduction

Superconductivity occurs proximate to a Mott insulator in a few materials. The most famous are the cuprate high-$T_c$ materials [71]; others include layered organic materials [72], certain fullerene superconductors [73], and some iron-based superconductors [74]. In these systems, there is a complex and often poorly understood relationship between the Mott insulator and the superconductor, which has spurred tremendous research activity in condensed matter physics in the last 30 years. Very recently, in some remarkable developments, both Mott insulating behavior
and proximate superconductivity have been observed in a very different platform: two layers of graphene that are rotated by a small angle relative to each other [33, 34].

Twisted bilayer graphene (TBG) structures have been studied intensely in the last few years [75–86]. The charge density is concentrated on a moiré pattern which forms (at least approximately) a triangular lattice [76, 77, 79, 82, 83]. The electronic states near each valley of each graphene monolayer hybridize with the corresponding states from the other monolayer. When the twisting angle is close to certain discrete values known as the magic angles, theoretical calculations show that there are two nearly flat bands (per valley per spin) that form in the middle of the full spectrum that are separated from other bands. [80] When the carrier density is such that the chemical potential lies within these nearly flat bands, interaction effects are expected to be enhanced. At a filling of 1/4 or 3/4 (denoted $\nu = -2$ and +2 respectively with full band filling denoted $\nu = +4$) of these nearly flat bands, Ref. [33] reported insulating behavior at very low temperatures. At such fillings band insulation is forbidden, which leads naturally to the expectation that these are correlation-driven (Mott) insulators. Doping the Mott insulator at 1/4 band filling - with either electrons or holes - reveals superconductivity at low $T$ [34].

A number of other striking observations have been made in Refs. [33, 34] about both the Mott insulator and the superconductor from transport studies in a magnetic field. The Mott insulation is suppressed through the Zeeman coupling of the magnetic field at a low scale $\approx 5T$ - roughly the same scale as the activation gap inferred from zero field resistivity. Quantum oscillations are seen in the hole doped state with a frequency set (in the hole doped side) by the density deviation from the Mott insulator. The degeneracy of the corresponding Landau levels is half of what might be expected from the spin and valley degrees of freedom that characterize electrons in graphene. The superconductivity occurs at temperatures that are high given the low density of charge carriers. Just like in other doped Mott insulators, there is a dome of superconductivity with $T_c$ reaching an “optimal” value at a finite doping. The superconductivity is readily suppressed in accessible magnetic fields - both perpendicular and parallel to the plane.

The observation of these classic phenomena in graphene gives new hope for theoretical progress in addressing old questions on Mott physics and its relationship to superconductivity. They also raise a number of questions. What is the nature of the insulators seen at these fractional fillings?
How are they related to the observed superconductivity? On the theoretical side, what is an appropriate model that captures the essential physics of this system?

In this paper we make a start on addressing these questions. The two nearly flat bands for each valley found in the band structure have Dirac crossings at the moiré K points (but not Γ). We argue that these Dirac crossings are protected by symmetries of the TBG system. We show that this precludes finding a real space representation of the nearly flat bands in terms of Wannier orbitals localized at the triangular moiré sites, in contrast to natural expectations. Thus, a suitable real space lattice model is necessarily different from a correlated triangular lattice model with two orbitals (corresponding to the two valleys) per site. We instead show that a representation that is faithful to the Dirac crossings is possible on a honeycomb lattice with two orbitals per site, but even this has some subtleties. First, one cannot implement a natural representation of all the important symmetries in the problem, which include spatial symmetries, time-reversal, and a separate conservation of electrons of each valley (which we dub \(U_{v}(1)\)). Second, since the charge density is concentrated at the moiré triangular sites (which appear as the centers of the honeycomb plaquettes), the dominant interaction is not an on-site Coulomb repulsion on the honeycomb sites. Rather it is a ‘cluster charging energy’ that favors having a fixed number of electrons in each honeycomb plaquette. This makes this model potentially rather different from more standard Hubbard models with on-site interactions.

Armed with this understanding of the microscopics, we can begin to address the experimental phenomenology. We propose that this system spontaneously breaks the valley \(U_{v}(1)\) symmetry - we call the resulting order as “Inter-Valley Coherent” (IVC). We discuss microscopic mechanisms that stabilize IVC symmetry breaking. We point out that even when the IVC is fully polarized it cannot, by itself, lead to a fully insulating state, but rather leads to a Dirac semi-metal. The development of a true insulator needs a further symmetry breaking (or some more exotic mechanism) to gap out the Dirac points. We show that once the valley symmetry is spontaneously broken, the physics at lower energy scales can be straightforwardly formulated in terms of a real space honeycomb lattice tight-binding model with a dominant cluster charging interaction, and other weaker interactions. We outline a number of different possible routes in which a true insu-
lator\(^1\) can be obtained in such a IVC ordered system. A concrete example is a state that further breaks \(C_3\) rotational symmetry. We show how doping this specific IVC insulator can explain the phenomenology of the experiments. We present a possible pairing mechanism due to an attractive interaction mediated by Goldstone fluctuations of the IVC phase. We describe and contrast features of other distinct routes by which the IVC state can become a true insulator at \(\nu = \pm 2\). We propose a number of future experiments that can distinguish between the different routes through which a IVC can become a true insulator.

2.2 \textbf{Electronic Structure of Twisted Bilayer Graphene: General Considerations}

2.2.1 Setup

First, to establish notation, let us consider a graphene monolayer, with lattice vectors \(\mathbf{A}_1\) and \(\mathbf{A}_2\) (see Appendix B.1 for details). The honeycomb lattice sites are located at \(r_{1,2} = \frac{1}{2}(\mathbf{A}_1 + \mathbf{A}_2) \mp \frac{1}{6}(\mathbf{A}_1 - \mathbf{A}_2)\), where the \(-\) and \(+\) signs are respectively for the sites labeled by 1 and 2.

Now consider the moiré pattern generated in the twisted bilayer problem. For concreteness, imagine we begin with a pair of perfectly aligned graphene sheets, and consider twisting the top layer by an angle \(\theta\) relative to the bottom one. Now we have two pairs of reciprocal lattice vectors, the original ones \(\mathbf{B}_a\) and \(\mathbf{B}_a' = R_\theta \mathbf{B}_a\). Like references \([75, 80]\) we approximate the moiré superlattice by the relative wavevectors, leading to a periodic structure with reciprocal lattice vectors \(\mathbf{b}_a = \mathbf{B}_a' - \mathbf{B}_a = (R_\theta - I)\mathbf{B}_a\). For small \(\theta\) we can approximate this by \(\mathbf{b}_a = \theta \hat{z} \times \mathbf{B}_a\). Thus, the moiré pattern also has triangular lattice symmetry, but it is rotated by 90\(^\circ\) and has a much larger lattice constant. Note, in this dominant harmonic approximation, questions of commensuration/incommensuration are avoided since no comparison is made between the other sets of harmonics of the moiré superlattice that are not commensurate to the dominant one.

\(^1\)In the literature a distinction is often made between Mott and Slater insulators. The former is used to denote situations in which there is a wide intermediate energy regime with insulating charge transport and well developed localized magnetic moments. In the twisted bilayer system it is not clear that such a regime exists. However it is clear that the insulating behavior is due to electron-electron interaction effects. Thus to avoid potential confusion it may be safer to refer to the insulator as a ‘correlated insulator’ to emphasize the importance of interaction effects without wading into the Mott vs Slater debate.
Let us now briefly review the low energy electronic structure of monolayer graphene to set the notation. Parameterizing \( k = \sum_{j=1,2} g_j B_j \) for \( g_j \in (-1/2, 1/2] \), the Bloch Hamiltonian for the nearest neighbor hopping model is:

\[
H(g) = t(e^{-i\frac{2\pi}{3}(g_1-g_2)} + e^{-i\frac{2\pi}{3}(g_1+2g_2)} + e^{i\frac{2\pi}{3}(2g_1+g_2)}) \sigma^- + \text{h.c.}
\]  

(2.1)

Note that, as a general property of our present choice of Fourier transform, the Bloch Hamiltonian is not manifestly periodic in the Brillouin zone. Rather, for any reciprocal lattice vector \( B \), we have \( H(k + B) = \eta_B H(k) \eta_B^\dagger \), where \( \eta_B = \text{diag}(e^{-iB \cdot r_a}) \). One can now pass to a continuum limit near each Dirac points \( K = (2B_1 - B_2)/3 \) and \(-K\). We then have the linearized Hamiltonian

\[
H(K + k) = -\hbar v_F k \cdot \sigma ;
\]

\[
H(-K + k) = \hbar v_F k \cdot \sigma^* ,
\]

(2.2)

where \( \hbar v_F = \sqrt{3}t a/2 \) in our simple nearest-neighbor model. Since \( H(q) \) is not periodic in the BZ, expanding about the other equivalent Dirac points will lead to a slightly modified form of the Hamiltonian (due to conjugation by some \( \eta \)). In second quantized notation we can write the continuum Hamiltonian:

\[
\hat{h}_+ = -\hbar v_F \int d^2k \tilde{\psi}_{+;k}^\dagger (k \cdot \sigma) \tilde{\psi}_{+;k} ;
\]

\[
\hat{h}_- = +\hbar v_F \int d^2k' \tilde{\psi}_{-;k'}^\dagger (k' \cdot \sigma^*) \tilde{\psi}_{-;k'} ,
\]

(2.3)

where the momentum integration is understood to be implemented near the Dirac point momentum by introducing a cutoff \(|k| \leq \Lambda\) and \( d^2k = \frac{dk_x dk_y}{(2\pi)^2} \). The symmetry implementation on the continuum fields is tabulated in Appendix B.1. For example, \( C_3 \)-rotation symmetry is represented as \( \hat{C}_3 \tilde{\psi}_{+;k} \hat{C}_3^{-1} = e^{\pm i\frac{2\pi}{3}\sigma_3} \tilde{\psi}_{+;C_3k} \), where \( \mu = t, b \) is the layer index.

Next, we couple the degrees of freedom in the two layers of graphene and arrive at a continuum theory for the twisted bilayer graphene system [75,80]. First, we note that the rotated Bloch Hamiltonian of a monolayer can then be identified as \( H_\varphi(k) = H(R_\varphi k) \). Linearizing about the
rotated K point $R_\varphi \mathbf{K}$, obtains $\hat{h}_\pm(\varphi)$, with $\hat{h}_\pm(\varphi)$ defined by replacing $\sigma \mapsto \sigma_\varphi$ in $\hat{h}_\pm$, where
\[
\sigma_\varphi \equiv e^{-i\varphi \sigma_3/2} \sigma e^{i\varphi \sigma_3/2}.
\] (2.4)

Focusing on a single valley, say $\mathbf{K}$, The continuum theory [75, 80] of the twisted bilayer graphene system is described by the Hamiltonian $\hat{H}_{\text{Cont.}} = \hat{H}_{\text{Dirac}} + \hat{H}_T$, where

\[
\hat{H}_{\text{Dirac}} = \hat{h}_+(\varphi_t) + \hat{h}_+(\varphi_b);
\]
\[
\hat{H}_T = \int_0^\Lambda d^2k \hat{\psi}_+^{\dagger, k} T_{\mathbf{q}_1} \hat{\psi}_{\pm t, k+\mathbf{q}_1} + \text{h.c.}
\] + symmetry related terms,

where $t$ and $b$ respectively denote the top and bottom layers, and we set $\varphi_t = \theta/2$ and $\varphi_b = -\theta/2$.

Here, we have introduced $\mathbf{q}_1 \equiv R_{-\theta/2} \mathbf{K} - R_{\theta/2} \mathbf{K}$, which characterizes the momentum transfer between the electronic degrees of freedom of the two layers [80]. Assuming $T_{\mathbf{q}_1}$ is real, the symmetries of the system, which we will discuss in the following subsection, constrain $T_{\mathbf{q}_1}$ [87] to take the form $T_{\mathbf{q}_1} = w_0 - w_1 \sigma_1$, where $w_{0,1}$ are real parameters \(^2\). Similarly, one can generate the omitted symmetry-related terms by applying symmetries on $T_{\mathbf{q}_1}$.

2.2.2 Symmetries of the Continuum Theory

Let us discuss how the symmetries of the graphene monolayer are modified in the twisted bilayer problem within the dominant harmonic approximation. We will see that in addition to the moiré translation symmetry, we have $C_6$ rotation, time reversal and a mirror symmetry. Furthermore, a U(1) valley symmetry that allows us to assign valley charge to the electrons emerges in the low-energy limit. The generator of $C_6$ rotation and time reversal will flip the valley charge, while reflection leaves it invariant.

Microscopically, the stacking pattern of the two layers can be specified as follows [80, 87, 88]: first, we align the two layers perfectly in a site-on-site manner, corresponding to the “AA stack-
ing” pattern, and then rotate the top and bottom layers about a hexagon center by angles $\theta/2$ and $-\theta/2$ clockwise respectively; second, we shift the top layer by a vector $d$ parallel to the plane. For generic values of $\theta$ and $d$ one expects that almost all of the spatial symmetries are broken.

However, within the dominant harmonic approximation it was found that, on top of possessing moiré lattice translation symmetries, the effective theory is also insensitive to $d$ [80]. This implies that, given $\theta$, the effective theory will at least possess all the exact symmetries for any choice of $d$. A particularly convenient choice is when we take $d = 0$. In this case, we can infer all the point-group symmetries of the system by focusing on the center of the hexagons (Fig. 2.1a). Aside from the rotational symmetries generated by the six-fold rotation $C_6$, we see that there is an additional mirror plane $M_y$, which, in fact, combines a mirror perpendicular to the 2D plane together with an in-plane mirror which flips the top and bottom layers. Strictly speaking, this leads to a two-fold rotation in 3D space, but when restricting our attention to a 2D system it acts as a mirror.

To summarize, the effective theory will at least have the following spatial symmetries: lattice translations, a six-fold rotation and a mirror. This allows one to uniquely identify its wallpaper group (i.e., 2D space group) as $p6mm$ (numbered 17 in Ref. [89]). Having identified the symmetries of the system, one can derive the model following a phenomenological approach by systematically incorporating all symmetry-allowed terms with some cutoff [87]. We have tabulated the explicit symmetry transformation of the electron operators in Appendix B.1.

In the effective theory the degrees of freedom arising form the microscopic K and K’ points are also essentially decoupled [75, 80]. This is because, for a small twist angle satisfying $|\sin \theta| \ll 1$, we have $|b_a| \ll |K|$, and therefore the coupling between the K and K’ points is a very high-order process. Hence, on top of the usual electron-charge conservation, the effective theory has an additional, emergent $U_v(1)$ conservation corresponding to the independent conservation of charge in the two valleys K and K’. Henceforth, we will refer to this as “valley conservation.” The valley charge operator is given by:

$$\hat{I}_z = \int d^2k \left( \hat{\psi}^{\dagger}_{+;k} \hat{\psi}_{+;k} - \hat{\psi}^{\dagger}_{-;k} \hat{\psi}_{-;k} \right)$$

(2.6)
Table 2.1: **Summary of key effective symmetries.** From top to bottom, the listed symmetries are time-reversal, moiré lattice translation, a perpendicular 2D mirror, three-fold rotation, combined symmetry of two-fold rotation and time-reversal, and valley $U_v(1)$ conservation. For any symmetry $g$, it either commutes ($\eta_g = +1$) or anticommutes ($\eta_g = -1$) with the valley charge operator $\hat{I}_z$.

<table>
<thead>
<tr>
<th>Symmetry</th>
<th>$\eta_g$</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{T}$</td>
<td>-1</td>
<td>Broken by valley polarization $\langle I_z \rangle \neq 0$</td>
</tr>
<tr>
<td>$t_a$</td>
<td>+1</td>
<td>–</td>
</tr>
<tr>
<td>$M_y$</td>
<td>+1</td>
<td>Broken by perpendicular electric field</td>
</tr>
<tr>
<td>$C_3$</td>
<td>+1</td>
<td>Pins Dirac points to $K_M$ and $K'_M$</td>
</tr>
<tr>
<td>$C_2\mathcal{T}$</td>
<td>+1</td>
<td>Protects the local stability of the Dirac points</td>
</tr>
<tr>
<td>$\exp(-i\theta \hat{I}_z)$</td>
<td>+1</td>
<td>–</td>
</tr>
</tbody>
</table>

Note that, as time-reversal $\mathcal{T}$ interchanges the K and K’ valleys, it is not a symmetry of a single valley. Similarly, $C_6$ also interchanges the two valleys, thus:

\[
\mathcal{T} \hat{I}_z \mathcal{T}^{-1} = -\hat{I}_z \quad (2.7) \\
\hat{C}_6 \hat{I}_z \hat{C}_6^{-1} = -\hat{I}_z \quad (2.8) \\
M_y \hat{I}_z M_y^{-1} = +\hat{I}_z \quad (2.9)
\]

We then see that their combined symmetry, $C_6\mathcal{T}$, is a symmetry in the single-valley problem.

In fact, one can check that the symmetries of single-valley problem is described by the magnetic space group 183.188 (BNS notation; Ref. [90]). We have tabulated the generating symmetries in Table 2.1.

### 2.3 Low Energy Theory: Two Band Projection

Formally, the continuum effective theory [75, 80] we described corresponds to an infinite-band problem for each valley. However, near charge neutrality it was found that, for some range of angles, the moiré potential can induce additional band gaps at certain commensurate filling of the moiré unit cell. The “nearly flat bands” identified near the magic angle correspond to two
Figure 2.1: Effective symmetries and constraints on band structures. (a) The effective symmetries of the twisted bilayer graphene system can be inferred by inspecting the point-group symmetry of a hexagon center in the real space, taken to be the rotation axis of the layers. (b) Schematic band structure along a high-symmetry path in the moiré Brillouin zone. (c-f) Effect of symmetry breaking. (c) Breaking the $C_2$ rotation will gap out the Dirac points. (d) An external perpendicular electric field breaks the mirror $M_y$ symmetry, which only modifies the energetics but cannot open a band gap at charge neutrality. (e) When $C_3$ rotation is broken, but the combined symmetry of two-fold rotation $C_2$ and time-reversal $T$ is preserved, the Dirac points remain protected, although unpinned from $K_M$ and $K'_M$. (f) When valley conservation $U_v(1)$ symmetry is broken, one can no longer label the bands using their valley index. The gaplessness at charge neutrality is no-longer symmetry-required, although, depending on detailed energeticss, there can still be remnant Dirac points. In contrast, at quarter filling relevant for the observed Mott physics, there are necessarily Dirac points present in this case. The other symmetry breaking patterns listed above also do not open band gaps at quarter filling.

bands per valley, separated from all other bands by band gaps, that form Dirac points at the $K_M$ and $K'_M$ points in the moiré BZ. These bands correspond to the relevant degrees of freedom for the correlated states observed in Refs. [33, 34], and in the following we will focus our attention to the properties of these bands. In this section, we will always focus on a single valley, say that corresponding to the K point in the microscopic description.
2.3.1 Symmetry-enforced Band Contacts

A salient feature of the effective theory is the presence of Dirac points at charge neutrality, whose velocity is strongly renormalized and approaches zero near the magic angle \([75,80]\). The stability of the Dirac points can be understood from symmetries: for a single valley, \(K_M\) is symmetric under the (magnetic) point-group generated by \(C_6 T\). In particular, \((C_6 T)^2 = C_3\), and therefore we can label each band at \(K_M\) by its \(C_3\) eigenvalue, which takes value in \(\{1, \omega = e^{-i \frac{2\pi}{3}}, \omega^*\}\).

In particular, a band with \(C_3\) eigenvalue \(\omega\) is necessarily degenerate with another with eigenvalue \(\omega^*\), as \((C_6 T)^2 \neq 1\) on these bands and enforces a Kramers-like degeneracy. The observed Dirac points at charge neutrality correspond precisely to this two-dimensional representation (Fig. 2.1b).

While we have alluded to the presence of \(C_6 T\) symmetry in explaining the stability of the Dirac points, these band contacts are actually locally stable so long as the symmetry \((C_6 T)^3 = C_2 T\) is kept. This can be reasoned by noting that \(C_2 T\) quantizes the Berry phase along any closed loop to \(0, \pi \mod 2\pi\), and a Dirac point corresponds precisely to the case of a nontrivial \(\pi\) Berry phase \([91]\).

Let us now consider the effect of breaking the various symmetries (spontaneously or explicitly) in the system. First, as \(C_2 T\) is crucial in protecting the local stability of the Dirac points, once it is broken the Dirac points can be immediately gapped out (Fig. 2.1c). However, as long as \(C_2 T\) symmetry is preserved, a small breaking of any other point-group symmetries will not lead to a gapped band structure at charge neutrality. For instance, the mirror \(M_y\) maps \(K_M\) to \(K_M'\), and its presence only ensures that the two inequivalent Dirac points are at the same energy. Therefore, even when a perpendicular electric field is externally applied such that \(M_y\) is broken, as in the setup of Refs. \([33, 34]\), it can only induce an energy difference between the two Dirac points \([75]\) (Fig. 2.1d). This should be contrasted with the case of Bernal-stacked bilayer graphene, whose quadratic band touching at charge neutrality can be gapped by an external electric field \(^3\). Alternatively, if \(C_3\) symmetry is broken the Dirac points are unpinned from \(K_M\)\(^3\). This can be understood by noting that, in the case of Bernal stacking, the system has an effective \(C_2\) symmetry which interchanges the top and bottom layer. A perpendicular electric field will therefore break this symmetry, and hence the band degeneracy at charge neutrality is lifted.

\(^3\)This can be understood by noting that, in the case of Bernal stacking, the system has an effective \(C_2\) symmetry which interchanges the top and bottom layer. A perpendicular electric field will therefore break this symmetry, and hence the band degeneracy at charge neutrality is lifted.
and $K'_M$ (Fig. 2.1e). As such, for a sufficiently strong $C_3$ breaking, a band gap might open at charge neutrality if the Dirac points could meet their oppositely charged partners and annihilate. (Though, as we will argue later, this is impossible without further symmetry breaking.) [92].

Now consider the case when valley conservation is spontaneously broken by an IVC, i.e., the valley charge $\hat{I}_z$ is no longer conserved. In this case, we should first consider the full four-band problem consisting of both valleys. At, say, $K_m$, the combined symmetry of $MpT$ ensures that the Dirac points from the two valleys are degenerate. While such degeneracy is lifted in the presence of an IVC, as long as the remaining symmetries are all intact we can only split the degeneracy according to $4 = 2 \oplus 2$ (Fig. 2.1f). This remaining two-fold degeneracy rules out an interpretation of the experimentally observed Mott insulator as a Slater insulator with a spatial-symmetry-respecting (ferro) IVC incorporated at the Hartree-Fock level. Instead, one must either introduce additional symmetry breaking, say that of $C_3$ or lattice translations, or consider an IVC which also breaks some additional spatial symmetries. We will elaborate on these points in Sec. 2.7. We also note that, an essentially identical argument holds for the case of spontaneously ferromagnetic order leading to fully spin-polarized bands of $I_z$ ordering. In this way it connects to the quarter-filled Mott insulator we will be interested in.

### 2.3.2 Triangular versus Honeycomb Lattice

A conventional route for understanding the correlated states observed in Refs. [33, 34] is to first build a real-space tight-binding model for the relevant bands, and then incorporate short-range interactions to arrive at, say, a Fermi-Hubbard model. Typically, the orbital degrees of freedom involved in the tight-binding model can be identified from either applying chemistry insight, or more systematically by studying the projected density of states for the relevant bands, both of which are inapplicable to the current moiré potential problem; furthermore, understanding on the structure of the wave-functions is required. Indeed, as is noted in Refs. [77, 88, 93], the local density states for the flat bands are well-localized to the AA regions of the moiré pattern, which form a triangular lattice. This theoretical prediction has also been confirmed experimentally [76, 79, 82]. Based on this observation, it is natural to consider a real-space model starting from effective orbitals centered at the AA sites, which corresponds to a tight-binding model de-
fined on the triangular lattice [33, 34]. In addition, by treating the two valleys separately, one envisions a model with two orbitals localized to each of the triangular sites (i.e., AA regions of the moiré pattern).

From symmetry representations, however, we can immediately rule out such a model. This can be readily inferred from the computed band structure [75, 80, 87, 88]: While the two bands are nondegenerate at Γ, as we have explained they form symmetry-protected Dirac points at K_M and K_M'. Using such pattern of degeneracies one can infer the possible symmetry representations at these high-symmetry points, and from a real-space analysis [94–96] one finds that a triangular-lattice model will always leads to the same symmetry representation at all three of the high-symmetry points, i.e., they are either all nondegenerate, or are all Dirac points. This is inconsistent with the observed pattern of degeneracies, which rules out all triangular-lattice models.

In fact, the degeneracy pattern described is familiar—it corresponds exactly to the monolayer graphene problem. One can further check that this is the only possible solution using the methods described in Refs. [94, 96]. Symmetry-wise, this implies that any tight-binding model must correspond to orbitals forming a honeycomb lattice. To reconcile with the predicted and observed local density of states [76, 77, 79, 82, 88], however, these orbitals must have nontrivial shapes: although each orbital is centered at a honeycomb site, which corresponds to the AB/BA region of the moiré pattern, the weight of the orbitals are mainly localized to the AA sites. Therefore, we expect the shape of the orbitals to resemble a (three-lobed) fidget spinner.

2.3.3 Obstructions to Symmetric Wannier States

Our symmetry analysis suggests that one should model the system by orbitals centered at the AB/BA regions of the moiré potential, which form a honeycomb lattice. A minimal tight-binding model of a single valley would then be

\[
\hat{H}_{\text{Minimal}} = \sum_{\rho_\ell} t_{\rho_\ell} e^{i\phi_{\rho_\ell}} \hat{c}_{\rho_\ell}^\dagger \hat{c}_{\rho_\ell} + \text{h.c.,}
\]  

(2.10)
where \( \hat{c}^\dagger_r \) is an electron creation operator centered at a honeycomb site (for a single valley), and \( \rho_i \) connects two \( i \)-th nearest neighbor sites. Given that this describes a single valley which breaks time reversal symmetry, the hoppings are in general complex unless constrained by a space-group symmetry.

A pedestrian approach would involve optimizing the parameters \( \{ t_{\rho_i}, \phi_{\rho_i} \} \) to reproduce the energy eigenvalues obtained from the continuum description. Would this be a good starting point for building up a real-space effective model upon which we can incorporate interaction terms? Contrary to usual expectations, we will argue that such an approach has a serious flaw in capturing certain essential properties. Specifically, we will show that while the energy eigenvalues may be well-approximated, the topology of the resulting Bloch wave-functions will necessarily be incorrect. This has important dynamical consequences, relating to the stability of band contacts under different symmetry assumptions, which in turn dictate whether an insulator will result at particular fillings. In particular, we found two symmetry obstructions to deriving a single-valley tight-binding model. The first concerns the symmetry representations of \( M_y \): we found that the two bands have opposite \( M_y \) eigenvalues of \( \pm 1 \), whereas, from a real-space analysis \([94–96]\), one can show that the two bands in a tight-binding model must have the same \( M_y \) eigenvalue.

There is a second, more serious, obstruction: aside from a quantized Berry phase of \( \pi \) for any closed loop encircling a single Dirac point, one can further define a \( \mathbb{Z} \)-valued winding number \([92]\). In contrast to the conventional case of graphene, the two inequivalent Dirac points in the single-valley model are known to have the same winding number \([33,92,97]\). As the net winding number of the Dirac points arising in any two-band tight-binding model would necessarily be zero, we can then conclude that there is an obstruction for a symmetric real-space description, i.e., there is an obstruction for constructing localized Wannier functions that reproduces just the two bands of interest, represents \( C_2T \) naturally, and preserves valley quantum numbers. A more detailed description of this obstruction, by relating it to the anomalous surface state of a three dimensional topological phase, is contained in the Appendix B.2. Essentially, this argument invokes three key ingredients: (i) a two band model and (ii) \( C_2T \) symmetry and (iii) net winding of the Dirac points in the Brillouin zone.
2.4 Mirror and chirality

Previously we have pointed out the existence of two obstructions to constructing exponentially localized Wannier functions for the two nearly flat bands of a single valley and spin in TBG [98]:

1. Mirror-eigenvalue obstruction: The mirror $M_y$ eigenvalues at $M$ (or $\Gamma$) is $\pm 1$.

2. Chirality obstruction: The entire Brillouin zone for a single valley has a nonzero net chirality. More precisely, the two Dirac points of the single-valley band structure have the same chirality.

In this section, we will elaborate on the relation between these two Wannier obstructions. It turns out that these two obstructions are equivalent in the context of TBG, as long as the system preserves the $M_y$ symmetry. When these obstructions are initially realized on a setting that has the $M_y$ symmetry, the chirality obstruction remains even if the $M_y$ symmetry is broken later, because the chirality is a discrete object that should not change upon breaking the mirror symmetry (at least weakly).

This observation is significant not only conceptually, but also practically. This is because there is always an intrinsic phase ambiguity associated with the Bloch wave function of a band structure, and it requires a smooth choice of the Bloch wave functions across the mBZ to determine the chirality. However, it takes some efforts to obtain such a smooth basis of Bloch wave functions. The above observation then greatly simplifies the problem of checking the chirality obstruction in TBG, since now one only has to check the mirror eigenvalues at high symmetry points in an $M_y$ symmetric setting, which does not require choosing a smooth basis across the entire mBZ.

Below we only sketch the logic to show the above statement, and leave the details in Appendix B.3. The chirality is only contributed by the gapless points in mBZ, so we can focus on the an open region of the mBZ that covers the gapless points. Unless very close to the magic angle, the only gapless points are the $K$ and $K'$ points. Upon approaching the magic angle, assuming the nearly-flat bands do not touch other bands, the net chirality will not change compared to the case before these satellite Dirac nodes appear, so we can always obtain the net chirality by look-
ing at the Dirac nodes at \( K \) and \( K' \).

In appendix B.3, we will show that there exists a smooth basis of Bloch wave functions so that the action of \( C_2T \) is

\[
\psi(k) \rightarrow \sigma_x K \psi(k)
\]  

(2.11)

where \( \psi(k) \) is a two-component operator that annihilates an electron at momentum \( k \) in the two nearly flat bands, and \( K \) stands for complex conjugation. In this basis, the first-quantized Hamiltonian can be written as

\[
H(k) = n_0(k) + n_1(k)\sigma_x + n_2(k)\sigma_y
\]  

(2.12)

The chirality is given by the winding of \( (n_1(k), n_2(k))^T \).

Furthermore, it is shown that in this basis the action of \( M_y \) can be chosen as

\[
\psi(k) \rightarrow \sigma_x \psi(k)
\]  

(2.13)

if the mirror eigenvalues at \( M \) are opposite, and as

\[
\psi(k) \rightarrow \eta_M \psi(k)
\]  

(2.14)

if the mirror eigenvalues at \( M \) are both \( \eta_M \). For two momenta related by \( M_y \), say, \( k \) and \( k' \), \( M_y \) requires

\[
n_1(k) = n_1(k'), \quad n_2(k) = -n_2(k')
\]  

(2.15)

if the mirror eigenvalues at \( M \) are opposite, and

\[
n_1(k) = n_1(k'), \quad n_2(k) = n_2(k')
\]  

(2.16)

if the mirror eigenvalues at \( M \) are identical.
To check the net chirality, now one can consider a small closed loop around each of $K$ and $K'$. It is straightforward to see that the windings around these two loops are the same if (2.15) holds, and they are opposite if (2.16) holds. This implies that having opposite (identical) mirror eigenvalues at $M$ is equivalent to having nonzero (zero) net chirality in mBZ.

The above claim stating that the mirror-eigenvalue obstruction and chirality obstruction are equivalent applies to any two-band system that has an odd number of mirror-related pairs of Dirac nodes. If there are an even number of mirror-related pairs of Dirac nodes, a nonzero net chirality still implies the mirror-eigenvalue obstruction. However, the converse is not true: the existence of the mirror-eigenvalue obstruction does not immediately imply a nonzero net chirality. To settle down the net chirality in this case, one can divide the entire Brillouin zone into two halves that are mirror-related, and then check the total chirality of one of the two halves.

2.5 Inter-valley coherent order: phenomenological motivation

We first describe some important clues from experiments [33,34] on the nature of both the Mott state and the superconductor. We begin with the observation that – at optimal doping – an in-plane magnetic field suppresses the superconductivity when the Zeeman energy scale is of order the zero field $T_c$. This shows that the superconductor has spin-singlet pairing. Upon hole doping the $\nu = -2$ insulator, quantum oscillations are seen with a frequency set by the density of doped holes in perpendicular $B$-fields exceeding $\approx 1T$. This tells us that the “normal” metallic state and the superconductor that emerges from it should be regarded as doped Mott insulators: the charge carriers that are available to form the normal state Fermi surface or the superconducting condensate are the doped holes. Thus the hole-doped superconductor retains information about the Mott insulator. In contrast, electron doping this Mott insulator leads very quickly to quantum oscillations with a high frequency that is set by the deviation of charge density from the charge neutrality point ($\nu = 0$). This may indicate a first order transition between a metal and Mott insulator on the electron doped side. It will be important to search for signs of hysteresis in transport experiments as the gate voltage is tuned. As the superconductor is better developed and characterized on the hole doped side, we will restrict attention to hole doping from now on.

A further important clue from the quantum oscillation data is that the Landau levels (per flux
quantum) is two-fold degenerate, whereas one would expect four-fold degeneracy coming from the spin and valley degeneracy. The doped holes have thus lost either their spin or valley quantum numbers (or some combination thereof). Losing spin makes it hard to reconcile with spin singlet pairing that can be suppressed with a Zeeman field. Thus, we propose instead that the valley quantum number is lost. The simplest option then is that the valley quantum number is frozen due to symmetry breaking, \( \langle I \rangle \neq 0 \). Here, we may define \( I \) using the electron operators \( \hat{c}(k) \) for the nearly flat band states:

\[
I = \sum_{a,b,n,\alpha,k} \hat{c}_{a\alpha}(k)\dagger \tau_{ab} \hat{c}_{b\alpha}(k),
\]

where \( a, b = \pm \) correspond to the valley index, \( \alpha \) is the spin index, \( n \) labels the two bands for each valley, and \( \tau \) denotes the standard Pauli matrices.

A non-zero expectation value for \( I^2 \) breaks time reversal symmetry. This will lead to a sharp finite temperature phase transition in 2d, and would likely have been detected in the experiments. Given the absence of any evidence of a sharp finite temperature transition we propose that the ordering is in the pseudospin \( xy \) plane. These phenomenological considerations therefore lead us to an IVC ordered state.

We note that, for IVC ordering to be useful to explain the quantum oscillations, it has to occur at a scale that is large compared to the scales set by the magnetic field. Specifically, the band splitting due to IVC ordering must be bigger than the Landau level spacing \( \approx 15 - 30K \) at the biggest fields used (of order 5\( T \)). This means that the IVC order is much more robust than the superconductivity and occurs at a higher temperature scale. We further need the IVC order to be present already in the Mott insulator, so that upon doping it can impact the quantum oscillations.

Thus, our view is that the first thing that happens as the sample is cooled from high temperature is IVC ordering. This order then sets the stage for other phenomena to occur at lower tem-

---

\(^4\)A more exotic option to explain the reduced Landau level degeneracy should also be kept in mind. Instead of losing the valley quantum number by symmetry breaking we lose it through fractionalization. For instance the electron could split into a fermion that carries its charge and spin but not the valley quantum number and a charge-0, spinless boson that carries its valley quantum number. If the boson is gapped while the fermion forms a fermi surface in the doped state we will get the reduced Landau level degeneracy. Of course such fractionalization will come hand in hand with an emergent gauge field.
perature (the Mott insulation, or the superconductivity).

2.6 Simple theory of the IVC ordered state

We now describe a mechanism that stabilizes IVC ordering, and describe the properties of the resultant state. Interestingly, to treat this stage of the problem it is sufficient to work within a momentum space formulation. This enables us to sidestep the difficulties elaborated in Sec. 2.3.3 with a real space tight-binding formulation.

Consider the nearly flat bands in the limit of strong Coulomb repulsion. Note that the dominant part of the interaction is fully $SU(4)$ invariant. We expect that the Coulomb interaction prefers an $SU(4)$ ferromagnetic state - similar to the $SU(4)$ ferromagnetism favored by Coulomb interaction in the zeroth Landau level in monolayer graphene [99–101], or in the extensive literature on flat band ferromagnetism [102]. Indeed, the difficulties with Wannier localization of the nearly flat bands also suggest that, when Coulomb interactions dominate, an $SU(4)$ ferromagnetic ground state will be favored. The band dispersion, however, is not $SU(4)$ symmetric, and hence there will be a selection of a particular direction of polarization in the $SU(4)$ space. To address this, we consider the energies of different orientations of the $SU(4)$ ferromagnet within a simple Hartree-Fock theory. Specifically we compare a spin polarized state, a pseudospin $I^z$ polarized state, and the IVC state with $I^z$ polarization.

Assume a Hamiltonian

$$H = H_0 + V$$  \hspace{1cm} (2.18)

with

$$H_0 = \sum_{\alpha a \mathbf{k}} \epsilon_{\alpha n}(\mathbf{k}) c_{\alpha n a}(\mathbf{k})^\dagger c_{\alpha n a}(\mathbf{k})$$  \hspace{1cm} (2.19)

Similarly to before, $a$ is the valley index, $\alpha$ is the spin index, and $n$ labels the two bands for each valley. The dispersion $\epsilon_{\alpha n}(\mathbf{k})$ is independent of the spin, and due to time reversal $\epsilon_{\alpha n}(\mathbf{k}) = -\epsilon_{\alpha n}(\mathbf{k})$. 

38
We assume a simple form of interaction:

\[ V = \frac{g}{N} \sum_{k_1, k_2, q} c_{\alpha a}^\dagger (k_1 + q) c_{\alpha a} (k_1) \]

\[ \cdot c_{\alpha' a' \alpha'}^\dagger (k_2 - q) c_{\alpha' a' \alpha'} (k_2) \]

(2.20)

where \( N \) is the number of \( k \)-points in the moiré Brillouin zone. Repeated indices are summed over here. This interaction actually has an \( SU(8) \) symmetry, but this is strongly broken down to \( SU(4) \) by the difference in dispersion between the two bands, and eventually down to \( U(2) \times U(2) \) by the asymmetry of the dispersion under \( k \rightarrow -k \). Each \( U(2) \) factor corresponds simply to independent \( U(1) \) charge and \( SU(2) \) spin conservation symmetries of the two valleys.

We also remark that Eq. (2.20) is overly simplified, for it does not incorporate form factors arising from the modulation of the Bloch wave-functions over the BZ when projecting onto the nearly flat bands. With such form factors included, the interaction projected onto the nearly flat bands should be written as

\[ V = \frac{g}{N} \sum_{k_1, k_2, q} \Lambda_{nn'}^a (k_1 + q, k_1) \Lambda_{mm'}^{a'} (k_2 - q, k_2) \]

\[ \cdot c_{\alpha a}^\dagger (k_1 + q) c_{\alpha a} (k_1) \cdot c_{\alpha' a' \alpha'}^\dagger (k_2 - q) c_{\alpha' a' \alpha'} (k_2) \]

(2.21)

with the form factors given by the Bloch wave functions of the states in the nearly flat bands via

\[ \Lambda_{nn'}^a (k_1, k_2) = \langle u_{an}(k_1) | u_{an'}(k_2) \rangle \]

(2.22)

where \( | u_{an}(k) \rangle \) is the Bloch wave function of a state in the nearly flat bands labelled by valley index \( a \), band index \( n \) and momentum \( k \) (it has no dependence on the spin indices). These form factors are potentially important for the present problem due to the nontrivial band topology present in the valley-resolved band structure. Our preliminary analysis suggests that the results of the Hatree-Fock calculation are modified in the ultra-flat-band limit, i.e., when the interaction term overwhelms the kinetic energy, whereas the key conclusions below are stable within a range of intermediate interaction strengths. In view of this, in the following we will first pursue the simplified Hatree-Fock theory, and leave the task of settling down the real ground state for future
(numerical) studies; it is an interesting question answering if the actual experimental systems demand a more sophisticated treatment.

The Hartree-Fock calculation is straightforward so its details will not be presented here. To summarize, we find that the IVC state has lower energy than both spin and $I_z$ polarized states. The physical reason is that, for both the spin and $I_z$ polarized states, the order parameter is conserved and hence there is a linear shift of the band when the order parameter is non-zero. In contrast, due to the $k \to -k$ dispersion anisotropy, the IVC order parameter does not commute with the Hamiltonian. IVC order thus does not simply shift the band but modifies it more significantly. Assuming a near full polarization in the Hartree-Fock Hamiltonian, the non-commutativity leads to an extra energy gain at second order in the IVC state compared to the spin polarized or $I_z$ polarized states.

Note that, in the presence of $U(2) \times U(2)$ symmetry, the spin singlet IVC state is degenerate with states that have spin triplet IVC ordering with an order parameter $I_z S$. The selection between the singlet and triplet IVC order has to occur due to other terms in the Hamiltonian that have been ignored so far. We will not attempt to pin down the details of this selection in this paper and will simply assume, as suggested by the phenomenology, that the spin singlet IVC is preferred, and discuss its consequences.

Next, we turn to a description of the properties of the IVC state. We assume that the order parameter is large, and first study its effects on the band structure. In the absence of valley ordering, at the two Dirac points there is a four-fold band degeneracy. As explained in Section 2.3.1. The valley ordering splits this four-fold degeneracy into two sets of two-fold degenerate Dirac points. When the order parameter is large, the four nearly flat bands split into two sets of two bands (Fig. 2.1f). At quarter filling, we fill the bottom most band. This, however, results in not a Mott insulator but a Dirac semi-metal. Thus, the IVC state by itself does not lead to a Mott insulator, and a further mechanism is needed. We discuss this in the next section. We note that the semimetals obtained from planar valley order versus $I_z$ order are rather different, the latter being similar to spin ordered states. Furthermore, while additionally breaking $C_3$ symmetry alone can eventually gap the Dirac points of the IVC semimetal, the same is not true of spin or $I_z$ ordered semimetals, which need further symmetry breaking due to their Dirac points.

40
carrying the same chirality.

Going beyond the mean field, the universal properties of the IVC ordered state are determined by its symmetry breaking. It will have a Goldstone mode with linear dispersion at the longest scales. Further, it will have a finite temperature BKT transition which will have weak signatures in standard experimental probes.\(^5\)

### 2.7 Inter-valley coherent Mott insulators: generalities and a concrete example

We saw that IVC ordering by itself only gives us a Dirac semi-metal and not a Mott insulator. We now consider the physics below the IVC ordering scale. First, we note that once \(U_v(1)\) symmetry is broken, there is no difficulty with writing down a real space tight-binding model for the two lowest bands. This model lives on the honeycomb lattice and must be supplemented with interactions. Naively we might imagine that the dominant interaction is an on-site Hubbard repulsion. However we know that the orbital shapes are such that the actual charge density is concentrated at the original triangular sites, \(i.e\), at the center of the hexagons of the honeycomb lattice. Now if there is an electron at a honeycomb site \(r\), its wavefunction is spread equally between the three hexagonal plaquettes that the site \(r\) is a part of. The integral of the modulus square of the wavefunction in any one such plaquette will be 1/3. The total charge that is localized at the center \(R\) of each hexagon (\(i.e\) the triangular moiré sites) is therefore

\[
Q_R = \sum_{r \in \text{hexagon}} \sum_{\alpha} \frac{n_{\alpha}(r)}{3} \tag{2.23}
\]

Now let us make the reasonable approximation that the primary Coulomb interaction is on-site on the triangular lattice sites \(R\). Then in terms of the honeycomb model the appropriate Coulomb interaction is a ‘cluster Hubbard’ term

\[
H_U = U \sum_R (Q_R - 2)^2 \tag{2.24}
\]

\(^5\)In principle, as discussed earlier the \(U_v(1)\) symmetry is only approximate and will be weakly broken by, say, small corrections to the band structures or by 3-body interaction terms. This will modify the physics very close to the phase transition.
We have also specialized to \( \nu = -2 \) when this honeycomb lattice is half-filled. This interaction penalizes charge fluctuations on each honeycomb plaquette. Thus, a suitable model Hamiltonian at scales much smaller than the intervalley coherence scale takes the form

\[
H = H_t + H_U
\]

\[
H_t = -\sum_{rr'}\sum_{\alpha} t_{rr'} c_{r\alpha}^\dagger c_{r'\alpha} + h.c
\]  

(2.25)

For the usual Hubbard model with a strong on-site repulsion, the Mott insulating state has the usual 2-sublattice Neel order. However, when the cluster charging energy is dominant, this is not obviously the case. We will therefore allow ourselves to consider a few different possibilities for the Mott insulator. Naturally, in all these options the charge gap of the insulator will be much lower than the scale of IVC ordering. In the experiments, the charge gap is estimated to be about 5\( K \). The IVC ordering should then occur at a much higher scale, consistent with what we already concluded based on the phenomenology. In this section, to be concrete, we focus on a particular Mott insulator where the \( C_3 \) rotation symmetry is spontaneously broken while preserving other symmetries (Fig. 2.2).

In passing we note that insulators driven by cluster charging have been considered in a number of different contexts before. For instance Refs. [103, 104] study models with extended Coulomb interactions as a route to access Wigner-Mott insulators. Cluster charging insulators have also been long studied [105–107] as a platform for various fractionalized insulating phases of matter. In contrast to these earlier works, where a dominant cluster charging interactions was simply postulated and the resulting physics explored, here we have identified a natural mechanism for such an interaction.

### 2.7.1 \( C_3 \) Broken Insulator

Breaking the \( C_3 \) symmetry allows gapping out the Dirac points and leads to an insulator. As the \( C_3 \) breaking order parameter increases, the two Dirac points will move towards each other (Fig. 2.1e) and eventually annihilate to produce a fully gapped insulator. This annihilation (and correspondingly the gap minimum just into the insulator) will occur either at the \( \Gamma \) or \( M \) point de-
pending on details of the dispersion. Note that, within this picture, the $C_3$ breaking also occurs at a scale bigger than the $\approx 5K$ charge gap of the Mott insulator. Clearly, the excitations above the charge gap are ordinary electrons, and their gap can be readily closed by a Zeeman field.

Upon doping this insulator, charge will enter as ordinary holes and form a small Fermi pocket. This pocket will be centered at either $\Gamma$ or $M$ depending on the location of the minimum insulating gap. In either choice, due to the absence of $C_3$ symmetry, there will just be a single such Fermi pocket which will accommodate the full density of doped holes. Due to the intervalley ordering these holes will be valley polarized in the $I^x$ direction. Naturally this explains the quantum oscillation experiments—the frequency will be set by the density of doped holes, and the Landau level degeneracy (per flux quantum) will only be two-fold (from the spin).

A natural pairing mechanism emerges from the coupling of the holes to Goldstone fluctuations of the intervalley order, as we now elaborate. In the presence of an intervalley condensate an appropriate effective action will be

\begin{align}
S &= S_0[\psi] + S_1[\psi, \theta] \\
S_0 &= \int d\tau \left( \int d^2x \bar{\psi} (\partial_\tau + \mu) \psi + \int d^2k \bar{\psi}_k h_k \psi_k \right) \\
S_1 &= \int d\tau d^2x \Phi_0 (e^{-i\theta} \bar{\psi}_+ \psi_- + c.c)
\end{align}

Here $\psi$ is a continuum electron field that represents the electrons in the low-energy nearly flat bands, $\theta$ is the phase of the intervalley condensate and $\Phi_0$ is its amplitude. Note that $h_k = \ldots$
\( \epsilon_s(k) + \epsilon_a(k) \tau^z \) is a \( 2 \times 2 \) matrix for each \( k \) point. We will allow for slow Goldstone fluctuations of the phase and obtain a convenient form of the electron-electron interaction induced by these fluctuations. To that end, we first define new fermion variables \( \chi \) through

\[
\psi = e^{i \frac{\theta z}{2}} \chi.
\] (2.29)

This removes the \( \theta \) dependence from \( S_1 \), but \( S_0 \) now takes the form

\[
S_0[\psi] = S_0[\chi] + S'_0[\chi, \theta]
\]

\[
S'_0[\chi, \theta] = \int_{x, \tau} i \partial_x \theta \bar{\chi} \tau^z \chi + \frac{1}{2} \partial_t \theta J^v_i(x)
\] (2.30)

Here \( J^v_i \) is the contribution to the \( U_v(1) \) current from the fermions. It is conveniently written down in momentum space as

\[
J^v_i(q) = \int d^2k \bar{\chi}_{k+q} \frac{\partial h_{ki}}{\partial k_i} \chi_k
\] (2.31)

Now we assume that \( \Phi_0 \) is near maximum polarization and diagonalize the \( \chi \) Hamiltonian obtained from \( S_0[\chi] + S_1[\chi] \). As discussed earlier, there are two sets of bands per spin (corresponding to \( I_x = \pm 1 \)) that are well separated from each other. The low energy electrons are those that have valley polarization \( I_x \approx 1 \). We wish to obtain the coupling of these electrons to the \( \theta \) fluctuations. For the bands with \( I_x \approx 1 \), we write

\[
\chi + \alpha = \chi - \alpha = d_\alpha
\] (2.32)

It follows that \( \bar{\chi} \tau^z \chi \approx 0 \) and similarly \( \bar{\chi} \frac{\partial \epsilon_{x}(k)}{\partial k} \tau^z \chi_k \approx 0 \). The only non-vanishing coupling there-

---

6Strictly speaking we should allow for all 4 bands per spin and work with a four component \( \psi \) and a corresponding \( 4 \times 4 \) Hamiltonian. However for the present discussion we will eventually only be interested in the modes in the vicinity of the Fermi surface after the flavor ordering. It is this sufficient to focus on the two lower bands that are split off by the flavor ordering. We therefore focus on just these two right from the start.
fore is to the contribution from $\epsilon_a(k)$. We get

$$J_i^\nu(q) \approx \int d^2k \; \bar{d}_{k+q} \frac{\partial \epsilon_a(k)}{\partial k_i} d_k \equiv \int d^2k \; v_i^a(k) \bar{d}_{k+q} d_k$$  \hspace{1cm} (2.33)

Now we assume we have integrated out the fermions everywhere except in the close vicinity of the Fermi surface. This gives a long wavelength, low frequency effective action for the $\theta$ fluctuations of the form

$$S_{\text{eff}}[\theta] = \int_{q,\omega} K \left( \frac{\omega^2}{v^2} + q^2 \right) |\theta(q, \omega)|^2$$  \hspace{1cm} (2.34)

Here $K$ is the phase stiffness of the $\theta$ field, and $v$ is the velocity of the linear dispersing $\theta$ fluctuations. We now integrate out $\theta$ to get an effective interaction between the $c$-electrons:

$$S_{\text{int}} = -\int_{q,\omega} \frac{q^2}{32K} \frac{\omega^2}{v^2} |J_i^\nu(q, \omega)|^2$$  \hspace{1cm} (2.35)

This is an attractive interaction. Anticipating that the important regime for pairing is $|\omega| \ll vq$ for an approximate treatment we set $\omega = 0$ in the prefactor to get a simplified effective interaction

$$S_{\text{int}} = -\frac{1}{32K} \int_{q,\omega} |J_i^\nu(q, \omega)|^2$$  \hspace{1cm} (2.36)

We emphasize again that - within our approximate treatment - the only contribution to $J_i^\nu$ comes from the \textit{antisymmetric} part of the “normal” state dispersion. This attractive interaction can now be treated within BCS mean field, and will lead to a superconducting state.

Note that, in real space, since the large repulsion will be on the hexagon center and not on the honeycomb site, there is no particular reason to disfavor on-site s-wave pairing. Though we will not give a detailed description of the pairing symmetry, the route to superconductivity sketched above naturally leads to a spin singlet superconductor. Further, it forms out of a ‘normal’ metal of ordinary holes through a BCS-like pairing mechanism. We expect then that Zeeman fields of order $T_c$ will efficiently suppress the superconductivity except possibly at very low doping (where
eventually phase fluctuations will kill $T_c$). At low doping, and when one is near a high symmetry point of the Brillouin zone (which is consistent with the fact that there is no additional degeneracy seen in quantum oscillations), the antisymmetric part of the dispersion is expected to constrained by symmetry to be small. For example near the $\Gamma$ point, it vanishes as the cube of the crystal momentum. This would lead to a small valley current (the derivative of the antisymmetric dispersion with respect to momentum) and hence a weakening of the coupling to valley Goldstone modes, as the doping is reduced. However, if $C_3$ rotation symmetry is broken, the antisymmetric dispersion can include a term that is linear in momentum, leading to a non-vanishing valley current at small doping.

It is also important to ask if a conventional pairing mechanism due to coupling to phonons might be operative. We note that the bandwidths of the nearly flat bands ($\approx 10\text{meV}$) are much smaller than the typical phonon energies in graphene. Thus the magic angle twisted bilayer graphene system is far from a regime in which phonon effects can be treated within the usual adiabatic approximation. Furthermore, the observation of superconductivity only in the vicinity of the correlated insulator appears unnatural within a phonon based theory. Nevertheless it is possible that phonon effects play a role in various aspects of our physics, and contribute to the effective interactions between the electrons. We will leave for future a proper treatment of the electron-phonon coupling in these systems, and focus instead here on electron-electron interaction effects which clearly must play an important role given the proximity of the superconductor to a correlation-driven Mott insulator.

2.8 Other possible Mott insulating states

The $C_3$ broken insulator is a concrete example of how an intervalley condensate of the twisted bilayer system can eventually become a Mott insulator. However, given the current experimental information, it is not clear that this is uniquely dictated. Therefore, we sketch a few different Mott insulating states and present some of their phenomenological consequences.

1. Translation broken insulator: Broken moiré translations - for instance Kekule ordering on the effective honeycomb lattice - can also gap out the Dirac points. The properties of this
state and its evolution into the doped superconductor will be similar to the $C_3$ broken insulator discussed above.

2. Antiferromagnetic insulator: This is the familiar Mott insulator of the usual honeycomb Hubbard model. Upon doping it is expected to evolve into a spin-singlet superconductor as seen in numerical studies of the $t-J$ honeycomb model [108]. The pairing symmetry appears to be $d + id$. It will be interesting to look for signatures of broken time reversal symmetry if this scenario is realized. Further, this state is known to have quantized spin and thermal Hall effects, and associated gapless edge states [109, 110]. Other properties related to this state are also discussed in the literature [111, 112].

3. Featureless Mott insulators: Given that the honeycomb lattice features two sites in the unit cell, it evades the Lieb-Shultz-Mattis theorem and allows for a featureless ground state (i.e. a gapped insulator with neither topological order nor symmetry breaking) at half filling [113–118]. Pictorially, this is viewed as a spin singlet Cooper pair of electrons being localized on orbitals composed of equal weight superpositions of the hexagons of the honeycomb lattice. While model wave functions of this phase have been constructed, the interactions that can drive a system into this phase remain to be understood.

4. Quantum spin liquids: The simplest possibility is a fully gapped quantum spin liquid. In this case there are neutral spin-$1/2$ excitations (spinons) in the insulator. Upon doping a natural possibility is that the charge goes in as bosonic holons (spinless charge-$e$ quasi-particles) whose condensation leads to superconductivity. This is the classic RVB mechanism [71] for superconductivity in a doped Mott insulator. However, in this scenario, at low doping the superconducting $T_c$ will not have anything to do with the spin gap (measured by the Zeeman scale needed to suppress pairing).

We do not attempt to decide between these different options in this paper. However, we will outline experiments that can distinguish between them in Sec. 2.9.
2.9 Proposed future experiments

As discussed in previous sections, the ideas presented in this paper suggest a number of experiments which will be extremely useful in revealing the physics. Here we reiterate and elaborate on some of these suggestions.

A crucial clue from the existing experiments is that an in-plane field suppresses the superconductivity - at optimal doping - when the Zeeman energy is of order the zero-field $T_c$. This indicates spin singlet pairing and that $T_c$ at optimal doping is associated with the loss of pairing. It will be extremely useful to study this systematically as a function of doping. For the doped $C_3$ broken insulator, the superconductivity may be driven by pairing of a small Fermi surface of electrons. Then (except perhaps at very small doping) $T_c$ and the critical Zeeman scale will continue to track each other as the doping is decreased. In contrast, if the pairing (in the form of singlet valence bond formation) already happens in the Mott insulator - as in the usual RVB theory, or with the featureless Mott insulator, then with decreasing doping $T_c$ and the critical Zeeman field should part ways significantly.

A second crucial clue from the experiments is the $2, 4, 6, 8, \ldots$ degeneracy pattern of the Landau fan emanating from the Mott insulator. We proposed that this was due to the freezing of the valley degree of freedom. This can be distinguished from the alternate possibility that there is spin freezing by studying the quantum oscillations in a tilted field. Zeeman splitting, if it exists, should show up in a characteristic way as a function of tilt angle.

Our proposal is the intervalley phase coherence at a scale higher than both the superconducting $T_c \approx 1.5K$ and the Mott insulating scale $\approx 5K$. The valley symmetry is as usual related to translational symmetry of the microscopic graphene lattice. In the twisted bilayer there is an approximate translation symmetry that holds at some short scale associated with translation by one unit cell of the microscopic graphene lattices. Under this approximate translation operation electrons at the different valleys get different phases. This is a $U_\nu(1)$ rotation. Therefore intervalley ordering will strongly break this approximate short translation symmetry. Within each moiré site the density of states will be uniform at the lattice scale when there is no intervalley ordering but will oscillate once this order sets in. This may be detectable through Scanning Tunneling
Microscopy (though if the bilayer graphene is fully encapsulated by Boron Nitride it may be challenging to see the graphene layer).

Assuming there is intervalley ordering, if the undoped Mott insulator develops antiferromagnetic order, it appears likely that the doped superconductor will be a spin singlet $d_{x^2-y^2} + id_{xy}$ superconductor. This spontaneously breaks time reversal symmetry. In contrast for a doped $C_3$ broken state, either $s$-wave or $d + id$ spin-singlet superconductivity seem possible. It will also be useful to directly search for broken $C_3$ or moiré translational symmetry in the experiments. Finally the very different behavior in quantum oscillations between electron and hole doping away from the Mott insulator suggests that there may be a first order transition into the Mott state as it is approached from the charge neutrality point. This will lead to hysteretic response as the gate voltage is tuned towards charge neutrality from the Mott insulator.

2.10 Conclusion

In this paper we addressed some of the theoretical challenges posed by the remarkable observations of Mott insulating states and proximate superconductivity in twisted bilayer graphene.

We proposed that both the Mott insulator and the superconductor develop out of a state with spontaneous intervalley coherence that breaks independent conservation of electrons at the two valleys. We described a mechanism for the selection of this order over other spin/valley polarized states owing to the peculiarities of the symmetry realization in the band structure. We showed that intervalley ordering by itself does not lead to a Mott insulator, and described possible routes through which a Mott insulator can develop at low temperature. A specific concrete example is a $C_3$ broken insulator. We showed how doping such an insulator leads to an understanding of the quantum oscillation data, and presented a possible pairing mechanism for the development of superconductivity. We described potentially useful experiments to distinguish the various possible routes to a Mott insulator from an intervalley coherent state.

Our work was rooted in a microscopic understanding of the twisted graphene bilayer. We showed that the momentum space structure of the nearly flat bands places strong constraints on real space descriptions. In particular, contrary to natural expectations, we showed that a real space lattice model is necessarily different from a correlated triangular lattice model with two
orbitals (corresponding to the two valleys) per site. This is due to a symmetry enforced obstruction to constructing Wannier functions centered at the triangular sites that can capture the Dirac crossings of the nearly flat bands. In our description of the intervalley ordered state and its subsequent low temperature evolution into the Mott/superconducting states, we sidestepped these difficulties by first treating the problem directly in momentum space and defining a real space model only at scales below the intervalley ordering (when the obstruction to a honeycomb representation is gone).
Faithful tight-binding models and fragile topology of magic-angle bilayer graphene in twisted bilayer graphene

3.1 Introduction

In strongly correlated materials such as transition metal oxides, which include the high-$T_c$ cuprate materials [71], the competition between kinetic energy and electron-electron interactions stabilizes remarkable phases such as Mott insulators and high-temperature superconductors. Their theoretical description traditionally begins with a tight-binding model which provides a real-
space representation of the relevant electronic bands. Interactions are then incorporated by means of a local $U$ term, leading to the Hubbard model.

Recently, another example of a correlated insulator in proximity to a superconductor has appeared—two adjacent graphene sheets that are twisted by a specific small angle relative to each other [33, 34]. Here, we will address the question of constructing a minimal model for twisted bilayer graphene, analogous to the square-lattice tight-binding model for cuprates. We will see that traditional approaches to this problem fail due to a form of band topology inherited from the underlying Dirac nature of the problem. Instead, a new approach is called for, which is developed in this paper.

Twisted bilayer graphene (TBG) structures have been studied intensely in the last decade [75–86, 88, 93, 97, 119–127]. To begin with, the two valleys of graphene are decoupled from one another, particularly in the limit of small twist angles, yielding a valley quantum number in addition to spin. The electronic states near each valley of each graphene monolayer hybridize with the corresponding states from the other monolayer. When the twist angle is close to certain discrete values known as the magic angles, e.g., at $\sim 1.05^\circ$, theoretical calculations show that there are two nearly flat bands (per valley and per spin) that form in the middle of the spectrum [80] and are separated from other bands [127]. The band gaps are also observed in experiments [33, 34]. These nearly flat bands contain Dirac nodes that intersect the chemical potential at neutrality.

Counting electron filling from neutrality, at fillings $\nu_T = \pm 4$ a band insulator is obtained. However, in experiments, correlated insulators are observed at partial band fillings of $\nu_T = \pm 2$ on cooling below a few Kelvin. Further doping this insulator at $\nu_T = -2$ with either electrons or holes reveals superconductivity at a $T_c \sim 1$ K [34]. A natural question that arises is: how similar is this phenomenon to that in the cuprates, and, relatedly, what is the minimal model that we should consider here, analogous to the square-lattice Hubbard model for the cuprates?

Ideally, a minimal tight-binding model for TBG would describe only the two nearly flat bands (per valley and per spin) and respect all symmetries. However, there is an interesting topological aspect to the nearly flat bands that obstructs finding such a model [98, 128]. This is intuitively seen by recognizing that the two Dirac points in the nearly flat bands originate from the
unperturbed Dirac cones which, while living in different layers, belong to the same valley. This suggests that they carry the same chirality [33, 92, 97, 122], i.e. the same Berry phase of $\pm \pi$ of the two Dirac cones, whose relative sign is well-defined. In contrast, any two-band tight-binding model would give Dirac cones with vanishing net chirality. This obstruction is also reflected in the symmetry eigenvalues of the bands, which cannot be captured in a minimal two-band tight-binding model [98, 128]. Given that a minimal tight binding model is forbidden, one must then proceed with one of the following two options (i) extend the model to include additional bands, or (ii) give up on some symmetries.

In this paper we will pursue the first option, namely, forgoing the minimality requirement and constructing models with more than two bands. Specifically, we will introduce a ten-band model in which the connected bands follow a 4-2-4 sequence, with the middle two bands representing the nearly flat bands of a single valley of TBG. Our model has the following advantages. First, all symmetries are respected and represented appropriately, and the two isolated bands incorporate the previously mentioned band topology of TBG. Second, the additional, complementary bands have a natural correspondence with the higher-energy bands in TBG. Finally, we can incorporate an approximate particle-hole symmetry in this description, which is known to be a good symmetry for the higher-energy states (although generally broken for the nearly flat bands). We also discuss the construction of a more minimal six-band model which retains a set of four excited bands only on one side of the nearly flat bands. Our models therefore pave the way to the derivation of a symmetric, interacting real-space description of TBG. Our solutions are reminiscent of ‘p-d’ models of the copper oxides, where correlated ‘d’ copper orbitals are augmented by oxygen ‘p’ orbitals [129]. An important difference here is that the topological obstruction prevents further downfolding that would eliminate the additional bands.

Furthermore, the complementary bands in our model are manifestly trivial, in that they can be smoothly deformed into an explicit atomic insulator. This is conceptually interesting, as it is in stark contrast with the familiar forms of band topology, say those exemplified by the Haldane [130] or Kane-Mele [131, 132] models, which are described by stable (K-theoretic) topological invariants [133–135]. There, the nontrivial topological indices must cancel when all sets of bands are accounted for, so a band with stable topology cannot be neutralized by additional triv-
ial bands. Rather, our model proves that the identified band topology in TBG falls into the class of “fragile topology” recently introduced in Ref. [136].

Let us also mention that, in Ref. [128], we provided a different recipe to extend the model. There, we constructed a four-band model where all the symmetries are implemented naturally, and the four bands split into 2-2 sets of isolated bands, each of which individually showcases the identified band topology of TBG. While the smaller number of bands is an advantage, the additional bands are disconnected from the physical degrees of freedom in TBG, and interactions can be reliably treated only when they are weak enough that interband mixing can be safely neglected. In contrast, we believe our present solution is superior in that the additional degrees of freedom correspond to physical excited states, and that it clarifies the fragile nature of the band topology.

The second option for constructing effective tight-binding models, which we recall by way of review, is to circumvent the Wannier obstructions by implementing some of the symmetries in a nontrivial manner [98]. This was done for valley symmetry in Ref. [98], which, however, entails a non-standard procedure to eventually recover the symmetry, unlike the option discussed here. Alternately, one can simply ignore some symmetries in the problem [137–139], or to adopt some different implementations of the symmetries [140–144]. An unintended consequence is a need for fine tuning. Hence, the theoretical predictions of these models are not automatically justified. For instance, in the symmetry setting of Refs. [137–139] a vertical electric field would lead to a band gap at charge neutrality, which is inconsistent with that dictated by the actual symmetries of the system [75].

We begin by briefly reviewing the symmetries and band topology relevant for small-angle TBG, before defining our ten- and six-band models, and providing a physical picture for their construction. Finally we discuss the fragile topology of the TBG flat bands and close with a discussion. Henceforth, we will focus on the single-valley problem with spin ignored.

### 3.2 Symmetry and topology of TBG

In the following, we focus on the symmetries of the continuum description [75, 80], which are also exact for the highest-symmetry commensurate lattice realizations [78, 81, 119, 128]. A more thor-
ough review of these topics can be found in Ref. [128].

The spatial symmetry group of TBG is generated by lattice translations, $C_6$ rotation, and a 2D mirror $M_y$ which flips the $y$ coordinate (more accurately, a layer-exchanging two-fold rotation in 3D)$^1$. It is also symmetric under time-reversal $T$. In addition, at small twist angle the two valleys are effectively decoupled, leading to an additional $U_v(1)$ valley charge conservation. This allows one to consider the electronic degrees of freedom residing in a single valley. Among the listed symmetries, only $C_6$ and $T$ exchange the two valleys; all the others, as well as the combinations like $C_6 T$, leave the valley charge invariant. Consequentially, the (magnetic) point group of the single-valley problem is generated by $C_6 T$ and $M_y$, and the problem is described by the magnetic space group [183.188 (in the BNS notation)]$^98$.

One can readily compute the symmetry representations furnished by the two nearly flat bands at different high-symmetry momenta, which we list in Table 3.1. Importantly, one can check that no atomic insulator with the same symmetries will have the same set of symmetry representations (Appendix C.1), and therefore there is an obstruction for constructing symmetric Wannier functions for the two nearly flat bands. This alone implies the two relevant bands are topologically obstructed from any tight-binding description that respects all symmetries$^98$. In addition, unlike the familiar case of monolayer graphene$^92$, the two Dirac points in the TBG band structures have the same chirality$^{33, 92, 97, 122}$. This is impossible in any two-band tight-binding model and constitutes another Wannier obstruction$^98$. Curiously, the two mentioned obstructions, derived respectively from the representations of $M_y$ and the net chirality of the Dirac points, are intertwined: it was shown in Ref. [128] that when the only Dirac points are pinned to the two (moiré) $K$ points, the mirror and chirality obstructions are equivalent.

$^1$Note that $M_y$ may not be present in the actual experiment, due to, for instance, different processing of the top and bottom layers. Nonetheless, as long as its (explicit) breaking is perturbative, it will be helpful to retain it as a theoretical crutch.
Table 3.1: Symmetry representations furnished by the nearly flat bands of twisted bilayer graphene [98,128], indicated by the eigenvalues of the generating symmetries of the point group. Eigenvalues from degenerate bands are grouped by parenthesis.

<table>
<thead>
<tr>
<th>Eig.</th>
<th>( \Gamma )</th>
<th>( K )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_3 )</td>
<td>1, 1 ( (\omega, \omega^* ) )</td>
<td></td>
</tr>
<tr>
<td>( M_y )</td>
<td>1, -1 not a symmetry</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Orbital content of the ten-band model. \( \tau \), \( \kappa \), and \( \eta \) respectively denote the triangular, kagome, and honeycomb sites. \( s \), \( p_z \), and \( p_\pm \) denote different orbital characters. One can also construct a more minimal six-band model using only the orbitals listed to the left of the double vertical line.

<table>
<thead>
<tr>
<th>Orbitals</th>
<th>( (\tau, p_z) )</th>
<th>( (\tau, p_\pm) )</th>
<th>( (\kappa, s) )</th>
<th>( (\eta, p_\pm) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of bands</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

3.3 Tight-binding models

3.3.1 A ten-band model

Let us now describe the mentioned ten-band model. In our present symmetry setting, one can label the orbitals as being either \( s \), \( p_z \), or \( p_\pm \). Both \( s \) and \( p_z \) orbitals transform trivially under a \( C_3 \) rotation, but \( p_z \) flips sign under \( M_y \) while \( s \) does not\(^2\). In contrast, the orbitals \( p_\pm \equiv p_x \pm ip_y \) are exchanged under \( M_y \) and form a doublet. A more systematic tabulation of the symmetry properties of the orbitals can be found in Appendix C.1. Our ten-band model comprises a \( p_z \) orbital and a pair of \( p_\pm \) orbitals localized to sites forming a triangular lattice (\( \tau \)), a \( s \) orbital on the kagome lattice (\( \kappa \)), and a pair of \( p_\pm \) orbitals forming a honeycomb lattice (\( \eta \)). For brevity, we will describe the orbital content using the notation (lattice, orbital), e.g., \( (\tau, p_z) \) denotes the \( p_z \) orbitals localized to the triangular site. Similarly, we denote the associated fermion operator by \( \hat{\tau}_{p_z} \). The described degrees of freedom are tersely summarized in Table 3.2 in this notation.

\(^2\)Recall, our “mirror” \( M_y \) is really a two-fold rotation in 3D.
guaranteed in our model by the following representation-matching equation:

\[
(\tau, p_\pm) \oplus (\tau, p_\pm) \oplus (\kappa, s) \overset{\text{rep.}}{=} (\eta, p_\pm) \oplus \text{(target)}, \tag{3.1}
\]

which one can verify using the comprehensive tabulation of the symmetry data in Appendix C.1. The physical meaning of Eq. (3.1) is that, representation-wise, it is possible to construct an atomic insulator, with the same symmetry properties as \((\eta, p_\pm)\), within the six-band sub-Hilbert space defined by the content on the left-hand side. In our ten-band model, we add another, explicit set of \((\eta, p_\pm)\) orbitals to both sides of the above equation, so that the full representation-matching equation reads

\[
(\tau, p_z) \oplus (\tau, p_\pm) \oplus (\kappa, s) \oplus (\eta, p_\pm)_0 \overset{\text{rep.}}{=} (\eta, p_\pm)_1 \oplus \text{(target)} \oplus (\eta, p_\pm)_2.
\tag{3.2}
\]

Note that we have added subscripts 0-2 to clarify that they represent different sets of physical degrees of freedom despite sharing identical symmetry properties.

Guided by this observation, one can construct a model with the targeted representations simply through the construction of the mentioned atomic insulator. Let \(\hat{c}_r^\dagger\) be the six-component fermion creation operators for the orbitals assigned to the unit cell at \(r\) (Fig. 3.1a). We want to construct a localized “quasi-orbital” wave function \(\hat{h}^{(l)}_{p_\pm; r}(x)\) such that \(\hat{h}^{(l)}_{p_\pm; r} \equiv \sum_x \hat{c}_x^\dagger \hat{h}^{(l)}_{p_\pm; r}(x)\) has the same symmetry properties as a \(p_\pm\) orbital centered at a honeycomb site, labeled by \(l = A, B\) in the unit cell \(r\). This can be achieved by using a trial wave function which vanishes everywhere except on the three nearest kagome and triangular sites surrounding the honeycomb site. Site symmetries reduce the freedom in the wave function to four real parameters\(^3\), which we denote by \(a\) through \(d\) (Fig. 3.1b). Once \(\hat{h}^{(l)}_{p_\pm; r}\) is specified, using symmetries one can generate \(\hat{h}^{(l)}_{p_-; r}\) centered at the same site, as well as those centered on the other sites. Note that we have not imposed orthogonality between the \(\hat{h}^{(l)}_{p_\pm; r}(x)\) wave functions, and so their associated fermion operators do not obey the canonical anti-commutation relations.

\(^3\)Or four complex parameters if we lift some fictitious symmetry constraints (Appendix C.2); one less if we also impose normalization.
Figure 3.1: Real-space orbitals. (a) Lattices and conventions. The shaded region indicates the relative positions of the sites assigned to the same unit cell. (b) The constructed quasi-orbital $h^{(B)}_{p+}$ in the real space. $\zeta = e^{i2\pi/6}$ and $\omega = \zeta^2$. Going from top to bottom, the entries in the three-component vectors attached to the triangular sites denote the amplitude for the $p_z$, $p_+$ and $p_-$ orbitals; that attached to the kagome sites denote the amplitude of their associated $s$ orbital.

We are now ready to define the ten-band model. Recall, in the above, we have not utilized the $(\eta, p_{\pm})_0$ orbitals in the system. Since they have identical symmetry properties as the $h^{(l)}_{p_{\pm}r}$ quasi-orbitals we constructed, we can couple the two sets in a minimal manner:

$$\hat{H}(t, \delta) = t \sum_{r,l=\alpha,\beta; \rho=p_{\pm}} \left( \hat{\eta}^{(l)}_{\rho,r} \hat{h}^{(l)}_{p_{\pm}r} + \text{h.c.} \right) + \delta \hat{V}, \quad (3.3)$$

where $t$ is a real parameter, $\hat{V}$ is a symmetry-allowed, local perturbation which we detail in Appendix C.2, and the dimensionless parameter $\delta \in [0, 1]$ controls the overall strength of the perturbation. Note that the finite range of the wave functions $h^{(l)}_{p_{\pm}r}$ implies $\hat{H}$ is local. If the $h^{(l)}_{p_{\pm}r}$ wave functions were orthonormal, $\hat{H}(t,0)$ would be on-site and could be readily diagonalized through the bonding and anti-bonding of the $\eta_{p_{\pm}}$ and $h_{p_{\pm}}$ (quasi-)orbitals. The actual band structure of $\hat{H}(t,0)$ is shown in Fig. 3.2a, and it can be readily interpreted as a slight dressing of the bonding/anti-bonding picture.

Importantly, by construction the band structure of $\hat{H}(t,0)$ includes two exactly flat bands pinned at zero energy (Fig. 3.2d; see also Appendix C.2), whose symmetry representations must match those of the nearly flat bands in TBG. Very briefly, these flat bands exist here for the same geometric reason as that of the Lieb lattice [145]. In addition, we have chosen the wave
function parameters $a$-$d$, listed in the caption of Fig. 3.2, to reproduce the broad energetic features of the higher-energy bands of TBG. Note that our model reproduces the approximate $E_k = -E_{-k}$ particle-hole symmetry of the higher energy states in TBG, although this is not a good symmetry of the nearly flat bands. With all the key properties built-in, we simply choose $\hat{V}$ such that $\hat{H}(t_0, 1)$ faithfully captures the energetics of the ten bands near charge neutrality in TBG. This leads to the band structure shown in Fig. 3.2b, which closely resembles that computed using the continuum theory of TBG (Fig. 3.2c, which was produced from the continuum model using parameters in Ref. [139]). In particular, the two bands near charge neutrality in Fig. 3.2e touches only at the Dirac points pinned at K and K’, just like that from the continuum theory (Fig. 3.2f). As they furnish the targeted symmetry representations in Table 3.1, from the results in Ref. [128] they must display both the mirror and chirality Wannier obstructions, i.e., this ten-band model serves as an explicit resolution of all the known Wannier obstructions of the nearly flat bands of TBG.

**Figure 3.2:** Band structures. (a,b) Bands from the ten-band Hamiltonian $\hat{H}(t_0, \delta)$. For both panels, we choose $t_0 \equiv 130$ meV, and the wave-function parameters $a = 0.110$, $b = 0.033$, $c = 0.033$ and $d = 0.573$. We set $\delta = 0$ in (a) and 1 in (b). (c) Bands obtained from the continuum theory for twisted bilayer graphene with a twist angle of $\theta = 1.05^\circ$. The ten bands around charge neutrality are highlighted. (d-f) A zoom-in of the two bands at charge neutrality for the corresponding panels in (a-c). The three-dimensional plots in (e,f) are plotted over the first Brillouin zone centered at Gamma, showing the presence of exactly two Dirac points pinned to K and K’ = −K. Note that (e) is generated from our tight-binding model, whereas (f) is generated from the continuum model.
3.3.2 A SIX-BAND MODEL

As the dominant term in the ten-band model in Eq. (3.3) can be viewed as a minimal coupling between the $\eta$ and $h$ degrees of freedom, one could imagine the consequences of “integrating out” the $\eta$ fermions, which results in a low-energy theory described in terms of the $h$ degrees of freedom. In our band-theory context, such a procedure can be done simply by adding an arbitrarily large chemical potential to $\eta$, which amounts to removing the four $\eta$ bands from the Hilbert space. The leads to a six-band low-energy Hilbert space with the orbital content on the left-hand side of Eq. (3.1), but with the dominant kinetic term involving only the four bands arising from the $h$ quasi-orbitals, i.e., there will again be two nearly flat bands near zero-energy.

![Figure 3.3: Band structures from a six-band model. (a) Color code for the orbital characters. (b) The broad energetic features can be set up using only the intra-orbital dispersion. (c) Band structure from the full model, with parameters detailed in Appendix C.2.1.](image)

While the preceding picture explains the existence of a six-band model, it is also desirable to construct such a model in a more conventional manner in terms of mostly nearest neighbor bonds. We will undertake this task below. Recall that the electron density of the nearly flat bands in TBG is localized to the “AA” regions, which form a triangular lattice at the moiré scale $[76, 77, 79, 82, 88, 93, 123]$. This suggests a tight-binding model with two orbitals placed on the triangular site. To capture the existence of Dirac points at K and K’, these orbitals should be $p_\pm$, and naturally we anticipate the nearly flat bands to overlap strongly with the $(\tau, p_\pm)$ orbitals in most of the Brillouin zone. However, the $(\tau, p_\pm)$ bands feature an additional quadratic touching\(^4\) at the Gamma point, which cancels the Dirac-point chirality. In contrast, in TBG the two

\(^4\)which could split into multiple Dirac cones when trigonal warping is incorporated
nearly flat bands are non-degenerate at \( \Gamma \), and so the \((\tau, p_{\pm})\) bands alone are incapable of capturing the \( \Gamma \)-point behavior \([98, 137]\). Therefore, we expect strong hybridization between the other orbitals in the vicinity of \( \Gamma \), such that the wave function of the two nearly flat bands correspond to the singlet representations in \((\tau, p_z)\) and \((\kappa, s)\).

Based on the above picture, we construct a six-band model which captures all the salient feature of TBG, as we show in Fig. 3.3 and elaborate on in Appendix C.2.1.

### 3.4 Fragile topology

We have presented two tight-binding models, with ten or six bands, each containing two isolated bands with all the known band topology in TBG. At a glimpse, this might appear to follow the general phenomenology of topological bands, which can only arise in a tight-binding model when the topological invariants are neutralized by complementary bands possessing the “opposite topology.” Paradoxically, the complementary bands in our model are constructed using quasi-orbitals \( h \) which correspond to an atomic insulator. This indicates that the complementary bands in the ten-band model \( \hat{H}(t_0, 1) \) are all trivial, in that they can be smoothly deformed into explicit atomic insulators. More precisely, we will demonstrate this following a trick described in Ref. \([136]\), which relies on a deformation Hamiltonian

\[
\hat{H}'_\mu = \hat{H} (f_\mu t_0, f_\mu^2) + \mu \sum_{r,l=A,B,\alpha=\pm} \left( \frac{1}{10} \eta_{p_{\alpha},r}^\dagger \eta_{p_{\alpha},r} - \hat{h}_{p_{\alpha},r}^\dagger \hat{h}_{p_{\alpha},r} \right),
\]

where we choose the dimensionless function \( f_\mu = \cos(\pi \mu / 2 \mu_0) \) such that \( f_0 = 1 \) and \( f_{-\mu_0} = f_{\mu_0} = 0 \), implying \( \hat{H}'_{\mu=0} = \hat{H}(t_0, 1) \). Note that the numerical factor of \( 1/10 \) is \textit{ad hoc} and is included simply to match the energy scales of the band gaps. Similarly, the precise form of \( f_\mu \), as well as the appearance of \( f_\mu^2 \), have little physical meaning; these are just convenient choices that suffice for our purpose. For \( \mu = \mu_0 > 0 \), the four highest bands coincide exactly with the atomic insulator arising from the full-filling of the \( \eta_{p_{\pm}} \) orbitals, and the same is true for the four lowest bands when \( \mu = -\mu_0 \). As shown in Fig. 3.4a, the two band gaps in the spectrum never collapse for all \( \mu \in [-\mu_0, \mu_0] \). This provides the needed adiabatic deformation to the explicit atomic limits.
Curiously, as both the full tight-binding model as well as the complementary bands correspond to atomic insulators, the band topology of the two nearly flat bands conform to the following equation:

\[
\text{(trivial)} = \text{(trivial')} \oplus \text{(nontrivial)},
\]  

(3.5)

where we say a set of bands is trivial if and only if they admit a full set of symmetric, localized Wannier functions, i.e., the full-filling of which leads to an insulator which can be smoothly deformed into a strict atomic limit \[94, 95, 146, 147\]. Eq. (3.5) is the defining property of “fragile topology” \[136, 148, 149\]. More concretely, we say the band topology of a set of gapped nontrivial bands is fragile if and only if one can append to the set another trivial set of bands such that, altogether, the augmented set is trivial; otherwise, we say the band topology is stable. As defined, stable and fragile topology are mutually exclusive concepts \[136\].

Since the band topology of our ten-band model conforms to Eq. (3.5), our model also serves to prove that the identified form of band topology is fragile in nature. For completeness, in Appendix C.3 we establish the band topology in our six-band model is also fragile. This suggests that the interesting correlated behavior observed in TBG \[33, 34\] could be related to interacting electrons occupying bands with an unconventional form of band topology.

As a conceptual remark, we note that the fragile phenomenology of our models will persist as long as we retain \(C_2T\) and lattice translations—the symmetries protecting the Dirac points. Therefore, upon the breaking of \(M_y\) and \(C_3\), our models provide an example of fragile topology not diagnosable using methods reliant on symmetry representations \[94, 95\]. \(C_2T\)-protected band topology has been studied in earlier works \[149–152\]. In particular, as a corollary from the fragile nature of our models, we remark that the 2D “Stiefel-Whitney insulators” proposed in Ref. \[152\] can be atomic.
Figure 3.4: Deformation to atomic limits. (a) The band gaps $\Delta$ below and above the two nearly flat bands stay open for all values of $\mu \in [-\mu_0, \mu_0]$ in Eq. (3.4). More than $4 \times 10^4$ momenta are sampled in the Brillouin zone in determining $\Delta$. We choose $\mu_0 = t_0 = 130$ meV. (b–d) Schematic band diagrams at various limits of the deformation. (b) When $\mu = -\mu_0$, the lowest four bands arise solely from the $(\eta, p_{\pm})$ orbitals and are therefore strictly atomic. Correspondingly, the upper six bands, altogether, are also in a strict atomic limit. (Dashed boxes indicate strictly atomic bands.) The same is true for the case of $\mu = \mu_0$ in (d), but with the role of the lowest and highest bands exchanged. Since the band gaps are maintained throughout the entire deformation, we can infer that all the (light and dark) purple blocks correspond to trivial, atomic bands. The nontrivial nearly flat bands at charge neutrality, therefore, must feature fragile topology.

3.5 Discussion

In this work, we prove that the identified form of band topology in TBG is “fragile” in nature [136], and could be resolved simply by the addition of trivial degrees of freedom. This is accomplished by the construction of explicit tight-binding models that faithfully capture all the key energetic, symmetry and topology features of the bands near charge neutrality. Our results are based on the observation encapsulated in Eq. (3.1) and its similarity with Eq. (3.5), where the latter is the defining phenomenology of fragile topology. However, we caution that such representation-matching equations are not unique (see Appendix C.1 for more details). For instance, one can interchange $s \leftrightarrow p_z$ in Eq. (3.1) and it still holds. Alternatively, one might also swap some of the orbitals used through the equality $(\tau, p_z) \oplus (\tau, p_{\pm}) \overset{\text{rep.}}{=} (\kappa, s)$ (and similarly with $s \leftrightarrow p_z$), despite the fact that the two sides of this equation are not adiabatically deformable into one an-
other [95,153]. As a concrete example, we utilize a different representation-matching equation to construct a five-band model in Appendix C.4. The smaller number of bands, however, comes at the cost of not faithfully capturing the representations of the higher-energy states.

The natural next step is to derive the dominant interaction terms in the problem by connecting our model to the microscopic degrees of freedom in TBG. Ideally, we would like to isolate the relevant bands in the continuum model and construct Wannier states which correspond to our tight-binding orbitals. We caution, however, that band crossings may occur at higher energies, so that isolating the relevant bands may require some judgement, although we do not expect this to affect the low-energy physics.

Another interesting future direction is to study how the unconventional nature of fragile topology might inform the physics of the interaction problem. Since the complementary trivial bands are fully filled, they correspond to an atomic insulator. In the limit where the band gap is much larger than the interaction strength, the problem should reduce to one involving certain local constraints on the Hilbert space, which has not previously been explored in this context. We leave these questions for future works.
Dimensional decoupling at continuous quantum critical Mott transitions

4.1 Introduction

Despite its great success in describing many metals, Fermi liquid theory fails to characterize many strongly correlated metals. Examples of “non-Fermi liquid” metals found experimentally include heavy fermion compounds, the cuprates, and some organic salts [26,35–37,154–156].

The central conceptual building block in Fermi liquid theory is the existence of long-lived electronic quasi-particles near a sharply defined Fermi surface in momentum space. The electron
spectral function $\mathcal{A}(k, \omega)$ near the Fermi surface takes the form

$$\mathcal{A}(k, \omega) = Z \delta(\omega - \epsilon(k))$$  \hspace{1cm} (4.1)

where $Z$, the quasiparticle residue, measures the overlap between the wave function of a quasiparticle and that of the original electron, and $\epsilon(k)$ is the (gapless) dispersion of quasiparticles. Metals in which this building block breaks down will show non-Fermi liquid properties in a number of experimental probes.

Our concern in this paper is on non-Fermi liquid metals in layered quasi-two dimensional systems near Mott metal-insulator (and other closely related) phase transitions. Specifically we will focus on metals at or near continuous (i.e quantum critical) Mott transitions in such systems. Previous work has demonstrated, for a single isolated two dimensional ($2d$) layer, the possibility of a second order quantum phase transition from a Fermi liquid metal to a quantum spin liquid Mott insulator with a Fermi surface of charge neutral spin-1/2 fermions [157]. These studies were motivated by the phenomenology of the quasi-2d triangular lattice organic materials and the cuprate metals [26, 156], and they have obtained support from numerical studies [158]. Here we study the effect of weak interlayer coupling on the fate of such continuous Mott transitions in the physical three dimensional material.

In the specific context of the cuprate metals it has long been appreciated that the combination of their layered quasi-two dimensional structure and their possible non-Fermi liquid properties could lead to peculiar interlayer transport. Specifically interlayer transport was argued to be related to the single particle spectrum in such layered materials, and hence probe very different physics from intralayer transport [159]. Interlayer transport is thus a very useful spectroscopic probe of a correlated quasi-two dimensional metal, particularly in situations where other direct probes like photoemission or tunneling is not feasible [160].

We show that, in the spin liquid Mott insulating phase, different layers decouple from each other so that the system behaves as a stack of two dimensional layers. We call this ‘dimensional decoupling’. In contrast, in the Fermi liquid, different layers recouple to form a coherent three dimensional Fermi surface (see Fig. 4.1). The transition is thus between a three dimensional
Fermi liquid and a dimensionally decoupled stack of two dimensional insulators. We show that this phase transition is continuous, and further that the dimensional decoupling happens already at the quantum critical point. In other words, universal critical properties are correctly obtained from a purely two dimensional theory though the metallic phase on one side is (at low energies) three dimensional. We discuss the implications for physical properties as the transition is approached from the metallic side. In particular we show that the interlayer conductivity as a function of decreasing temperature in the nearly critical metal has a ‘coherence’ peak at a crossover scale determined by the distance from the quantum critical point, and determine the detailed universal temperature dependence in various regimes.

Very recent experiments have studied interlayer transport in the doped triangular lattice $\kappa$–$ET$ organic metals [37], and our results should be a useful guide to their interpretation. Ref. [37] found the intralayer transport appears to be non-Fermi liquid like at ambient pressure and becomes Fermi liquid like under high enough pressure. As the temperature is decreased in the non-Fermi liquid regime, the interlayer resistivity first increases when the temperature is high. But it starts to decrease when the temperature is low enough, so it has a peak as temperature changes. This ‘interlayer coherence’ peak becomes broadened and is shifted to higher temperatures as the pressure becomes higher. Further this peak seems to occur in a regime where the intralayer transport is non-Fermi liquid like. It is thus timely to study issues related to interlayer coupling near continuous Mott transitions.

In the cuprate context previous theoretical work on doped Mott insulators identified, within a slave boson framework, an “Incoherent Fermi Liquid” (IFL) regime that has some phenomenological appeal as a description of the strange metal normal state [161]. Our results describe the interlayer transport of this Incoherent Fermi Liquid. Our results also carry over straightforwardly to Kondo breakdown transitions (of the kind studied in Refs. [162]) in Kondo lattice models in layered systems.

We emphasize that dimensional decoupling is not guaranteed to describe all two dimensional non-Fermi liquids. To illustrate this we consider a class of non-Fermi liquids that develop at metallic quantum critical points associated with onset of broken symmetry. A good and topical example is the onset of Ising nematic ordering from a symmetry preserving metal. In Ap-
In Appendix D.1 we show, within the existing theory of this transition, that at the quantum critical point there is no dimensional decoupling. We provide arguments that this is likely the case at all Landauesque quantum critical points driven by broken symmetry order parameter fluctuations (coupled to the metallic electrons).

In contrast continuous Mott transitions or the Kondo breakdown transitions are driven by electronic structure fluctuations which cannot be captured through Landau order parameters. On approaching this kind of quantum critical point from the metallic Fermi liquid side the quasiparticle residue $Z$ is expected to vanish continuously. Right at the critical point the quasiparticle is thus destroyed everywhere. Despite this, Ref. [27] argued that the quantum critical point is characterized by a sharply defined Fermi surface (dubbed a “critical Fermi surface”). Concrete examples which illustrate the general arguments of Ref. [27] are in Refs. [157, 162, 163] based on slave-particle gauge theories.

Clearly such phase transitions driven by electronic structure fluctuations require a different conceptual framework from more conventional order parameter driven ones. The corresponding quantum critical phenomenology will also be very different. Our results add to the growing list of distinctions between these two classes of metallic quantum critical phenomena.

**Figure 4.1:** Dimensional decoupling across the phase transition. On the side with $g > g_c$ we have a 3D Fermi liquid, and at the quantum critical point and on the other side we have a state that behaves as a stack of many decoupled layers at low energies.
To address the possibility of dimensional decoupling in these systems, it is necessary to examine the effects of all possible interlayer interactions. The simplest amongst these is electron tunneling between different layers. Other potentially important couplings is that between slowly fluctuating bosonic order parameters obtained as bilinears made out of the underlying electrons. The phenomenon of dimensional decoupling requires that, in the renormalization group (RG) sense, none of the possible interlayer interactions is relevant at the decoupled fixed point with no interlayer coupling. For the order parameter driven phase transitions, the coupling between order parameter fluctuations in different layers becomes relevant and destabilizes the decoupled fixed point. However this does not happen at the continuous Mott transition.

In passing, we note that our starting point is that the coupling between different layers is weak at the lattice level, so it is legitimate to first consider the system as a stack of decoupled layers and then study the effects of the interlayer couplings. In contrast, other works, such as Ref. [164], study the situation where the interlayer interactions are as strong as intralayer interactions, and in this case considering a 3D system from the beginning is more appropriate.

The rest of this paper is organized as follows. We begin with a brief review of some general (semi-) quantitative aspects of a critical Fermi surface and continuous Mott transitions in Sec. 4.2. Before diving into the more complicated case of continuous Mott transitions, in Sec. 4.3 we will warm up by illustrating the phenomenon of dimensional decoupling in a simpler context: a quantum phase transition between a Fermi liquid and an orthogonal metal. We will first consider a single 2D layer and review the nature of the orthogonal metal state and the phase transition in 2D, then we consider a stack of many such 2D layers and examine the effects of interlayer interactions. Following the same strategy, we will go into the case of continuous Mott transitions and demonstrate dimensional decoupling thereof in Sec. 4.4. After this, in Sec. 4.5 we will calculate the interlayer electric conductivity induced by electron tunneling between different layers in the QC regimes of the various cases of interests. Because of the rich crossover structure predicted by the theory of continuous Mott transitions, in order to identify some of its experimental signatures, in Sec. 4.6 we study the interlayer electric conductivity as the system crosses over from its QC regimes to the Fermi liquid regime. Finally, we conclude with some discussions on experiments in Sec. 4.7.
4.2 Preliminaries

In this section we collect together some previous results on critical Fermi surfaces and continuous Mott transitions that will be used extensively in the rest of this paper.

4.2.1 Critical Fermi surface

A continuous Mott transition from a Fermi liquid metal requires a sudden death of the metallic Fermi surface. The transition to the Mott insulator occurs without the vanishing of the free carrier density, and hence the Fermi surface cannot simply shrink to zero. Thus continuous Mott transitions necessarily involve the death of an entire Fermi surface of some fixed size. Ref. [27] argued that at the corresponding quantum critical point there will be a sharp critical Fermi surface but without well-defined Landau quasiparticles. To discuss interlayer coupling effects in different systems in a unified manner, we will need some very general scaling properties of the electron spectral function at such a critical Fermi surface which we now summarize.

Suppose one can access such a quantum critical point by tuning a parameter $g$. It is natural that at temperature $T$, the electron spectral function around the critical Fermi surface near such a critical point has the scaling form:

$$A(K, \omega; g, T) \sim \frac{c_0}{|\omega|^{\alpha/2}} F \left( \frac{\omega}{T}, \frac{c_1 k_{||}}{T^{\nu}}, k_{||}^{\beta} \right)$$

(4.2)

where $\alpha$ and $z$ are two universal exponents, and $F$ is a universal function. $k_{||}$ is the deviation of momentum $K$ from the Fermi surface. Constants $c_0$ and $c_1$ are non-universal. $\xi$ is the correlation length of the system at the same $g$ but at zero temperature, and it is a measure of the deviation of $g$ from the quantum critical point $g_c$ since $\xi^{-1} \sim |g - g_c|^\nu$, with a universal zero temperature correlation length exponent $\nu$. This scaling form of the spectral function applies to each patch of the Fermi surface, and, in general, $\alpha$, $z$, $\nu$, $c_0$ and $c_1$ can all depend on the position of the patch on the Fermi surface where (4.2) is applied. We will take the convention that the Fermi liquid is on the $g > g_c$ side throughout this paper.
According to (4.2), the number of fermions with a certain momentum $k$ is

$$n(k) = \int_{-\infty}^{0} d\omega A_c(k, \omega) \sim |k||^{-\alpha}$$  \hspace{1cm} (4.3)

Because the fermion number is upper bounded, we must have

$$z \geq \alpha$$  \hspace{1cm} (4.4)

Again, we notice that (4.2) and (4.4) apply to each patch of the Fermi surface, and in general the exponents can depend on the position of patch. The same scaling considerations apply to all systems with a critical Fermi surface.

### 4.2.2 Theory of a continuous $2d$ Mott transition

We consider two types of continuous Mott transitions: chemical potential tuned and bandwidth tuned, for which in a single $2d$ layer there are well developed theories. Theoretically, it is expected that these Mott transitions can be realized by the Hubbard-type Hamiltonian on a triangular lattice

$$H = -t \sum_{ij,\sigma} (c_{i\sigma}^\dagger c_{j\sigma} + \text{h.c.}) + U \sum_{i} n_{i\uparrow}n_{i\downarrow} - \mu \sum_{i,\sigma} n_{i\sigma}$$  \hspace{1cm} (4.5)

where $c_{i\sigma}$ annihilates an electron with spin $\sigma$ at site $i$, and $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$ is the number operator of the fermion. There is numerical evidence for a regime described by a Mott insulator with a spinon Fermi surface [15, 165–167]. Experimentally the organic compound $\kappa -(ET)_{2}Cu_{2}(CN)_{3}$, which is believed to be well described by a one-band Hubbard model on a triangular lattice, also exhibits signatures of these transitions [35, 36, 156].

The transition is conveniently accessed by formally writing the electron operator $c_{i\sigma}$ as

$$c_{i\sigma} = f_{i\sigma} \cdot b_{i}$$  \hspace{1cm} (4.6)
where the fermionic spinon $f$ carries spin-$\frac{1}{2}$, and the boson carries unit physical charge. The physical electron operator is invariant under a local $U(1)$ gauge transformation $f_{i\sigma} \to f_{i\sigma} e^{i\phi_i}$ and $b_i \to b_i e^{-i\phi_i}$, which leads to an emergent $U(1)$ gauge field at low energies. Consider a situation where spinons form a Fermi surface. If the boson $b$ condenses, the gauge field is Higgsed out and we get a Fermi liquid of the original electron. If $b$ is gapped, we get a spin liquid with a spinon Fermi surface coupled to a $U(1)$ gauge field. The longitudinal component of the gauge field is screened by the spinon Fermi surface, but the transverse components remain strongly interacting with the spinons. The Mott transition from the Fermi liquid is driven by losing the condensate of $b$.

For both the chemical potential and bandwidth tuned transitions, Ref. [157] showed that the quantum critical fluctuations of $b$ (at $T = 0$) are dynamically decoupled from the spinon-gauge system. However, the dynamics of the latter is affected by the criticality of $b$. This leads to a tractable theory of these continuous Mott transitions and a number of universal physical properties have been computed. An interesting feature shared by both transitions is that the crossover from the quantum critical metal to the Landau Fermi liquid on the metallic side occurs in two stages. The charge sector crosses over at an energy scale parametrically larger than the spin sector. At intermediate energies a non-Fermi liquid metallic regime is reached which is distinct from the quantum critical non-Fermi liquid.

We will study the effects of interlayer coupling on these transitions, and determine the nature of interlayer transport both in the quantum critical non-Fermi liquid and in the intermediate energy non-Fermi liquid that exists before the emergence of the fully coherent Landau Fermi liquid.

Notice in principle this $U(1)$ gauge field should be taken to be compact, but as pointed out in Ref. [25], the effect of instantons of this compact $U(1)$ gauge field will be suppressed by the spinon Fermi surface, so it is adequate to just consider a noncompact $U(1)$ gauge field. It is also sufficient for our purposes to treat this non-compact $U(1)$ gauge field under a random-phase-approximation (RPA), which can be formally justified by a controlled expansion in its leading order [168].
4.3 **WARM-UP: DIMENSIONAL DECOUPLING IN A QUANTUM PHASE TRANSITION BETWEEN A FERMI LIQUID AND AN ORTHOGONAL METAL**

To examine effects of interlayer coupling at continuous Mott transitions we need to confront the full theory of the the critical $b$ fluctuations and the spinon-gauge system. In this section we warm up to this task by considering a different problem which has some of the same ingredients. Rather than studying the transition from a Fermi liquid to a Mott insulator, we study the transition to a different phase dubbed the orthogonal metal. We demonstrate the phenomenon of dimensional decoupling at this quantum phase transition.

An orthogonal metal is a state where the electron is fractionalized into a fermion $f_{i\sigma}$ that carries both the charge and spin of the electron and a discrete degree of freedom $s_i$ that is gapped [163]. This fractionalization is accompanied by a deconfined discrete gauge field to which both $f_{i\sigma}$ and $s_i$ are coupled. The $f_{i\sigma}$ forms a Fermi surface, and the system has metallic charge/spin transport. However, the gapless $f_{i\sigma}$ have zero overlap with the physical electrons (they are orthogonal). Thus despite the metallic charge transport single particle-probes like tunneling or photoemission will see insulating behavior. The orthogonal metal is the simplest non-Fermi liquid in spatial dimension $d \geq 2$, and its universal properties are easily computed. Some lattice models that can realize this state are provided in Ref. [163].

Formally we write the electron operator $c_{i\sigma}$ as

$$c_{i\sigma} = f_{i\sigma} \cdot s_i$$  \hspace{1cm} (4.7)

Assuming the fermions $f$ are in their Fermi liquid phase, if $s$ condenses, the resulting state is a Fermi liquid of the original electrons. However, if $s$ is gapped, we get an orthogonal metal, whose nature will be determined by the nature of the discrete variable $s$.

In passing, we remark that all systems discussed in this paper have a Fermi surface, either of the physical electrons or of some emergent fractionalized fermions. It is well-known that a Fermi surface may potentially suffer from an instability towards Cooper pairing, but because this instability occurs only at very low temperatures, we will assume our systems are in a regime free
of pairing instability through out this paper. We also notice that in some systems with a Fermi surface coupled to a gauge field, the pairing instability is suppressed [169].

Now consider a single 2D layer first. One can drive a transition from a Fermi liquid of electrons, a condensate of $s$, to an orthogonal metal by destroying the condensate. As discussed in Ref. [163], in 2D the transition to a $Z_2$ orthogonal metal in a lattice needs fine-tuning of the parameters in the Hamiltonian, and the simplest example where a generic second-order phase transition can occur in a 2D lattice is between an electronic Fermi liquid and a $Z_4$ orthogonal metal. The critical theory for such a transition is described by the following Lagrangian

$$L = b^* \partial_\tau b + \frac{1}{2m_b} |\nabla b|^2 + t |b|^2 + \frac{u}{4} |b|^4 + \frac{v}{4!} [b^4 + (b^*)^4]$$

(4.8)

where the $b$ transforms as $b \rightarrow ib$ under the $Z_4$ transformation. Notice this is not a symmetry-breaking phase transition because $b$ itself is not gauge invariant. Instead, this is a transition associated with electron fractionalization.

Figure 4.2: The schematic phase diagram and crossover structure of an orthogonal metal transition. When $g > g_c$ we have a Fermi liquid (FL), while we have an orthogonal metal (OM) when $g < g_c$. At finite temperature, there is a quantum critical regime (QC) where the system is non-Fermi liquid like.

At the critical point, the quasiparticle residue of the electron vanishes and the excitations are
incoherent. The spectral function of the $Z_4$ spins at the critical point is $A_s(q, \Omega) \sim \delta(\Omega - \frac{q^2}{2m_b})$. Since the spinons are in their Fermi liquid phase, they have spectral function $A_f(k, \omega) \sim \delta(\omega - v_F k_\parallel)$. Convolving them we get the electron spectral function at the quantum critical point:

$$A_c(k, \omega) = \int_q \int_0^\omega d\Omega A_s(q, \Omega) A_f(k - q, \omega - \Omega)$$

$$\sim \left(\omega - \frac{k_\parallel^2}{2m_b}\right) \frac{1}{2} \theta \left(\omega - \frac{k_\parallel^2}{2m_b}\right)$$

This has the form of (4.2), and we see that this orthogonal metal transition has a critical Fermi surface with $\alpha = -1$ and $z = 2$. As we can also see, (4.4) is indeed satisfied. The large and negative value of $\alpha$ means in the quantum critical regime (QC) above the quantum critical point, the system is highly non-Fermi liquid like. If $g \neq g_c$, starting from the QC regime, as the temperature is decreased, the system crosses over to a Fermi liquid metal or an orthogonal metal, depending on the relative magnitude of $g$ compared to $g_c$ (see Fig. 4.2).

Now consider a stack of such 2D systems, with each layer going through a phase transition between a Fermi liquid and a $Z_4$ orthogonal metal. When the interlayer interactions are absent, in terms of RG, the critical point is described by a fixed point where each layer corresponds to an individual fixed point and all these fixed points are decoupled. We will call such a fixed point a “decoupled fixed point.” Let us examine the effect of all possible interlayer interactions on this decoupled fixed point. Note that the decoupled fixed point has separate conservation of physical electric charge in different layers corresponding to independent global $U(1)$ symmetry rotations in each layer. In writing (4.7), we have introduced a $Z_4$ gauge redundancy on each layer, so the interlayer interactions should be invariant under a local $Z_4$ gauge transformation within each layer. The most obvious physical coupling is simply electron tunneling between different layers - this breaks the infinite number of $U(1)$ symmetries associated with conservation of electric charge separately in each layer to a single common global $U(1)$. We will first however focus on interlayer couplings that preserve this infinite $U(1)$ symmetry.

The most important such interactions consistent with gauge invariance and global symmetries
are the coupling between energy densities of different layers, which is of the form

\[ \delta L_1 = \sum_\alpha \int d\tau d^2 x g_{1\alpha\beta} |b_\alpha|^2 |b_\beta|^2 \]  

(4.10)

and the coupling between the energy density of one layer and the collective excitation around the spinon Fermi surface of the other layer, which is of the form

\[ \delta L_2 = \sum_{\alpha\beta,\sigma} \int d\tau d^2 x g_{2\alpha\beta} |b_\alpha|^2 |f_{\beta\sigma}^\dagger f_{\beta\sigma}| \]  

(4.11)

Here we use \( \alpha \) and \( \beta \) to index the layers.

Perturbing the decoupled fixed point with \( \delta L_1 \), we get its RG equations

\[ \frac{dg_{1\alpha\beta}}{dl} = -C g_{1\alpha\beta}^2 \]  

(4.12)

with a constant \( C > 0 \) (see Appendix D.2). From this RG equation, we see \( \delta L_1 \) is (marginally) irrelevant.

On the other hand, to consider the effect of \( \delta L_2 \) on the decoupled fixed point, it is convenient to integrate out the degrees of freedom from the spinons, and the most relevant interaction generated from this procedure has the following Landau damping form

\[ \sum_\alpha g'_{2\alpha\beta} \int \omega, q \Pi(q, \omega) \cdot |b_\alpha|^2(q, \omega) \cdot |b_\beta|^2(q, \omega) \]  

(4.13)

with

\[ \Pi(q, \omega) \sim \frac{|\omega|}{|q|} \]  

(4.14)

for small frequencies and momenta. Because the decoupled fixed point has dynamical exponent \( z = 2 \), the dependence of \( \Pi \) on the frequency and momentum \( \frac{|\omega|}{|q|} \sim q \) makes this interaction irrelevant.

As for other interactions, one should in principle consider the couplings between charge densities, spin densities, charge currents and spin currents, and the most relevant ones of these cou-
plings in this case are of the form of four-fermion interactions of the spinon \( f \). It is well-known that in the presence of a Fermi surface, most of the four-fermion interactions are strongly irrelevant, except for the forward scattering and the BCS scattering. The former is marginal and will not modify the physical properties of the system qualitatively, and they can be described by a set of Landau parameters, while the latter is marginally irrelevant and can induce pairing instability at very low temperatures [170–172]. In our case, most of these interlayer four-fermion interactions are also strongly irrelevant due to the kinematic constraint of the Fermi surface, and the analog of forward scattering will renormalize the in-plane Landau parameters, but they will not give rise to any qualitative change of the physics. The analog of BCS scattering can in principle induce interlayer pairing at very low temperatures, but as declared before, we will ignore it.

We now return to the important effect of interlayer electron tunneling. To discuss these we will use a scaling argument that is somewhat different from the one above.

Consider a general action for the interlayer electron hopping. At low energies it is appropriate to work with electronic modes near the critical Fermi surface. The hopping term can then be written as

\[
\delta S_{\text{tunneling}} = -\int d\omega dk_\parallel d\theta \sum_{\gamma\delta} t_{\gamma\delta}(\theta) \cdot \left( c^{\dagger}_\gamma(k_\parallel, \theta, \omega) c_\delta(k_\parallel, \theta, \omega) + \text{h.c.} \right)
\]

(4.15)

Now for any critical Fermi surface consider a scaling transformation that renormalizes toward the Fermi surface. We let \( k_\parallel \rightarrow k_\parallel' = k_\parallel s \), and \( \omega \rightarrow \omega' = \omega s^\alpha \). The two point correlation function of \( c(k_\parallel, \theta, \omega) \) satisfies

\[
\langle c(k_\parallel_1, \theta, \omega_1) c^{\dagger}(k_\parallel_2, \theta, \omega_2) \rangle = \delta(k_\parallel_1 - k_\parallel_2) \delta(\omega_1 - \omega_2) G(k_\parallel_1, \theta, \omega_1)
\]

If the electron spectral function for a 2D layer has the scaling form (4.2), then the electron operator transforms as \( c(k_\parallel, \theta, \omega) \rightarrow c'(k_\parallel', \theta, \omega') = s^{\frac{\alpha+1}{2}} c(k_\parallel, \theta, \omega) \). It follows that the scaling of the hopping parameter is \( t'(\theta) = t(\theta) s^\alpha \). If \( \alpha < 0 \), we expect that the interlayer electron hopping is irrelevant.
As discussed above the electron spectral function for a 2D layer at the critical point for the Fermi liquid to orthogonal metal transition satisfies the scaling form of Eqn. 4.2 with \( \alpha = -1 \) and \( z = 2 \), and thus the inter electron hopping indeed scales to zero at low energy.

We point out a caveat in this scaling analysis. In the example of orthogonal metal transition discussed in this section (and the chemical potential tuned Mott transition discussed later), the boson sector has dynamical exponent \( z_b = 2 \) while the fermion sector has dynamical exponent \( z_f = 1 \), and the electron Green’s function has dynamical exponent \( z_c = 2 \). So when we do a scaling analysis, how should we scale space and time? Notice the above dynamical exponents indicate the important dynamical regions are \( \omega \sim k^2 \) and \( \omega \sim k \), and the former governs the important electron dynamics. Moreover, the former is a slower regime compared to the latter. Therefore, to consider the lowest energy physics which is also pertinent to the electrons, we choose \( z = 2 \) in the above scaling analysis.

Therefore, none of the interlayer interactions is relevant at the decoupled fixed point, and different layers do decouple at the quantum critical point. If we go into the orthogonal metal side, since \( \mathbf{b} \) is gapped there, the above interactions that involve \( \mathbf{b} \) will be more irrelevant and the other interactions stay as irrelevant as they are at the critical point. So as long as the system leaves the Fermi liquid phase, it behaves as a stack of many decoupled layers, as shown in Fig. 4.1. This is our simplest example of dimensional decoupling.

Before moving on, we emphasize the important role of the emergent gauge invariance in obtaining dimensional decoupling. Gauge invariance strongly constrains the possible form of the interlayer interactions, and in the absence of this constraint there will be relevant interactions in general (see Appendix D.1).

### 4.4 Dimensional decoupling in continuous Mott transitions

Warmed up with the example of the orthogonal metal transition, now we are ready to deal with the more complicated problem of continuous Mott transitions.
4.4.1 Chemical potential tuned continuous Mott transition

If the system is not at half-filling and the transition is accessed by changing the electron doping, or equivalently, by tuning the chemical potential, the low-energy effective Lagrangian of this transition is

\[
\mathcal{L} = \mathcal{L}_b + \mathcal{L}_f + \mathcal{L}_{\text{gauge}} \tag{4.16}
\]

with

\[
\mathcal{L}_b = \bar{b} \left[ \partial_\tau - i a_0 - \mu - \frac{\nabla - ia}{2m_b} \right] b + V(|b|^2)
\]

\[
\mathcal{L}_f = \bar{f} \left[ \partial_\tau + i a_0 - \frac{\nabla + ia}{2m_f} - \mu_f \right] f \tag{4.17}
\]

\[
\mathcal{L}_{\text{gauge}} = \frac{1}{4e^2} (\epsilon^{\mu\nu\lambda} \partial_\nu a_\lambda)^2
\]

The potential \( V(|b|^2) \) can be taken to be of the conventional form \( r|b|^2 + g|b|^4 \). Notice we did not include a direct interaction between bosons and spinons, since this can be shown to be irrelevant [157]. This model has been extensively studied. Closely related models appear in theories of the cuprates [173] and of the Kondo breakdown phenomenon in Kondo lattice systems [162].

There are several interesting features associated with this transition [157]. First of all, it is shown that at the quantum critical point the bosons are dynamically decoupled from the spinon-gauge-field sector. Therefore, the transition is in the universality class of a dilute (nonrelativistic) bose gas. Using this, one can calculate the electron spectral function at the critical point and show there is indeed a critical Fermi surface, and \( \mathcal{A}(K, \omega) \) in this case has the same form as (4.9). Since the details of this calculation were not given in Ref. [157], we present it in Appendix D.3. As shown in Fig. 4.3, there is a QC regime above the zero temperature quantum critical point, where the system is strikingly non-Fermi liquid like, just as the QC regime of the \( Z_4 \) orthogonal metal transition.

Moreover, when the system crosses over out from the QC regime to the Fermi liquid, the system has to first go through an intermediate regime. In particular, we can choose the boson phase stiffness \( \rho_s \) as the characteristic energy scale on the Fermi liquid side. The system is in its quan-
Figure 4.3: Phase diagram and crossover structure of a chemical potential tuned Mott transition. The critical point corresponds to $\mu = 0$ and $T = 0$. The $\mu < 0$ side corresponds to a spin liquid Mott insulator (MI), and the $\mu > 0$ side corresponds to a Fermi liquid metal (FL). The quantum critical regime is highly non-Fermi liquid like (QC nFL). In crossing over to the Fermi liquid, the system has to first go through an intermediate incoherent Fermi liquid regime (IFL).

... quantum critical (QC) regime if $T \gg \rho_s$. If we decrease the temperature from QC so that $T \ll \rho_s$, the bosons appear to condense. However, the spinon-gauge-field sector has not yet felt the Higgs effect and it behaves as if it is still in its own quantum critical regime. The system in this regime is also a non-Fermi liquid, and it is dubbed “incoherent Fermi liquid” (IFL) [161]. When the temperature is further decreased so that $T \ll \rho_s^{\frac{3}{2}}$, the Higgs effect is manifested and the spinon-gauge-field sector also crosses over out of its quantum critical regime, and the system appears as a Fermi liquid. These results can be shown, for example, by calculating the electron spectral function in various regimes.

On the Fermi liquid side, the propagator of the transverse components of the gauge field in Coulomb gauge under RPA is

$$D_c(q, i\Omega) = \frac{1}{k_0 |\Omega| + \chi_d q^2 + \rho_s}$$  \hspace{1cm} (4.18)

where $k_0$ is of the order of a typical Fermi momentum of the spinon Fermi surface, and $\chi_d q^2$ is
the diamagnetic term. Upon approaching the quantum critical point from the Fermi liquid side, the superfluid density vanishes as \( \rho_s \sim \xi^{-z} \sim (g - g_c)^{z\nu} \) up to a logarithmic correction that we will ignore \([174]\), with \( z = 2 \) and \( \nu = \frac{1}{2} \) in the universality class of a dilute bose gas. After the system passes the critical point and enters the spin liquid Mott insulator phase, the RPA gauge field propagator in Coulomb gauge becomes

\[
D_c(q, i\Omega) = \frac{1}{k_0 \frac{|\Omega|}{|q|} + (\chi_d + \frac{1}{2}) q^2}
\]  
(4.19)

where \( \Delta \) is of the order of the boson gap.

Now let us examine the effects of interlayer interactions. Similar to the warm-up example in Sec. 4.3, the interlayer interactions should also obey the gauge invariance within each layer. In the present case the gauge field structure is \( U(1) \). Apart from electron tunneling between layers, the other most important interlayer interactions here include the coupling between energy densities of different layers of the form (4.10) and the coupling between the energy density of one layer and the collective excitation of the spinon Fermi surface of the other layer of the form (4.11).

Since the chemical potential tuned Mott transition is in the universality class of a dilute bose gas, the interaction (4.10) is already seen to be irrelevant due to similar reasons as discussed in Sec. 4.3. Also, although the spinons are in their own non-Fermi liquid state at the critical point, the factor \( \Pi(q, \omega) \) obtained by integrating the spinons out still has the form given by (4.14) \([175]\).

Again, because the dynamical exponent \( z = 2 \), this interaction is also irrelevant.

However, due to the presence of a gapless \( U(1) \) gauge field, there is a coupling of the form

\[
\delta \mathcal{L}_3 = \sum_{\alpha\beta} \int d\tau d^2x g_{3\alpha\beta} (\nabla \times a_\alpha) \cdot (\nabla \times a_\beta)
\]

(4.20)

In momentum space, this coupling has the form \( \delta \mathcal{L}_3 = \sum_{\alpha\beta} \int d\omega d^2q g_{3\alpha\beta} q^2 a(q, i\omega) \cdot a(-q, -i\omega) \).

At the critical point, \( \rho_s = 0 \), and comparing (18) and the momentum-space representation of \( \delta \mathcal{L}_3 \), we see this coupling is marginal. As for any physics that only involves quantities within the same layer, the effect of this coupling is merely to modify the effective diamagnetic susceptibility. In particular, under the assumption that the interlayer interactions are weak, this coupling is not
able to change the sign of $\chi_d$, so it will not modify the physics qualitatively. Its most important effect on physics involving different layers may be that it can potentially enhance the interlayer pairing (see Appendix D.4). Since in this paper we consistently assume we are in a regime away from any pairing instability, we will also ignore it here and (4.20) will not modify the physics.

One should in principle also consider the couplings between electric currents, spin currents and spin densities between different layers. The most relevant coupling between electric currents contains two more derivatives compared to (4.10), so it is even more irrelevant than the latter. The most relevant couplings between spin currents and spin densities are all of the form of four-fermion interaction, which is not relevant for similar reasons as in the case of the orthogonal metal transition.

Of course potentially the most important interlayer coupling is through electron tunneling between the different layers. We can discuss it within the framework introduced to analyze the same issue for the orthogonal metal transition. As the exponent $\alpha < 0$ here as well the interlayer hopping will scale to zero at low energies.

So we see that at the quantum critical point of a chemical potential tuned continuous Mott transition, no interlayer interaction is relevant with respect to the decoupled fixed point. Going into the spin liquid Mott insulator side, because the bosons are gapped, the interlayer interactions that involve charges become more irrelevant while other interactions stay as irrelevant as they are at the critical point, so different layers are still decoupled. Therefore, we conclude that at the quantum critical point and in the Mott insulator phase of a chemical potential tuned Mott transition, the system behaves as a stack of many decoupled 2D layers and hence exhibits dimensional decoupling. Again, we notice the emergent gauge invariance plays an important role in obtaining this phenomenon.

4.4.2 Bandwidth controlled continuous Mott transition

If the electron filling is fixed to be at half, we can access the bandwidth controlled Mott transition by tuning the ratio of the electron bandwidth to the interaction strength, which can be done, for example, by tuning the pressure. The effective field theory of this phase transition is similar to (4.16), except that now the bosons are relativistic and described by the following La-
because of the emergent particle-hole symmetry of the boson. This difference has been realized since the study of the transition between a bosonic Mott insulator and a superfluid \[176\].

\[
L_b = |(\partial_\mu - ia_\mu)b|^2 + V(|b|^2)
\]  \hspace{1cm} (4.21)

As in the case of the chemical potential tuned Mott transition, the boson sector is again dynamically decoupled from the spinon-gauge-field sector, and the transition is therefore of the universality class of the 3D XY model. Using this, it is shown that at the quantum critical point the system is non-Fermi liquid like and electron spectral function is

\[
\mathcal{A}_c(k, \omega) \sim \frac{\omega^n}{\ln \frac{\omega}{\Lambda}} f \left( \frac{\omega \ln \frac{\Lambda}{\omega}}{v_F k_\parallel} \right)
\]  \hspace{1cm} (4.22)
where $\eta$ is the anomalous dimension of 3D XY model and the universal function $f$ is

$$f(x) = \left(1 - \frac{1}{x}\right)^\eta \theta(x - 1) \quad (4.23)$$

Fitting this spectral function into the form (4.2), we find $\alpha = -\eta$ and $z = 1^+$, with the understanding that expression such as $\omega^{\frac{1}{z}}$ should be interpreted as $\omega \ln \frac{1}{\omega}$. Again, we see (4.4) is satisfied.

In addition, the crossover out of QC regime to the Fermi liquid regime again involves an intermediate regime. Choosing the boson phase stiffness $\rho_s$ as the characteristic energy scale of the Fermi liquid, when the temperature is such that $T \gg \rho_s$, the system is in its QC regime and displays non-Fermi liquid behaviors. When the temperature is decreased so that $T \ll \rho_s$, the bosons behave as if they already condensed (i.e Higgsed), but the spinon-gauge-field sector does not feel the Higgs effect until the temperature is further lowered to the order of $\rho_s^2$. In this intermediate regime, the system behaves as a marginal Fermi liquid (MFL) state that was originally proposed by Varma et al. to describe the optimally doped cuprates [177]. Only when the temperature is further lowered so that $T \ll \rho_s^2$ does the spinon-gauge system notice the the boson condensation and the system behaves as a Fermi liquid. The difference in the powers of $\rho_s$ below which scale a Fermi liquid results in the chemical potential tuned and bandwidth tuned Mott transitions reflects that the two transitions are in different universality classes.

In this case on the Fermi liquid side the RPA gauge field propagator in Coulomb gauge is

$$D_b(q, i\Omega) = \frac{1}{k_0 \frac{|\Omega|}{|q|} + \sigma_0 \sqrt{\Omega^2 + q^2 P \left(\frac{\sqrt{\Omega^2 + q^2}}{\rho_s}\right)}} \quad (4.24)$$

where the universal function $P(x)$ satisfies $\lim_{x \to 0} P(x) \sim \frac{1}{x}$ and $\lim_{x \to \infty} P(x) = 1$, and $\sigma_0 \sim \frac{e^2}{h}$ is a universal conductance. On the Mott insulator side, it becomes

$$D_b(q, i\Omega) = \frac{1}{k_0 \frac{|\Omega|}{|q|} + \sigma_0 \sqrt{\Omega^2 + q^2 Q \left(\frac{\sqrt{\Omega^2 + q^2}}{\Delta}\right)}} \quad (4.25)$$

where $\Delta$ is on the order of the boson gap and the universal function $Q(x)$ satisfies $\lim_{x \to 0} Q(x) \sim$
Suppose we have a stack of such 2D layers, let us now examine the effects of interlayer interactions on the decoupled fixed point. These interlayer interactions must be invariant under the local $U(1)$ gauge transformation within each layer. We will follow closely the strategy used in previous sections. Interlayer electron tunneling is potentially the most important coupling. Because in this case $\alpha = -\eta < 0$, according to the similar arguments in the previous sections it is irrelevant. The other important interlayer interactions are again the coupling between energy densities on different layers and the coupling between the energy density in one layer and the collective excitations of the spinon Fermi surface in another layer, which still has the form of (4.10) and (4.11), respectively. However, because this transition is in another universality class, the previous argument should be modified. In particular, because the energy density $|b|^2$ has scaling dimension $3 - \frac{1}{p}$, the scaling dimension of $g_{\alpha\beta}$ is $\frac{2}{p} - 3$. It is known that $\nu > \frac{2}{3}$ for the 3D XY model, so $g_{\alpha\beta}$ has negative scaling dimension and (4.10) is irrelevant. On the other hand, integrating out the degrees of freedom from the spinons in (4.11), the most relevant interaction we get is again of the form (4.13). Since we have dynamical exponent $z = 1$ in this case, the scaling dimension of coupling constant of this resulting interaction is the same as that of the coupling between energy densities, which is negative as discussed above. Therefore, this interaction is also irrelevant.

Notice in this case, the coupling of the form (4.20) is simply irrelevant at the critical point, because there the gauge field propagator (4.24) is

$$D_b(q, i\Omega) = \frac{1}{k_0 |q| + \sigma_0 |\Omega|^2 + \sqrt{q^2} + \sigma_0}$$  \hspace{1cm} (4.26)$$

For the gauge field the most important fluctuations are the modes with $\Omega \sim q^2 \ll q$, so we can further approximate (4.26) as

$$D_b(q, i\Omega) = \frac{1}{k_0 |q| + \sigma_0 |q|}$$  \hspace{1cm} (4.27)$$

The coupling of the form (4.20) contains two spatial derivatives, so it is irrelevant. However, be-
cause of the structure of (4.25), as in the case of the chemical potential tuned Mott transition, (4.20) is marginal deep in the Mott insulator side. But similar arguments as there indicates this coupling does not modify the physics.

For similar reasons as in the case of chemical potential tuned Mott transition, the couplings between charge currents, spin currents and spin densities in different layers are also not relevant. As for interlayer electron tunneling, the critical Fermi surface in this problem has $\alpha = -\eta < 0$. Thus the interlayer hopping will renormalize to zero at the critical point at low energies. In the Mott insulator phase, because the boson is gapped, the interlayer interactions involving charge become more irrelevant while other interactions stay as irrelevant as they are at the critical point, so different layers are still decoupled.

Therefore, we conclude that none of the interlayer interactions at the quantum critical point of the bandwidth controlled Mott transition and in the Mott insulator phase is relevant, and the system exhibits dimensional decoupling. It is again the emergent gauge invariance that constrains the possible forms of the interlayer interactions and give rise to this phenomenon.

4.5 Interlayer Conductivity

In this section we derive the interlayer electric conductivity $\sigma$ in perturbation theory in the electron tunneling. As is well known, this leads to a formula for the interlayer conductivity in terms of the electron spectral function in each layer. We then we apply the formula to the QC regimes of the various phase transitions discussed above, and show in all three cases $\sigma \to 0$ as $T \to 0$. This is of course consistent with that the interlayer tunneling is irrelevant at the decoupled fixed point.

Suppose the total Hamiltonian of the layered three dimensional system is

$$H = \sum_a H_a + H_{\text{tunneling}}$$

(4.28)

where $H_a$ is the Hamiltonian for the ath layer that may take the form of (4.5), and the tunneling
Hamiltonian $H_{\text{tunneling}}$ takes the following specific form

$$H_{\text{tunneling}} = -t \sum_{a,i} \left( c_{i,a+1}^\dagger c_{i,a} + \text{h.c.} \right)$$

where $a$ labels the layer and $i$ labels the position of the site on a given layer.

As shown in Appendix D.5, the interlayer DC electric conductivity to the second order of interlayer tunneling amplitude $t$ is

$$\sigma = -\pi N (ted)^2 \sum_k \int d\omega (A(k,\omega))^2 \frac{df}{d\omega}$$

where $N$ is the number of layers, $d$ is the interlayer spacing, $A(k,\omega)$ is the electron spectral function on each layer, and $f(x) = \frac{1}{e^{x}+1}$ is the Fermi-Dirac function.

For systems with a critical Fermi surface, in the quantum critical regime where the zero temperature correlation length $\xi$ is so large that it drops out from the universal function, we can write the scaling form of the spectral function (4.2) as

$$A(k,\omega) = \frac{c_0}{|\omega|^2} F_c \left( \frac{\omega}{T}, \frac{c_1 k_{||}}{T^2} \right)$$

with another universal function $F_c$ simply related to the original universal function $F$ in (4.2) via $F_c(x,y) = F(x,y,\infty)$. Applying (4.31) to (4.30), we get

$$\sigma \sim \int d\theta \frac{c_0(\theta)^2}{c_1(\theta)} T^{\frac{1-2\alpha(\theta)}{\xi(\theta)}}$$

where $\theta$ denotes the angular position of a patch on the Fermi surface, and the integral is over all these patches. Therefore, as long as $\alpha < \frac{1}{2}$ for each patch, the interlayer electric conductivity vanishes as the temperature goes to zero. Notice the condition for the interlayer electron hopping to be irrelevant, i.e. $\alpha < 0$, is stronger than this condition, which is of course expected. Both conditions are satisfied for all three cases discussed above, and this confirms our previous assertion that $\sigma \to 0$ as $T \to 0$.

Recall that we have a general inequality (4.4) for a system with a critical Fermi surface, from
which we obtain

\( \frac{1 - 2\alpha}{z} \geq \frac{1}{z} - 2 > -2 \) \hspace{1cm} (4.33)

This implies for such systems, in this perturbative regime, the scaling behavior of the interlayer electric conductivity with respect to temperature cannot be more singular than that of a Fermi liquid

\[ \sigma \lesssim \frac{1}{T^2} \] \hspace{1cm} (4.34)

Now we apply (4.32) to the examples discussed in the previous sections. For all examples, the exponents \( \alpha \) and \( z \) do not depend on its angular position \( \theta \) (even though the non-universal constants \( c_0 \) and \( c_1 \) can depend on \( \theta \) in general), and (4.32) simplifies to \(^1\)

\[ \sigma \sim T^{1-2\alpha/z} \] \hspace{1cm} (4.35)

For the QC regimes of the chemical potential tuned Mott transition (and the related problem of the orthogonal metal transition) \( \), we have \( \alpha = -1 \) and \( z = 2 \). Therefore

\[ \sigma \sim T^\frac{3}{2} \] \hspace{1cm} (4.36)

For the QC regime of the bandwidth controlled Mott transition, we have \( \alpha = -\eta \) and \( z = 1^+ \), so up to logarithms we have

\[ \sigma \sim T^{1+2\eta} \] \hspace{1cm} (4.37)

Before continuing, we comment on the in-plane electric conductivity in these QC regimes. For orthogonal metals, as discussed before, because the fermion \( f \) carries charge and it is in its Fermi liquid state, the in-plane electric conductivity is Fermi liquid like. In the QC regime of the chemical potential tuned Mott transition, the behavior of the in-plane conductivity is non-Fermi liquid

\(^1\)One can similarly get the zero temperature AC conductivity to be \( \sigma(\omega) \sim \omega^{\frac{1-2\alpha}{2}} \).
like. In particular, this conductivity behaves as $\ln\left(\frac{1}{T}\right)$ at the lowest temperatures according to the conventional slave-particle theories [162]. As for the QC regime in the bandwidth controlled Mott transition, if $g > g_c$, the in-plane resistivity has a universal crossover from a finite value to the impurity-induced residual resistivity of the Fermi liquid [157,178].

4.6 Crossovers out of criticalities

As reviewed in Sec. 4.4, a very interesting feature of the continuous Mott transitions is that the crossovers out from the QC regimes to the Fermi liquid regime involve an intermediate regime: there is an IFL regime for chemical potential tuned Mott transition and an MFL regime for bandwidth controlled Mott transition. It is clearly interesting to study the interlayer electric transport properties in these regimes.

As long as at the lattice level the interlayer interactions are much weaker than the intralayer interactions, it may still be legitimate to apply (4.30) to calculate the interlayer electric conductivity. However, unlike that in the quantum critical regime where the zero temperature correlation length $\xi$ drops out from the universal function and we can get the scaling behavior of the conductivity simply by using the scaling form of the electron spectral function $A(k,\omega)$, here we no longer have a simple scaling form for it and we need to calculate it explicitly.

This is done by recalling that the physical electron operator is written as

$$c_{i\sigma} = f_{i\sigma} b_i$$

In both IFL and MFL regimes, despite the Mermin-Wagner theorem (stating that there is no true long-range order at any finite temperature in 2D), as long as $T \ll \rho_s$, the correlation length and correlation time are still extremely large, and the bosons behave as if they already condense. Therefore, to calculate the electron spectral function, we only need to calculate the spinon spectral function and multiply it by the condensate magnitude $|\langle b \rangle|^2 \sim (g - g_c)^{2\beta}$, where $\beta$ is the order parameter exponent of the transition.

The spinon spectral function is determined by the imaginary part of the spinon Green's func-
\[ \mathcal{A}_f(k, \omega) = -\frac{1}{\pi} \text{Im} \mathcal{G}_f(k, i\omega \rightarrow \omega + i\eta) \]  

(4.39)

Writing the (real frequency) spinon Green’s function in terms of spinon self-energy \( \Sigma_f(k, \omega) \), we get

\[ \mathcal{G}_f(k, \omega) = \frac{1}{i\omega - \epsilon_f(k) - \Sigma_f(k, \omega)} \]  

(4.40)

with the bare spinon dispersion

\[ \epsilon_f(k) = v_F k_\parallel + \kappa k_\perp^2 \]  

(4.41)

where \( v_F \) is a (finite) bare Fermi velocity, and \( \kappa \) is determined by the Fermi surface curvature. So the spinon spectral function is given by

\[ \mathcal{A}_f(k, \omega) = \frac{\Sigma''_f(k, \omega)}{-\frac{1}{\pi} \left( \omega - \epsilon_f(k) - \Sigma'_f(k, \omega) \right)^2 + \left( \Sigma''_f(k, \omega) \right)^2} \]  

(4.42)

where \( \Sigma'_f(k, \omega) \) and \( \Sigma''_f(k, \omega) \) are the real and imaginary parts of \( \Sigma_f(k, \omega) \), respectively.

### 4.6.1 IFL in a Chemical Potential Tuned Mott Transition

According to (4.42), to calculate the spinon spectral function, we need to calculate the spinon self-energy. In the IFL regime, the imaginary frequency self-energy of the spinons near a certain patch on the spinon Fermi surface is given by (see Fig. 4.5)

\[ \Sigma_f(k, i\omega) = v_F^2 T \sum_{\Omega_n} \int \limits_q D_c(q, i\Omega) \mathcal{G}^{(0)}_f(k - q, i(\omega - \Omega)) \]  

(4.43)
where the gauge field propagator $D_c(q, iΩ_n)$ takes the form of (4.18), and the bare spinon propagator $G_f^{(0)}(k, iω)$ is given by

$$G_f^{(0)}(k, iω) = \frac{1}{iω - ϵ_f(k)} \tag{4.44}$$

The summation is over the bosonic Matsubara frequencies $Ω = 2πnT$, $n = 0, ±1, ±2, \cdots$.

**Figure 4.5:** Spinon self-energy. The wavy line represents a gauge boson with momentum and frequency $(q, iΩ)$, and the dashed line represents a spinon with momentum and frequency $(k - q, i(ω - Ω))$.

Due to the structure of the bare spinon propagator, the important region of integral involves $q_∥ \sim q_⊥^2 \ll q_⊥$, so we can ignore the $q_∥$ dependence in the gauge field propagator $D_c(q, iΩ)$. After this, the integral over $q_∥$ can be done and we find the spinon self-energy does not have a singular dependence on $k_∥$. This implies, because of the form of the spinon spectral function (4.42) and the form of the interlayer electric conductivity (4.30), having $Σ_f''(k, ω)$ is sufficient to obtain $σ$.

We calculate $Σ_f''(k, ω)$ and find its leading singular part is (see Appendix D.6)

$$Σ_f''(k, ω) \sim \begin{cases} -\frac{T}{√ρ_s}, & |ω| \lesssim T \\ -|ω|^\frac{2}{3}, & |ω| \gtrsim T \end{cases} \tag{4.45}$$

Plugging this result into (4.42) and (4.30), we find the leading dependence of $σ$ on the temperature in IFL regime to be

$$σ \sim t^2|⟨b⟩|^4T^{-\frac{2}{3}} \tag{4.46}$$

This result means in the IFL regime, if the temperature is fixed, when we get closer and closer to the QC regime, the interlayer electric conductivity decreases as $(g - g_c)^{4β}$. On the other hand,
if we fix $g$, which experimentally corresponds to fixing the doping, as the temperature is lowered, the interlayer electric conductivity increases as $T^{-\frac{2}{3}}$, so the out-of-plane transport appears to be conducting. As discussed in Ref. [161], the in-plane conductivity of the IFL regime is non-Fermi liquid like, which behaves as $T^{-\frac{2}{3}}$. Combining these and previous results for the QC regime, this implies on the $g > g_c$ side, in the QC regime, the in-plane conductivity is non-Fermi liquid like, while the out-of-plane electric transport appears to be insulating. When the temperature is decreased and the system enters the IFL regime, the in-plane conductivity is still non-Fermi liquid like, but the out-of-plane electric transport already seems to be conducting but non-Fermi liquid like. If the temperature is further decreased, both the in-plane and out-of-plane electric transports behave like a Fermi liquid (see Fig. 4.6). This phenomenon can be viewed as an experimental signature of the crossover structure displayed in Fig. 4.3.

![Figure 4.6: Schematic crossover behaviors of the resistivity with respect to temperature for chemical potential tuned Mott transition (left) and for bandwidth controlled Mott transition (right). The solid line represents the out-of-plane conductivity and the dashed line represents the in-plane conductivity.](image)

4.6.2 MFL in a bandwidth controlled Mott transition

Similar to the chemical potential tuned Mott transition, we again just need the imaginary part of the spinon self-energy to calculate $\sigma$. The difference between these two cases is that now the gauge field propagator is given by (4.24). Plugging (4.24) into (4.43), we find the leading singular part of $\Sigma_f''(k, \omega)$

$$
\Sigma_f''(k, \omega) \sim \begin{cases} 
-T \ln \frac{T}{\rho_s^2}, & |\omega| \lesssim T \\
-(|\omega| + \lambda T \ln \frac{|\omega|}{\rho_s^2}), & |\omega| \gtrsim T 
\end{cases}
$$

(4.47)
with $\lambda \sim O(1)$ a non-universal constant. To the best of our knowledge, the result for $|\omega| \lesssim T$ is new. Combining this result, (4.42) and (4.30), we get the leading dependence of $\sigma$ on the temperature in MFL regime to be

$$\sigma \sim t^2 |\langle b \rangle|^4 \left( T \ln \left( \frac{T}{\rho_s} \right) \right)^{-1}$$

(4.48)

If we ignore the logarithmic correction in the above result, just as the case of the IFL regime in the chemical potential tuned Mott transition, when the system gets closer to the QC regime from the MFL regime in the bandwidth controlled Mott transition with temperature fixed, the interlayer electric conductivity becomes smaller and smaller as if it vanishes as $(g - g_c)^{4\beta}$. And if we decrease the temperature with $g$ fixed, which experimentally corresponds to fixing the pressure, the interlayer electric conductivity increases as $T^{-1}$, so the out-of-plane transport appears to be conducting. As shown in Appendix D.6, the in-plane electric conductivity in the MFL regime behaves as $T^{-2}$, which is Fermi liquid like. Combining these results and the discussion on the QC regime, we see on the $g > g_c$ side, in the QC regime the in-plane resistivity has a universal jump, while the out-of-plane electric transport seems insulating. When the temperature is lowered and the system enters the MFL regime, the in-plane conductivity already behaves as that of a Fermi liquid, while the out-of-plane conductivity is conducting but non-Fermi liquid like. If the temperature is further lowered so that the system is in the Fermi liquid regime, the temperature dependence of the in-plane conductivity does not change qualitatively, and the out-of-plane conductivity eventually becomes Fermi liquid like (see Fig. 4.6). This phenomenon serves as an experimental signature of the crossover structure displayed in Fig. 4.4.

We end this section by commenting on the validity of the perturbative calculation of $\sigma$. We first discuss the case of the IFL regime. Notice (4.30) is derived perturbatively up to the second order of the interlayer tunneling amplitude $t$, and (4.46) shows $\sigma \to \infty$ as $T \to 0$, so one may wonder whether the higher order terms in $t$ should be included. However, in the IFL regime, we are working in a temperature regime $T \gg \rho_s^2$ and we will not go into an arbitrarily low temperature. As argued in Appendix D.7, in this regime the higher order contributions are expected to be indeed small compared to this leading order contribution, so the perturbative result (4.46) is
valid.

Now we turn to the MFL regime. Again, in the MFL regime $T \gg \rho_s^2$ and we will not go to an arbitrarily low temperature. As argued in Appendix D.7, in most parameter regimes of experimental interests, the perturbative calculation is expected to be valid as long as the interlayer electron tunneling amplitude $t$ is small. If the system is extremely close to the quantum critical point, there may be a narrow window where the perturbative calculation breaks down, and the temperature dependence of $\sigma$ is expected to be more singular than (4.48), but it should not be more singular than the Fermi liquid form $T^{-2}$. Since this window is very narrow, it may not be too significant experimentally.

4.7 Discussion

In this paper, we have demonstrated the phenomenon of dimensional decoupling at continuous Mott transitions between a Fermi liquid metal and a Mott insulator in a multilayered quasi-2D system. At low energies, in the Mott insulating phase as well as right at the quantum critical point, the system behaves as a stack of many decoupled 2D layers, while it behaves as a 3D Fermi liquid in the metallic side. Experimentally, for example, this implies the interlayer electric transport will become insulating in the quantum critical regimes of these transitions. We emphasize the importance role that electron fractionalization plays in obtaining dimensional decoupling, and we also point out, under reasonable assumptions, this phenomenon cannot occur in a non-Fermi liquid obtained near a conventional quantum critical point that is associated with a spontaneous breaking of internal symmetries.

By calculating the temperature dependence of the interlayer electric conductivity $\sigma$ induced by electron tunneling between different layers, we have systematically explored the crossover behavior of the interlayer transport in Sec. 4.6. We find for these continuous Mott transitions the interlayer conductivity vanishes as the temperature goes to zero in the QC regimes, which is consistent with that the interlayer electron tunneling is irrelevant at low energies. Intuitively, this is because an electron will be split into a boson and a fermion across the transitions. To have an electron tunnel from one layer to the other, both the boson and the fermion have to tunnel collectively. However, since the boson is to become gapped across the transitions, this process is
suppressed. One of the interesting features of these continuous Mott transitions is the existence of an intermediate regime when the system crosses over from the QC regime to the Fermi liquid, and we also derived the scaling behavior of $\sigma$ in these regimes and showed $\sigma$ increases as temperature decreases. In particular, in the IFL regime of the chemical potential tuned Mott transition we find $\sigma \sim T^{-\frac{2}{3}}$, and in the MFL regime of the bandwidth controlled Mott transition we find $\sigma \sim (T \ln \frac{1}{T})^{-1}$. This metallic behavior of interlayer transport is because of the partial recombination of the boson and the fermion. When the temperature is low enough so that the system enters the Fermi liquid regime, the boson and the fermion will fully combine and become the original electron, so the system behaves as a coherent 3D metal where both the intralayer and interlayer conductivities behave as $T^{-2}$. Therefore, there will be a coherence peak in the temperature dependence of the interlayer resistivity (see Fig. 4.6). The position of the peak is around where the system crosses out from the QC regimes into the intermediate regimes, so the deeper the system is in the Fermi liquid side, the higher the temperature of this peak. These predictions serve as useful guidance to compare this theory to experiments.

Recently out-of-plane transport of a layered quasi-2D doped organic compound has been studied experimentally [37]. This material is suggested to be a candidate of a doped spin liquid, in the sense that charge and spin will be separated in the absence of doping. As reviewed in Sec. 4.1, this material shows non-Fermi liquid like transport behavior at ambient pressure, and becomes Fermi liquid like under high enough pressure. So pressure can be viewed to drive this crossover from non-Fermi liquid to Fermi liquid, and the higher the pressure is, the deeper the system is in the Fermi liquid side. It is found that there is indeed a regime where the in-plane transport is non-Fermi liquid like, and the out-of-plane transport behaves insulating at high temperatures but becomes metallic at low temperatures. The coherence peak of the interlayer resistivity seems to occur at a temperature that is small compared to the relevant lattice energy scales, and it shifts to higher temperatures as the pressure is increased. All these features agree with the predictions of our theory on the metallic side qualitatively. Ref. [37] also suggested, similar to the considerations in this paper, that the metallic behavior of out-of-plane transport when the in-plane transport is non-Fermi liquid like is due to the recombination of charge and spin.

Finally, we note, as discussed in Sec. 4.5, for a layered quasi-2D non-Fermi liquid, as long as
the electron spectral function is singular enough \((\alpha > \frac{1}{2})\), the out-of-plane transport may be metallic. This scenario can occur in non-Fermi liquids obtained near a conventional quantum critical point where the transition is driven by fluctuating local order parameter, and a possible example is discussed in Appendix D.1. However, in this scenario there is not expected to be a peak in the temperature dependence of the interlayer resistivity as long as the system is in the scaling regime of the transition. Therefore, the observation made in Ref. [37] is unlikely to fall into this scenario.
Field-induced neutral Fermi surfaces and QCD$_3$-Chern-Simons quantum criticalities in Kitaev materials

5.1 Introduction

Exploring exotic quantum phases and phase transitions is one of the frontiers of physics. Traditionally, quantum phases of matter are often distinguished by different patterns of spontaneous symmetry-breaking, and the low-energy excitations in these phases can be understood as long-lived quasi-particles. In the traditional wisdom, quantum phase transitions are often character-
ized by fluctuating order parameters of certain symmetries \cite{1, 2}.

Over the last decades, exotic quantum phases and phase transitions beyond the conventional paradigm are uncovered. The most well established examples may be the fractional quantum Hall (FQH) states, which are nontrivial phases of matter not associated with spontaneous symmetry-breaking. The low-energy excitations in the FQH states are still long-lived quasi-particles, but these quasi-particles are fractionalized, i.e., they can carry fractional statistics and fractional charge \cite{1}.

Another class of long-sought exotic phases are quantum spin liquids (QSLs) \cite{1}. A QSL is a spin system with a nontrivial structure of long-range entanglement in the ground state. QSLs can be broadly classified into two groups: gapped QSLs and gapless QSLs. A gapped QSL has an energy gap that separates its ground state(s) from its excited states, and its low-energy excitations also contain fractionalized long-lived quasi-particles, just as in the FQH states. In a gapless QSL, the dynamics is more complex because the excitations are gapless, and it may not be possible to view these excitations as any type of well-defined quasi-particles.

Despite the great progress in understanding QSLs conceptually and the tremendous effort in searching them experimentally, so far no solid-state material has been unambiguously identified as a QSL. There are two main difficulties in finding a QSL experimentally: it often requires strong and certain particular form of interactions in the system to stabilize a QSL, and it is often difficult to find a smoking-gun signature to decide whether a system is a QSL. Because of these difficulties, it is of vital importance to build and realize realistic models that give rise to a QSL, and to develop effective methods to detect it.

One particularly interesting model of QSLs was proposed by Kitaev \cite{9}. This is a model of localized spin-1/2 particles on a honeycomb lattice, and it features a rather special form of spin-spin interactions. By tuning the parameters in this model, Kitaev showed that this model can realize at least three different types of QSLs: a gapless QSL, a gapped Abelian QSL, and a gapped non-Abelian QSL. These gapped QSLs, if realized, are potentially useful in performing topological quantum computation \cite{179}.

The gapped non-Abelian QSL, more precisely speaking, realizes an Ising topological order (ITO), which has two types of fractionalized excitations: a non-Abelian anyon $\sigma$ and a Majo-
rana fermion. This state can be viewed as a $p + ip$ superconductor where the Bogoliubov quasi-particles therein are coupled to a dynamical $Z_2$ gauge field, and $\sigma$ plays the role of the $\pi$ flux in this superconductor. Just as a $p + ip$ superconductor, the ITO must break time reversal and mirror symmetries. In fact, it has a chiral edge mode of Majorana fermions, and it is supposed to have a quantized thermal Hall conductance $\kappa_{xy} = 1/2$ in units of $(\pi/6)(k_B^2T/h)$. This property offers an experimentally feasible method to detect the ITO.

Because of the peculiar form of the spin-spin interaction in Kitaev’s model, it seems rather unrealistic to be realized in a material. Remarkably, however, it was shown that some Mott insulators with strong spin-orbit coupling can potentially realize some variants of this model [180–182]. Since then, searching for Kitaev spin liquids in these materials becomes an active area. For recent reviews, see Refs. [183–185].

Among the various Kitaev materials, $\alpha$-RuCl$_3$ has received significant attention recently. It is found that the ground state of $\alpha$-RuCl$_3$ is magnetically ordered in the absence of a Zeeman field [186, 187]. Specifically, the spins are ordered in a zigzag pattern [188–190], as shown in Fig. 5.1. The Heisenberg and $\Gamma$ interactions (introduced in Eq. (5.1)) are believed to be responsible for the zigzag order [191, 192], although the signs and strengths of these interactions in the real material are not fully settled down. It is also suggested that the $\Gamma$ term may help to stabilize a QSL [193]. Notice a similar zigzag order has also been found in another Kitaev material, Na$_2$IrO$_3$ [194–196].

Upon applying a Zeeman field, the zigzag magnetic order in $\alpha$-RuCl$_3$ melts [197–200]. If the Zeeman field is strong enough, the system will become a trivial polarized state. Remarkably, the measured thermal Hall conductance in certain range of field strengths is quantized exactly at the value predicted for the ITO [38], and this phenomenon is similar to that in the pure Kitaev model under a Zeeman field. Although there is some subtlety in interpreting this experiment [201–203], and the results therein need to be confirmed by further studies, this discovery has triggered great excitement.

These results give hope that $\alpha$-RuCl$_3$ and other Kitaev materials may realize an ITO under a Zeeman field. Two interesting questions immediately arise:

1) By increasing the field, is the transition from the ITO to the trivially polarized state direct, or is there an intermediate phase?
Figure 5.1: Under a Zeeman field, a Kitaev material may go through four different phases (scenario A): a zigzag ordered state, a non-Abelian chiral quantum spin liquid with Ising topological order (ITO), a neutral Fermi surface (NFS) coupled to an emergent dynamical $U(1)$ gauge field, and a trivially polarized state. The phase transitions between these phases are described by the 1) QCD$_3$-Chern-Simons theory, 2) gauged pairing transition, and 3) gauged Lifshitz transition. Depending on the microscopic details (e.g., direction and strength of the Zeeman field), there can also be a direct continuous quantum phase transition between the ITO state and the trivial polarized state, and the critical theory is described by another QCD$_3$-Chern-Simons theory (scenario B). The QCD$_3$-Chern-Simons transitions require a high symmetry of the system to be stable, which can be satisfied if the magnetic field is in the $ac^*$ plane. All other phases and transitions do not rely on the high symmetries of the system.

2) Is there a theory that describes the quantum phase transition from the ITO to the magnetic order (e.g., zigzag) upon decreasing the field?

We will address these two questions in this paper, and our main results are summarized in Fig. 5.1. Most surprisingly, we find that another QSL with a neutral Fermi surface (NFS) appears between the ITO and the trivially polarized state, as also suggested by Refs. [204–206]. This phase is described by a dynamical $U(1)$ gauge field coupled to Fermi pockets of fermionic spinons around $\Gamma$ and $\pm K$ points of the Brillouin zone (BZ). Notice this dynamical $U(1)$ gauge field is emergent due to the nontrivial entanglement pattern of this state, and it should not be confused with the ordinary electromagnetic gauge field. The spinons, emergent fermionic excitations, are neutral under the ordinary electromagnetic gauge field but interact strongly with the emergent dynamical $U(1)$ gauge field. Initially motivated by the physics of cuprate high-
temperature superconductors [207], this state has been studied extensively over the years [20–29, 208]. It is a very interesting type of QSL: it is gapless and its low-energy excitations cannot be viewed as well-defined quasi-particles, however, it is believed to be stable under any local perturbations, even if there is no symmetry to protect it. This property is very special, because usually a gapless state is stable if it is protected by some symmetry or all its excitations can be viewed as well-defined quasi-particles. We note that another material that also potentially realizes this state is the organic compound, \( \kappa-(\text{BEDT-TTF})_2\text{Cu}_2(\text{CN})_3 \) [209]. Recently, it was argued that YbMgGaO\(_4\) also realizes this state [210], but this proposal is under debate [211].

The NFS state can serve as a mother state of the ITO state and the trivially polarized state. In particular, the ITO state can be obtained from the NFS state by pairing the spinons, and a critical theory for this transition has been discussed in Ref. [29]. The trivially polarized state can potentially be obtained from the NFS state by shrinking the Fermi pockets to zero size. This phase transition is a gauged Lifshitz transition, \( i.e., \) a Lifshitz transition of fermions coupled to a dynamical \( U(1) \) gauge field. Both these transitions are beyond the conventional paradigm because of the coupling to the dynamical gauge field. We will discuss the NFS state and these transitions both numerically and theoretically in Sec. 5.3.

Another extremely intriguing aspect of this system is the possible exotic quantum phase transitions from the ITO state to the zigzag order and to the trivially polarized state, which can be realized by tuning the strength of the Zeeman field [38, 212]. We find that such quantum phase transitions are strikingly different from the conventional phase transitions, owing to the emergence of some deconfined non-Abelian gauge fields. In particular, these quantum critical points mimic the QCD theories in 2+1 dimensions, which have emergent quarks and gluons that are strongly interacting with each other. Moreover, these critical theories have interesting duality properties, namely, they can be described either by critical bosons interacting with a \( U(2) \) Chern-Simons gauge field, or by gapless Dirac fermions interacting with a \( U(2) \) Chern-Simons gauge field. Dualities of interacting gauge theories has generated huge theoretical enthusiasm in both the condensed matter and the high energy community [213–218], and the Kitaev materials may be one of the few experimental platforms [219] to study the duality proposals.

The rest of the paper is organized as follows. In Sec. 5.2, we first review the global symmetries
Figure 5.2: The honeycomb lattice. The $x$-bond, $y$-bond and $z$-bond are along the direction of $a_1$, $a_2$ and $a_3$, respectively. The $a$- and $b$-axes are shown in the figure, and the $c^*$-axis is perpendicular to the paper and pointing outwards.

of some representative Kitaev materials, including $\alpha$-RuCl$_3$, Na$_2$IrO$_3$, etc. This will guide us to study different models in the later parts of the paper. In Sec. 5.3, we first numerically study an idealized model, the Kitaev model under a magnetic field. We find strong evidence supporting the existence of the NFS state under an intermediate field. Then, through a parton construction, we provide theoretical understanding of this state and its transitions to the trivial polarized state and the ITO state. Notice at the time of writing this thesis, some details of this section are still being finalized. To see the final results related to the NFS state, see Ref. [220]. In Sec. 5.4, we discuss the quantum phase transitions from the ITO state to the zigzag phase and to the trivial polarized state, and we find they are described by emergent QCD$_3$-Chern-Simons gauge theories. Finally, we summarize our results and discuss some future directions in Sec. 8.5. The appendices contain various technical details.
5.2 Symmetries of the materials and models

A general Hamiltonian for a Kitaev material under a magnetic field up to nearest-neighbor coupling can be written as

\[
H = \sum_{\langle ij \rangle \in \alpha} K_\alpha S_i^\alpha S_j^\alpha - \sum_i h \cdot S_i + J_H \sum_{\langle ij \rangle \in \alpha} S_i \cdot S_j + \sum_{\langle ij \rangle \in \alpha} \Gamma_\alpha S_i^\beta S_j^\gamma. \tag{5.1}
\]

where the second and third terms are the familiar Zeeman field term and Heisenberg interaction, respectively. The \(K_\alpha S_i^\alpha S_j^\alpha\) and \(\Gamma_\alpha S_i^\beta S_j^\gamma\) terms are often referred to as the Kitaev term and \(\Gamma\) term, respectively. On the bond \(\langle ij \rangle_x\), the \(x\)-bond connecting site \(i\) and site \(j\), they denote terms \(K_x S_i^x S_j^x\) and \(\Gamma_x (S_i^y S_j^y + S_i^z S_j^z)\), respectively. And the similar notation is used for the \(y\)- and \(z\)-bonds. We denote the field direction \(h = (h_x, h_y, h_z)\) as \([h_x, h_y, h_z]\), and the field direction \(h = (-h_x, h_y, h_z)\) as \([\bar{h}_x, h_y, h_z]\).

This Hamiltonian is exactly solvable if only the Kitaev term is present. In this case, it hosts two spin liquid ground states. In particular, if \(|K_x| < |K_y| + |K_z|\), etc., the ground state is a gapless \(Z_2\) QSL with two Majorana cones. Under a small magnetic field, as shown by Kitaev, the Hamiltonian gives rise to an ITO, a non-Abelian chiral QSL ground state. Under a very large magnetic field, the spins are trivially polarized \cite{9}.

Once perturbed away from the pure Kitaev model, the Hamiltonian is no longer exactly solvable. It is still an open issue about the precise values of the coupling constants of each interaction term. In this paper, we will be primarily concerned with the symmetries of the system, and will not worry about the microscopic interaction strengths.

Some representative Kitaev materials, including \(\alpha\)-RuCl\(_3\), Na\(_2\)IrO\(_3\), etc., are layered quasi-two-dimensional materials with point group symmetry \(C_2/m\) \cite{185}.\(^1\) The \(C_2/m\) symmetry constrains that \(K_x = K_y\), \(\Gamma_x = \Gamma_y\). It is expected that \(K_z \sim K_{x,y}\) and \(\Gamma_z \sim \Gamma_{x,y}\), but it is not crucial.

\(^1\)We note that there is debate on the precise symmetry of \(\alpha\)-RuCl\(_3\). Some recent papers claim the symmetry is \(C2/m\) \cite{189, 190}, while some others claim the symmetry should be \(R\bar{3}\) \cite{221, 222}. Here we assume the symmetry is \(C2/m\), and the general method presented in this paper can also be straightforwardly adapted to other symmetry settings.
**Table 5.1:** Symmetries of some representative Kitaev materials (including $\alpha$-RuCl$_3$, Na$_2$IrO$_3$, etc.) under the Zeeman field along different field direction $\mathbf{h} = (h_x, h_y, h_z)$.

<table>
<thead>
<tr>
<th>$\mathbf{h}$</th>
<th>$T_{1,2}$</th>
<th>$C_2$</th>
<th>$\mathcal{T}$</th>
<th>$\sigma^*$</th>
<th>$\mathcal{T}\sigma^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbf{h} = 0$</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>$\mathbf{h} \parallel [11x]$, in $ac^*$ plane</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>$\mathbf{h} \parallel [\bar{1}10]$, parallel to $b$</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>$\mathbf{h} \not\parallel [\bar{1}10], [11x]$</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

for our discussion. Without the Zeeman field, the Hamiltonian enjoys the translation symmetry $T_{1,2}$ along $n_{1,2}$, inversion symmetry $C_2$, pseudo-mirror symmetry $\sigma^*$ (with the mirror axis perpendicular to the $z$ bond, or, equivalently, along the dashed line in Fig. 5.2), and time-reversal symmetry $\mathcal{T}$ ($\mathbf{S} \rightarrow -\mathbf{S}$). The pseudo-mirror symmetry $\sigma^*$ is the conventional mirror symmetry followed by a spin rotation symmetry $e^{i\pi S_y} e^{i\pi/2 S_z}$,

$$\sigma^*: \quad S_r^x \rightarrow -S_{\sigma r}^y,$$  \hfill (5.2)  

$$S_r^y \rightarrow -S_{\sigma r}^x,$$  \hfill (5.3)  

$$S_r^z \rightarrow -S_{\sigma r}^z.$$  \hfill (5.4)

We remark that the spin flip symmetries $e^{i\pi S_\alpha}$ are broken due to the $\Gamma$ terms.

Under a finite Zeeman field, the time reversal symmetry will be broken. The pseudo-mirror symmetry is also broken by a field along a generic direction. There are special directions along which the pseudo-mirror symmetry or its combination with the time-reversal symmetry is preserved. The details are summarized in Table 5.1.

The time reversal symmetry $\mathcal{T}$ or the pseudo-mirror symmetry $\sigma^*$ forbids a finite thermal Hall conductance. Therefore, if the Zeeman field is parallel to the $b$ axis ([\bar{1}10] direction), one cannot have an ITO unless $\sigma^*$ is spontaneously breaking. On the other hand, if the Zeeman field is on the $ac^*$ plane, as is done in the thermal Hall experiments [38], there is no symmetry that forbids the ITO. However, it does not mean that we should expect an ITO for a field in a generic direction on the $ac^*$ plane. After all, if one rotates the field on the $ac^*$ plane, there should be a
phase transition between the ITO and its time-reversal partner. This transition can be direct and continuous, or there can be an intermediate phase, e.g., a $Z_2$ toric code phase, as one rotates the Zeeman field on the $ac^*$ plane. In this paper, we will not pursue in this direction.

Interestingly, the combination of time-reversal and pseudo-mirror symmetry, $T\sigma^*$, is preserved for the field on the $ac^*$ plane ($h \parallel [11x]$). This symmetry is crucial for the stability of the QCD$_3$-Chern-Simons quantum critical points, as we will discuss in Sec. 5.4. The NFS, on the other hand, is not sensible to any symmetry breaking perturbation. We note that Refs. [190, 223] reported that the zigzag order is on the $ac^*$ plane, which means this order preserves $\sigma^*$ and breaks $T\sigma^*$.

5.3 Field-induced neutral Fermi surface

To discuss the main features of the NFS state, we consider an idealized model, i.e., an isotropic Kitaev model ($K_x = K_y = K_z = K$) under a Zeeman field along the $[111]$ direction $h = h(1,1,1)$. Meanwhile, we neglect the $\Gamma$ and Heisenberg interactions. For this special choice of parameters, the system has an extra symmetry, namely, $C_6^* = C_6 \cdot e^{i2\pi/3(S_x+S_y+S_z)/\sqrt{3}}$. Since the NFS is rather insensitive to various perturbations, the main features in this ideal model will still hold after including other realistic interactions and symmetry breaking effects in materials. However, a higher symmetry puts more constraints on the microscopic theory of the gapless spin liquid, and yields more clear predictions that can be tested numerically.

5.3.1 Numerical results

We use the infinite DMRG algorithm [224–226] to simulate the Hamiltonian Eq. (5.1) with $K_x = K_y = K_z = 1, h_x = h_y = h_z = h$, and $\Gamma = J = 0$. There is previous numerical work studying similar Hamiltonians [204–206, 227]. We wrap the honeycomb lattice on a cylinder (as shown in the Fig. 5.3a). There are $L_y$ unit cells around the cylinder ($e_1$ direction), and an infinite number of unit cells along the cylinder ($e_2$ direction). Due to the lack of any conserved quantum number in the Hamiltonian, the simulation is more expensive compared to the familiar spin models (e.g., the Heisenberg model). We carry out simulations for two different sizes, $L_y = 2, 3$ unit cells (i.e.,
Figure 5.3: (a) Cylinder geometry used in DMRG simulations. Here we plot an example with $L_y = 3$ unit cells (6 sites). The sites labeled by the same letters are identified. (b), (c) The dependence of the correlation length $\xi$ on the magnetic field strength $h$ and bond dimension $\chi$ on the (b) $L_y = 2$ cylinder and the (c) $L_y = 3$ cylinder.

$L_y = 4, 6$ sites), and we keep the bond dimension of DMRG up to $\chi \approx 1400 \sim 2000$. The truncation error is reasonably small, $\sim 10^{-10}$ for $L_y = 2$ unit cells and $\sim 10^{-7}$ for $L_y = 3$ unit cells, signifying the convergence of DMRG simulations.

Fig. 5.3(b)-(c) show the behavior of correlation length $\xi$ (along $e_2$) as the magnetic field $h$ changes. Clearly the correlation length is pretty large in a finite region of intermediate field strengths, $h \in (0.31, 0.41)$ for $L_y = 2$ unit cells and $h \in (0.24, 0.36)$ for $L_y = 3$ unit cells. Moreover, the correlation length $\xi$ increases significantly with the bond dimension $\chi$, which is typical for a gapless state in the infinite DMRG simulation. In contrast, in the smaller field (gapped ITO) or larger field region (polarized state), the correlation length does not show any dependence on the bond dimension. This simple observation provides convincing evidence for the existence of a gapless phase in the region of intermediate field strengths. The gapless state is supposed to have a divergent correlation length in the 2D limit, but in the simulations the finite bond dimension $\chi$ introduces a cutoff. Meanwhile, for a given bond dimension, the system with a larger size gets a larger truncation, hence the cylinder with $L_y = 3$ has a smaller correlation length than the one with $L_y = 2$ with the same bond dimension. We think it is the finite size effect that makes the parameter regimes for this gapless phase different on the cylinders with $L_y = 2$ and $L_y = 3$. In the following, we will show that this gapless phase is consistent with an NFS phase, i.e., Fermi surfaces coupled to a $U(1)$ gauge field. For numerical studies on the gapped ITO, we refer the readers to Refs. [204–206].

On a cylinder, the momenta $k_1$ (along $e_1$) are quantized with a spacing $2\pi/L_y$ due to the finite
Therefore, instead of having a genuine 2D state in our system, we expect a number of 1D wires of fermionic spinons labeled by $k_1$, as shown in Fig. 5.4(a)-(b). In the 2D limit, the spinons are forming Fermi surfaces around high symmetry points. On the cylinder, these Fermi surfaces would then be cut by the 1D wires (according to their momenta), yielding gapless modes. In the case of free fermions, the number of gapless modes are simply the number of cuts between the 1D wires and Fermi surfaces. In our case, the fermionic spinons are interacting with a dynamical $U(1)$ gauge field. As shown in Ref. [228], the $U(1)$ gauge field will change the system into a set of coupled Luttinger liquids with central charge $c = N_w - 1$, where $N_w$ is the central charge of the naive free fermion limit.

To extract the central charge of the gapless phase, we use the finite entanglement scaling introduced in Ref. [229]. As shown in the Fig. 5.4(c)-(d), the central charge is $c = 1$ on the $L_y = 2$ cylinder and is $c = 2$ on the $L_y = 3$ cylinder. This indicates that the 1D wires cut the Fermi surfaces twice on the $L_y = 2$ cylinder, and three times on the $L_y = 3$ cylinder. One consistent scenario is shown in Fig. 5.4(a)-(b). On the $L_y = 2$ cylinder, both the $K$ and $-K$ Fermi pockets are cut once by the wires $k_1 = \pi/2$ and $k_1 = -\pi/2$ respectively. On the $L_y = 3$ cylinder, the Fermi pocket around the $\Gamma$ point is cut once by the wire with $k_1 = 0$; while the Fermi pocket around the $K$ ($-K$) point is cut once by the wires with $k_1 = 2\pi/3$ ($k_1 = -2\pi/3$). One may note that the quantized momenta on the $L_y = 2$ and $L_y = 3$ are corresponding to different boundary conditions of fermionic spinons. This is made possible by the fact that the spinons are interacting with a $U(1)$ gauge field. Consequently, a gauge flux $\phi$ can in principle be trapped in the cylinder [230, 231]. Moreover, due to the $C_2$ symmetry, $\phi$ should be quantized to either 0 or $\pi$. In other words, the fermionic spinons have either a periodic or an anti-periodic boundary condition, depending on the microscopic energetics.

Having established a field-induced NFS state in the Kitaev model, it is important to know which operators can detect the $2k_f$ scattering modes of the Fermi surfaces. We calculate the static structure factor of spin operators,

$$D_{xx}(q_1, q_2) = \sum_r e^{-iq_1 r_1 - iq_2 r_2} S^x_{0} S^x_{r}.$$  \hspace{1cm} (5.5)
Figure 5.4: (a), (b): Schematic illustration of Brillouin zone (BZ) and Fermi pockets. The dashed line represent the momenta points accessible on (a) the $L_y = 2$ cylinder and (b) the $L_y = 3$ cylinder. The spinons have anti-periodic boundary condition on the $L_y = 2$ cylinder, and periodic boundary condition on the $L_y = 3$ cylinder. (c), (d): Central charge fitting for the (c) $L_y = 2$ cylinder and (d) $L_y = 3$ cylinder.

where $q_{1,2}$ and $r_{1,2}$ are the components of $q$ and $r$ along $e_{1,2}$, respectively. Fig. 5.5 shows $|D_{xx}(q_1,q_2 = 0)|$ on the $L_y = 3$ cylinder, which clearly has singular peaks at the finite momenta $q_2 = \pm 2k_f$. By detailed examination of the spectra of correlation lengths [231, 232], we conclude these singular peaks correspond to a $2k_f$ scattering mode from the Fermi pockets around the $\pm K$ points. Ideally, there are other scattering modes, such as the $2k_f$ modes at the $\Gamma$ pocket as well as the scattering modes between the $\pm K$ pockets and the $\Gamma$ pocket. However, they do not have any sharp features in our numerical data of static structure factor. We think this is due to the possibility that those modes have a much smaller correlation length, hence their signatures are washed out.

Therefore, correlation functions of the single spin operator $S^\alpha$ can detect the gapless modes of the NFS state. It is in sharp contrast to the $Z_2$ gapless Majorana cone state in the isotropic Kitaev model, in which the simple spin excitation is gapped due to the gapped visons [233]. We expect that experimental techniques such as neutron scattering is able to directly measure the Fermi surfaces of spinons.

Our numerical data could also be consistent with a scenario with only one Fermi pocket around
the $\Gamma$ point. However, such a state at the mean field level will violate the Luttinger volume law, for which we are not able to find a theory. On the other hand, theoretically we could also imagine there are Fermi pockets only around $\pm K$ points, or around both the $\pm K$ and $M_{1,2,3}$ points. However, these are not consistent with numerical results. In the former case, for example, one would always expect an odd central charge, which contradicts with $c = 2$ we obtained on the $L_y = 3$ cylinder.

At last, we remark that, if the size of the $\pm K$ pockets are large, the $k_1 = \pm \pi/2$ wires on the $L_y = 2$ cylinder may also cut through them. If this happens, we expect a larger central charge, for example $c = 3$. Similar scenario could happen for the $L_y = 3$ cylinder, where the $k_1 = \pm 2\pi/3$ wires could also cut the $\Gamma$ pocket if the latter is large enough. However, we have not observed this in our numerical simulations.

Below we turn to understanding our numerical results theoretically.

5.3.2 Parton theory of the NFS state

From Kitaev’s original analysis, it is known that the isotropic Kitaev model perturbed by a weak magnetic field hosts a non-Abelian ITO [9]. In Sec. 5.3.1, we have demonstrated that a differ-
ent QSL state that has a NFS can be stabilized when the magnetic field becomes stronger. This state persists in a range of magnetic fields, and finally becomes a trivial polarized state. These phenomena can be understood in the following picture.

Previous studies indicate that an NFS state can become an ITO through a continuous quantum phase transition by condensing Cooper pairs of fermionic spinons [29], which is a natural scenario in our system when the magnetic field is decreased so that the NFS state transits into the ITO. On the other hand, when the magnetic field is increased, the Fermi surface of the NFS state may shrink and disappear. Then the gauge fluctuations will render the system a short-range entangled state [234], which can be the trivial polarized state. To gain more insights, in this subsection we provide a parton mean field theory for the NFS state and its transitions out. Note to get the full theory beyond the mean field, one must incorporate the coupling between the spinons and the dynamical $U(1)$ gauge field.

Before presenting our parton mean field theory, let us pause for a moment to explain two principles to find such a theory. First, because we would like to have both the NFS and ITO states preserve translation, $C_6^*$ and $T\sigma^*$ symmetries, and to have a scenario that condensing spinon pairs in the NFS state leads to the ITO state, all symmetries of the ITO mean field must also be present in the NFS mean field. Second, we would like to have a parameter regime in which there are Fermi pockets around $\Gamma$, $K$ and $-K$ points of the BZ, respectively.

We will use the standard $SU(2)$ parton construction, which writes the spin operator in terms of fermionic spinon operators [1, 71]:

$$S_i = \frac{1}{2} f_i^\dagger \sigma f_i$$

where $S_i$ represents the spin operator at site $i$, $f_i = (f_{i1}, f_{i2})^T$ is a two-component fermionic spinon operator, and $\sigma$ denotes the usual Pauli matrices. The physical states in the Hilbert space satisfy the constraint

$$f_i^\dagger f_i = 1$$

This parton construction has an $SU(2)$ gauge structure [1, 71], but for our purpose the $SU(2)$
gauge structure will be broken down to $U(1)$, where the gauge transformation reads $f_i \rightarrow f_i e^{i\theta_i}$. We will use the terminology that $f_i$ carries gauge charge 1. Notice this is the charge under an emergent dynamical $U(1)$ gauge field, which should not be confused with the ordinary $U(1)$ electromagnetic field. In fact, these spinons are neutral under the usual electromagnetic field.

This formalism has been applied to study the Kitaev model [235, 236]. In particular, the implementation of symmetries in the Kitaev model is discussed in Ref. [236], where it is shown that $f$ becomes a linear combination of $f$ and $f^\dagger$ under the relevant symmetry transformations, i.e., the symmetry transformations mix positive and negative gauge charges. This type of symmetry actions is unnatural to describe a $U(1)$ quantum spin liquid with an NFS, where a more natural symmetry implementation should either preserve the gauge charge (i.e., take $f$ to $f$) or flip the gauge charge (i.e., take $f$ to $f^\dagger$). If the symmetries flip the gauge charge, it would mean that whenever there is a Fermi pocket around the $\Gamma$ point, there needs to be a hole pocket around the $\Gamma$ point as well. Similarly, there will also be two pockets around $K$ and $-K$ with opposite gauge charges. This leads to more than three pockets in the BZ, which is inconsistent with our numerical results that indicate there are only three pockets, around $\Gamma$, $K$ and $-K$, respectively. Therefore, we will look for a symmetry implementation that preserves the gauge charge of the spinons.

Although the gauge charges are mixed under the symmetry actions in the $SU(2)$ parton formulation described in Ref. [236], these symmetry actions themselves can be changed by utilizing the gauge freedom in the parton construction (5.6). In particular, in Appendix E.1 we present a gauge transformation under which the symmetry actions of the Kitaev model become those in Table 5.2, where we can see explicitly that the gauge charge is preserved under the symmetry actions.

Table 5.2: Symmetry actions on the spinons. Adopting the notations in Ref. [236], $\sigma_{C_6} \equiv \frac{1+i(\sigma_1+\sigma_2+\sigma_3)}{2}$ and $\sigma_{T\sigma^*} \equiv e^{-i\sigma_3\frac{\pi}{4}}$. Notice multiplying these transformation rules by a $U(1)$ phase factor gives rise to the same physical symmetry action in the NFS state, because this does not change the symmetry action on any physical (gauge invariant) operator.

<table>
<thead>
<tr>
<th>$T_{1,2}$</th>
<th>$C_6^*$</th>
<th>$\mathcal{T}\sigma^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f(\mathbf{r}_i)$</td>
<td>$f(\mathbf{r}<em>i + \mathbf{n}</em>{1,2})$</td>
<td>$e^{i\frac{\pi}{4}}\alpha_{C_6}^T f(C_6\mathbf{r}_i)$</td>
</tr>
</tbody>
</table>
After performing this gauge transformation, the Kitaev model rewritten in terms of the spinon
operators $f$ contains two types of terms:

$$H_K = H_{\text{hopping}} + H_{\text{pairing}}$$  \hspace{1cm} (5.8)

The detailed form of these terms are given in Appendix E.2, and here we only note that $H_{\text{hopping}}$
contains only hopping operators of the spinons, \textit{i.e.}, each term in $H_{\text{hopping}}$ preserves the $U(1)$
gauge charge of the spinon. On the other hand, $H_{\text{pairing}}$ contains purely pairing of the spinons,
\textit{i.e.}, each term in $H_{\text{pairing}}$ breaks the $U(1)$ gauge structure to a $Z_2$ gauge structure. Furthermore,
both $H_{\text{hopping}}$ and $H_{\text{pairing}}$ preserve the symmetries in Table 5.2.

The mean field Hamiltonian that we will construct for the NFS state is given by

$$H_{\text{NFS}} = H_{\text{hopping}} + H_h$$  \hspace{1cm} (5.9)

where

$$H_h = -h \sum_i f_i^\dagger (\sigma_1 + \sigma_2 + \sigma_3) f_i$$  \hspace{1cm} (5.10)

represents the Zeeman coupling to a magnetic field in the (111) direction.\footnote{We note that $H_h$ should not be literally taken as the full effect of a Zeeman field, because a Zeeman field can also renormalize the parameters in other terms of the Hamiltonian.} This Hamiltonian also preserves all symmetries in Table 5.2.

The representative dispersions of this four-band mean field Hamiltonian are shown in Fig. 5.6. As we can see, for a range of magnetic fields, an NFS state is realized, where the chemical potential cuts through the second and the third bands (in the first three figures of Fig. 5.6). Furthermore, there are in total three Fermi pockets that are around $\Gamma$ and $\pm K$, respectively. The Fermi pockets around $K$ and $-K$ carry the same gauge charge, which is consistent with the $C_6^*$ and $T\sigma^*$ symmetries. The Fermi pocket around $\Gamma$ carries opposite gauge charge compared to the two at $\pm K$, which is consistent with the Luttinger theorem applied to the mean field Hamiltonian (5.9) subject to the constraint $\langle f_i^\dagger f_i \rangle = 1$. We note that our state is different from the spinon-
Figure 5.6: Representative dispersions of the Hamiltonian (5.9). The parameters we use are $J'_1/J_1 = 3.75$, $J_2/J_1 = -0.75$, and $h/J_1 = 1, 2, 3, 4$ for the four dispersions, respectively. The definitions of the parameters $J_1$, $J_2$ and $J'_1$ are given in Appendix E.2. In all these dispersions, the chemical potential is tuned so that the gauge constraint (5.7) is satisfied at the level of expectation values: $\langle f_i^\dagger f_i \rangle = 1$.

Fermi-surface state in Ref. [237], where the spinons are coupled to a dynamical $Z_2$ gauge field.

The NFS state with gauge fluctuations taken into account is still a stable quantum phase of matter, and its properties have been studied extensively [22, 25–29, 208]. Below we discuss the quantum phase transitions from the NFS state to the trivial polarized state and to the ITO state.

**Transition into the Trivial Polarized State**

Upon increasing the magnetic field, the pockets change in size and shape, and eventually disappear. We note in this process there can be Dirac crossings of the four bands, but such Dirac
crossings are found to always occur away from the chemical potential, so they do not induce a phase transition. When the magnetic field exceeds some critical value where the pockets disappear, the dynamical $U(1)$ gauge field will be confined due to the proliferation of monopoles [234]. The precise nature of the confining state will depend on the quantum numbers of the monopole, which in turn depends on the mean field Hamiltonian of the spinons. In this regime where the pockets disappear upon increasing the field, the two lower filled bands can be smoothly deformed into a trivial atomic limit on the honeycomb lattice with $\hbar \to \infty$. Following the argument in Ref. [238], this means the quantum numbers of the monopole are trivial, and the resulting state after the proliferation of monopoles is consistent with a trivial symmetry-preserving short-range entangled state, i.e., the trivial polarized state.

Notice that as the Fermi volume decreases, there can potentially be other instabilities (towards, e.g., exciton condensation). In this case, there will be an intermediate state between the NFS state and the trivial polarized state.

Therefore, our theory is potentially capable of describing the quantum phase transition from the NFS state to the trivial polarized state by increasing the magnetic field. The putative critical theory for this transition is a gauged Lifshitz transition, where fermions coupled to a dynamical $U(1)$ gauge field are undergoing a Lifshitz transition. We conjecture that the confining state is indeed the polarized state, and this transition can indeed be continuous. More detailed studies of this phase transition are left to the future.

**Transition into ITO**

Now we turn to the transition from NFS to ITO. A usual Fermi liquid is unstable to an infinitesimal attractive interaction, which will induce pairing. According to Ref. [29], however, the gauge fluctuations make an NFS state stable against an infinitesimal attractive interaction among the spinons, and only if this attractive interaction is stronger than a critical value does the NFS state undergo a pairing transition. In our case, the strength of the attractive interaction among the spinons depends on the microscopic details, such as the external magnetic field. It is natural to expect that, at a certain Zeeman field, the attractive interaction becomes critical and the pairing transition takes place.
To gain more insight into this transition, consider adding a pairing term so that the Hamiltonian becomes

\[ H = H_{\text{NFS}} + \lambda_\Delta H_{\text{pairing}} \tag{5.11} \]

where \( \lambda_\Delta \in [0,1] \). This pairing term is supposed to be induced by some attractive interaction among the spinons.

We check the effect of the pairing term numerically. For concreteness, let us focus on the case where \( h = J_1 \) (panel (a) in Fig. 5.6). We find that an infinitesimal \( \lambda_\Delta \) is able to open a gap for the Hamiltonian Eq. (5.11). After this gap is opened, we calculate the Chern number of the ground state using the method in Ref. [239], and find a Chern number consistent with that of a \( p \pm ip \) superconductor.\(^3\) This means by this pairing transition our NFS state can become an ITO, and this phase transition can be a generic second-order transition according to Ref. [29].

We caution the reader that so far we have only shown that after the above pairing transition, the NFS state becomes a non-Abelian ITO, but we have not proved that this ITO is identical to that of the Kitaev model. This is because the mean field state described by Eq. (5.11) may differ from Eq. (5.8) by a symmetry-protected topological (SPT) state under the symmetries in Table 5.2, which, after taking into account the coupling between the spinons and gauge field, can potentially lead to a symmetry-enriched topological (SET) state that is distinct from the one described by the Kitaev model.\(^4\)

To settle down this potential ambiguity, we design an interpolating family of Hamiltonians that connect Eq. (5.11) and Eq. (5.8). More precisely, consider the Hamiltonian family parameterized by \( \lambda_\Delta \) and \( \lambda_h \):

\[ H(\lambda_\Delta, \lambda_B) = H_{\text{hopping}} + \lambda_h H_h + \lambda_\Delta H_{\text{pairing}} \tag{5.12} \]

Clearly, \( H(\lambda_\Delta, 1) \) gives Eq. (5.11), and \( H(1, 0) \) gives Eq. (5.8). Also, notice this entire family of

---

\(^3\)We note that, in the discussion of this type of mean field theory, the difference between \( p + ip \) and \( p - ip \) superconductors is inessential, because one can always apply a parity transformation to the Hamiltonian (5.11) to convert one of them into the other, in a way that the Zeeman field term (5.10) is invariant.

\(^4\)We note there are also examples of distinct phases that become identical when coupled to a dynamical gauge field. For example, see [240–243].
Hamiltonians respect the symmetries in Table 5.2. To connect (5.11) to (5.8), we first increase \( \lambda_\Delta \) from 0 to 1 while fixing \( \lambda_h = 1 \). After this is done, we decrease \( \lambda_h \) from 1 to 0 while fixing \( \lambda_\Delta = 1 \). We find that, apart from the initial point with \( (\lambda_\Delta, \lambda_h) = (0, 1) \), an energy gap is always finite in the entire interpolation, which establishes that Eq. (5.11) and Eq. (5.8) indeed describe the same quantum phase. Therefore, by condensing pairs of spinons our NFS state can indeed transit into the Kitaev spin liquid.

5.4 QCD\(_3\)-Chern-Simons quantum criticalities

In the previous section, we have discussed the quantum phase transition between the NFS state and the polarized state, as well as the transition between the NFS state and the ITO state. Both transitions are beyond the conventional paradigm. In this section, we study the possible exotic quantum phase transitions from the ITO state to the zigzag phase and to the polarized phase. These two states have no topological order, so these transitions can be viewed as confinement transitions of the ITO.

At a first glance, such confinement transitions are rather nontrivial if they can be continuous. To appreciate this, first notice the anyonic excitations in the ITO only include the non-Abelian Ising anyon \( \sigma \) and the Majorana fermion. One common way to confine a topological order is to condense some of its anyonic excitations that have bosonic self-statistics and proper mutual statistics with other anyons. In ITO, however, there is no obvious such (bosonic) anyon that can condense. One may also try to describe the transition in terms of gapless fermions coupled to gauge fields. As mentioned in the introduction, the ITO can be understood as a \( Z_2 \) gauge field coupled to Majorana fermions in a topological band with Chern number \( C = 1 \). To confine the ITO, one needs to first change the Chern number of the Majorana fermions from \( C = 1 \) to \( C = 0 \). This process yields a pure deconfined \( Z_2 \) gauge theory, which is the more familiar \( Z_2 \) toric code state [9]. To get a topologically trivial state, one needs to further confine the pure \( Z_2 \) gauge theory. In other words, one needs two separate transitions to confine the ITO. The first transition is described by a single Majorana cone coupled to a \( Z_2 \) gauge field, and the second transition is the confinement transition of the pure \( Z_2 \) gauge theory, which can be described by an Ising order
parameter coupled to a $Z_2$ gauge field \cite{244}.

The way to make progress, as we will discuss in the following, is to consider dual topological
quantum field theory (TQFT) descriptions of the ITO. More precisely, we will find other gauge
theories that are capable of describing the ITO, such that the confinement transitions of these
gauge theories can be understood either by critical bosons or gapless fermions coupled to the
gauge fields.

5.4.1 **Topological aspects and bosonic critical theories**

To apply this strategy to our case, first recall that the ITO can be viewed as a $p + ip$ superconductor coupled to a dynamical $Z_2$ gauge field that corresponds to the fermion parity symmetry, *i.e.*, the ITO is a gauged $p + ip$ superconductor. Furthermore, there is a 16-fold-way classification
of $2 + 1$ D gapped superconductors coupled to such a $Z_2$ gauge field, where the ITO corresponds
to the state with an index $\nu = 1$ \cite{9}. Suppose we take the superconductor with $\nu = 3$ together
with another superconductor with $\nu = -2$, and weakly hybridize the fermions in these two superconductors, the resulting state is the one with $\nu = 1$.

This observation is useful because it is known that the state with $\nu = 3$ can be described by
an $SU(2)_2$ Chern-Simons theory coupled to a boson. This theory also has two nontrivial anyons:
a non-Abelian anyon $\sigma'$ and a Majorana fermion. In addition, the state with $\nu = -2$ can be described by a $U(1)_{-4}$ Chern-Simons theory coupled to a boson, and this theory has three nontrivial Abelian anyons, with one of them a Majorana fermion. Therefore, we can arrive at the ITO state by taking an $SU(2)_2$ theory and a $U(1)_{-4}$ theory, and hybridizing the Majorana fermions in these two theories. More formally, this hybridization of the Majorana fermions can be viewed as a process of anyon condensation, where the bound state of the Majorana fermions from the $SU(2)_2$ and $U(1)_{-4}$ theories are condensed. In the language of TQFT, the resulting coupled the-

\footnote{We note that due to the coupling to a dynamical $Z_2$ gauge field, the first transition is in a distinct universality class compared to a single free gapless Majorana fermion, and the second transition is also in a distinct universality class compared to the 3D Ising transition.}
ory is denoted as $U(2)_{2,-2}$, and we have derived a known duality \[ U(2)_{2,-2} = \frac{SU(2) \times U(1)_{-4}}{\mathbb{Z}_2}. \] (5.13)

The Lagrangian of the $U(2)_{2,-2}$ theory can be written as

$$L_{\text{CS}} = -\frac{2}{4\pi} \text{Tr}(b db - \frac{2i}{3} b^3) + \frac{2}{4\pi} (\text{Tr}b)(\text{Tr}b).$$ (5.14)

where $b = b + \tilde{b} 1$ is a 2-by-2 $U(2)$ gauge field, with $b$ an $SU(2)$ gauge field and $\tilde{b}$ a $U(1)$ gauge field. 7

This $U(2)$ gauge field is coupled to dynamical bosonic matter fields $\Phi$, so that the total Lagrangian is

$$L = L_{\Phi,b} + L_{\text{CS}} + L_{\text{Maxwell}} - \frac{1}{2\pi} B d(\text{Tr}b) + \cdots$$ (5.15)

Here $\Phi$ may have different flavors, and each flavor can be thought of as a two-component (corresponding to the color index) complex boson, $\Phi = (\phi_a, \phi_b)^T$, which are in the fundamental representation of the $U(2)$ gauge group. The third term $L_{\text{Maxwell}}$ is the standard Maxwell Lagrangian of the gauge field, and at long distances it is less relevant compared to the topological part, Eq. (5.14).

Before proceeding, we pause to comment on the global symmetries of the theory Eq. (5.15) in the absence of the last $\cdots$ term. As a quantum field theory in the continuum, besides the Poincare symmetry, CPT symmetry, etc., this theory also enjoys a $U(1)$ symmetry corresponding to the gauge flux conservation of $\tilde{b}$, as well as an $SU(N_f)$ flavor symmetry. These symmetries may not be present in the physical system, but it is nevertheless helpful to keep track of them.

The microscopic symmetries of the physical system must be embedded into these symmetries, but, a priori, the precise embedding pattern can only be determined after we have a concrete microscopic construction where this field theory emerges at long distances. When specifying to the

---

6In Ref. [218], this theory is denoted as $U(2)_{2,-2}$. In Ref. [245], it is denoted as $U(2)_{2,-4}$.

7This theory can also be written as,

$$L_{\text{CS}} = -\frac{2}{4\pi} \text{Tr}(b db - \frac{2i}{3} b^3) + \frac{4}{4\pi} \tilde{b} d \tilde{b}.$$

118
physical system, we will add appropriate \cdots terms to Eq. (5.15) to break its full symmetries to the physical symmetries. For example, we can add monopole operators of $\tilde{b}$ to break the $U(1)$ flux conservation symmetry, and add certain quartic interactions to break this $SU(N_f)$ flavor symmetry. To keep track of the $U(1)$ symmetry, we have added the fourth term, where $B$ is the probe gauge field of this $U(1)$ symmetry.

The dynamics of the bosonic field $\Phi$ is described by the standard $\phi^4$ theory, with $N_f = 1, 2$ flavors,

$$
\mathcal{L}_{[\Phi, b]} = \sum_{I=1}^{N_f} |(\partial^\mu - i b^\mu)\Phi_I|^2 - m \sum |\Phi_I|^2 - V(\Phi).
$$

(5.16)

where $V(\Phi)$ is the symmetry-consistent quartic potential term.

If the $\Phi$ fields are gapped, they are dynamically trivial and hence can be simply neglected. The theory is then described by the $U(2)_{2-f}$ theory, which is nothing but the ITO. On the other hand, if the $\Phi$ fields are condensed, the $U(2)$ gauge field will be Higgsed, which destroys the ITO. The mass of $\Phi$ is the tuning parameter for this phase transition. In the continuum field theory, it is straightforward to understand the phases when $\Phi$ is condensed.

When $N_f = 1$, the $U(2)$ gauge field is Higgsed down to $U(1)$. Without loss of generality, let us suppose the first color component of $\Phi$ gets a nonzero vacuum expectation value, then only $b_{22}$ is an active gauge field. In the absence of the \cdots term in Eq. (5.15), the Lagrangian describing this remaining gauge field is

$$
\mathcal{L} = -\frac{2}{4\pi} b_{22} db_{22} + \frac{2}{4\pi} b_{22} db_{22} + \mathcal{L}_{\text{Maxwell}} - \frac{1}{2\pi} B db_{22}
$$

$$
= \mathcal{L}_{\text{Maxwell}} - \frac{1}{2\pi} B db_{22}.
$$

(5.17)

So we end up with a $2 + 1$D $U(1)$ Maxwell theory, which is nothing but a Goldstone phase with the $U(1)$ flux conservation symmetry spontaneously broken [246]. In the Kitaev materials, this $U(1)$ symmetry should be explicitly broken, and the monopole operators responsible for this symmetry breaking will gap out the Goldstone mode. Physically, it may be tempting to identify this phase as the zigzag phase. However, the precise nature of this confined state depends on the quantum numbers of the monopoles, which we will discuss in the next subsection.
When $N_f = 2$, the $U(2)$ gauge field will generically be completely Higgsed. The gauge sector is trivial, and the precise nature of the resulting confined state is determined by whether the condensate of $\Phi$ spontaneously breaks any symmetry. Before the condensation, the system has an flavor rotation symmetry between $\Phi_1$ and $\Phi_2$, which can maximally be $SU(2)$. The condensation pattern of $\Phi_{1,2}$ is dependent on the form of the quartic potential $V(\Phi)$ in Eq. (5.17). Specifically, if $V(\Phi)$ is $SU(2)$ invariant, it should have the form $\rho \text{Tr} M^2 + \lambda (\text{Tr} M)^2$, with $M_{IJ} = \sum_a \phi_{Ia} \phi_{Ja}^\dagger$. Here $I, J$ are the flavor indices, and $a$ is the color index. If $\rho, \lambda > 0$, $\Phi$ will condense in the $SU(2)$ invariant channel. In practice, the $SU(2)$ flavor symmetry is absent, but it is still possible that the condensation pattern of $\Phi$ does not break any physical symmetry, depending on the microscopic details. So we can end up with a completely trivial state with no topological order or spontaneous symmetry breaking.

Therefore, we have reached two continuum field theories for the confinement transitions of the ITO, with $N_f = 1, 2$, respectively. In both theories, in order to determine the symmetries of the confining states, we need to understand how the physical symmetries are embedded into the emergent symmetries of Eq. (5.15). Also, we need to know whether the physical symmetries are sufficient to forbid all other possibly relevant operators with respect to these critical theories. In order to do this, a concrete microscopic construction of the critical theory is needed. It turns out to be easier to achieve this goal with a dual fermionic description to Eq. (5.15), as we will discuss below.

Before leaving this subsection, we point out an interesting relation between the theory Eq. (5.15) and the bosonic integer quantum Hall states (BIQH) [247, 248], although this relation is not of vital relevance for the discussions in this paper. The BIQH states are often viewed as bosonic SPTs protected by a $U(1)$ symmetry, but they are in fact compatible with a $U(2)$ symmetry. These states can be labelled by their Hall conductance under the $U(1)$ gauge field corresponding to the protecting $U(1)$ symmetry, $\sigma_{xy} = 2n$ with $n$ an integer (in units such that the state described in Ref. [247] has $n = 1$). The response of the state with $n = -2$ to the $U(2)$ gauge field corresponding to the $U(2)$ symmetry is precisely given by (5.14) [247, 248]. In other words, the ITO can be obtained by gauging two copies of the BIQH states in Ref. [247], which is indeed similar to that the Abelian chiral spin liquid is a gauged (one-copy) BIQH state.
Therefore, the theory Eq.(5.15) can be understood as condensing the bosons in the gauged BIQH states.

5.4.2 Symmetry properties and dual fermionic theories

The bosonic critical theory turns out to be dual to a fermionic critical theory,

$$\mathcal{L} = \sum_{I=1}^{N_f} \bar{\Psi}_I (\partial - i a) \Psi_I + m \sum \bar{\Psi}_I \Psi_I + \mathcal{L}_{\text{top}},$$

$$\mathcal{L}_{\text{top}} = \frac{2 - N_f/2}{4\pi} \text{Tr} \left[ a d a - \frac{2i}{3} a^3 \right] + (4 - N_f) CS_g$$

$$+ \frac{2}{4\pi} \beta d \beta - \frac{1}{2\pi} \beta d(B - \text{Tra}).$$

Here $a$ is a $U(2)$ gauge field, $CS_g$ denotes the gravitational Chern-Simons term, $\beta$ is a dynamical $U(1)$ gauge field, and $B$ is a probe gauge field of the global $U(1)$ symmetry as in the bosonic critical theory. In our convention, when the coefficient of $CS_g$ is 1, the theory has thermal Hall conductance $\kappa_{xy} = 1$ in units of $(\pi/6)(k_B^2 T / \hbar)$, or, in other words, it has an edge with chiral central charge $c_- = 1$. The fermion field $\Psi$ is in the fundamental representation of the $U(2)$ gauge group, and its flavor number can be $N_f = 1, 2$. This duality can be derived using the level-rank duality [218] (see Appendix E.3), and it was also presented in Ref. [251]8.

Here the singlet mass of Dirac fermions $m \sum \bar{\Psi}_I \Psi_I$ is the tuning parameter of the confinement transition. When $m \ll -1$ (in proper units), integrating out the Dirac fermions gives a non-Abelian Chern-Simons theory,

$$\mathcal{L} = \frac{2}{4\pi} \text{Tr} \left[ a d a - \frac{2i}{3} a^3 \right] + 4CS_g + \frac{2}{4\pi} \beta d \beta + \frac{1}{2\pi} \beta d(\text{Tra}).$$

(5.18)

This theory indeed describes the ITO. One might be confused about this statement, since the Chern-Simons levels here look rather distinct from those in Eq. (5.15). However, it is inappropriate to directly compare the Chern-Simons levels between these two theories, because here $a$ is coupled to fermions, while $b$ in Eq. (5.15) is coupled to bosons. After taking into account the difference in the matter fields, we can show that the topological order of Eq. (5.18) is exactly

8The duality only holds for $N_f = 1, 2$ [218].
the same as \( U(2)_{-2} \) Chern-Simons theory, \( i.e. \), the Ising TQFT. We can also do a quick self-consistent check by examining the gravitational response, whose coefficient corresponds to the physical thermal Hall conductance. The non-Abelian Chern-Simons term in Eq. (5.18) can be roughly considered as \( U(2)_{-2} \times U(1)_{-2} \), and integrating out them yields a gravitational Chern-Simons term \(-\frac{7}{2} \text{CS}_g \) \( U(2)_{-2} \) contributes \(-\frac{5}{2} \text{CS}_g \) and \( U(1)_{-2} \) contributes \(-\text{CS}_g \)\). Combined with the \( 4\text{CS}_g \) term in (5.18), the total gravitational response is \( \frac{1}{2} \text{CS}_g \), which is the identical to that of the ITO.

On the other hand, when \( m \gg 1 \), the ITO will be destroyed. As in the bosonic theories, the phases that ITO is confined to depend on the fermion flavor number \( N_f \) and the actions of the physical symmetries in these critical theories. More precisely, when \( N_f = 1 \), the theory will be confined to a pure \( U(1) \) Maxwell theory, in which the monopole will proliferate and the nature of the resulting phase depends on the quantum numbers of the monopole. When \( N_f = 2 \), all the gauge fields will be confined without breaking any symmetry. This gives a trivially polarized state as long as there is no other relevant perturbation that can destroy the critical point. To understand the final fates of the confined states, we need to have concrete microscopic construction of the critical theories.

The above fermionic critical theories motivate a parton construction for the ITO and its confinement transitions. With such an explicit construction, we are able to directly work out the symmetry properties of the field theories. In particular, using our parton constructions, we will show that,

i) The confined phase in the theory with \( N_f = 1 \) can indeed be the zigzag magnetic order.

ii) The confined phase in the theory with \( N_f = 2 \) can indeed be a trivial state with all symmetries preserved.

iii) The symmetries of the representative Kitaev materials (listed in Table 5.1) are sufficient to forbid the most obvious relevant operators in both critical theories (with \( N_f = 1 \) and \( N_f = 2 \), respectively).
To make the symmetries of Kitaev materials manifest, we consider a rotated spin basis,

\[ e S_x = S_x + S_y + S_z \sqrt{3}, \]
\[ e S_y = S_x + S_y - 2S_z \sqrt{6}, \]
\[ e S_z = S_x - S_y \sqrt{2}. \]

Here \( e S_x, e S_y, e S_z \) are parallel to the \( c^*, a \) and \( b \) axis. The parton construction is \([252]\):

\[ \tilde{S}^+ = \phi^\dagger f_a^\dagger f_b^\dagger, \quad \tilde{S}^z = \frac{n_\phi + n_{f_a} + n_{f_b}}{3} - \frac{1}{2}, \]

with a constraint \( n_\phi = n_{f_a} = n_{f_b} \). This parton construction has a \( U(2) \) gauge invariance: \( \Psi = (f_a, f_b)^T \) is in the \( U(2) \) fundamental representation, and it is interacting with a \( U(2) \) gauge field \( a; \phi \) carries charge under the diagonal part of the \( U(2) \) gauge field and a global \( U(1) \) charge (of the \( \tilde{S}_z \) rotation).

To get the ITO, we can put the bosonic parton \( \phi \) into a Laughlin state at \( \nu = -1/2 \), and put the fermionic partons \( f_i \) into a topological band with Chern number \( C = 2 \). This gives exactly the Chern-Simons theory in Eq. (5.18): the fermionic parton contributes a \( U(2)_{-2} \) Chern-Simons term for \( a \), while the bosonic parton is described by a \( U(1)_{-2} \) Chern-Simons term of the \( U(1) \) gauge field \( \beta \).

The confinement transition of ITO can be triggered by changing the Chern number of fermionic partons. Specifically, for a transition from \( C = 2 \) to \( C = 1 \), we get a critical theory with \( N_f = 1 \), while for a transition from \( C = 2 \) to \( C = 0 \), we get a critical theory with \( N_f = 2 \). In Appendix E.4, we provide the concrete mean-field ansatz for these two Chern-number-changing transitions. We only consider a Zeeman field on the \( ac^* \) plane, which is the direction of Zeeman field reported in Ref. [38]. In this case, the symmetries of the system include translation \( T_{1,2} \), inversion \( C_2 \), as well as the combination of time-reversal and pseudo-mirror \( \mathcal{T}\sigma^* \) (see Table 5.1). We also work out how those symmetries are implemented in the critical theories using the mean-field ansatz.

In the theory with \( N_f = 1 \), besides the singlet mass (tuning parameter of the transition), the
Table 5.3: Symmetries of operators in the $N_f = 1$ critical theory. $d(Tra)$ happens to have the same quantum number as $\Psi \gamma^\mu \Psi$.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$C_2$</th>
<th>$\mathcal{T} \sigma^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{M}$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-\mathcal{M}^\dagger$</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^0 \Psi$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^1 \Psi$</td>
<td>$1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^2 \Psi$</td>
<td>$1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
</tr>
</tbody>
</table>

most relevant operators are the monopole operator $\mathcal{M}$, conserved current, $d(Tra)$ and $\Psi \gamma^\mu \Psi$. Their quantum numbers are shown in Table 5.3, and all of them are disallowed by symmetries.

The minimally allowed monopole operator is a two-fold monopole, which may or may not be relevant in the infrared. If it is irrelevant, we may have a stable critical point with an emergent $U(2)$ gauge field. Furthermore, the monopole has exactly the same quantum number as the zigzag magnetic order, assuming the magnetic moments are ordered on the $ac^*$ plane in the zigzag phase, as suggested by Refs. [190, 223]. Therefore, the proliferation of monopoles in the theory with $N_f = 1$ results in precisely the zigzag magnetic order.

We now turn to the critical theory with $N_f = 2$. Besides the $U(1)$ flux conservation, the critical theory also has an $SU(2)$ flavor rotation symmetry. The most relevant operators are the monopole operator $\mathcal{M}_{1,2,3}$, $SU(2)$ adjoint mass $\bar{\Psi} \tau^\alpha \Psi (\tau$ acts on the flavor index), and conserved currents, $d(Tra)$, $\bar{\Psi} \gamma^\mu \tau^\alpha \Psi$ and $\bar{\Psi} \gamma^\mu \Psi$. Here the monopoles are in the adjoint representation of the $SU(2)$ flavor symmetry, and it has three components. Again, we want to work out the quantum numbers of these operators to see if they are forbidden by symmetries. There turn out to be three different cases, depending on the locations of Dirac cones. Constrained by symmetries, the two Dirac points have to stay at the high symmetry points/lines (see Fig. 5.7):

1. The two Dirac cones are at the $M_{1,2}$ points, $(k_1, k_2) = (\pi, 0), (0, \pi)$.

2. The two Dirac cones are at $(k_1, k_2) = (k, k), (-k, -k)$ points ($k$ is an arbitrary number), which are on the high symmetry line $K - K'$.

3. The two Dirac cones are at $(k_1, k_2) = (k, -k), (-k, k)$ points ($k$ is an arbitrary number),
In Appendix E.4, we provide the mean-field ansatz for all the three possibilities, and the quantum numbers of operators are summarized in Table E.1-E.3.

In case (1) (the nodes are located at $M_{1,2}$ points), there is one symmetry allowed operator, $\bar{\Psi}\gamma^0\tau^z\Psi = \Psi_1^\dagger\Psi_1 - \Psi_2^\dagger\Psi_2$. This operator will destabilize the quantum critical point: it will dope the Dirac cones at the $M_1, M_2$ points and generate particle and hole pockets. These two Fermi pockets are interacting with a $U(2)$ gauge field, which may or may not be stable. In cases (2) and (3), again there is one symmetry allowed relevant operator: $\bar{\Psi}\gamma^1\tau^z\Psi$ and $\bar{\Psi}\gamma^2\tau^z\Psi$, respectively. However, different from the first situation, these operators will not destroy the quantum critical points. Instead, they will just move the Dirac points along the high symmetry lines (along either the $K-K'$ or the $M_3-M_3$ line). Therefore, the quantum critical point between the ITO and polarized state may be stable if the two Dirac nodes are staying at the high symmetric lines.

5.5 Discussion

Motivated by the recent theoretical and experimental progress in the research on Kitaev materials, we have studied the phase diagram of Kitaev materials. Within the Kitaev model subject
to a Zeeman field, we have numerically demonstrated that, in addition to the trivial polarized state and the Ising topologically ordered (ITO) state, another exotic quantum spin liquid state with neutral Fermi surfaces (NFS) appears in certain range of the field strengths. This state can be described by Fermi pockets of emergent electrically neutral spinons coupled to an emergent dynamical $U(1)$ gauge field. We have presented a parton construction of this NFS state, and described the quantum phase transitions from this state to the trivial polarized state and to the ITO state. These phase transitions can be understood in terms of Lifshitz transition and pairing transition of the spinons coupled to the emergent $U(1)$ gauge field, respectively.

The real Kitaev materials have very complex structures for the spin-spin interactions [253, 254], and it will be certainly very helpful to understand the ground states in more realistic models. Although our numerical model is based on an idealized antiferromagnetic Kitaev model under a Zeeman field, we expect the qualitative features of our results to hold for various real Kitaev materials, especially the ones that can realize ITO under a finite Zeeman field. Therefore, it is important to experimentally determine whether the real materials realize the exotic phases discussed here.

One of the experimental signatures of the NFS state is that the specific heat and longitudinal thermal conductance (with phonon contributions subtracted) at the lowest temperatures should increase significantly when the system enters this phase from the ITO state. Within this phase, these quantities should qualitatively have power-law temperature dependence with some anomalous exponents. Because of the existence of a neutral Fermi surface, the spin-spin correlation functions will show signatures of $2k_f$ singularities. Therefore, neutron scattering can be used to detect this NFS state, as is done in the material YbMgGaO$_4$ [210]. This can be a method to distinguish the NFS state and the $Z_2$ gapless Majorana cone state in the pure Kitaev model [74, 233, 255]. Perhaps more striking phenomena of this state are the possible quantum magnetic oscillations and charge Friedel oscillations in this electric insulator [256–258]. These oscillating behaviors were first discussed in the context of the organic compound $\kappa$-(BEDT-TTF)$_2$Cu$_2$(CN)$_3$, which is known to be a weak Mott insulator, in the sense that it is close to a metal-insulator transition [259]. On the other hand, Kitaev materials such as $\alpha$-RuCl$_3$ are strong Mott insulators [260, 261], so we expect quantitative differences on these oscillating effects therein from the
We have further studied the transitions from the ITO state to the zigzag order and to the trivial polarized state. We find that these transitions can be described by QCD$_3$-Chern-Simons theories. In particular, the transition between the ITO and the zigzag order (the trivial polarized state) can be described by a dynamical $U(2)$ gauge field coupled to $N_f = 1$ ($N_f = 2$) critical fermions. We have checked that the symmetries of some representative Kitaev materials (listed in Table 5.1) are sufficient to forbid the most obvious relevant operators of these putative critical theories. Therefore, these transitions can potentially be generic direct continuous quantum phase transitions. We note that our method can also be adopted to study the transitions between the ITO and magnetic orders other than the zigzag type.

We also note that these critical theories are dual to $N_f = 1$ ($N_f = 2$) species of critical bosons coupled to a dynamical $U(2)$ gauge field. There is an interesting relation between these critical theories with bosonic integer quantum Hall (BIQH) states. The quantum phase transitions from the BIQH states to a superfluid and to a trivial insulator have been widely studied in recent years [262–264], and it may be worth relating these transitions to the transitions from the ITO state to other states.

The critical theories presented in this paper are all beyond the conventional paradigm of phase transitions, and it is worth studying them in more depth. On the experimental front, it is helpful to examine the phase diagrams of the Kitaev materials more closely, and identify different phases and study the phase transitions. Numerically, it is important to study the phase diagram of more realistic lattice models. In the purely theoretical direction, it will be of interest to study the low-energy dynamics of these critical theories to determine whether these transitions can be continuous, and what the critical exponents are. These studies will provide further insights to the experimental studies of the Kitaev materials. Also, given the similarities among the critical theories between different pairs of phases, it may be interesting to look for a theory of a multicritical point that becomes these phases and critical theories upon adding perturbations. This will potentially lead to unified understanding of the rich structures of the quantum magnetism in these systems.

Finally, we remark that dualities and emergent non-Abelian gauge theories similar to ours may
be useful tools to understand other types of exotic quantum phases and phase transitions in con-
densed matter systems, and we expect more applications of related ideas will arrive in the future
and be proved helpful.
6

Spurious long-range entanglement and replica correlation length

6.1 Introduction

Topologically ordered states are nontrivial gapped states, which are beyond Landau’s symmetry breaking paradigm. [1] Prototypical examples include gapped spin liquid states and fractional quantum Hall states. These states exhibit robust ground state degeneracy depending on the topology of the system, and fractionalized excitations. [265, 266] Even more intriguing is its potential of being a fault-tolerant quantum information processing platform. [267, 268] Recently, long-range entanglement has been appreciated to understand the topological order, as
topologically nontrivial states cannot be connected to a product state by local unitary transformations. [269–272]

Detecting topological order, however, has still been a challenge. The ground state degeneracy and the fractional quantum numbers of the excitations are difficult to measure even in the numerical calculations, let alone experimental situations. Instead, the so-called topological entanglement entropy (TEE) [273–278] is being recognized as an important quantity especially in numerics, for the purpose of distinguishing topologically ordered states from topologically trivial states.

It is believed that the bipartite entanglement entropy of the ground state of a gapped system in two spatial dimensions obeys an “area” law with a constant sub-leading correction. Specifically, the entanglement entropy for a disk of circumference \( L \) is given by

\[
S = \alpha L - \gamma + \cdots
\]

where \( \alpha \) is a model-specific non-universal coefficient, \( \gamma \) is a sub-leading correction, and the ellipses represent terms that vanish in the large \( L \) limit. This constant correction \( \gamma \) is the universal TEE of the state. It is shown that \( \gamma = \log D \), where \( D \) is the so-called total quantum dimension of the system, determined by the anyon content of the topological order that the state represents. Roughly speaking, \( D \) counts the types of fractionalized particles. Since \( D > 1 \) implies the state supports fractionalized excitations, \( \gamma \) is regarded as a smoking-gun signature of 2D topological order.

Remark that the definition of \( \gamma \) through Eq. (6.1) is inherently ambiguous; it depends on fine details of a regularization scheme for the calculation of \( S \). On a lattice, the circumference \( L \) and therefore the subleading term \( \gamma \) vary according to how one counts the number of sites along the boundary of the disk; for example, the circumference of a rectangle that encloses \( L \times L \) sites may be counted as \((L + 2)^2 - L^2 = 4L + 4\) or \( L^2 - (L - 2)^2 = 4L - 4\). To resolve this ambiguity by eliminating the boundary term, Ref. [275, 276] take a linear combination of entanglement entropies for
Figure 6.1: Two methods of extracting TEE. Left: Kitaev-Preskill prescription divides the system into four parts and extract TEE by using (6.2). Right: DMRG calculations put the system on an infinite cylinder and divide the system into two parts, then calculate the entanglement entropy $S(L)$ between the two parts for different circumferences $L$ of the cylinder. Fitting the results into (6.1), TEE is identified as $-S(L = 0)$.

various regions. Concretely, Ref. [275] proposes the following combination.

$$\gamma = S_{AB} + S_{BC} + S_{CA} - S_A - S_B - S_C - S_{ABC} \quad (6.2)$$

where the subscripts refer to the regions specified in Fig. 6.1. In fact, it is this linear combination that enables one to argue that $\gamma$ is a robust quantity under small changes in the Hamiltonian and the region sizes; the combination contains an equal number of terms of opposite signs for each subregion so that a small change in any subregion may be canceled overall. We will refer to this proposal as the Kitaev-Preskill prescription hereafter. Note that it is not too important at long distances whether one uses von Neumann entropy or Rényi entropy. [277,278]

Despite of its conceptual importance, the Kitaev-Preskill prescription of extracting TEE is not of great practical use because it requires each partition be much larger than the correlation length of the system. [279] This is very challenging especially in density-matrix-renormalization-group (DMRG) methods. Alternatively, by exploiting the fact that the DMRG algorithm systematically produces minimally entangled states, it is proposed that one can simply put the system on infinite cylinders with various circumferences $L$ (see Fig. 6.1), and extrapolate the data using (6.1) to read off $\gamma$. [280] The mentioned ambiguity of defining the circumference of a disk on lattices, does not apply here since the circumference of a cylinder is simply well-defined. We
will refer to this method as the **cylinder extrapolation method**. An important advantage of this method is that one can regard the region size as large as $L$, rather than some small fraction of $L$, and hence can expect finite size effects to be very small. This method might be useful in light of recent experimental developments as there are proposals to measure entanglement entropies, and experiments have already been performed on simple cold atom systems. [281,282]

However, the cylinder extrapolation method seems to yield inconsistent results. By applying the cylinder extrapolation method to the $J_1$-$J_2$ antiferromagnetic Heisenberg model on a square lattice, it is found that $S(L = 0) \simeq -\ln 2$ in a certain parameter regime, [283] and thus the ground state is identified as a topologically ordered spin liquid. However, this result was later objected by another DMRG study from an independent group, revealing a plaquette valence-bond order. [284] Furthermore, Ref. [285] studies the Heisenberg model on the honeycomb lattice, and reports $S(L = 0) \simeq -\ln 2$ by the cylinder extrapolation method, but observes a plaquette valence-bond order, which leads to a suspicion of finite size effects on $S(L = 0) < 0$. These numerical results question the validity of the cylinder extrapolation method.

In this paper, we point out one scenario under which the cylinder extrapolation method can be proven to be invalid. While we do not attempt to address specific reasons behind the discrepancy between Ref. [283] and [284], we will give a sufficient condition for topologically trivial 2D states, under which the cylinder extrapolation method must give a nonvanishing sub-leading term for any Rényi entropy calculations, leading to a spurious TEE.

We start with a general observation that when a two-dimensional state is topologically trivial, the entanglement computation in Fig. 6.1(b) reduces to that of a one-dimensional state with respect to an extensive bipartition. We show that whenever the derived one-dimensional state exhibits a symmetry-protected topological (SPT) order under a product group, which we define precisely below, then the cylinder extrapolation method must output a nonzero sub-leading term.

Under generic perturbations that break the symmetry of the reduced one-dimensional states in our examples, the sub-leading term is suppressed exponentially in the system size. Furthermore, if one applies the Kitaev-Preskill prescription in the bulk, one still obtains a value that is consistent with the total quantum dimension of the underlying topological particle content. Hence, it is improper to say that the notion of topological entanglement entropy is invalidated. Rather,
our examples make it clear that one has to be careful in interpreting results from the cylinder extrapolation method.

To decide when the results from the cylinder extrapolation method can be trusted, we consider a length scale, termed replica correlation length $\xi_\alpha$. The ratio $L/\xi_\alpha$ determines the magnitude the sub-leading term in the cylinder extrapolation method. This replica correlation length may be arbitrarily large compared to the usual correlation length of the 2D state. The usual correlation length of a state $\rho$ is the decay rate of (the connected part of) a two-point function $\text{Tr}(\rho O O')$ as a function of the distance between local operators $O$ and $O'$. In contrast, the replica correlation length is the decay rate of a two-point replica correlation function $\text{Tr}(\rho O_1 O'_1 \rho O_2 O'_2)/\text{Tr}(\rho^2)$ where the unprimed operators and primed operators are far separated.

At the first glance on its definition, the replica correlation length may seem difficult to calculate. We propose a relatively simple way of measuring the replica correlation function in numerics, as a natural extension of the measurement of Rényi entropies using swap operations. [286] We find that a 2D cluster state (graph state [287]), which is topologically trivial, has an infinite replica correlation length, while certain representative wave functions of the $\mathbb{Z}_N$ gauge theory and the double-semion theory have replica correlation length zero. We conjecture that the $\gamma$ value from the cylinder extrapolation method is given by the total quantum dimension whenever the replica correlation length is small.

Some of our examples can be adapted to three or higher dimensions [288] and give similar effects. It would also be interesting to consider thermal states. [289–291] We will comment on these in the discussion section.

The rest of the paper is organized as follows. In Sec. 6.2, we study the 2D cluster state on a triangular lattice, an exactly soluble model of topologically trivial 2D states, and show that the sub-leading term of the entanglement Rényi entropy calculated by the cylinder extrapolation method is nonzero. This example illustrates important points and serves as a warm-up for the more general consideration. In Sec. 6.3, we study the Rényi entropies of a class of topologically trivial 2D states and show that the sub-leading term is strictly negative, if it is calculated using the cylinder extrapolation method. In particular, we will attribute the nonzero sub-leading term to a SPT order under a product group of the derived one-dimensional state. In Sec. 6.4,
we consider generic non-symmetric states and introduce the replica correlation length, that is responsible for the non-vanishing sub-leading term of entanglement entropy. We calculate the replica correlation length to show that it is infinite for the 2D cluster state on triangular lattice, while it is zero for some ideal wave functions of the $Z_N$ gauge theories and double-semion theory. We conclude in Sec. 7.9 with discussion on higher dimensions and thermal states. Appendices include further considerations. In App. F.1, we provide an example where the sub-leading term calculated by the cylinder extrapolation method oscillates with the system size. In App. F.2, we discuss lattice symmetries and time reversal symmetries, and find that these symmetries are not responsible for a robust non-vanishing sub-leading term by the cylinder extrapolation method. App. F.3 contains calculation of the mutual information of the cluster state at nonzero temperature.

6.2 Example: 2D cluster state on a triangular lattice

In this section we study an example of topologically trivial 2D state that nevertheless exhibits nonvanishing sub-leading term of entanglement entropy under the cylinder extrapolation method. We start with a triangular lattice with one spin-1/2 (qubit) $\{|0\rangle, |1\rangle\}$ per lattice site, governed by a Hamiltonian

$$H_0 = -\sum_j \sigma_j^z$$

(6.3)

The ground state $|\psi_0\rangle$ of $H_0$ is a product state and there is no entanglement. Clearly, this state is topologically trivial. Next, we apply to $|\psi_0\rangle$ a layer of local unitary transformations $U_{jk}$ for each pair $\langle jk \rangle$ of nearest neighbor qubits. The two-qubit unitary $U = U_{jk} = U_{kj}$ is most conveniently defined in a basis where $\sigma^z$ is diagonal:

$$U |a\rangle |b\rangle = \begin{cases} -|1\rangle |1\rangle & \text{if } a = b = 1, \\ +|a\rangle |b\rangle & \text{otherwise.} \end{cases}$$

(6.4)

Since $U_{jk}$ are simultaneously diagonal in a basis, they commute with each other

$$U_{jk}U_{j'k'} = U_{j'k'}U_{jk}.$$  

(6.5)
Hence, there is no ambiguity in the formula

\[ U = \prod_{(jk)} U_{jk}, \quad (6.6) \]

and we define a state \( |\psi\rangle = U |\psi_0\rangle \), which is called the cluster state. Using the identity

\[ U_{jk}(\sigma_j^x \otimes I)U_{jk}^\dagger = \sigma_j^x \otimes \sigma_k^z, \quad (6.7) \]

we see that the cluster state is a ground state of a Hamiltonian

\[ H = U H_0 U^\dagger = - \sum_j \left( \sigma_j^x \prod_{k:(jk)} \sigma_k^z \right). \quad (6.8) \]

Graphically, each term in the summation of the new Hamiltonian (6.8) is the 7-spin interaction as shown in Fig. 6.2.

**6.2.1 Entanglement entropy on a cylinder**

Since we obtain the cluster state from a product state by a small depth quantum circuit, the cluster state is also topologically trivial. One may expect its bipartite entanglement entropy has a vanishing sub-leading term. This is indeed the case if we use the Kitaev-Preskill prescription
to extract the sub-leading term. Now let us examine it using the cylinder extrapolation method.

We put our state on an infinite cylinder by imposing a periodic boundary condition along one of three directions parallel to any side of a triangle. If the circumference is $L$, the number of bond cuts is $2L$. We will compute the entanglement entropy between the two sides $A, B$ divided by this circumference (see Fig. 6.3).

![Figure 6.3: Triangular lattice with an entanglement cut parallel to one of the sides of a triangle. The upper region is $A$, and the lower is $B$.](image)

Even though we started with a 2D state, the entropy computation reduces to that of a 1D chain with an extensive bipartition. To see this, recall that the entanglement entropy is invariant under any unitary that acts exclusively on either side of the bipartition:

$$S = S(\rho_A) = S(U_A \rho_A U_A^\dagger) = S(\rho_B) = S(U_B \rho_B U_B^\dagger)$$

where the subscript $A$ or $B$ denotes the region on which the operator is supported. In particular, we can choose $U_A$ to be the product of all $U_{jk}$ where the edge $\langle jk \rangle$ belongs to $A$. Since $U_{jk}^2 = I$, this amounts to disentangling $|\psi\rangle$ on the region $A$. A similar disentangling unitary can be applied on $B$. What is left is a zig-zag 1D chain that straddles two regions along the cut, and some completely disentangled qubits in the product state. See Fig. 6.4. It remains to compute the en-
tanglement entropy of the 1D chain $|\psi_1\rangle$, which is a ground state of

$$H_1 = -\sum_{j=1}^{L} \sigma^x_{j,A} \sigma^z_{j-1,B} \sigma^z_{j,B} - \sum_{j=1}^{L} \sigma^x_{j,B} \sigma^z_{j,A} \sigma^z_{j+1,A}$$

$$= -\sum_{k=1}^{2L} \sigma^x_{k-1} \sigma^z_{k} \sigma^z_{k+1}$$  \hspace{1cm} (6.10)

where a periodic boundary condition is imposed, i.e., $j = L + 1$ site is equal to $j = 1$ site. Note that since the Hamiltonian is commuting and the ground state is non-degenerate, any correlation function is identically zero beyond distance 2, and hence the correlation length vanishes.

Figure 6.4: The reduced 1D chain of the 2D cluster state. The red circles represent qubits in region A and the green circles represent qubits in region B.

One might guess that the entanglement entropy of $|\psi_1\rangle$ is just proportional to the number of bond cuts, because the entangling unitary $U_{jk}$ makes the product state $|+\rangle_j |+\rangle_k$ into a maximally entangled state, where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. However, a more careful and direct computation reveals that

$$S(A) = (\log 2) L - \log 2.$$  \hspace{1cm} (6.11)

This computation can be done by exploiting the fact that the Hamiltonian $H_1$ consists of commuting tensor products of Pauli matrices. It is known that the eigenvalue spectrum of the reduced density matrix $\rho_A$ for any set of qubits $A$ consists of a single nonzero value (flat entanglement spectrum), and the number of nonzero eigenvalues is always $2^k$ for some integer $k \geq 0$. The exponent $k$ depends only on the number of operators $P$ that are products of Pauli matrices and stabilize the state, $P |\psi\rangle = |\psi\rangle$; more precisely, $k$ equals the number of qubits in the region $A$. 

137
minus the logarithm of the order of the stabilizer group $\mathcal{G}_A$ supported on $A$,

$$k = |A| - \log_2 |\mathcal{G}_A|.$$  \hfill (6.12)

See Section 6.5.1 for a simple proof. In the present example, there are $L$ qubits in $A$, and there is a single non-identity stabilizer $\prod_{j=1}^{L} \sigma^x_{j,A}$ supported on $A$, and hence Eq. (6.11) follows. From Eq. (6.11), one may mistakenly conclude that $\gamma = \log 2$ from the cylinder extrapolation method and that the original 2D state is topologically ordered.

In the next section, we explain why such a topologically trivial state can give rise to a nonzero sub-leading term in the entanglement entropy calculated by the cylinder extrapolation method. We will attribute the nonzero sub-leading correction term $\gamma = \log 2$ to the nontriviality of the state $|\psi_1\rangle$ as a SPT order under symmetry $\mathbb{Z}_2 \times \mathbb{Z}_2$, where the first $\mathbb{Z}_2$ factor acts by $\sigma^x$ on the red side of the entanglement cut, and the second $\mathbb{Z}_2$ factor by $\sigma^x$ on the green side of Fig. 6.4.

### 6.2.2 Reduction to 1D state and its symmetries

Before proceeding, we remark that for any state constructed from a product state by a small-depth quantum circuit the entanglement entropy calculation reduces to that of a 1D chain with an extensive partition. Therefore, our consideration of 1D chains is appropriate and general to study the bipartite entanglement entropy of 2D topologically trivial states. The proof of this remark is simple. One can remove all entangling unitaries except for those near the cut without changing the entanglement. The unitaries that cannot be removed are supported on a strip whose width is proportional to the number of layers in the quantum circuit. Hence, one finally arrives at a quasi-1D system with a bipartition along the extended direction. Note that this simple argument proves the area law of entanglement entropy for such 2D states.

We also emphasize that the symmetry of the resulting 1D state is not necessarily the symmetry of the original 2D state. The disentangling deformations have no reason to obey any symmetry of the 2D state; their role is purely to transform the 2D state to a 1D chain immersed in a product state background. This means that even if the resulting 1D state is symmetric or close to symmetric, this symmetry is not necessarily visible in the 2D state. As an example, one
can consider the deformation of the 2D triangular lattice cluster state in the form discussed in Sec. 6.4. One could deform the bonds arbitrarily except those that are crossed by the entanglement cut such that there is no on-site symmetry. After the disentangling transformations for the entanglement entropy evaluation, one can find $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry of the 1D chain.

Because the symmetry of the reduced 1D state is obscure from the original 2D system, we will introduce a replica correlation length in Sec. 6.4, 6.5 below. It is distinct from the usual correlation length but can be checked directly for the original 2D system without investigating a hidden 1D state.

### 6.2.3 Bravyi’s example

Bravyi has considered the cluster state on a circular zig-zag chain on a plane, where an entanglement cut is chosen such that exactly half of the chain is inside the cut.[292] The rest of the plane is assumed to be occupied by qubits in a trivial product state. Dividing the disk into three circular sectors (see Figure 6.5), and taking the Kitaev-Preskill combination, one will find that $\log 2$ remains. This is, as far as we know, the first example in which Kitaev-Preskill combination can be nonzero for a topologically trivial state. Levin-Wen combination gives no different answer. The state is highly inhomogeneous, and the partition must be introduced very carefully. In contrast, our example is manifestly translation-invariant, and more relevant to current DMRG methods. Bravyi’s example and ours are of course closely related: If we cap off one end of the cylinder to turn it into a topological plain, then our zig-zag chain and partition becomes those of Bravyi’s.

### 6.3 A sufficient condition for a nonzero sub-leading term of entanglement entropy under the cylinder extrapolation method

In this section, we will provide a sufficient condition for a class of topologically trivial 2D states under which these states nevertheless give a nonzero sub-leading term of entanglement entropy under the cylinder extrapolation method. Since the bipartite entanglement entropy of any topologically trivial 2D states is identical to that of a reduced 1D chain under an extensive bipar-
Figure 6.5: Bravyi’s example. There is a qubit on each vertex of the zig-zag chain, and the dashed line represents an entanglement cut that divides the chain into two halves. If the chain is in the cluster state of Eq. (6.10), the entanglement entropy is equal to \( \frac{1}{2} \#(\text{bond cuts}) - 1 \) in the units of \( \log 2 \).

We will show for any nontrivial 1D SPT under a product group symmetry \( G_1 \times G_2 \) (defined more precisely below), there must be a nonzero negative sub-leading term to the entanglement entropy with respect to a bipartition where the partition \( i = 1, 2 \) includes all degrees of freedom acted on by \( G_i \). In Fig. 6.4, for example, \( G_1 \) acts on the red sites and \( G_2 \) acts on the green sites. Here the entanglement entropy is measured by the Rényi entropy

\[
S_\alpha(\rho) = \frac{1}{1 - \alpha} \log \text{Tr}(\rho^\alpha),
\]

and our state is assumed to have a matrix-product state (MPS) representation with a finite virtual bond dimension. We restrict ourselves to Rényi entropy of integer indices \( \alpha = 2, 3, 4, \ldots \). The von Neumann entropy \( (S_{\alpha \to 1}) \) will not be treated explicitly.

6.3.1 Nontrivial SPT

First, we specify what a nontrivial SPT under a product group is. Recall that 1D SPTs are classified by the second group cohomology \( H^2(G; U(1)) \), which enumerates all equivalence classes of
factor systems

\[ \omega : G \times G \rightarrow U(1) \]  

(6.14)

obeying the cocycle condition

\[ \frac{\omega(b, c)\omega(a, bc)}{\omega(ab, c)\omega(a, b)} = 1 \quad \text{for all } a, b, c \in G \]  

(6.15)

up to exact cycles defined by \( \delta \lambda(a, b) := \lambda(a)\lambda(b)/\lambda(ab) \) for some function \( \lambda : G \rightarrow U(1) \). \[293,294\]

Now, suppose the symmetry group is \( G = G_1 \times G_2 \), and each component acts on different physical qudits. Graphically, \( G_1 \) acts on the red sites in Fig. 6.4 and \( G_2 \) acts on the green sites there. We say an SPT state under a product group is nontrivial if its associated factor system \( \omega \) admits

\[ \Omega := \frac{\omega(a, b)}{\omega(b, a)} \neq 1 \quad \text{for some } a \in G_1, b \in G_2. \]  

(6.16)

Note that \( \Omega \) is independent of the cohomology representative \( \omega \) since multiplying \( \omega \) by an exact cycle does not change \( \Omega \). Since any factor system gives rise to a projective representation \( V \), \( \Omega \neq 1 \) means that a commuting pair of elements \( a, b \) of \( G \) are represented by a pair of non-commuting unitaries:

\[ V_a V_b = \frac{1}{\omega(a, b)} V_{ab} = \frac{1}{\omega(a, b)} V_{ba} = \frac{\omega(b, a)}{\omega(a, b)} V_b V_a \neq V_b V_a \]  

(6.17)

In the matrix product state (MPS) representation, this projective representation appears in how the symmetry action is implemented on the virtual level. That is, if the local tensor for a translation-invariant SPT state is given by

\[ M = \sum_{p, v, w} M_{vw}^{(p)} |p\rangle \otimes |v\rangle \langle w|, \]  

(6.18)

the state is acted on by the symmetry such that

\[ \sum_{p'} (U_g)^p_{p'} M^{(p')} = \eta_g V_g^{\dagger} M^{(p)} V_g \]  

(6.19)

for any \( g \in G \), where \( \eta_g \) is a phase factor. \[295\] Thus, if the SPT is nontrivial, the commuting
symmetry actions are lifted to non-commuting unitaries on the virtual level.

Eq. (6.19) is most conveniently expressed in diagrams. If $U$ is a matrix representing the symmetry action of $G_1$, it is

$$U = V' V$$

(6.20)

where we have drawn two lines for the physical qudit, to emphasize that we assumed two components of the symmetry group act on distinct physical qudits. We have omitted the phase factor $\eta_g$. If $U$ were representing $G_2$, the box of $U$ on the left-hand side would have been on the second vertical line. Typically in literature, a box is inserted at the intersection of lines to signify a tensor $M$, but we omitted it.

### 6.3.2 Transfer matrix

Consider a 1D chain on a ring. Suppose there are $L$ physical sites with two physical qudits per site, and the two qudits on each site will be acted on by $G_1$ and $G_2$, respectively. Tracing out one qudit for every site, we have a reduced density matrix $\rho$. In a diagram, $\rho$ is depicted as Figure 6.6(a). An integral power of the reduced density matrix can also be represented by a diagram. For example, $\rho^2$ is depicted in Figure 6.6(b).

For Rényi entropy computations, we need to evaluate $\text{Tr}(\rho^\alpha)$ for a positive integer $\alpha$. This amounts to contracting all top vertical bonds with the bottom vertical bonds in Figure 6.6. It is instructive to look at $\rho^2$. Due to the 1D structure, it is useful to analyze transfer matrix $T$ defined by the left-most diagram in Figure 6.7. For each integer Rényi index $\alpha$, there is a transfer matrix $T_\alpha$, which is independent of system size $L$. Note that the rows and columns of the reduced density matrix $\rho$ are indexed by the physical qudits, whereas those of the transfer matrix are indexed by the virtual bonds. We have a trivial yet useful identity:

$$\text{Tr}(\rho^\alpha) = \text{Tr}(T_\alpha^L).$$

(6.21)

If the eigenvalues of $T$ are $\{\lambda_i\}$, then $\text{Tr}(T^L) = \sum_{i=1}^{D^{2\alpha}} \lambda_i^L$, where $D^{2\alpha}$ is the size of the matrix $T$.

There is no guarantee that $\lambda_i$ are all positive; indeed, in Appendix F.1 we give an example
Figure 6.6: (a) $\rho$ obtained by tracing out one physical qudits for every site. (b) $\rho^2$. Connected bonds are contracted. Wiggling lines represent complex conjugation. The horizontal virtual bonds are contracted due to the periodic boundary condition, which is not drawn. $\text{Tr}(\rho)$ and $\text{Tr}(\rho^2)$ are computed by contracting the upper vertical wiggling bonds with the lower straight ones. For this purpose, it is enough and more efficient to consider the transfer matrix designated by the dotted rectangle; see Eq. (6.21).
whose nonzero eigenvalues of $T_2$ are $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}$. In this case, the entanglement entropy is

$$S_2(L) = L \log 2 - \begin{cases} 
\log 4 & \text{if } L \text{ is even,} \\
\log 2 & \text{otherwise.}
\end{cases} \quad (6.22)$$

Summarizing, the Rényi entropy for $L$ sites is

$$S_\alpha = \frac{\log(1/\lambda_1^{(\alpha)})}{\alpha - 1} L - \frac{\log m}{\alpha - 1} + \cdots \quad (6.23)$$

where $\lambda_1^{(\alpha)}$ is the largest eigenvalue of $T_\alpha$, which is necessarily real positive, $m = m(L)$ is an integer which may depend on $L$, and $\cdots$ represents vanishing terms in the large $L$ limit. The number $m$ is usually the degeneracy of the largest eigenvalue of $T$.

Figure 6.7: Transfer matrix and its symmetry. The left-most diagram of (b) represents the transfer matrix $T_3$ for $\rho^3$. The ensuing equalities are direct consequences of the symmetry lifted to the virtual level. This implies that the transfer matrix $T_\alpha$ has degeneracy $q^{\alpha-1}$ for some integer $q > 1$.

6.3.3 Degeneracy of the Transfer Matrix

We now use the nontrivial SPT to show that the transfer matrix $T_\alpha$ has degeneracy $m \geq q^{\alpha-1}$ for some integer $q > 1$. The degeneracy bound is uniform to every eigenvalue. This implies that

$$\gamma_\alpha = \frac{\log m}{\alpha - 1} \geq \log q > 0. \quad (6.24)$$

To this end, we rewrite the symmetry lifting Eq. (6.20) as in Figure 6.7(a), by which we define the unitaries $V$ and $W$ up to phase factors. The diagrams in Figure 6.7(b) follows at once. The
nontrivial SPT implies that

\[ WV = \Omega V W, \quad \Omega = \exp(2\pi ip/q) \neq 1 \quad (6.25) \]

where \( p \) and \( q > 1 \) are coprime.

Let us define \( X_j = W_{2j-1}^{*} \otimes W_{2j} \) and \( Y_j = V_{2j} \otimes V_{2j+1}^{*} \), where we have indexed the virtual bonds from the top to the bottom by integers modulo \( 2\alpha \). The index \( j \) takes values \( 1, \ldots, \alpha \). The symmetries \( X_j, Y_j \) of \( \mathbb{T} \) form an algebra obeying the following commutation relations.

\[ X_j Y_j = \Omega Y_j X_j \quad (j = 1, \ldots, \alpha) \quad (6.26) \]

\[ X_{j+1} Y_j = \Omega^{-1} Y_j X_{j+1} \quad (X_{\alpha+1} = X_1). \]

All other commutators among \( X_i, Y_j \) are vanishing. They can be rearranged as follows to determine a minimal representation. Since \( X_1 X_2 \cdots X_\alpha \) and \( Y_1 Y_2 \cdots Y_\alpha \) are in the center of the algebra, we take out \( X_\alpha \) and \( Y_\alpha \) from the generating set, and do not consider them any more.

\[ Z_j := X_1 X_2 \cdots X_j \quad \text{for } j = 1, \ldots, \alpha - 1 \]

\[ Z_j Y_j = \Omega Y_j Z_j \quad (6.27) \]

\[ [Z_j, Y_{j'}] = 0 \quad \text{if } 1 \leq j \neq j' \leq \alpha - 1 \]

Since \( Z_j, Y_j \) generate the same algebra as \( X_j, Y_j \) do, it is clear that the minimal representation of the symmetry algebra generated by \( X_i, Y_i \) has dimension \( q^{\alpha-1} \) where \( q \) is the multiplicative order of \( \Omega \). It follows from the nontrivial SPT assumption that \( q > 1 \). We have proved Eq. (6.24) for any integer \( \alpha > 1 \). Therefore, we have found the promised condition that if the reduced 1D chain is a nontrivial SPT under a product group, the cylinder extrapolation method will give a nonzero sub-leading term of entanglement entropy.

Note that we have not used all physical symmetry elements. Sufficient is only one pair of commuting physical symmetry operators that are lifted to non-commuting virtual unitaries. The actual degeneracy may be even larger.
Let us apply our general analysis to the previous example of the cluster state. It turns out that the 1D cluster state has $\mathbb{Z}_2 \times \mathbb{Z}_2$ symmetry under which it is a nontrivial SPT, and the nonzero $\gamma$ will be a consequence of this. This example is the simplest possible. We will need to construct an MPS representation, and identify the projective symmetry on the virtual level.

To find an MPS representation, we write the wave function of the 1D cluster state in $\sigma^z$ basis using Eq. (6.6).

$$
\langle \cdots s_{j-1}s_js_{j+1}\cdots | \psi \rangle = (-1)^{\sum_j s_js_{j+1}}/\sqrt{2^{2L}}
$$

where $s_j = 0, 1$ and $2L$ is the number of qubits. It is therefore sufficient for the local tensor $A_j$ at site $j$ to take value $-1$ if $s_j = s_{j+1} = 1$ and $+1$ otherwise. The following tensor satisfies this condition.

$$
M = \sum_{s=0,1} (-1)^{ss'} |s\rangle \otimes |s\rangle \langle s'|
$$

The physical qubit at site $j$ is synchronized with the left virtual bond, and thus $s'$ is the state of the physical qubit at site $j+1$ upon contracting the virtual bonds. The local tensor $M$ correctly describes the cluster state, but the symmetry action will not be on-site. So, we block two neighboring physical qubits as one super-site and write the MPS representation as

$$
\mathcal{M} = \sum_{s_L,s_R=0,1; s'=0,1} (-1)^{s_Ls_R+s_Rs'}/2 |s_Ls_R\rangle \otimes |s_L\rangle \langle s'|
$$

It is easier to determine the symmetry once we rewrite $\mathcal{M}$ as a collection of matrices.

$$
\mathcal{M}^{(++)} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathcal{M}^{(+\ -)} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathcal{M}^{(-\ +)} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \mathcal{M}^{(\ -\ -)} = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},
$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. Notice if there was no $\mathcal{M}^{(++)}$, the remaining three tensors form an
MPS description of the AKLT state. [296] The action of $\sigma^x$’s on every other physical qubit forms a group $G = \mathbb{Z}_2 \times \mathbb{Z}_2$, where the first component $\mathbb{Z}_2$ is implemented as

$$M^{(\leftrightarrow)} \leftrightarrow M^{(\leftrightarrow)}, \quad M^{(-\star)} \leftrightarrow -M^{(-\star)},$$

and the second $\mathbb{Z}_2$ is implemented as

$$M^{(\leftrightarrow)} \leftrightarrow M^{(\leftrightarrow)}, \quad M^{(-\star)} \leftrightarrow -M^{(-\star)}.$$  \tag{6.33}

These transformations can be enacted by conjugations by $\sigma^x$ and $\sigma^z$. The conjugations on the virtual level do not change the state at all, and therefore $G$ is a symmetry of the cluster state.

It is evident now that the symmetry we just identified is in accordance with Eq. (6.19) with $\eta_g$ being trivial. The commuting symmetry action $G$ on the physical level is lifted to a noncommuting symmetry $D_4 = \langle \sigma^x, \sigma^z \rangle$ on the virtual level. In fact, it is known $H^2(\mathbb{Z}_2 \times \mathbb{Z}_2; U(1)) = \mathbb{Z}_2$, and this representation is precisely a projective representation of the group $\mathbb{Z}_2 \times \mathbb{Z}_2$ and the cluster state is a nontrivial SPT associated with it. The phase factor of Eq. (6.25) is $\Omega = -1$, and the result of the previous section implies that $\gamma \geq \log 2$. Indeed, we have $\gamma = \log 2$ in Eq. (6.11).

6.4 Generic behavior of $\gamma$: replica correlation length

In the previous section we attributed the subleading term $\gamma$ of the entanglement entropy to the degeneracy of the transfer matrix for Rényi entropies. Generically when there is no degeneracy, the entanglement entropy would be

$$S_\alpha = \frac{\log \lambda_1}{\alpha - 1} L + O(r^L), \quad r = \left| \frac{\lambda_2}{\lambda_1} \right| = e^{-1/\xi_\alpha} < 1.$$  \tag{6.34}

where $\lambda_2$ is the second largest eigenvalue of the transfer matrix. The subleading term converges to zero with a characteristic length scale $\xi_\alpha$. This indicates that without a symmetry protection the subleading term $\gamma$ should be zero for long chains. However, it is important to remark that $\xi_\alpha$ has little to do with the correlation length $\xi$. Indeed, being the ground state of a commuting Hamiltonian, the cluster state has the correlation length $\xi = 0$. Nonetheless, there is degeneracy
in the transfer matrix for Rényi entropy, which means that \( \xi_\alpha = \infty \) for any \( \alpha = 2, 3, 4, \ldots \).

To understand the generic finite size effect more closely, we consider a deformed cluster state \(|\theta\rangle\) specified by an angle \( \theta \in [0, \pi] \). The deformation is achieved by replacing the \(-1\) with \( e^{i\theta} \) in the MPS representation Eq. (6.30). This amounts to transforming the state \( \otimes_i |+\rangle \) by a two-qubit unitary \( U(\theta) = \text{diag}(1, 1, 1, e^{i\theta}) \) instead of \( U \) in Eq. (6.4). It is anyway a transformation by a small-depth quantum circuit from a state of no correlation, the resulting state \(|\theta\rangle\) has correlation length identically zero. Clearly, \( \theta = \pi \) reproduces the previous cluster state.

The tensor network in Figure 6.6 for \(|\theta\rangle\) can be explicitly evaluated, although the computation becomes more complicated as \( \alpha \) is increased. The eigenvalues \( \lambda_1, 2, \ldots \) of the transfer matrix \( T_\alpha \) for \( \rho^2 \) can be easily computed by a computer algebra system. The result is that

\[
\begin{align*}
\lambda_1 &= \frac{3 + y^2 + (y + 1)\sqrt{(y-1)^2 + 4}}{8} \\
\lambda_2 &= \frac{3 + y^2 - (y + 1)\sqrt{(y-1)^2 + 4}}{8} \\
\lambda_3 &= \frac{y^2 - 1}{4}, \quad \text{where } y = \cos \theta
\end{align*}
\]

(6.35)

and all other eigenvalues vanish for any \( \theta \). The largest eigenvalue \( \lambda_1 \) is non-degenerate unless \( \theta = \pi \), at which the symmetry \( G = \mathbb{Z}_2 \times \mathbb{Z}_2 \) is restored. The \( \theta = 0 \) point is also \( G \)-symmetric; however, the symmetry is lifted to an abelian virtual symmetry, and hence the state is a trivial SPT. The ratio of the second largest eigenvalue to \( \lambda_1 \) can be any value between 0 and 1. In other words, the length scale \( \xi_{\alpha=2}(\theta) \) interpolates from 0 to \( \infty \) continuously, while the correlation length \( \xi \) is held at zero.

Kitaev and Preskill [275] gave an argument that the subleading term \( \gamma \) can be robustly defined by taking a linear combination of entanglement entropies. There, it was essentially used that a small change in a region \( A \) far from a region \( B \) leaves the following combination invariant:

\[ \Delta S(A) - \Delta S(A \cup B) \simeq 0. \]

This is false when \( A \cup B \) happen to include exactly a half of a nontrivial 1D SPT chain as in Bravyi’s example in Sec. 6.2.3. From our consideration, generically, the distance between the re-
gion $B$ and the region at which the change occurs should be compared to the replica length scale $\xi_\alpha$ of Eq. (6.34), not to the usual correlation length. It should be made clear that we did not prove that a small $\xi_\alpha$ implies that the cylinder extrapolation method or Kitaev-Preskill prescription gives the total quantum dimension of the topological particle content. Rather, we showed a short correlation length does not imply that TEE that results from the cylinder extrapolation method gives the total quantum dimension. We gave an evidence for the conjecture that a short replica correlation length would imply the validity of the cylinder extrapolation method.

6.5 Replica correlation functions

Figure 6.8: Measuring replica correlation functions. Observables are inserted in the circles. Even if the global state $\rho_{AB}$ has a short correlation length, the positive semi-definite operator $\rho_A^2$, treated as a normalized state $\rho_A^2/\Tr(\rho_A^2)$, may have much longer correlation length. The latter length scale, which we call as the replica correlation length, can be simply measured in numerical calculations, and is the relevant length scale for the subleading term in the entanglement entropy.

We can actually probe the replica correlation length $\xi_\alpha$ by a replica correlation function on the original 2D state, without reducing it into a 1D chain. Let $\psi = |\psi\rangle \langle \psi|$ denote the density matrix of a state. Given a two copies of a state $\psi^{\otimes 2}$, and a bipartition $A \sqcup B$ of the system, we are formally provided with four subsystems $A_1, B_1, A_2, B_2$. Define $F_A$ to be the swap operator between $A_1$ and $A_2$. If $\psi_B = \Tr_A[\psi]$ is the reduced density matrix for $B$, it holds that

$$\Tr[\psi^{\otimes 2} F_A(O_{B_1} \otimes O_{B_2})] = \Tr[\psi_B O_{B_1} \psi_B O_{B_2}]$$

$$= \Tr[\psi_B^2 \langle O_B \rangle_{\alpha=2}] \quad (6.36)$$
for an arbitrary observable $O_B = O_{B_1} \otimes O_{B_2}$ on the subsystem $B$. By Eq. (6.36), we have defined $\langle O_B \rangle_{\alpha=2}$. For an observable $O_B = O_{B_1} \otimes O_{B_2}$, we define a replica correlation function

$$\text{Cor}_{\alpha=2}(O_B)(x) := \langle O_B(i=0)O_B(i=L) \rangle_{\alpha=2} - \langle O_B(i=0) \rangle_{\alpha=2} \langle O_B(i=L) \rangle_{\alpha=2}. \tag{6.37}$$

The slowest possible decay of $\text{Cor}_{\alpha=2}$ is determined by the length scale $\xi_2$ of Eq. (6.34), and this is actually achievable.

We prove this claim by an example. We will show the cluster state $|\theta = \pi \rangle$ has a nonzero constant replica correlation function that does not decay at all as a function of the distance between observables.

In addition, we will calculate the replica correlation functions for the $\mathbb{Z}_N$ gauge theory ground state in two dimensions, and the double semion ground state also in two dimensions. We find that the replica correlation function decays abruptly; the replica correlation length is zero. The purpose of this computation is to show that genuinely topologically ordered phases that give nonzero $\gamma$ do have representative wave functions with fast decaying replica correlation functions. We expect that every Levin-Wen wave function [297] would have a small replica correlation length. This strongly suggests that the replica correlation function can be used to determine when one should rely on the $\gamma$ value obtained from, e.g., DMRG computation.

In a DMRG calculation for a (isotropic) 2D state, we propose a measurement of $\xi_{\alpha=2}$ by the following steps. One prepares two copies of the state, and inserts the swap operator on the half strip of the state. Then, one measures the correlation function for a pair of point observables inserted near the region where swap operator is applied. See Fig. 6.9.

6.5.1 Flat entanglement spectrum

Before we delve into the replica correlation function calculations, we review a technique and expression for special reduced density matrices that are proportional to projectors. [298, 299] This includes the cluster state and the ground states of $\mathbb{Z}_N$ gauge theory that are eigenstates of string...
Figure 6.9: Replica correlation function calculation. One prepares two copies of the state, and apply the swap operator on the shaded region; insert observables in the circles, and compute the overlap with the original, unswapped state. The overlap is generally exponentially small in the boundary length of the shaded region, but after normalization this reveals the replica correlation length.

operators. The technique here will be used crucially in later calculations. We follow Proposition 9 and Corollary 10 of Ref. [300].

Define $N \times N$ matrices as

$$X_N = \sum_{j \in \mathbb{Z}_N} |j + 1\rangle \langle j|, \quad Z_N = \sum_{j \in \mathbb{Z}_N} e^{2\pi ij/N} |j\rangle \langle j|.$$  \hfill (6.38)

Any product $P = X_N^n Z_N^m$ of these matrices has a property that

$$\text{Tr } P = \begin{cases} N & \text{if } n = m = 0 \in \mathbb{Z}_N, \\ 0 & \text{otherwise}. \end{cases}$$  \hfill (6.39)

Now, consider any multiplicative group $\mathcal{G}$ generated by tensor products of these matrices together with the phase factor $e^{2\pi i/N}$ on $(\mathbb{C}^N)^{\otimes n}$. Examples are the multiplicative group generated by the term of the Hamiltonian of the cluster state, and of the $\mathbb{Z}_N$ gauge theory as in (6.49) below. If $\mathcal{G}$ is abelian, there exists a common eigenstate $|\psi\rangle \in (\mathbb{C}^N)^{\otimes n}$ of all elements of $\mathcal{G}$.

Suppose that

- there is a unique eigenstate $|\psi\rangle$ of eigenvalue $+1$ for all $g \in \mathcal{G}$.

Then, $\mathcal{G}$ cannot contain any pure scalar element $\eta \neq 1$ because such a scalar can only have eigen-
value that is not 1. The projector onto $|\psi\rangle$ can be written as

$$
|\psi\rangle \langle \psi| = \Pi_G := \frac{1}{|G|} \sum_{g \in G} g.
$$

(6.40)

Because of (6.39), any non-identity element of $G$ has zero trace. Hence, taking the trace on both sides we see

$$
1 = \frac{1}{|G|} \text{Tr}(I) = \frac{N^n}{|G|};
$$

(6.41)

that is, the order of the group must be the full dimension of the Hilbert space $(\mathbb{C}^N)^\otimes n$.

Now, divide the system into two subsystems $A$ and $B = A^c$ so that $(\mathbb{C}^N)^\otimes n = (\mathbb{C}^N)^\otimes |A| \otimes (\mathbb{C}^N)^\otimes |B|$. Tracing out $A$ from the density matrix $|\psi\rangle \langle \psi|$, we obtain the reduced density matrix for $B$:

$$
\rho_B = \text{Tr}_A \Pi_G = \frac{1}{N^{|A|} + |B|} \sum_{g \in G} \text{Tr}_A(g)
$$

(6.42)

Again due to (6.39), $\text{Tr}_A(g)$ is zero unless $g$ acts on $A$ by the identity (supported on $B$), in which case $\text{Tr}_A(g) = (g|B)N^{|A|}$. The elements of $G$ that are supported on $B$ form a subgroup $G_B$, and we can write

$$
\rho_B = \frac{N^{|A|}}{N^{|A|} + |B|} \sum_{g \in G_B} g = \frac{|G_B|}{|G_B|} \frac{1}{|G_B|} \sum_{g \in G_B} g.
$$

(6.43)

This implies that $\rho_B$ is proportional to the projector $\Pi_{G_B}$ of rank $N^{|B|}/|G_B|$. In other words, the entanglement spectrum (the eigenvalues of $\rho_B$) is flat and the entanglement entropy (von Neumann or Rényi) is

$$
S(B) = |B| \log N - \log |G_B|.
$$

(6.44)

This expression for the entropy is also derived in Ref. [273,301].
6.5.2 Cluster state

In this subsection we calculate the replica correlation length of the 2D cluster state and its deformed cousins, by considering the behavior of the replica correlation functions. We will find observables that achieve the slowest possible decay of replica correlation functions. Since a local observable is mapped to a local observable under finite depth quantum circuits (local unitaries), we may study the replica correlation function after simplifying the state by local unitaries. This means that we can focus on the 1D cluster state that is reduced from the 2D cluster state.

Consider the cluster state $|\theta = \pi\rangle$ on $2L$ spins with the periodic boundary condition. As before, let $B$ be the region that contains every other spin, in total of $L$ spins. The reduced density matrix $\psi_B$ has a flat eigenvalue spectrum; $\psi_B$ is proportional to a projector.

$$\psi_B = \frac{1}{2L} \left( I + \prod_{i \in B} \sigma_x^i \right) \quad (6.45)$$

where the number of nonzero eigenvalues is $M = 2^{L-1}$. See the previous subsection 6.5.1 for a derivation.

Let $O_B(i) = \sigma_z^i \otimes \sigma_z^i$ be an observable for two copies of the state. The normalization factor $\text{Tr}[\psi_B^2]$ in Eq. (6.36) is equal to $1/M$. Thus,

$$\langle O_B(0)O_B(i) \rangle_{\alpha=2} = \frac{1}{M} \text{Tr}[\psi_B \sigma_0^z \sigma_i^z \psi_B \sigma_0^z \sigma_i^z] = \frac{1}{M} \text{Tr}[\psi_B \sigma_0^z \sigma_i^z \sigma_0^z \sigma_i^z \psi_B] \quad (6.46)$$

$$= \frac{1}{M} \text{Tr}[\psi_B^2] = 1$$
where the second equality is because $\sigma_0^z \sigma_i^z$ commutes with $\psi_B$. On the other hand,

\[
\langle O_B(j) \rangle_{\alpha=2} = M \text{Tr}[\psi_B^2 \psi_B^2] = 0
\]

because $\psi_B \sigma_j^z \psi_B = 0$. Therefore, the replica correlation function reads

\[
\text{Cor}_{\alpha=2}(O_B(0), O_B(x)) = 1
\]

independent of the separation $x$.

For generic values of $\theta$, we numerically checked that replica correlation functions of generic observables on the state $|\theta\rangle$ decay according to the finite replica correlation length calculated from (6.35). We emphasize once again that the usual correlation length $\xi$ measured by

\[
\langle \theta | K_1 K_2 | \theta \rangle - \langle \theta | K_1 | \theta \rangle \langle \theta | K_2 | \theta \rangle
\]

is identically zero for any observables $K_1$ and $K_2$ for the state $|\theta\rangle$ for any $\theta$.

6.5.3 $\mathbb{Z}_N$ Gauge Theory

The replica correlation length/function is introduced to pick up a fine detail of a state, and therefore we calculate it for a particular ground state of an exactly soluble model of the $\mathbb{Z}_N$ gauge theory in the deconfined phase. Unlike the 2D cluster state, abelian discrete gauge theory ground states, as well as the double semion state of the next subsection, requires deep quantum circuit to disentangle, [269, 271, 272] so we are forced to work with the 2D state directly in the replica correlation functions. The lattice of the model is not too important, but we consider the square lattice in two dimensions. The Hamiltonian is sum of star terms (gauge transformation), and pla-
quete terms (flux):

\[
H_{Z_N} = - \sum_s X_{s,\text{east}} X_{s,\text{north}} X_{s,\text{west}} X_{s,\text{south}}^\dagger - \sum_p Z_{p,\text{south}} Z_{p,\text{east}} Z_{p,\text{north}} Z_{p,\text{west}}^\dagger.
\]

(6.49)

where \( s \) denotes a site (vertex) and \( p \) denotes a plaquette (face), and \( X = X_N, Z = Z_N \) of (6.38).

The ground state is an equal amplitude superposition of “loop” configurations, where the loops come in \( N \) types and obey the group law of \( \mathbb{Z}_N \).

When put on a thin torus as in Fig. 6.9, the Hamiltonian \( H_{Z_N} \) has an \( N^2 \)-fold degenerate ground space. As the DMRG algorithm is biased to states with minimal entanglement across the circumferential cut, \cite{280} we consider the state \( |\psi\rangle \) that has +1 eigenvalue of the \( Z \)-type string operator and +1 of \( X \)-type string operator along the circumference (the shortest nontrivial loop).

For this state \( |\psi\rangle \), which is the unique common \((+1)\)-eigenvector of a commuting set of tensor products of \( X \) and \( Z \) matrices, the entanglement spectrum for any bipartition is flat, and the reduced density matrix \( \psi_B \) is proportional to a projector

\[
\Pi = \frac{1}{|\mathcal{G}_B|} \sum_{g \in \mathcal{G}_B} g.
\]

(6.50)

Here, the group \( \mathcal{G}_B \) is a subgroup of the multiplicative group \( \mathcal{G} \) that is generated by all Hamiltonian terms.\(^1\) \( \mathcal{G}_B \) consists of all elements of \( \mathcal{G} \) that is supported on \( B \). (See the previous subsection 6.5.1.)

In fact, \( \mathcal{G}_B \) is generated precisely by the Hamiltonian terms supported on \( B \). To see this, suppose \( g \in \mathcal{G}_B \). Since \( \mathcal{G} \) is abelian, the operator \( g \) can be written as a product of closed \( Z \)-loop operators and closed \( X \)-loop operators. These loop operators are contractible and supported on \( B \), so each \( X \)- or \( Z \)-loop can be deformed to vanish by multiplying the smallest loop operators, which is exactly the Hamiltonian terms on \( B \). This implies that \( g \) is a product of the Hamiltonian terms on \( B \). \( \mathcal{G}_B \) having a local generating set is an important difference from the cluster state, for which the density matrix (6.45) is the sum over a group with no local generators.

\(^1\)For other ground states that are eigenstates of string operators on the orthogonal loop, this is not the case and one has to include the topologically nontrivial string operators in \( \mathcal{G} \).
Note that since $\Pi$ is a sum of all elements of a group, we see

$$\Pi g = \Pi = g \Pi \quad (6.51)$$

for any element $g \in G_B$. We claim that for any operator $O$ there exists $\tilde{O}$ of the *same* support such that

$$\Pi O \Pi = \Pi \tilde{O} \Pi, \quad \text{and} \quad [\Pi, \tilde{O}] = 0. \quad (6.52)$$

To construct $\tilde{O}$, let $G_i \leq G_B$ be the group generated by the Hamiltonian terms that overlap the site (or region) $i$ on which $O$ is supported. Let

$$\tilde{O} = \frac{1}{|G_i|} \sum_{g \in G_i} g O g^{-1}. \quad (6.53)$$

The support of $\tilde{O}$ is the same as that of $O$ because every element $g \in G_i$ is a tensor product unitary operator. Any tensor component of $g$ that acts outside of the support of $O$ cancels in the expression $g O g^{-1}$. Hence, every summand $g O g^{-1}$ has the same support as $O$, so is the sum $\tilde{O}$. Note that $\tilde{O}$ is zero if, e.g., $O$ anti-commutes with some $g \in G_i$.

Next, we verify (6.52).

$$\Pi \tilde{O} \Pi = \frac{1}{|G_i|} \sum_{g \in G_i} \Pi g O g^{-1} \Pi$$

$$= \frac{1}{|G_i|} \sum_{g \in G_i} \Pi O \Pi$$

$$= \Pi O \Pi, \quad (6.54)$$

where we used (6.51). In addition, $g \tilde{O} g^{-1} = \tilde{O}$ if a Hamiltonian term $g \in G_i$ overlaps $i$ by (6.53). If $h \in G_B \setminus G_i$ is a Hamiltonian term that does not overlap $i$, then $h$ commutes with $g \in G_i$ because $G_B$ is abelian, and $h$ also commutes with $O$ trivially, so $h$ commutes with $\tilde{O}$. Since $G_B$ is generated by Hamiltonian terms, $\tilde{O}$ commutes with every element of $G_B$.\(^2\)

\(^2\)Even if a topologically nontrivial string operators are in $G_B$, this is still true as the long string generator $s$ of $G$ can always chosen to be commuting with $O$ so that $s O s^{-1} = O$. Note that a Hamiltonian term
is the sum of all group elements of $\mathcal{G}_B$, it follows that

$$[\Pi, \tilde{O}] = 0.$$ \hfill (6.55)

Moreover, if $O$ and $O'$ are far separated so that no generator of $\mathcal{G}_i$ or $\mathcal{G}_{i'}$ overlaps with both $O$ and $O'$, then

$$\tilde{O}\tilde{O}' = \frac{1}{|\mathcal{G}_i| \cdot |\mathcal{G}_{i'}|} \sum_{g \in \mathcal{G}_i} \sum_{h \in \mathcal{G}_{i'}} (gOg^{-1})(hO'h^{-1})$$

$$= \frac{1}{|\mathcal{G}_i| \cdot |\mathcal{G}_{i'}|} \sum_{g \in \mathcal{G}_i} \sum_{h \in \mathcal{G}_{i'}} ghOO'(gh)^{-1}$$ \hfill (6.56)

$$= \frac{1}{|\mathcal{G}_i \times \mathcal{G}_{i'}|} \sum_{g \in \mathcal{G}_i \times \mathcal{G}_{i'}} gOO'g^{-1}$$

$$= \tilde{O}\tilde{O'}.$$

Consider arbitrary observables $O = \sum_a P_a \otimes Q_a$ near a site $i$ and $O' = \sum_b P'_b \otimes Q'_b$ near a site $i'$ on the two copies of the state. Suppose $i$ and $i'$ are sufficiently separated, say by 5 lattice spacing, so that no term in the Hamiltonian $H_{Z_N}$ overlaps simultaneously with $O$ and $O'$. The normalization factor of (6.36) is given by $M = \text{Tr}[\psi_B^2]^{-1} = \text{Tr}[\Pi] = |\mathcal{G}_B|$. Using tilde operators, that overlaps $i$ but is not in $\mathcal{G}_B$ may not commute with $\tilde{O}$. 

157
the replica correlation function becomes a usual correlation function:

\[
\langle OO' \rangle_2 - \langle O \rangle_2 \langle O' \rangle_2
= |G_B|^{-1} \sum_{a,b} \text{Tr}[\Pi P_a P'_b Q_a Q'_b]
- |G_B|^{-2} \sum_{a,b} \text{Tr}[\Pi P_a Q_a] \text{Tr}[\Pi P'_b Q'_b]
\]

\[
= |G_B|^{-1} \sum_{a,b} \text{Tr}[\tilde{\Pi} \tilde{P}_a \tilde{P}'_b Q_a Q'_b]
- |G_B|^{-2} \sum_{a,b} \text{Tr}[\tilde{\Pi} \tilde{P}_a Q_a] \text{Tr}[\tilde{\Pi} \tilde{P}'_b Q'_b]
\]

\[
= \text{Cor} \left( \sum_a \tilde{P}_a Q_a, \sum_b \tilde{P}'_b Q'_b \right)
\]

\[
= 0
\]

where in the second equality \( \tilde{P}_a \tilde{P}'_b = \tilde{P}_a \tilde{P}'_b \) is because they act on separated spins, and the last equality is because \( H_{Z_N} \) is commuting with locally indistinguishable ground space.

The replica correlation function is not identically zero when the site \( i \) and \( i' \) are close. This is simple. The replica correlation function becomes the usual correlation function when \( O = O \otimes I \).

Consider \( O = X_{s,\text{east}} X_{s,\text{north}} + \text{h.c.} \) and \( O' = X_{s,\text{west}} X_{s,\text{south}}^\dagger + \text{h.c.} \) where \( s \) is in the interior of \( B \). Then, \( \text{Tr}[OO' \psi_B] = 2 \), but \( \text{Tr}[O \psi_B] = \text{Tr}[O' \psi_B] = 0 \), so \( \text{Cor}(O, O') = 2 \).

In conclusion, we have shown that the replica correlation function for \( H_{Z_N} \) is not identically zero but decays to zero after separation distance 5. Therefore, the replica correlation length is zero. This is in contrast to the cluster state calculation where the replica correlation function is nonzero and does not decay at all.

### 6.5.4 Double Semion Model

Using the similar technique as in the previous subsection, we will calculate the replica correlation function for a version of double semion model on the honeycomb lattice, \([297, 302]\) and find that it decays abruptly in the same ways as for the \( Z_N \) gauge theory.

The lattice consists of 2 spin-\( \frac{1}{2} \)'s at each edge. The two spins on an edge is going to be (ener-
getically) “synchronized”. The Hamiltonian is

\[ H_{DS} = - \sum_p - \prod_{e,e' \in p} \sigma^x_e \sigma^x_{e'} \prod_{e \in \partial p} \sqrt{\sigma^z_e} + h.c. \]

\[ - \sum_v \prod_{e \in v} \sigma^z_e - \sum_e \sigma^z_e \sigma^z_{e'} g_e \]

The first term \( g_p \) has the minus sign, and is defined for every hexagon \( p \) where \( e, e' \in p \) means all 12 spins on the plaquette, and \( e \in \partial p \) means the 6 spins on the legs of the plaquette that are immediate neighbors of the plaquette. The second term is defined for every trivalent vertex \( v \), and \( e \in v \) means the 3 spins that are immediate neighbors of \( v \). The third term is defined for every edge that contains two spins \( e \) and \( e' \); this ferromagnetic term “synchronizes” the two spins on the edge. Our Hamiltonian (6.58) is slightly different from those in Ref. [297, 302], but splitting the edge spin into two spins has appeared in Ref. [276]. We are considering the split version because of the simplicity of reduced density matrix expression.

The plaquette terms \( g_p \) commute with the vertex terms \( g_v \) and edge terms \( g_e \), but they do not commute among themselves; however, they do commute on the constrained subspace where every vertex and edge term takes +1 eigenvalue \( (g_v = g_e = 1) \). This is equivalent to the following. Let \( G^P \) denote the nonabelian multiplicative group generated by all \( g_p \)'s, and let \( G^Z \) denote the abelian multiplicative group generated by all \( g_v \)'s and \( g_e \)'s. Also, let \( G = G^P G^Z \) denote the group generated by all terms in the Hamiltonian. Then,

\[ g z g^{-1} z^{-1} = 1 \quad \forall g \in G^P, z \in G^Z \]

\[ g h g^{-1} h^{-1} \in G^Z \quad \forall g, h \in G^P. \] (6.59)

(See also Ref. [303].)

The ground space of \( H_{DS} \) on a torus is four-fold degenerate, and the minimally entangled states are eigenstates of string operators along a shortest topologically nontrivial loop. The string operators whose end points, if open, corresponds to semions are not too simple (Eq. 9 of Ref. [302]),
Figure 6.10: Double-semion model defined in (6.58). (a) shows the lattice configuration and the location of the degrees of freedom. Each edge of the hexagon accommodates two spins, denoted by the blue dots on the edge. The dashed line is the entanglement cut. (b-d) pictorially show the three terms in the Hamiltonian for a plaquette, vertex and edge. The symbol near a spin denotes the operator acting on this spin, where $X$ means $\sigma_x$, $Z$ means $\sigma_z$, $\sqrt{Z} = \text{diag}(1, i)$, and $\text{Id}$ is the identity matrix. The total Hamiltonian is the summation over all plaquettes, vertices and edges, with appropriate sign factors defined in (6.58).

but satisfies an important property that the partial trace is zero. Let us fix the entanglement cut that passes an array of plaquettes along a straight line that is orthogonal to an edge. This entanglement cut passes in between the two spins on the intersecting edges, which is considered in Ref. [276]. The purpose of this special cut is to have

$$\text{Tr}_A[g] = 0 \text{ if } g \in \mathcal{G}^P \text{ overlaps with } A \quad (6.60)$$

since any such $g$ has a $\sigma^x$ or $\sigma^z$ tensor component in addition to $\sqrt{\sigma^z}$ within $A$. Then, it follows
by the same reasoning as in Sec. 6.5.1 that

\[ \psi_B = \text{Tr}_A[\psi] \propto \sum_{g \in G_B} g \tag{6.61} \]

where \( G_B \leq G = G^P G^Z \), consists of the elements supported on \( B \).\(^3\) The entanglement spectrum is flat.

The subgroup \( G_B \) is generated by the local Hamiltonian terms supported on \( B \). The reason is similar to that for the previous \( Z_N \) theory. Observe that \( \sigma^x \sqrt{\sigma^z} \sigma^x = i(\sqrt{\sigma^z})^\dagger \). Hence, any element \( g \in G = G^P G^Z \) can be written as a product of \( \prod_i \sigma_i^x \) and \( \prod_i \sqrt{\sigma_i^{z_m}} \) up to an overall phase factor. The first factor \( \prod_i \sigma_i^x \) has to form a closed loop since it arises from \( g_p \) terms. The closed loop of \( \sigma^x \) must be entirely contained in \( B \), and we can eliminate it by multiplying \( g_p \) operators on \( B \) to \( g \). Therefore, it suffices for us to show that any “diagonal” element \( z \in G \) (a product of \( \sqrt{\sigma^z} \) supported on \( B \), is given by a product of \( g_v \) and \( g_e \) on \( B \). Note that any diagonal element of \( G^P \) arises from \( g_p^2 \), and \( g_p^2 \) belongs to \( G^Z \).\(^{[303]}\) The group \( G^Z \) can be viewed as the group of null-homologous \( \mathbb{Z}_2 \)-loops on the triangular lattice (the dual lattice of the honeycomb lattice). Therefore, if \( z \in G^Z \) is supported on \( B \), then \( z \) is a product of \( g_v \) and \( g_e \) on \( B \). This implies that \( g \in G \) supported on \( B \) is a product of Hamiltonian terms on \( B \) up to a phase factor. The phase factor must be 1 because the group \( G \) does not contain any nontrivial phase factor. This completes the reasoning.

As in the previous section, we can turn any operator \( O \) supported on \( B \) into \( \hat{O} \) such that

\[
\text{support}(O) = \text{support}(\hat{O}),
\]

\[
\psi_B O \psi_B = \psi_B \hat{O} \psi_B,
\]

\[
[\hat{O}, \psi_B] = 0,
\]

\[
\hat{O} \hat{O}' = \hat{O} \hat{O}' \quad \text{if separated.}
\]

Let \( G^P_i \) be the group generated by \( g_p \) that overlaps \( i \) on which \( O \) is supported, and let \( G^Z_i \) be the

\(^3\)The string operator that wraps around the shortest topologically nontrivial loop of the torus does not enter, because its partial trace is zero.

161
group generated by $g_e$ and $g_v$ on spins where $G_i^P$ is supported. The choice of group $G_i^Z$ is to have

\[ gzg^{-1}z^{-1} = 1 \quad \forall g \in G^P, z \in G_i^Z \]

\[ ghg^{-1}h^{-1} \in G_i^Z \quad \forall g \in G^P, h \in G_i^P. \]  \hspace{1cm} (6.63)

Now, let $G_i = G_i^P G_i^Z$, and define $\tilde{O}$ by the formula (6.53). $\tilde{O}$ has the same support as $O$ because $G_i$ is a group of tensor product unitary operators. That $\psi_B O \psi_B = \psi_B \tilde{O} \psi_B$ follows from a similar equation as (6.54) since $\psi_B$ for the double semion state is also a sum over a group $G_B \supset G_i$. As for $[\tilde{O}, \psi_B] = 0$, we see $h\tilde{O}h^{-1} = \tilde{O}$ if $h \in G_i^P$ is one of the plaquette terms $g_P$ that overlap $i$, by definition of $\tilde{O}$. If $k \in G^P \setminus G_i^P$ is a plaquette term that does not overlap $i$, then

\[
k\tilde{O}k^{-1} = \sum_{g \in G_i} kgOg^{-1}k^{-1}
= \sum_{g \in G_i} g(k^{-1}g^{-1}kg)O(k^{-1}g^{-1}kg)^{-1}(gk)^{-1}
= \sum_{g \in G_i} g(k^{-1}g^{-1}kg)O(k^{-1}g^{-1}kg)^{-1}g^{-1}
= \sum_{g \in G_i} gOg^{-1}
= \tilde{O}
\]

(6.64)

where in the third equality we used (6.63) and that $[k, O] = 0$, and in the fourth equality we redefined the dummy variable $g$ since $(k^{-1}g^{-1}kg) \in G_i$. If $z$ is one of $g_e$ or $g_v$ terms, then a similar calculation, simpler than (6.64), shows $z\tilde{O}z^{-1} = \tilde{O}$. As we have shown that $\psi_B$ is a sum over a group generated by the terms of $H_{DS}$, the proof that $[\psi_B, \tilde{O}] = 0$ is complete. The last property $\tilde{O}O' = \tilde{O}\tilde{O}'$ follows from similar equations as (6.56).

In conclusion, Eq. (6.57) holds without any modification using (6.62) for the double semion model under our bipartition. Therefore, the replica correlation function reduces to a usual correlation function, and the replica correlation length is zero.
6.6 Discussion

In this paper we have studied the behavior of the sub-leading term of the bipartite entanglement
entropy of topologically trivial 2D states calculated by the cylinder extrapolation method, and
found a sufficient condition under which a topologically trivial state will give a nonvanishing sub-
leading term under this method. In particular, we showed the bipartite entanglement entropy of
a such 2D state can be reduced to that of a 1D chain under an extensive bipartition. If this 1D
chain is in a nontrivial SPT state under a product group $G = G_1 \times G_2$, where $G_1$ and $G_2$ act
exclusively on the two sides of the bipartition, then a nonvanishing sub-leading term appears in
the cylinder extrapolation method.

Our result does not necessarily invalidate the notion of topological entanglement entropy. In
fact, the examples in this paper that are translation-invariant yield the correct total quantum
dimension of the topological phase under the Kitaev-Preskill or Levin-Wen prescription. Rather,
our finding makes it clear that the cylinder extrapolation method may give a different answer
than the Kitaev-Preskill or Levin-Wen prescription in the bulk.

Notice the above condition requires the state be in a nontrivial SPT state of $G$ in the way we
described above. This requirement, where the nontriviality of the state is protected by $G_1$ and
$G_2$ simultaneously, and $G_1$ and $G_2$ act exclusively on the two sides of the bipartition, is stronger
than the general condition of 1D SPT based on group cohomology. [293,294] A nontrivial SPT in
the general sense can be protected by $G_1$ or $G_2$ alone, but this is not sufficient to yield a nonvan-
ishing sub-leading term from the cylinder extrapolation method.

We have introduced the replica correlation length/function. In Sec. 6.5 we gave an operational
meaning to it and demonstrated that it can be determined numerically. Though we only dis-
cussed $\alpha = 2$ replica correlation length/function, it is straightforward to consider $\alpha > 2$ cases
by considering cyclic permutation operators instead of the swap operator. Our result suggests a
conjecture that the cylinder extrapolation method on minimally entangled states yields the total
quantum dimension of the topological particle content if the replica correlation length is small
compared to the system size. All our examples of the SPT states under a product group, the $\mathbb{Z}_N$
lattice gauge theory, and the double semion model should be read as evidences in favor of this
conjecture.

Analogues of TEE for the ground states of gapped Hamiltonians in 3 or higher dimensions have been proposed. Ref. [288] studies various solid torus geometries and identifies multitude of TEEs that are associated with Betti numbers of the region for which entanglement entropy is calculated. Our examples can be generalized to this setting using graph states [287], and indeed modifies the subleading constant term of the entanglement entropy. If one tries a (hyper-)cylinder extrapolation method to read off the subleading term, then our ideas here give translation-invariant states with a modified subleading term. As remarked before, our examples in this higher dimensional generalization will be fine-tuned, but the length scale where the finite size effect is relevant can be arbitrarily larger than usual correlation lengths.

Besides, it is natural to consider topological entropy for thermal states. An immediate problem is that the entropy of the reduced density matrix of a thermal state obeys a volume law. This is easily overcome by using mutual information, which obeys an area law at any nonzero temperature. In 2D, while every known ground state with a nonzero TEE requires a local unitary transformation (quantum circuit) of large depth (linear in system size) in order to be transformed to a product state, every Gibbs state at any nonzero temperature of any commuting Hamiltonian can be transformed by a quantum circuit of small depth (logarithmic in system size) into a Gibbs state of a classical Hamiltonian. This is consistent to the calculation of topological entropy of the 2D \( \mathbb{Z}_2 \) gauge theory at nonzero temperature where the subleading term of mutual information is shown to vanish.

In 3D, the entropies on solid torus of \( \mathbb{Z}_2 \) lattice gauge theory at nonzero temperature have been calculated. It is observed that at low nonzero temperature, a nonzero subleading term survives when a certain linear combinations of mutual information is used to cancel extensive parts. But Ref. also shows that the Gibbs state of this model at nonzero temperature can be connected to the Gibbs state of a classical Hamiltonian by a small-depth quantum circuit. (See also Ref. [305].) Therefore, on the contrary to the 2D case, this value being nonzero is not related to topological order in the sense of generating quantum circuits of large depth. This means that if we accept the complexity of the generating quantum circuit as the definition of topological order for thermal states, then we should conclude that the subleading term of mutual
information does not give an ‘order parameter’ for topological order for thermal states in 3 or higher dimensions.

At nonzero temperature, it appears that our examples only give a contribution that is exponentially small in the system size to the subleading term (although the length scale of our effect is still different from the usual correlation length). The mutual information of the 2D cluster state Eq. (6.8) at finite inverse temperature $\beta$ with respect to the bipartition in Fig. 6.3 is

$$I(A : B) = 2L(\log 2 + O((1 - t) \log(1 - t))) - O(t^{2L})$$

(6.65)

where $t = \tanh \beta$ is close to but smaller than 1. The detail of the calculation can be found in Appendix F.3.

Finally, we note there exists a notion of localizable entanglement and associated entanglement length, [306] whose divergence is connected to a string order parameter. [307] It is shown that a subclass of our nontrivial 1D SPT can be used as a perfect quantum repeater. [308] However, it remains unclear how the entanglement length is related to our replica correlation length $\xi_\alpha$. A technical difference is that in our definition $\xi_\alpha$ carries an (artificial) index $\alpha$, as it is defined by the eigenvalues of the transfer matrix $T_\alpha$ for $\rho^\alpha$. 
Symmetry enriched U(1) quantum spin liquids

7.1 Introduction

Symmetry and entanglement both play important roles in understanding quantum phases of matter. It is by now well known that the ground states of quantum many-body systems may be in phases characterized by long-range entanglement between local degrees of freedom. Global symmetry may be realized in interesting ways in such long-range entangled phases. The simplest (and best understood) cases are gapped topologically ordered quantum phases as exemplified by the fractional quantum Hall states. The long-range entanglement in the fractional quantum Hall ground state wavefunctions enables gapped quasiparticle excitations showing fractional statistics and fractional charge. The fractional statistics is a fundamental phenomenon that follows from
the topological order, while the fractional charge describes the implementation of the global $U(1)$ charge conservation symmetry in this state.

Another prototypical class of states that possess long-range entanglement are quantum spin liquid phases of insulating magnets. A wide variety of quantum spin liquids have been described theoretically. Their universal low energy physics is (in most known examples) described by a deconfined emergent gauge theory coupled to matter fields. In the presence of global symmetries it is necessary to also specify the symmetry implementation in this low energy theory. Indeed two phases with the same structure of long-range entanglement (eg, same low energy gauge theory) can still be sharply distinguished by their symmetry implementations. This leads to a symmetry protected distinction between symmetry unbroken phases of matter (as is familiar from the theory of topological band insulators).

It is useful to distinguish two very broad classes of spin liquids. The simplest and best understood are ones in which all excitations are gapped. These gapped spin liquids are topologically ordered - they have well defined quasiparticle excitations with non-local ‘statistical’ interactions, ground state degeneracies on topologically non-trivial manifolds, etc. Global symmetries can be implemented non-trivially in topologically ordered phases. For instance a symmetry may be fractionalized. Topological phases in the presence of global symmetries have been dubbed “Symmetry Enriched Topological” (SET) matter. Thus symmetry protected distinctions between different SET phases may be much more striking than in topological band insulators. Though much of the early work on spin liquids dealt with SET phases, it is only in the last few years that there has been tremendous and systematic progress in understanding their full structure and classification in two dimensional systems [309–319]. Some limited progress has been made for three dimensional SET phases as well [320–322]. A different broad class of spin liquid phases have gapless excitations. These are much less understood theoretically though they have tremendous experimental relevance.

In this paper, we study a particularly simple class of quantum spin liquids in three spatial dimensions (3D) with an emergent gapless photon excitation. Their low energy dynamics is described by a deconfined $U(1)$ gauge theory. Microscopic models for such phases were described in Refs [13–19]. The emergence of the photon is necessarily accompanied by the emergence of quasi-
particles carrying electric and/or magnetic charges that couple to the photon. We will restrict attention to phases where these ‘charged’ matter fields are all gapped\(^1\). One of our main focuses is on the realization of such \(U(1)\) quantum spin liquids in 3D magnets with spin \(SO(3)\) and time reversal \(\mathcal{T}\) symmetries. After warming up with this example, we will describe a general framework to classify symmetry enriched \(U(1)\) quantum spin liquids with a large class of symmetries. Then we will apply this framework to the more complicated case where the symmetry is \(Z_2 \times \mathcal{T}\). We will also briefly discuss such \(U(1)\) quantum spin liquids enriched by some other symmetries.

In previous work [241] (see also Ref. [316]) we described the various such phases when time reversal is the only global internal symmetry. The extension to \(SO(3) \times \mathcal{T}, Z_2 \times \mathcal{T}\) and other symmetries is non-trivial and requires some conceptual and technical advances which we describe in detail in this paper.

For the case with \(SO(3) \times \mathcal{T}\) symmetry, we find that there are 15 families of such \(U(1)\) quantum spin liquids which may be distinguished by the symmetry realizations on the gapped electric/magnetic excitations. We describe the physical properties of these states. We will show that there are two such quantum spin liquids which have a “fractional” response to a background external \(SO(3)\) gauge field. For this reason we dub them “Fractional Topological Paramagnets”. They are closely analogous to the fractional topological insulators discussed theoretically.

Each of the 15 families is further refined when the quantum spin liquid phase is combined with a Symmetry Protected Topological (SPT) phase of the underlying spin system protected by the same \(SO(3) \times \mathcal{T}\) symmetry. This does not change the bulk excitation spectrum but manifests itself in different boundary properties. As described in our previous work [241] this refinement can be non-trivial: some but not all SPT phases can be “absorbed” by the spin liquid and not lead to a new state of matter. Including this refinement we find a total of 168 different such \(U(1)\) quantum spin liquids with \(SO(3) \times \mathcal{T}\) symmetry.

For the case with \(Z_2 \times \mathcal{T}\) symmetry, we find there are 38 distinct of such \(U(1)\) quantum spin liquids based on the properties of the bulk fractional excitations. We also obtain the classification

\(^1\)The problem of gapless matter fields coupled to a (compact) \(U(1)\) gauge field is an interesting and extensively studied problem. For some representative papers from the condensed matter literature see Refs. [21, 25, 157, 165, 173, 208, 323, 324]. A full classification of such phases with gapless matter fields is beyond the reach of currently available theoretical tools.
for such spin liquids with some other symmetries.

Studying symmetry enriched $U(1)$ quantum spin liquids is of conceptual and practical importance not only for quantum magnetism, but has far reaching connections to many other topics in modern theoretical physics. First as emphasized in previous papers [241], there is a very useful connection to the theory of Symmetry Protected Topological (SPTs) insulators of bosons/fermions. It is very helpful to view these $U(1)$ quantum spin liquids as the gauged version of some SPTs with a $U(1)$ symmetry, i.e. these quantum spin liquids can be obtained by coupling the relevant SPTs to a dynamical $U(1)$ gauge field. There are actually two distinct ways in which the same $U(1)$ QSL can be viewed as a gauged $U(1)$ SPT - either as a gauged SPT of the electric charge or a gauged SPT of the magnetic monopole. This leads to a generalization of the standard electric-magnetic duality of three dimensional Maxwell theory which incorporates the realization of global symmetry [29, 241, 325, 326]. In the presence of a boundary, this $3 + 1$-dimensional “symmetry-enriched” electric-magnetic duality implies interesting and non-trivial dualities between $2 + 1$-dimensional quantum field theories [29, 326–328]. This line of thinking has proven to be very powerful in studying difficult problems in strongly-correlation physics in two space dimensions. Examples include quantum hall systems, especially the half-filled Landau level [213, 325, 329–332], interacting topological insulator surfaces [29, 327, 328], quantum electrodynamics in $2 + 1$ dimensions [326, 333] and a class of Landau-forbidden quantum phase transitions known as deconfined quantum criticality [334]. The lower dimensional dualities are also interesting on their own as nontrivial results in $2 + 1$ dimensional quantum field theory [217, 245, 335, 336]. Therefore, we discuss in detail the relation between different symmetry enriched $U(1)$ quantum spin liquids and various SPTs.

The rest of the paper is organized as follows. In Sec. 7.2, we enumerate all possible $SO(3) \times T$ symmetric $U(1)$ quantum spin liquid states based on the properties of their bulk fractional excitations. However, we will find that 11 of them are anomalous in the sense that these states cannot be realized in any three dimensional spin system with time reversal and $SO(3)$ spin rotational symmetries. We will present various ways of understanding the 15 non-anomalous families. In particular, we describe their physical properties and their construction as gauged SPTs. In Sec. 7.3, we explain why the other 11 states are anomalous. In Sec. 7.4 we discuss the topo-
logical response of the $U(1)$ spin liquids to an $SO(3)$ probe gauge field, which leads to the notion of “fractional topological paramagnets”. In Sec. 7.5, we combine the non-anomalous $U(1)$ quantum spin liquids with 3D bosonic SPTs with the same symmetry, and discuss how the presence of the SPTs further enriches the classification of the quantum spin liquids. After warming up with the example of $SO(3) \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids, in Sec. 7.6 we describe a general framework to classify symmetry enriched $U(1)$ quantum spin liquids for a large class of symmetries. We will apply the general framework to classify $Z_2 \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids in Sec. 7.7, and to classify $U(1)$ quantum spin liquids with some other symmetries in Sec. 7.8. Finally, we conclude in Sec. 7.9. The appendices contain some supplementary details, and some contents there are interesting and important, albeit rather technical.

7.2 $U(1)$ Quantum Spin Liquids Enriched by Time Reversal and $SO(3)$ Spin Rotational Symmetries

We will start by considering systems of interacting spins on a lattice with $SO(3) \times \mathcal{T}$ symmetries. The microscopic Hilbert space thus has a tensor product structure. Further all local operators in this Hilbert space will transform under linear representations of the $SO(3) \times \mathcal{T}$ symmetry (i.e integer spin and Kramers singlet). Also, these local operators can only create bosonic excitations. Our goal is to classify and characterize $U(1)$ quantum spin liquids that can emerge in such systems with the simplifying assumption that only the emergent photon is gapless.

A first cut understanding of the different possible such $U(1)$ spin liquids is obtained by focusing on the properties of the gapped matter excitations, such as their statistics and their quantum numbers under the relevant symmetries [241]. In three dimensions, the statistics of particles can be either bosonic or fermionic. Under time reversal symmetry, they can be Kramers doublets or non-Kramers. Under $SO(3)$, they can either be in a linear representation (spin-1) or its projective representation (spin-1/2). Note that any excitation with integer spin can be reduced to spin-1 by binding local excitations (i.e excitations created by local operators), and half-integer spin excitations can be similarly reduced to ones with minimal spin-1/2. Thus the only distinction is between linear and projective realizations of the global symmetry.
In the presence of time reversal symmetry, it is helpful to integrate out the gapped matter fields and consider the effective theory of the photon field. In general this effective theory has the form

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{Maxwell}} + \frac{\theta}{4\pi^2} E \cdot B$$  \hspace{1cm} (7.1)$$

where $\mathcal{L}_{\text{Maxwell}}$ represents the usual Maxwell Lagrangian, and the second term is of topological character. It is customary to define time-reversal transform such that the electric charge is invariant, namely under time reversal $E \rightarrow E$ and $B \rightarrow -B$. This definition also implies that $\theta \rightarrow -\theta$ under time-reversal. It is also known that $\theta$ is periodic with a period $4\pi$ if the elementary electric charge is a boson, or a period $2\pi$ if the elementary electric charge is a fermion (see Refs. [337, 338] for arguments from a condensed matter perspective). In all cases, the possible electric and magnetic charges of excitations form a two-dimensional lattice, and there are only two distinct configurations of this charge-monopole lattice, i.e. $\theta = 0 \pmod{2\pi}$ and $\theta = \pi \pmod{2\pi}$, as shown in Fig. 7.1 and Fig. 7.2, respectively. For notational simplicity, we will denote these two cases by $\theta = 0$ and $\theta = \pi$, respectively. Notice we take the normalization that the elementary electric charge is 1, and the minimal magnetic charge is such that it emits $2\pi$ flux seen by the elementary charge.

It is natural to ask whether time-reversal can act on the charge-monopole lattice in more complicated ways. Some examples were discussed in Ref. [245], in which the charge-monopole lattices undergoes a rotation (also known as S-duality transform) under time-reversal. However, in those examples the theories can be redefined, through appropriate electric-magnetic duality transforms, into the conventional form with the canonical time-reversal transform ($E \rightarrow E$ and $B \rightarrow -B$). In general, such a redefinition should always be possible if the theory, while preserving time-reversal symmetry, has a weakly coupled limit (with gauge coupling $e^2 \ll 1$).

Denote an excitation with electric charge $q_e$ and magnetic charge $q_m$ by $(q_e, q_m)$. When $\theta = 0$, the lattice of charge-monopole excitations is generated by the two particles $(1, 0)$ which we denote $E$ and $(0, 1)$ which we denote $M$. Then the distinct possibilities for the statistics and quantum numbers of $E$ and $M$ will correspond to distinct $U(1)$ quantum spin liquids. Under time
reversal an excitation with nonzero magnetic charge is transformed to another excitation that differs from the original one by a nonlocal operation. It is then meaningless to discuss whether these excitations are Kramers doublet or not, because \( T^2 \) is not a gauge invariant quantity for them \([316,339]\). On the other hand, all the pure electric charges should have well-defined \( T^2 \), and they are either Kramers singlet (\( T^2 = 1 \)) or Kramers doublet (\( T^2 = -1 \)). More details can be found in Appendix G.1.

For the case with \( \theta = \pi \), time reversal interchanges \((\frac{1}{2}, 1)\) and \((\frac{1}{2}, -1)\), which generate the entire charge-monopole lattice. In this case, we will still denote \( E \) as the \((1, 0)\) excitation, but we denote \( M \) as the \((0, 2)\) excitation.

\[ \text{Figure 7.1: Charge-monopole lattice at } \theta = 0 \text{ (mod } 2\pi). \]

\[ \text{Figure 7.2: Charge-monopole lattice at } \theta = \pi \text{ (mod } 2\pi). \]

7.2.1 Quantum Spin Liquids with \( \theta = 0 \)

We start with phases where \( \theta = 0 \). Let us consider the distinct possibilities for the \( E \) and \( M \) particles. Note that U(1) quantum spin liquids with both \( E \) and \( M \) fermionic are anomalous, i.e.,
they cannot be realized in a strictly three dimensional bosonic system but they can be realized as the surface of some four dimensional bosonic systems. [339–341] We will therefore restrict to situations in which at most one of $E$ and $M$ is a fermion. Consider the case where $E$ is a boson. Naively then $E$ may have $SO(3)$ spin $S = 0$ or $S = 1/2$, and may be Kramers singlet or doublet, while $M$ may be either a boson or fermion, and may have $S = 0$ or $1/2$. This gives 16 distinct possibilities. If instead $E$ is a fermion, it may again have $S = 0$ or $1/2$, and $T^2 = \pm 1$ while $M$ must be a boson but may have $S = 0$ or $1/2$, corresponding to 8 distinct possibilities. In total this gives 24 distinct possibilities for the $E$ and $M$ particles which each correspond to a distinct symmetry enriched $U(1)$ QSL (see Figure 7.3). However we will argue below that of these 10 are anomalous (i.e. the symmetry implementation is inconsistent in a strictly $3+1$-D system and is only consistent at the boundary of a $4+1$-dimensional SPT phase). We will discard these so that there are only 14 distinct possibilities for the $E$ and $M$ particles at $\theta = 0$. These will describe 14 distinct families of $U(1)$ QSLs.

![Figure 7.3: Symmetry protected distinctions among symmetry enriched $U(1)$ quantum spin liquids. For example, with $SO(3) \times T$ symmetries, two phases, $E_bM_b$ and $E_{bT^{1/2}}M_{f^{1/2}}$ in this example, cannot be connected without crossing a phase transition. When the symmetry is broken, they can be connected without crossing a phase transition.](image)

In Table 7.1, we list these distinct possible families, and introduce labels for them that we will use in the rest of the paper. The rest of this subsection will explain how to obtain these 14 spin liquids and Sec. 7.3 will show that the other 10 spin liquids are anomalous.

Among the 14 quantum spin liquids, the 6 of them in which none of $E$ or $M$ carries spin-$1/2$ have been described in detail previously [241]². Below we demonstrate how the other 8 can be...
Table 7.1: List of U(1) quantum spin liquids at $\theta = 0$. The subscripts “b” and “f” refer to bosonic or fermionic statistics of the associated particle, respectively. $T^2_E = 1$ ($T^2_E = -1$) means the electric charge is a Kramers singlet (doublet). $S_E$ and $S_M$ refer to the spin of the corresponding particle under $SO(3)$.

In this table the spin liquids with both $E$ and $M$ fermions are not listed, because they are known to be anomalous. We identified ten more anomalous spin liquids, and they are divided into three classes. More details can be found in the main texts.

<table>
<thead>
<tr>
<th>$T^2_E$</th>
<th>$S_E$</th>
<th>$S_M$</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_bM_b$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_bM_f$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_mM_b$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>$E_mM_f$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>$E_{bt}M_b$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_{bt}M_f$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_{bt}^2M_b$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>$E_{bt}^2M_f$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>$E_{ft}M_b$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_{ft}M_f$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_{ft}^2M_b$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>$E_{ft}^2M_f$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T^2_E$</th>
<th>$S_E$</th>
<th>$S_M$</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bf}M_{b\frac{1}{2}}$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{bf}M_{f\frac{1}{2}}$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{ft}M_{b\frac{1}{2}}$</td>
<td>-1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{ft}M_{f\frac{1}{2}}$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
</tr>
<tr>
<td>$E_{bt}M_{b\frac{1}{2}}$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{bt}M_{f\frac{1}{2}}$</td>
<td>-1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{bf}M_{b\frac{1}{2}}$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{bf}M_{f\frac{1}{2}}$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{ft}M_{b\frac{1}{2}}$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{ft}M_{f\frac{1}{2}}$</td>
<td>-1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>
constructed. Many of these spin liquids can be obtained simply. Specifically if either $E$ or $M$ is a trivial boson \( i.e \) has $S = 0$ and (for $E$ particles) $T_E^2 = 1$, then the corresponding spin liquid is obtained by gauging a trivial insulator of the other particle. For instance, to obtain $E_b^{\frac{1}{2}} M_b$, start with a trivial insulator formed by bosons with $S = 1/2$ and a conserved $U(1)$ charge that is even under time reversal. Coupling this charge to a dynamical $U(1)$ gauge field produces a quantum spin liquid which is precisely $E_b^{\frac{1}{2}} M_b$. If instead we wanted to obtain $E_b M_b^{\frac{1}{2}}$, we begin with a trivial insulator of a boson with $S = 1/2$ and a conserved $U(1)$ charge that is odd under time reversal. Gauging this insulator produces $E_b M_b^{\frac{1}{2}}$. This kind of construction clearly works for 6 of the 8 phases where one of $E$ or $M$ is a trivial boson while the other has $S = 1/2$. It is instructive to also understand these phases from a different ‘dual’ perspective where we will need to gauge the $U(1)$ symmetry of some SPTs with symmetries that contain a $U(1)$ subgroup. We explain this first below. This will also set the stage to understand the two interesting remaining cases where neither $E$ nor $M$ is a trivial boson (these are $E_b T_\frac{1}{2} M_f^{\frac{1}{2}}$ and $E_f T_\frac{1}{2} M_b^{\frac{1}{2}}$).

1. $E_b M_b^{\frac{1}{2}}$

From the point of view of $M$ (that is, viewing $M$ as the gauge charge), $E_b M_b^{\frac{1}{2}}$ can be viewed as a gauged trivial bosonic insulator with symmetry \(((U(1) \times SU(2))/Z_2) \times \mathcal{T}\). From the point of view of $E$, it can be viewed as a gauged SPT with symmetry \((U(1) \times \mathcal{T}) \times SO(3)\), where the microscopic boson is a Kramers singlet. This SPT is denoted by $e C m^1_2$, which means that it can have a surface topologically ordered (STO) state with $Z_2$ topological order, where the topological sectors, \((1, e, m, \epsilon)\), have $e$ carrying charge-1/2 under the $U(1)$ symmetry and $m$ carrying spin-1/2 under the $SO(3)$ symmetry. This SPT is discussed in more details in Appendix G.2.

2. $E_b M_f^{\frac{1}{2}}$

From the point of view of $M$, $E_b M_f^{\frac{1}{2}}$ can be viewed as a gauged trivial fermionic insulator with symmetry \(((U(1) \times SU(2))/Z_2) \times \mathcal{T}\). From the point of view of $E$, it can be viewed as a gauged bosonic SPT with symmetry \((U(1) \times \mathcal{T}) \times SO(3)\), where the microscopic boson is exchanges $e$ and $m$, while their neutral bound state $\epsilon$ is a Kramers singlet. This surface state is labelled as \((e C m C) T_\epsilon\).
a Kramers singlet. We denote this SPT by $eCmC_{\frac{1}{2}}$, which can be viewed as a combination of $eCm_{\frac{1}{2}}$ and $eCmC$, a well-known SPT with symmetry $U(1) \times T$ or $U(1) \ltimes T$. [316,337,338] In fact, $eCmC$ is still a nontrivial SPT even if there is additional $SO(3)$ symmetry that commutes with $U(1) \times T$ or $U(1) \ltimes T$.

3. $E_{b\frac{1}{2}}M_b$

From the point of view of $E$, $E_{b\frac{1}{2}}M_b$ can be viewed as a gauged trivial bosonic insulator with $((U(1) \times T) \times SU(2))/Z_2$, where the microscopic bosons are Kramers singlets. From the point of view of $M$, it can be viewed as the gauged $eCm_{\frac{1}{2}}$, but with symmetry $U(1) \times T \times SO(3)$.

4. $E_{bT\frac{1}{2}}M_b$

From the point of view of $E$, $E_{bT\frac{1}{2}}M_b$ can be viewed as a gauged trivial bosonic insulator with $((U(1) \times T) \times SU(2))/Z_2^2$, where the microscopic bosons are Kramers doublets. From the point of view of $M$, it can be viewed as a gauged SPT $eCmT_{\frac{1}{2}}$ under symmetry $U(1) \times T \times SO(3)$. This SPT can be viewed as a combination of $eCm_{\frac{1}{2}}$ and $eCmT$, another well-known SPT with symmetry $U(1) \times T$. [316,337] It can be shown that this is still a nontrivial SPT even if there is additional $SO(3)$ symmetry that commutes with $U(1) \times T$.

5. $E_{f\frac{1}{2}}M_b$

From the point of view of $E$, $E_{f\frac{1}{2}}M_b$ can be viewed as a gauged trivial fermionic insulator with $((U(1) \times T) \times SU(2))/Z_2$, where the microscopic fermions are Kramers singlet. From the point of view of $M$, it can be viewed as a gauged $eCmC_{\frac{1}{2}}$ with symmetry $U(1) \times T \times SO(3)$.

6. $E_{fT\frac{1}{2}}M_b$

From the point of view of $E$, $E_{fT\frac{1}{2}}M_b$ can be viewed as a gauged trivial fermionic insulator with $((U(1) \times T) \times SU(2))/Z_2^2$, where the microscopic fermions are Kramers doublets. From the point of view of $M$, it can be viewed as a gauged $eCmCT_{\frac{1}{2}}$ under symmetry $U(1) \times T \times SO(3)$. This SPT can be viewed as a combination of $eCm_{\frac{1}{2}}$, $eCmC$ and $eCmT$. 176
We now turn to the last 2 cases $E_{bT^1_2} M^1_2$ and $E_{f^1_2} M^1_2$. As both the $E$ and $M$ are non-trivial in these spin liquids, in both the electric and magnetic pictures they should be viewed as gauged SPTs. We state their construction here and describe their properties in greater detail later. We will see that they should be viewed as “Fractional Topological Paramagnets”.

7. $E_{bT^1_2} M^1_2$

From the point of view of $M$, $E_{bT^1_2} M^1_2$ can be viewed as a gauged $n = 2$ topological superconductor with symmetry $((U(1) \times SU(2))/Z_2) \times \mathcal{T}$. From the point of view of $E$, it can be viewed as a gauged bosonic $\theta = 2\pi$ SPT under symmetry $((U(1) \times SU(2))/Z_2^2)$, where the microscopic bosons are Kramers doublets.

8. $E_{f^1_2} M^1_2$

From the point of view of $E$, $E_{f^1_2} M^1_2$ can be viewed as a gauged $n = 2$ topological insulator of fermions with symmetry $((U(1) \times SU(2))/Z_2)$, where the microscopic fermions are Kramers singlets. From the point of view of $M$, it can be viewed as a gauged bosonic $\theta = 2\pi$ SPT with symmetry $((U(1) \times SU(2))/Z_2) \times \mathcal{T}$. The properties of this SPT is described in Ref. [331].

7.2.2 Quantum spin liquids with $\theta = \pi$

Now we turn to $U(1)$ quantum spin liquids with $\theta = \pi$. At $\theta = \pi$, the charge-monopole lattice is shown in Fig. 7.2. Because time reversal symmetry exchanges $(\frac{1}{2}, 1)$ and $(\frac{1}{2}, -1)$, they should have the same statistics and quantum numbers. Further, they have $\pi$ mutual braiding statistics. This implies that $E$, the bound state of $(\frac{1}{2}, 1)$ and $(\frac{1}{2}, -1)$, has to be a Kramers doublet spin-1 fermion. [339] Also, because $(-\frac{1}{2}, 1)$ dyon is the antiparticle of $(\frac{1}{2}, -1)$ dyon, it has the same properties as $(\frac{1}{2}, 1)$. Due to the mutual $\pi$ braiding between $(\frac{1}{2}, 1)$ and $(-\frac{1}{2}, 1)$, their bound state, $M$, is also a fermion that carries spin-1 and is non-Kramers. Similar thoughts imply that the statistics and quantum numbers of $(\frac{1}{2}, 1)$ will determine the statistics and quantum numbers of all gapped excitations. So the classification of $U(1)$ spin liquids with $\theta = \pi$ is equivalent to the classification of the statistics and quantum numbers of the $(\frac{1}{2}, 1)$ dyon.
Table 7.2: List of U(1) quantum spin liquids at $\theta = \pi$. $S_D = 1$ ($S_D = \frac{1}{2}$) represents the case where the $(\frac{1}{2}, 1)$ dyon carries spin-1 (spin-1/2).

<table>
<thead>
<tr>
<th>$E_{fT}M_f\theta$</th>
<th>$S_D$</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{fT}M_f\theta$</td>
<td>1</td>
<td>E: TI, M: n=1 TSC</td>
</tr>
<tr>
<td>$E_{fT}M_f\theta_{\frac{1}{2}}$</td>
<td>$\frac{1}{2}$</td>
<td>anomalous, class II</td>
</tr>
</tbody>
</table>

It is known that $(\frac{1}{2}, 1)$ must be a boson. [241, 339, 340] Under time reversal symmetry, $T^2$ is not a gauge invariant quantity for $(\frac{1}{2}, 1)$, so it is non-Kramers. Under $SO(3)$, it can carry either spin-1 or spin-1/2. We will denote the former by $(E_{fT}M_f)^\theta$ and the latter by $(E_{fT}M_f)^{\theta\frac{1}{2}}$. These states are summarized in Table 7.2.

$(E_{fT}M_f)^\theta$ has been described in detail in Ref. [241]. From the point of view of $E$, it can be viewed as a gauged free fermion topological insulator with symmetry $((U(1) \times T))/Z_2 \times SO(3)$, where the microscopic fermions are Kramers doublets. From the point of view of $M$, it can be viewed as a gauged $n = 1$ free fermion topological superconductor with symmetry $U(1) \times T \times SO(3)$.

In Sec. 7.3, we will show that $(E_{fT}M_f)^{\theta\frac{1}{2}}$ is anomalous.

7.3 Anomalous quantum spin liquids with $SO(3) \times T$ symmetry

In the enumeration in Sec. 7.2, 11 states are claimed to be anomalous, where 10 of them have $\theta = 0$ and 1 has $\theta = \pi$. In this section we will provide arguments to demonstrate these anomalies. We start with the 10 with $\theta = 0$.

7.3.1 Anomalous states with $\theta = 0$

The 10 anomalous quantum spin liquid states at $\theta = 0$ are grouped into three classes, such that within each class any one of them can be obtained by coupling another in the same class and some non-anomalous quantum spin liquids. For illustration, let us demonstrate how to obtain $E_fM_{\theta\frac{1}{2}}$ by coupling $E_{b\frac{1}{2}}M_{b\frac{1}{2}}$ and $E_{f\frac{1}{2}}M_{b\frac{1}{2}}$, a non-anomalous quantum spin liquid. To do this, one can couple $E_{b\frac{1}{2}}M_{b\frac{1}{2}}$ and $E_{f\frac{1}{2}}M_{b\frac{1}{2}}$ and condense the bound state of the monopole of $E_{b\frac{1}{2}}M_{b\frac{1}{2}}$ and the anti-monopole of $E_{f\frac{1}{2}}M_{b\frac{1}{2}}$. This bound state is a trivial boson, so condensing it will not
break any symmetry. After this condensation, the electric charge of $E_{b\frac{1}{2}}M_{b\frac{1}{2}}$ and that of $E_{f\frac{1}{2}}M_{b\frac{1}{2}}$ will be confined together, and the resulting bound state is a fermion that carries no nontrivial quantum number. This is precisely $E_{f}M_{b\frac{1}{2}}$.

The above example shows the relation between the two anomalous quantum spin liquids in class I. We denote this relation by

$$E_{b\frac{1}{2}}M_{b\frac{1}{2}} \xleftarrow{E_{f\frac{1}{2}}M_{b\frac{1}{2}}} E_{f}M_{b\frac{1}{2}}$$

(7.2)

The relations among the two other classes are listed here:

class II:

$$E_{b\frac{1}{2}}M_{f\frac{1}{2}} \xleftarrow{E_{bT}M_{f\frac{1}{2}}} E_{bT\frac{1}{2}}M_{f\frac{1}{2}}$$

$$\xleftarrow{E_{bT\frac{1}{2}}M_{f\frac{1}{2}}} E_{bT\frac{1}{2}}M_{b\frac{1}{2}} \xleftarrow{E_{f\frac{1}{2}}M_{b\frac{1}{2}}} E_{fT}M_{b\frac{1}{2}}$$

(7.3)

class III:

$$E_{b\frac{1}{2}}M_{f\frac{1}{2}} \xleftarrow{E_{bT}M_{f\frac{1}{2}}} E_{bT\frac{1}{2}}M_{f\frac{1}{2}}$$

$$\xleftarrow{E_{bT}M_{f\frac{1}{2}}} E_{bT\frac{1}{2}}M_{b\frac{1}{2}} \xleftarrow{E_{f\frac{1}{2}}M_{b\frac{1}{2}}} E_{fT}M_{b\frac{1}{2}}$$

(7.4)

Because of these relations, given that the other 14 quantum spin liquids can be realized in strictly three dimensional bosonic systems, showing that any one of the states of a certain class is anomalous is sufficient to show the entire class is anomalous. Below we will show that $E_{b\frac{1}{2}}M_{b\frac{1}{2}}$, $E_{bT\frac{1}{2}}M_{b\frac{1}{2}}$ and $E_{bT}M_{b\frac{1}{2}}$ are anomalous.

States of matter that realize a global symmetry non-anomalously allow a consistent coupling of background gauge fields. In our context a non-anomalous realization of $SO(3)$ symmetry thus implies that we can consistently couple background $SO(3)$ gauge fields. Conversely anomalous states can be detected by finding inconsistencies when such background gauge fields are turned on.

Let us therefore couple our spin liquids to a probe $SO(3)$ gauge field. Because $\pi_1(SO(3)) = Z_2$, there are monopole configurations of this $SO(3)$ gauge field that are classified by $Z_2$. [342]
One explicit expression of a nontrivial $SO(3)$ monopole configuration is:

$$A_1^\mu = A_2^\mu = 0, \ A_3^\mu = A_{U(1),\mu}$$ (7.5)

where $A_\mu = \sum A_\mu^a T^a$ is the Lie algebra valued $SO(3)$ gauge field with $T^a$ the generators. $A_{U(1),\mu}$ is the $U(1)$ gauge field configuration of a $U(1)$ monopole. [342] One of the physical consequences of this $SO(3)$ monopole is a Berry phase factor of an excitation going around a closed loop around it:

$$\exp \left( i \frac{\Omega}{2} S^z \right)$$ (7.6)

where $\Omega$ is the solid angle of the closed loop with respect to the monopole and $S^z$ is the representation of one of the generators of $SO(3)$. For spin-1 particles, $S^z$ can be taken to be $S^z_{S=1} = \text{diag}(1, 0, -1)$. For spin-1/2 particles, $S^z$ can be taken to be $S^z_{S=1/2} = \text{diag}(1/2, -1/2)$. This formula can be easily obtained by borrowing the well-known result of the Berry phase factor of a $U(1)$ charge moving around a $U(1)$ monopole and using (7.5).

Now consider a Dirac string that ends at this monopole. According to (7.6), moving around an infinitesimal loop around the Dirac string, a spin-1 particle will get a unit phase factor, which seems normal. But a spin-1/2 particle will see a phase factor of $-1$, which is unphysical. To cancel this phase factor, another defect that also gives a $-1$ phase factor to spin-1/2 particles around the Dirac string needs to be trapped at the $SO(3)$ monopole. We will denote such an $SO(3)$ monopole (with this defect included) by $M_{SO(3)}$.

Next we argue that the defect trapped at another $SO(3)$ monopole with

$$A_1^\mu = A_2^\mu = 0, \ A_3^\mu = -A_{U(1),\mu}$$ (7.7)

can be essentially the same as the one trapped at the previous $SO(3)$ monopole, $M_{SO(3)}$. This is because this new $SO(3)$ monopole can be obtained by performing on $M_{SO(3)}$ a $\pi$-rotation around any axis on the $xy$-plane. In the presence of $SO(3)$ symmetry, the defect trapped by it should be the same as that trapped in $M_{SO(3)}$ up to a spin rotation. We denote this $SO(3)$ (anti)monopole
(with the same defect included) by \( \mathcal{M}'_{SO(3)} \). Notice if an \( \mathcal{M}_{SO(3)} \) and an \( \mathcal{M}'_{SO(3)} \) are fused together, the \( SO(3) \) gauge field background will be cancelled, and what remains will be an excitation of the original system without any \( SO(3) \) gauge field. These imply that the defect trapped in these \( SO(3) \) monopoles can be viewed as “half” of an excitation of the quantum spin liquid.

For the quantum spin liquid states \( E_{b\frac{1}{2}} M_{b\frac{1}{2}} \) and \( E_{bT\frac{1}{2}} M_{b\frac{1}{2}} \), the Dirac string of a bare \( SO(3) \) monopole will give any excitation with spin-1/2 a \(-1\) phase factor. These excitations all satisfy \( \Delta q = q_e - q_m \) is odd. For these excitations, a \((Q_e, Q_m)\) dyon with odd \( Q_e \) and \( Q_m \) will give rise to a phase \( 2\pi(Q_e q_m - Q_m q_e) \) around an infinitesimal loop around the Dirac string, so “half” of such a \((Q_e, Q_m)\) dyon will give a phase, which is an odd multiple of \( \pi \), that exactly cancels the \(-1\) phase factor due to the bare \( SO(3) \) monopole. One can also check this \(-1\) phase factor cannot be cancelled by “half” of any other type of excitations, where at least one of \( Q_e \) and \( Q_m \) is even.

According to the argument above, fusing a \( \mathcal{M}_{SO(3)} \) with \( \mathcal{M}'_{SO(3)} \) here should give rise to an \((Q_e, Q_m)\) dyon, with both \( Q_e \) and \( Q_m \) odd integers. That is,

\[
\mathcal{M}_{SO(3)} \times \mathcal{M}'_{SO(3)} \sim D_{(Q_e, Q_m)}
\]  

However, the \((Q_e, Q_m)\) dyon is a fermion as long as both \( Q_e \) and \( Q_m \) are odd, \([343]\) and this is inconsistent: \( \mathcal{M}' \) and \( \mathcal{M} \) cannot have any nontrivial mutual Berry phase since they differ merely by a continuous gauge rotation, so the bound state of the two cannot be a fermion. Therefore, the above fusion rule is physically impossible. This shows that all states in class I and class II are anomalous.

The anomalies in the states discussed above do not involve time reversal symmetry in an essential way, but this is not the case for \( E_{bT} M_{b\frac{1}{2}} \). For \( E_{bT} M_{b\frac{1}{2}} \), the analogous fusion rule we will obtain is

\[
\mathcal{M}_{SO(3)} \times \mathcal{M}'_{SO(3)} \sim D_{(Q_e, Q_m)}
\]

with an odd \( Q_e \) and an even \( Q_m \). As \( Q_m \) is even, we can always bind \(-Q_m/2 \) \( U(1) \) monopoles to \( \mathcal{M}_{SO(3)} \) and \( \mathcal{M}'_{SO(3)} \) to cancel their magnetic charges. Thus \( Q_m \) can be taken to be zero in the above fusion rule. In this case, the time reversal partners of the \( \mathcal{M}_{SO(3)} \) (and \( \mathcal{M}'_{SO(3)} \) ) will
differ from itself only by a local operator. This implies that they have a well-defined value for $T^2$. However this is also seen to be impossible: first note that a $(Q_e, 0)$ dyon with odd $Q_e$ is a Kramers doublet in this case and all microscopic degrees of freedom are Kramers singlet. Suppose the fusion rule in Eqn. 7.9 is possible, then $\mathcal{M}_{SO(3)}$ and $\mathcal{M}'_{SO(3)}$ should satisfy $T^4 = -1$. The argument in Appendix G.1 shows this is impossible unless there are microscopic Kramers doublets, which is absent by assumption. Therefore, $E_b T M_{\frac{1}{2}}$ and hence all states in class III are anomalous.

7.3.2 Anomalous state with $\theta = \pi$

Now we show $(E_f T M_f)_{\theta \frac{1}{2}}$ is also anomalous. The simplest way to see this is to first ignore time reversal symmetry, then from the point of view of $(\frac{1}{2}, 1)$ and $(\frac{1}{2}, -1)$ dyons, this spin liquid is just $E_{\frac{1}{2}} M_{\frac{1}{2}}$. We have shown $E_{\frac{1}{2}} M_{\frac{1}{2}}$ is anomalous with SO(3) symmetry alone even without using time reversal symmetry, this implies $(E_f T M_f)_{\theta \frac{1}{2}}$ must be anomalous. Another way to see the anomaly is to notice the relation

\[
(E_f T M_f)_{\theta \frac{1}{2}} \xrightarrow{(E_f T M_f)_{\theta \frac{1}{2}}} E_f T M_{\frac{1}{2}}
\]  

(7.10)

This also shows $(E_f T M_f)_{\theta \frac{1}{2}}$ is anomalous, and in the presence of time reversal symmetry its anomaly belongs to class II.

A more direct argument similar to the ones used above goes as follows. In this case, all $(q_e, q_m)$ dyons with $q_e$ an half-odd-integer and $q_m$ an odd integer carry spin-1/2. This implies the following fusion rule

\[
\mathcal{M}_{SO(3)} \times \mathcal{M}'_{SO(3)} \sim D_{(Q_e, Q_m)}
\]  

(7.11)

with $Q_e = 2n$ and $Q_m = 4m + 2$, or $Q_e = 2n + 1$ and $Q_m = 4m$, where $n$ and $m$ are integers. One can check this dyon must be a fermion, which in turn shows that this spin liquid is anomalous.
7.3.3 Some comments

The above arguments show that the 11 quantum spin liquids cannot be realized in strictly three dimensions made of bosons if the symmetry $SO(3) \times T$ is present. Careful readers may have noticed that the descendent states of these anomalous states will still be anomalous if the symmetry is broken down to $(U(1) \times Z_2) \times T \cong O(2) \times T$, where $U(1)$ is the spin rotation around one axis, say, the $z$ axis, and $Z_2$ is a discrete $\pi$-spin rotation around an axis perpendicular to the $z$ axis. In this case, we can couple the system to a $U(1)$ gauge field corresponding to the spin rotational symmetry around the $z$ axis, then $M_{SO(3)}$ and $M'_{SO(3)}$ become the monopoles of this $U(1)$ gauge field, and the analogous equations of (7.8) and (7.9) still hold. These two monopoles are mapped into each other by the $Z_2$ transformation. Because this unitary $Z_2$ transformation flips both $S_z$, the spin component along the $z$ direction, and the field value of the $U(1)$ gauge field corresponding to $S_z$ rotational symmetry, there is no mutual statistics between these two monopoles. Therefore, all previous arguments still apply. In fact, we conjecture even if the symmetry is broken down to $Z_2 \times Z_2 \times T$, the descendant states of these anomalous states will still be anomalous\(^3\). In Appendix G.3 we will show they can be realized as the surface of some four dimensional short-range entangled bosonic systems. In particular, four dimensional bosonic SPT states with only $SO(3)$ symmetry were discussed in Ref. [344] using group cohomology, where the SPT states have a $Z_2$ classification. This is consistent with our result: the only anomalous $U(1)$ spin liquid with $SO(3)$ symmetry is $E_{b_1} M_{b_{1}} \rangle$.

If these states were not anomalous, they could also be viewed as some gauged SPTs. So their anomalies imply the impossibilities of some SPTs, which is discussed in more general terms in Sec. 7.6.2. One such example is given in Appendix G.2.

We would also like to mention that, although the anomalies of these states are shown by examining the $SO(3)$ monopoles, an alternative argument independent of the $SO(3)$ monopoles is sketched in Sec. 7.6.

In passing, we notice that the fact that “half” of a dyon is confined by itself does not invalidate our arguments. In fact, the phenomenon where a defect is unphysical unless it traps a con-

\(^3\)However, if the symmetry is broken down to $U(1) \times T$, the descendants of all the anomalous states will become non-anomalous (see Sec. 7.8.1).
fined object is familiar. The most familiar example may be that in a conventional two dimensional superconductor obtained by condensing charge-1 bosonic chargons from a spin-charge separated described by a $Z_2$ gauge theory, a $\pi$-flux always appears with a vison, which is confined by itself in the superconducting phase. [109]

7.4 Fractional topological paramagnets

In this section we study the topological response of the spin liquid phases to an external $SO(3)$ gauge field that couples with the $SO(3)$ spin degrees of freedom. In particular we show that the two phases $E_{bT}^{1/2}M_{f_1}^{1/2}$ and $E_f^{1/2}M_{b_2}^{1/2}$ exhibit nontrivial fractionalized topological response, due to which we dub them “fractional topological paramagnets”.

We start with non-fractionalized (short-range entangled) bosonic phases with $SO(3) \times T$ symmetry, coupled with a background $SO(3)$ gauge field $A_\mu$. Since the bulk dynamics is trivial by assumption, we can integrate out all the bulk degrees of freedom and ask about the effective response theory for the $SO(3)$ gauge field $A$. The simplest topological response is a theta-term:

$$S_\Theta = \frac{\Theta}{16\pi^2} \int \text{tr}_{SO(3)} F \wedge F,$$

(7.12)

where $F$ is the $SO(3)$ field strength. The normalization is chosen so that if the $SO(3)$ symmetry is broken down to $U(1) \sim SO(2)$, the term becomes a theta-term for the $U(1)$ gauge field with familiar normalization.

It is important to realize that the period of $\Theta$ is $4\pi$ for purely bosonic systems, in contrary to fermionic systems where the period is $2\pi$. In fact a bosonic short-range entangled phase with $\Theta = 2\pi$ is a nontrivial SPT state protected by $SO(3) \times T$. The physics behind is what is known as the “statistical Witten effect” [338]: consider inserting a monopole configuration of $A$ of the form of Eq. (7.5), we can ask about the $SO(3)$ charge carried by this monopole. But since the monopole configuration already breaks the symmetry down to $SO(2) \sim U(1)$, we can only ask about the $U(1)$ charge it carries. The standard Witten effect implies that the monopole carries $U(1)$ charge $q_s = \Theta/2\pi = 1$. We can bind a gauge charge to the monopole to neutralize the gauge charge, but this converts the monopole to a fermion [343].
The above argument also shows that for short-range entangled bosonic phases with $SO(3) \times T$, the minimal nontrivial $\Theta$-angle is $2\pi$ since under time-reversal $\Theta \to -\Theta$. However, it is also known that for long-range entangled (fractionalized) phases, time-reversal symmetry could be compatible with smaller $\Theta$-angles [345, 346]. This is because the effective period of $\Theta$ is reduced due to the presence of fractionalized excitations. More formally, in the presence of emergent dynamical gauge fields, it is more appropriate to integrate out only the gapped matter fields and keep the low energy dynamics of the gauge field explicit. The response theory is then correctly captured by a $\Theta$-term and a dynamical term

$$\tilde{S}_\Theta = \frac{\Theta}{2\pi^2} \int \text{tr}_{SO(3)} F \wedge F + S'_\Theta[a_\mu, A_\mu],$$

(7.13)

where the second term involves the dynamical gauge field $a_\mu$. It is this $\tilde{S}_\Theta$ that has a reduced period of $\Theta$. We will explain this point in more concrete examples at the end of Sec. 7.4.2. However, to understand the physics, it suffices to simply study the properties of an $SO(3)$ magnetic monopole (the Witten effect) carefully – we will mainly focus on this approach here.

We argue below, in the context of $U(1)$ spin liquids, that the effective period of $\Theta$ is reduced to $\pi$ when spin-1/2 excitations are allowed in the bulk. This allows, in principle, time-reversal symmetric phases with $\Theta = \pi/2 \pmod{\pi}$.

7.4.1 Triviality of $\Theta = \pi$

First, we need to show that $\Theta = \pi$ is in some sense trivial if (and only if) there are spin-1/2 excitations (either $E$ or $M$ particle). Our argument proceeds by carefully studying the Witten effect. Consider again a monopole of $A$ of the form of Eq. (7.5), denoted by $\mathcal{M}_{SO(3)}$. In general it could carry both the $SO(2)$ charge $q_s = \Theta/2\pi = 1/2$, and the electric-magnetic charge of the dynamical $U(1)$ gauge field $(q_e, q_m)$. We denote this object as $\mathcal{M}_{SO(3)} = (q_e, q_m, q_s, q_M) = (q_e, q_m, 1/2, 1)$. Time-reversal symmetry implies that the object $\tilde{\mathcal{M}}_{SO(3)} = (q_e, -q_m, -q_s, q_M) = (q_e, -q_m, -1/2, 1)$ must also exist in the spectrum, and it must have the same statistics with $\mathcal{M}_{SO(3)}$. One can think of $\tilde{\mathcal{M}}_{SO(3)}$ as $\mathcal{M}_{SO(3)}$ attached with a $(0, 2q_m, 1, 0)$ particle (which implies that this particle should exist in the excitation spectrum). Notice that if $q_m = 0$, this at-
tachment will change the statistics of $M_{SO(3)}$ from boson to fermion (or vice versa), and $M_{SO(3)}$ cannot have the same statistics with $\tilde{M}_{SO(3)}$ – this is precisely why in the absence of fractionalization, $\Theta = \pi$ cannot be time-reversal symmetric for a bosonic system. Now with nonzero $(q_m, q_e)$, the issue can be cured by another statistical transmutation if

$$2q_e q_m + S_{(0, 2q_m, 1, 0)} = 1 \pmod{2},$$

(7.14)

where $S_{(0, 2q_m, 1, 0)} = 0$ if the $(0, 2q_m, 1, 0)$ particle is a boson, and $S_{(0, 2q_m, 1, 0)} = 1$ if it is a fermion.

Furthermore, any excitations in the (ungauged) $U(1)$ spin liquid $(q'_e, q'_m, q'_s, 0)$ should satisfy the general Dirac quantization condition with respect to $M_{SO(3)}$:

$$q_e q'_m - q_m q'_e - q'_s = 0 \pmod{1}.$$  

(7.15)

The two conditions Eq. (7.14) and (7.15), together with the existence of $(0, 2, 1, 0)$ in the spectrum, strongly constrains the allowed values of $(q_e, q_m)$ for $M_{SO(3)}$ and the allowed spectra of the $U(1)$ spin liquids. For example, the $(E_f T M_f)_\theta$ spin liquid could satisfy Eq. (7.14) with $q_e = 0, q_m = 1$, but this choice inevitably violates Eq. (7.15) with $q'_e = 1/2, q'_m = 1, q'_s = 0$. A related phase $(E_f T M_f)_{\theta_{1/2}}$ could satisfy all conditions since the test particle for Eq. (7.15) should have $q'_e = 1/2, q'_m = 1, q'_s = 1/2$ – the problem is that this phase is anomalous and cannot be realized in three dimensions on its own. It can be seen after some careful examination, that among the anomaly-free $U(1)$ spin liquids, only those with either $E$ or $M$ particle (but not both) carrying spin-$1/2$ are allowed for $\Theta = \pi$. These include $(E_b M_{b\frac{1}{2}}, E_b M_{f\frac{1}{2}}, E_b M_{b}, E_{bT\frac{1}{2}} M_b, E_{f\frac{1}{2}} M_b, E_{fT\frac{1}{2}} M_b)$. The values of $q_e$ and $q_m$ for $M_{SO(3)}$ are chosen in the following way: if $E$ particle carries spin-$1/2$, then $q_e = 1 \pmod{2}$ and $q_m = 1/2 \pmod{1}$; if $M$ carries spin-$1/2$, then $q_e = 1/2 \pmod{1}$ and $q_m = 1 \pmod{2}$. This choice is needed to satisfy the Dirac quantization condition Eq. (7.15).

It is also easy to check that Eq. (7.14) is satisfied (for those states without anomaly).

It is now easy to see why $\Theta = \pi$ should be considered trivial. In those spin liquids where $E$ particles carry spin-$1/2$, we can bind an $E$ particle to $M_{SO(3)}$. This gives another, equally legitimate, $SO(3)$ monopole with $q_s = 0$ and $q_e = 0$. We still have $q_m = 1/2$ for the monopole, but this is simply a consequence of the spin-$1/2$ carried by $E$ particle which should be true re-
gardless of what value $\Theta$ takes. Therefore one can equivalently view this phase as having $\Theta = 0 \pmod{2\pi}$ (notice that both $q_s = 0$ and $q_e = 0$ for the redefined monopole are important to draw this conclusion). The argument is identical if $M$ particles carry spin-$1/2$ instead. There is still the ambiguity whether the redefined $q_s = 0$ monopole is a boson or a fermion, but this is simply about whether $\Theta = 0$ or $\Theta = 2\pi \pmod{4\pi}$ – or whether a boson SPT state has been stacked on top of the $U(1)$ spin liquid. We therefore conclude that for a $U(1)$ spin liquid, $\Theta = \pi$ is trivial.

7.4.2 $\Theta = \pi/2$: Fractional Topological Paramagnets

We now argue that the two $U(1)$ spin liquid phases $E_{bT\frac{1}{2}} M_{f\frac{1}{2}}$ and $E_{f\frac{1}{2}} M_{b\frac{1}{2}}$ effectively have $\Theta = \pi/2$, and hence can be called “Fractional Topological Paramagnets”.

Again we consider a monopole $\mathcal{M}_{SO(3)}$. In general it could carry both an $SO(2)$ charge $q_s$, and the electric-magnetic charge of the dynamical $U(1)$ gauge field $(q_e, q_m)$. Since both the fundamental electric and magnetic excitations ($E$ and $M$) of the two spin liquids carry spin-$1/2$, according to the argument in Sec. 7.3.1 we require $(q_e, q_m) = (1/2, 1/2)$ for $\mathcal{M}_{SO(3)}$, up to integer shifts. We denote this object as $\mathcal{M}_{SO(3)} = (q_e, q_m, q_s, q_M) = (1/2, 1/2, q_s, 1)$. Time-reversal symmetry implies that the object $\tilde{\mathcal{M}}_{SO(3)} = (q_e, -q_m, -q_s, q_M) = (1/2, -1/2, -q_s, 1)$ must also exist in the spectrum. Now take the anti-particle of $\tilde{\mathcal{M}}_{SO(3)}$ and bind it together with $\mathcal{M}_{SO(3)}$, we get an object $(0, 1, 2q_s, 0)$. Since this object does not carry magnetic charge of the $SO(3)$ gauge field, it must exist in the $U(1)$ spin liquid phase before coupling to $\mathcal{A}^{SO(3)}$. But in the spin liquid phase any particle with $q_m = 1$ and $q_e = 0$ must carry spin-$1/2$. Therefore $2q_s = 1/2 \pmod{1}$ and $q_s = 1/4 \pmod{1/2}$. This implies an effective $\Theta = \pi/2 \pmod{\pi}$.

One can also ask whether $E_{bT\frac{1}{2}} M_{f\frac{1}{2}}$ and $E_{f\frac{1}{2}} M_{b\frac{1}{2}}$ are the only two ($T$-invariant) $U(1)$ spin liquids with $\Theta = \pi/2$. An argument similar to that in Sec. 7.4.1 for $\Theta = \pi$ shows that these two are indeed the only $U(1)$ spin liquids with $\Theta = \pi/2$.

The fractional value of $\Theta$ for the two spin liquids can be understood quite easily if they are viewed as some gauge SPT states (as discussed in Sec. 7.2.1). For concreteness we take $E_{f\frac{1}{2}} M_{b\frac{1}{2}}$ as example (the logic will be parallel for the other state). This state can be obtained from $E_{f\frac{1}{2}} M_{b}$ by putting the fermionic $E$ particles into a topological band. The corresponding surface state for $E$ will have two Dirac cones – one for each spin. It is well known [347] that this state, when cou-
pled to an $SU(2)$ gauge field, induces a theta-term for the $SU(2)$ gauge field at $\Theta_{SU(2)} = \pi$. This implies $\Theta = \pi/2$ for the $SO(3)$ gauge field.

The Witten effect is also easy to study in this picture: an $SO(3)$ monopole $\mathcal{M}_{SO(3)}$ is viewed by the spin-$1/2$ $E$ particles as a half-monopole. Therefore it should bind with a magnetic charge $q_m = 1/2 \pmod{1}$. Let’s choose $q_m = 1/2$. The monopole is then viewed as a $q_m = 1$ monopole by the spin-up fermion $f^\uparrow$, and a $q_m = 0$ object viewed by the spin-down fermion $f^\downarrow$. Since each fermion (up or down) has one Dirac cone on the surface, similar to the usual topological insulator, the $\mathcal{M}_{SO(3)}$ monopole will trap half of the charge of an $f^\uparrow$ fermion, which gives $q_e = 1/2$ and $q_s = 1/4$, in agreement with what was obtained earlier using a direct argument.

Alternatively, one can obtain the $E_{f^\frac{1}{2}}M_{b^\frac{1}{2}}$ state from $E_b M_{b^\frac{1}{2}}$ by putting the spin-$1/2$ boson $M$ into a bosonic topological insulating state. The result should be identical, even though the bosonic state is harder to picture due to the lack of non-interacting limit.

We can make the picture slightly more precise by writing down the response theory. We first consider the electric picture, viewing the state as a gauged fermion SPT. This is the more convenient choice if the gauge coupling for the Maxwell term $e^2$ is weak. Integrating out the fermion matter field gives (on a general oriented manifold $Y^4$)

$$\tilde{S}_{\Theta=\pi/2} = \frac{\pi}{2} \left( \frac{1}{16\pi^2} \int \text{tr}_{SO(3)} F \wedge F + \frac{1}{2\pi^2} \int f \wedge f + \frac{1}{2} \cdot 24\pi^2 \int \text{tr} R \wedge R \right),$$

where $f = da$ is the field strength for the dynamical gauge field, and $R$ is the Riemann curvature tensor. The first term comes from the $\Theta_{SU(2)} = \pi$ response of the fermion topological band. The second and third terms are the $U(1)$ and gravitational theta-terms induced by the fermions. The gauge field strength $f$ satisfies the cocycle condition

$$\int \left( \frac{f}{\pi} + w_2^{TM} + w_2^{SO(3)} \right) = 0 \pmod{2},$$

where $w_2^{TM}$ is the second Stiefel-Whitney class of the tangent bundle on $Y^4$, $w_2^{SO(3)}$ is the second Stiefel-Whitney class of the $SO(3)$ gauge bundle (physically it measures the $\mathbb{Z}_2$-valued $SO(3)$ monopole number and serves as an obstruction to lifting the gauge bundle to an $SU(2)$ one), and the integration is taken on arbitrary 2-cycles on $Y^4$. The Maxwell term for $f$ is suppressed in the
above equation for simplicity. Physically this cocycle condition simply represents the fact that charge-1 objects under $a_\mu$ must carry spin-1/2 of the global $SO(3)$ symmetry and must also be a fermion. When $w_2^{TM}$ is trivial, this requires an $SO(3)$ monopole to be accompanied by a half $U(1)$ magnetic-charge, a conclusion we have drawn previously in less formal terms.

To show that Eq. (7.16) is time-reversal invariant, we only need to show that $2\tilde{S}_{\Theta=\pi/2}$ is trivial (mod $2\pi$). This was shown explicitly in Ref. [334] (Sec. VII A therein). This also provides an explicit example, in which $\Theta = \pi$ is trivial in the sense that $\tilde{S}_{\Theta=\pi} = 2\tilde{S}_{\Theta=\pi/2}$ is trivial.

Similar result can also be obtained in the magnetic picture (with an inverted Maxwell coupling $e^2$). Integrating out the bosonic ($M$) degrees of freedom gives

$$\tilde{S}'_{\Theta=\pi/2} = \frac{\pi}{2} \left( \frac{1}{16\pi^2} \int \text{tr} F \wedge F - \frac{1}{2\pi^2} \int \tilde{f} \wedge \tilde{f} \right),$$

(7.18)

where the inverted sign of the $U(1)$ theta-term and the absence of the gravitational term is simply reflecting the fact that for a bosonic integer quantum hall state in two dimensions with $U(2) = U(1) \times SU(2)/\mathbb{Z}_2$ symmetry, the spin and charge hall conductance are opposite in sign and the net thermal hall conductance is zero [348, 349]. The cocycle condition for the dual field strength is now

$$\int \left( \frac{\tilde{f}}{\pi} + w_2^{SO(3)} \right) = 0 \pmod{2}.$$

(7.19)

Following an argument similar to that in Ref. [334] (Sec. VII A therein), one can show that $\tilde{S}'_{\Theta=\pi} = 2\tilde{S}'_{\Theta=\pi/2}$ is trivial (mod $2\pi$). Therefore the effective theory in the magnetic picture is also time-reversal invariant.

### 7.4.3 Surface states

Perhaps the most striking property of a topological insulator is the presence of protected surface states. It is natural then to ask about the physics at the surface of the Fractional Topological Paramagnets. Specifically we consider an interface between the vacuum and a material in a Fractional Topological Paramagnet phase. The gauged SPT point of view then makes it natural that
both $E_{bT} M_{f\frac{1}{2}}$ and $E_{f\frac{1}{2}} M_{b\frac{1}{2}}$ have protected states at such an interface.

Protected surface states for $U(1)$ quantum spin liquids with time reversal were described in Ref. [241]. As discussed there, in states where both $E$ and $M$ are non-trivial (i.e. not simply a boson transforming trivially under the global symmetry) the surface to the vacuum necessarily has protected states. Of the 15 families of $U(1)$ quantum spin liquids with $SO(3) \times T$, only $E_{bT} M_{f\frac{1}{2}}$ and $E_{f\frac{1}{2}} M_{b\frac{1}{2}}$ therefore necessarily have protected surface states. In both these cases the parent SPTs (either in the $E$ or $M$ points of view) are such that the surface exhibits the phenomenon of Symmetry Enforced Gaplessness, i.e., there is no symmetry preserving gapped surface even with topological order. Symmetry preserving surfaces are necessarily gapless. For the Fractional Topological Para-magnets a gapless surface state is readily described from the fermion point of view. Both states then have 2 gapless surface Dirac cones (one for each spin) that is coupled to the bulk $U(1)$ gauge field. Time reversal acts differently on the surface Dirac fermions in the two states (the time reversal is inherited from that on the bulk fermionic quasiparticle).

7.5 COMBINING $U(1)$ QSLS AND BOSONIC SPTs PROTECTED BY $SO(3) \times T$

We have thus far described the distinct possible realizations of symmetry for the bulk excitations of $U(1)$ quantum spin liquids with time reversal and $SO(3)$ spin rotational symmetries. However, strictly speaking, this is not the complete classification of such spin liquids. We can in principle obtain distinct spin liquids with the same symmetry fractionalization patterns by simply combining spin liquids with SPT states protected by the global $SO(3) \times T$ symmetry. This was demonstrated for time reversal invariant $U(1)$ spin liquids in Ref. [241]. Further it was shown that not all SPTs remain non-trivial when combined with a spin liquid. In other words some SPTs can “dissolve” into some spin liquids without leading to a distinct state. Determining the distinct spin liquids that result when SPTs are combined with spin liquids is a delicate but unavoidable task that is part of any classification of symmetry enriched spin liquids. In this section we undertake this task for the $SO(3) \times T$ symmetric $U(1)$ QSLs of primary interest in this paper. We will show that each of the 15 families of such $U(1)$ spin liquids described so far is further refined to give a total of 168 distinct phases. We expect that this is the complete classification of $U(1)$ QSLs enriched with $SO(3) \times T$ symmetry.
Table 7.3: Properties of the surface $Z_2$ topological orders of the four root states of bosonic SPTs with symmetry $SO(3) \times T$.

<table>
<thead>
<tr>
<th></th>
<th>$T_e^2$</th>
<th>$T_m^2$</th>
<th>$S_e$</th>
<th>$S_m$</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$eTmT$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$efmf$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$e$ and $m$ are fermions</td>
</tr>
<tr>
<td>$e^2m^2$</td>
<td>1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td></td>
</tr>
<tr>
<td>$e^2mT$</td>
<td>1</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Bosonic SPTs with symmetry $SO(3) \times T$ are classified by $Z_4^2$. The four root states all admit surface $Z_2$ topological order $\{1, e, m, \epsilon\}$, with different assignments of fractional quantum numbers to the anyons $e, m, \epsilon$ (notice that $e, m$ here denote the anyons in the $2d$ surface topological order, which are not to be confused with $E, M$ in earlier sections denoting electric and magnetic charges in the $3d$ bulk $U(1)$ gauge theory). These surface $Z_2$ topological orders realize symmetries in a way that is impossible for a purely two dimensional system (see Table 7.3. More details can be found in Appendix G.6). The surface topological order provides a non-perturbative characterization of these SPTs; we therefore label the SPTs themselves by their surface $Z_2$ topological orders. The four root states generate in total 16 distinct SPTs, and each can be viewed as a combination of some of the root states. For example, if $efmf$ and $e^2m^2$ are taken as two root states, weakly coupling them produces a new SPT denoted by $efmf \oplus e^2m^2$. In this example the notation of the state can be simplified because a surface phase transition can be induced such that the bound state of the $\epsilon$’s from the surface $efmf$ and $e^2m^2$ is condensed. This condensation will not change the bulk property, but the surface now has $Z_2$ topological order $ef^2mf^2$, where both the $e$ and $m$ are spin-1/2 fermions. So for simplicity $efmf \oplus e^2m^2$ can be denoted by $ef^2mf^2$.

Below in Sec. 7.5.1 we use the same strategy as in Ref. [241] to determine if these nontrivial SPTs are trivial or still nontrivial in the presence of the excitations in the quantum spin liquids. Then in Sec. 7.5.2, we apply these results to obtain the enriched classification of $U(1)$ quantum spin liquids combined with SPTs.
Table 7.4: Triviality of the root states of bosonic SPTs with symmetry $SO(3) \times T$ in the presence of nontrivial bosonic excitations. The rows represent the nontrivial SPT states, and the columns represent the quantum numbers of the bosonic excitation. $C^2$ means the elementary boson carries electric charge 1, and $\tilde{C}^2$ means it carries magnetic charge 1. Notice electric (magnetic) charge is even (odd) under time reversal. $T$ means the elementary boson is a Kramers doublet, and $\frac{1}{2}$ means it carries spin-$\frac{1}{2}$. A cross (hook) means the topological order is anomalous (non-anomalous) in the presence of the excitation from the quantum spin liquid.

<table>
<thead>
<tr>
<th></th>
<th>$C^2$</th>
<th>$C^2$</th>
<th>$C^2T$</th>
<th>$C^2\frac{1}{2}$</th>
<th>$C^2\frac{1}{2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$eTmT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\sqrt{}$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efmf$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mf\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\sqrt{}$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efTmfT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$eT\frac{1}{2}mT\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$eT\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$ef\frac{1}{2}mf\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efmf \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efT\frac{1}{2}mT\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efTmfT \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$eT\frac{1}{2}mT \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$ef\frac{1}{2}mf\frac{1}{2} \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$ef\frac{1}{2}mf\frac{1}{2} \oplus eT\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
</tbody>
</table>

7.5.1 SPTs in the presence of nontrivial excitations

Table 7.4 and table 7.5 show whether the nontrivial SPTs are trivial or nontrivial in the presence of fractional excitations with all possible statistics and relevant quantum numbers. Below we explain the reasons for the entries of these tables. The notations that will be used below are defined in the captions of these tables.
SPTs with component $efmf$ always enrich the classification of the quantum spin liquids

When time reversal symmetry is broken on its surface, $efmf$ has surface thermal Hall conductance $\kappa_{xy} = 4 \text{ (mod) } 8$ in units of $\frac{e^2 k_B}{\hbar}$. Thus it always enriches the classification of the quantum spin liquids [241, 337]. The same is true for all SPTs that are obtained by combining $efmf$ and other root states. Besides $efmf$, these include $efTmfT$, $ef\frac{1}{2}mf\frac{1}{2}$, $efmf \oplus e\frac{1}{2}mT$, $efT\frac{1}{2}mfT\frac{1}{2}$, $efTmfT \oplus e\frac{1}{2}mT$, $ef\frac{1}{2}mf\frac{1}{2} \oplus e\frac{1}{2}mT$ and $ef\frac{1}{2}mf\frac{1}{2} \oplus eT\frac{1}{2}mT$.

SPTs with a component $Z_2$ topological order where both $e$ and $m$ carry spin-$\frac{1}{2}$ are anomalous in the presence of the nontrivial excitations. These SPTs include $e\frac{1}{2}m\frac{1}{2}$, $eT\frac{1}{2}mT\frac{1}{2}$, $ef\frac{1}{2}mf\frac{1}{2}$, $e\frac{1}{2}mT\frac{1}{2}$, $efT\frac{1}{2}mfT\frac{1}{2}$, $eT\frac{1}{2}mT\frac{1}{2} \oplus e\frac{1}{2}mT$, $ef\frac{1}{2}mf\frac{1}{2} \oplus e\frac{1}{2}mT$ and $ef\frac{1}{2}mf\frac{1}{2} \oplus eT\frac{1}{2}mT$. In this case, the $SO(3) \Theta = 2\pi$ (see Appendix G.6). But as discussed in Sec. 7.4, none of the spin liquids have $\Theta = 2\pi$. So coupling these SPTs with spin liquids cannot change $\Theta$ from $2\pi$ to $0 \text{ (mod } 4\pi)$, and all these surface states will remain anomalous even when coupled to a spin liquid. Below we provide more physical reasoning to demonstrate their anomalies in the presence of the excitations from the spin liquids.

To see their anomalies, we can first assume that such a state can exist in a purely two dimensional system. Then in the case where the nontrivial excitations carry quantum numbers $C^2$, $\tilde{C}^2$ or $C^2T$, we can tunnel an $SO(3)$ monopole through the system, which leaves a flux. This is a local process, but due to the spin-$1/2$ of $e$ and $m$, both $e$ and $m$ will see a $-1$ phase factor around the flux, no matter how far they are away from it. To cancel this nonlocal effect, an $\epsilon$ has to be generated in the process of tunneling the monopole. As shown in Appendix G.7, this process will not induce any polarization charge or spin because of the symmetry of the system. Therefore, this local process generates a single neutral and spinless fermion in the system, which is impossible.

4To characterize the SPT $efmf$ more formally, one can consider its response to a change of the background metric. Then this SPT is characterized by a bulk gravitational response term given by $\frac{1}{16\pi} \int \text{tr} R \wedge R$, where $R$ is the Riemann curvature tensor. In this formal language, because none of the $U(1)$ quantum spin liquids discussed has a gravitational response term that can cancel this one, this SPT cannot be “absorbed” by any of these spin liquids.

5More formally, this is characterized by $\frac{\Theta}{16\pi} \int \text{tr} SO(3) F \wedge F$, a response term to a background $SO(3)$ gauge field, where $F$ is the $SO(3)$ field strength and $\Theta = 2\pi$ for these states.
ble and shows these states are still anomalous even in the presence of the nontrivial excitations.

When the quantum numbers of the nontrivial excitations are $C^2 \frac{1}{2}$, $\bar{C}^2 \frac{1}{2}$ or $C^2 T \frac{1}{2}$, tunneling an $SO(3)$ monopole is no longer a local process, but tunneling a bound state of an $SO(3)$ monopole and half of a $U(1)$ monopole is still local. Again, this process will generate a single neutral and spinless fermion in the system, which is impossible and shows the anomalies of these states in the presence of these nontrivial excitations.

$eTmT$ is anomalous in the presence of non-Kramers bosons

It is known that the anomaly of $eTmT$ only comes from time reversal symmetry, so the presence of bosons with quantum numbers $C^2$, $\bar{C}^2$, $C^2 \frac{1}{2}$ or $\bar{C}^2 \frac{1}{2}$ will not remove the anomaly.

$e \frac{1}{2} mT$ and $eT \frac{1}{2} mT$ are anomalous in the presence of nontrivial excitations with quantum numbers $C^2$, $\bar{C}^2$, $C^2 T$ and $\bar{C}^2 \frac{1}{2}$

It turns out that $e \frac{1}{2} mT$ and $eT \frac{1}{2} mT$ are also still anomalous in the presence of nontrivial excitations with quantum numbers $C^2$, $\bar{C}^2$, $C^2 T$ or $\bar{C}^2 \frac{1}{2}$. To see this, for the cases where excitations carry quantum numbers $C^2$, $\bar{C}^2$ or $C^2 T$, we can again tunnel an $SO(3)$ monopole through the system. This will leave a flux such that $e$ and $\epsilon$ see a $-1$ phase factor no matter how far they are away from it. To cancel this nonlocal effect, an $m$ will have to be generated in the process. As argued in Appendix G.7, this process cannot induce any polarization charge or spin. Because the $SO(3)$ flux background is invariant under time reversal, such a local process generates a Kramers doublet that carries no other nontrivial quantum number. But there are no such local degrees of freedom in these cases, so this is impossible. Thus these $Z_2$ topological orders are still anomalous.

If the excitations carry quantum number $\bar{C}^2 \frac{1}{2}$, one can tunnel a bound state of an $SO(3)$ monopole and half of a $U(1)$ monopole. Similar argument shows an $m$ needs to be produced in the process. Again, because both $SO(3)$ and $U(1)$ commute with $T$, the flux background left on the system is time reversal invariant. This again shows that a local process generates a Kramers doublet with no other nontrivial quantum number and thus it is impossible.
Denote $eTmT$ in the presence of bosons with quantum number $C^2T$ by $(eTmT, bC^2T)$. It turns out this is non-anomalous. [241] To see this, we can attach a boson with quantum number $C^2T$ to the $e$ particle, then $eTmT$ will be relabelled as $eC^2mT$. This is a non-anomalous state. To construct it, one can first construct $eC^2T$, which is non-anomalous because the topological order can be confined by condensing $m$ without breaking any symmetry. Then putting the $\epsilon$ into a quantum spin Hall state makes it $eC^2mT$. [350,351]

Similarly, with parallel notations, $(eTmT, bC^2T)$, $(eTmT, fC^2T)$, $(eTmT, fC^2)$, and $(eTmT, fC^2)$ are also non-anomalous.

**Other entries in Table 7.4 and 7.5 are anomalous**

For other entries in Table 7.4 and Table 7.5, the arguments utilized above do not apply. However, they are still expected to be anomalous. Below we sketch the logic to show this, and more details can be found in Appendix G.8.

Suppose any of these $Z_2$ topological orders is non-anomalous, that is, it can be realized in a purely two dimensional system, it must allow a physical edge, i.e. a boundary that separates this state and the trivial vacuum. It is believed that the $K$-matrix formalism can describe all two dimensional Abelian topological orders, and in particular, $K$-matrix theory naturally allows a physical edge. [1] So if no $K$-matrix description of a $Z_2$ topological order exists, it should not be edgeable, i.e. it is anomalous. We note the $K$-matrix formalism has already been applied to check edgeability or to classify SPTs and symmetry-enriched topological orders in the literature. [309,315,316,348]

Indeed, in Appendix G.8 we will show all other entries are not edgeable. This implies they are still anomalous.
Table 7.5: Trivialness of the root states of bosonic SPTs with symmetry $SO(3) \times T$ in the presence of nontrivial fermionic excitations. The rows represent the nontrivial SPT states, and the columns represent the quantum numbers of the fermionic excitation. $C^2$ means the elementary fermion carries electric charge 1, and $\tilde{C}^2$ means it carries magnetic charge 1. Notice electric (magnetic) charge is even (odd) under time reversal. $T$ means the elementary fermion is a Kramers doublet, and $\frac{1}{2}$ means it carries spin-$\frac{1}{2}$. A cross (hook) means the topological order is anomalous (non-anomalous) in the presence of the excitation from the quantum spin liquid.

<table>
<thead>
<tr>
<th></th>
<th>$C^2$</th>
<th>$C^2$</th>
<th>$C^2T$</th>
<th>$C^2\frac{1}{2}$</th>
<th>$C^2\frac{1}{2}$</th>
<th>$C^2T\frac{1}{2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$eTmT$</td>
<td>$\sqrt{}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efmf$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efTmftT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\sqrt{}$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efm\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efmf \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mftT\frac{1}{2}$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$efTmftT \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$e\frac{1}{2}mT\frac{1}{2} \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$ef\frac{1}{2}ms\frac{1}{2} \oplus e\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>$ef\frac{1}{2}ms\frac{1}{2} \oplus e\frac{1}{2}T\frac{1}{2}mT$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
</tbody>
</table>
7.5.2 Enriched classification of quantum spin liquids combined with SPTs

In the previous subsection, we have shown when the nontrivial bosonic SPTs are in an environment with some nontrivial particles, which ones are still nontrivial and which ones become trivial. In most cases, the nontrivial SPTs remain nontrivial. Then each of the quantum spin liquids can become $2^4 = 16$ distinct ones after being weakly coupled with the bosonic SPTs. In the presence of the excitations in the quantum spin liquids, the cases where nontrivial SPTs become trivial are when $eTmT$ coupled with $E_{bT}M_b$, $E_{bT}M_f$ or $E_fM_b$, when $e\frac{1}{2}mT$ coupled with $E_{b\frac{1}{2}}M_b$ or $E_{fT}\frac{1}{2}M_b$ and when $eT\frac{1}{2}mT$ coupled with $E_{bT}\frac{1}{2}M_b$, $E_{bT}\frac{1}{2}M_f\frac{1}{2}$, $E_{f}\frac{1}{2}M_b$ or $E_{f}\frac{1}{2}M_b\frac{1}{2}$. For these quantum spin liquids, each can become $2^3 = 8$ distinct ones after weakly coupled with the bosonic SPTs. All these SPT-enriched quantum spin liquids are different from each other. Therefore, when weakly coupled with bosonic SPTs with time reversal and $SO(3)$ spin rotational symmetries, there are in total $6 \times 16 + 9 \times 8 = 168$ distinct $U(1)$ quantum spin liquids.

7.6 A general framework to classify symmetry enriched $U(1)$ quantum spin liquids

The above discussion on the classification of $SO(3) \times T$ symmetric $U(1)$ quantum spin liquids provides a good example. In this section we describe a general framework to classify symmetry enriched $U(1)$ quantum spin liquids. It involves three steps: enumerating all putative states, examining the anomalies of these states, and coupling these states to 3D bosonic SPTs with the same symmetry. This framework is physics-based. After discussing this framework, we will briefly discuss a supplementary formal approach to classify such states, which can be potentially more useful for thinking about these problems more abstractly.

7.6.1 Enumerate putative states

We begin with the first step: enumerating all putative states. As discussed earlier, different symmetry enriched $U(1)$ quantum spin liquids are distinguished by the properties of their excitations, and to determine the phase, we need to specify the statistics and symmetry quantum numbers of the excitations.
We start with the simpler case where the symmetry $G$ is unitary and connected, that is, all elements in the symmetry group are unitary and they can all be continuously connected to the identity element. In this case, the symmetry cannot exchange the type of the fractional excitations. Also, one can tune $\theta$ such that the charge-monopole lattice is of the $\theta = 0$-type, and both $E$ and $M$ are bosons (this is shown more explicitly in the examples in Sec. 7.8.2). To fully determine the properties of the excitations, we just need to specify the symmetry quantum numbers of $E$ and $M$. More precisely, we need to assign (projective) representations to $E$ and $M$, which are classified by the second group cohomology $H^2(G, U(1))$. While doing this, we also need to keep in mind that $E$ and $M$ are equivalent in this case, so that, for example, $E_b M_b$ and $E_0 M_b$ are the same $SO(3)$ symmetric phase. When this is done, all putative states will be listed.

Next we go to the more complicated case where the symmetry is $G \times T$ (or, more generally, $G \rtimes T$), with $G$ a connected unitary group. Again, the elements in $G$ will not change the type of fractional excitations. However, time reversal will necessarily change some types of fractional excitations, and we will always take the convention that the emergent electric (magnetic) field is even (odd) under time reversal. Then there are two types of charge-monopole lattice, with $\theta = 0$ and $\theta = \pi$, respectively.

Consider the states at $\theta = 0$ first. Then the $U(1)$ quantum spin liquids are classified by the statistics and symmetry quantum numbers carried by $E$ and $M$. As for statistics, the only constraint at this point is just that $E$ and $M$ cannot both be fermions. Below we discuss the symmetry quantum numbers, or in other words, projective representations.

Let us start from the case with $\theta = 0$. Here we need to distinguish two types of projective representations: the electric (standard) one and the magnetic (twisted) one. The electric projective representations are applicable to $E$, and they are classified by $\mathcal{H}^2(G \times T, U_T(1))$, where $G \times T$ acts on the $U(1)$ coefficient by taking the complex conjugate if the group element is anti-unitary. This is the standard classification of the projective representations of a group with anti-unitary elements. However, another type of projective representations apply to $M$, which are classified by another group cohomology $\mathcal{H}^2(G \times T, U_T^M(1))$ (see Appendix G.9 for more details). This group cohomology differs from the standard one in the group action on the $U(1)$ factors, and this difference comes from the convention that the magnetic (electric) field is odd (even) under time.
reversal. After assigning statistics and symmetry quantum numbers to $E$ and $M$ as above, all putative $G \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids with $\theta = 0$ will be listed.

As for states with $\theta = \pi$, the properties of all excitations are determined by the properties of the $(\frac{1}{2}, \pm 1)$ dyons, which must be bosons. So to list all putative states, we only need to assign symmetry quantum numbers to these two dyons. As discussed in Appendix G.9, these symmetry quantum numbers are given by the dyonic (mixed) projective representations, which are classified by another group cohomology $H^2(G \times \mathcal{T}, U^D(1) \times U^D(1))$. After assigning symmetry quantum numbers to the $(\frac{1}{2}, \pm 1)$ dyons, all putative $G \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids with $\theta = \pi$ will be listed.

If the symmetry group is $G$ or $G \times \mathcal{T}$, where $G$ is unitary but not connected, the elements in $G$ can also permute fractional excitations, as we will see in examples below. The putative states in this more complicated scenario can be listed in a similar manner as above: one has to fix the shape of the charge-monopole lattice, specify the statistics of the relevant excitations, and specify the symmetry quantum numbers of the relevant excitations. The first two steps are identical as the previous cases, but the classification of symmetry quantum numbers will be more complicated in this case, and in this paper we do not attempt to give a mathematical framework for this step, although it can be done in a case-by-case manner.

7.6.2 EXAMINE THE ANOMALIES

After enumerating all putative symmetry enriched $U(1)$ quantum spin liquids, we need to examine which ones are anomalous. A general way of doing this is to consider whether the corresponding SPT of this spin liquid can exist. Denote the symmetry of the $U(1)$ quantum spin liquid by $G$, which is supposed to be a completely general on-site symmetry in this subsection (it can contain anti-unitary elements, and its unitary elements do not need to be continuously connected with identity). By the corresponding SPT of a spin liquid, we mean an SPT protected by a $U(1)$ central extension of $G$,\footnote{In the more standard terminology, this $U(1)$ central extension of $G$ is the projective symmetry group of $G$, which was first introduced in Ref. [352].} which becomes this spin liquid once this $U(1)$ symmetry is gauged. Clearly, by definition, as long as this SPT can exist, the spin liquid state must
be anomaly-free. If this SPT is intrinsically inconsistent, then the corresponding spin liquid state must be anomalous. To see this, suppose this SPT is problematic but the corresponding spin liquid is anomaly-free, then we will show this leads to a contradiction, because of a systematic method to ungage the gauge theory and generate the corresponding SPT.

More precisely, suppose this spin liquid can be realized, one can bring in a trivial insulator made of particles that have the same properties (same statistics and quantum numbers) as the particles that make up this corresponding SPT, and condense the bound state between the particle in this trivial insulator and the electric charge or magnetic monopole of the spin liquid. This is a systematic method to ungage the gauge theory: it will confine the dynamical $U(1)$ gauge field, and the resulting state will be precisely the corresponding SPT of the spin liquid state (an example is shown in Figure 7.4). [325] This leads to a contradiction to the original assumption that such SPT is problematic. Therefore, a sufficient and necessary condition for a symmetry enriched $U(1)$ quantum spin liquid to be anomaly-free is that its corresponding SPT is consistent.

How do we check whether the corresponding SPT is consistent? One way is to consider whether it has a consistent surface state. This condition - known as “edgeability” - was defined in Ref. [316]. Assuming such SPT is consistent, one can first condense certain charges on the surface of this SPT and get a surface superfluid. Then one can try to condense certain vortices to restore the symmetry on the surface. If the symmetric surface state is consistent (but possibly anomalous), one can build up the three dimensional bulk SPT (for example, by a layer construction or a Walker-Wang type construction). If the putative symmetric surface state is inconsistent, then this SPT is inconsistent, because it has an invalid edge state.

In summary, a systematic physical way to examine whether a putative symmetry enriched $U(1)$ quantum spin liquid is anomalous is to check whether its corresponding SPT can have a legitimate surface state. If so, this spin liquid state is non-anomalous. Otherwise, it is anomalous. These relations is sketched in Figure 7.5.

This method of anomaly detection applies to any symmetry enriched $U(1)$ quantum spin liquids, but for some particular cases, there are more physical ways of doing it by focusing on the spin liquid state itself, instead of its corresponding SPT. For example, we have used the $SO(3)$ monopole to detect the anomaly of some putative $SO(3)$ symmetric spin liquid states in Sec. 7.3.
Figure 7.4: There is a systematic ungauging procedure that takes a symmetry enriched $U(1)$ quantum spin liquid to its corresponding SPT. Consider the time reversal symmetric $U(1)$ quantum spin liquid $(E_{fT} M_f)_\theta$ (the upper left system), and we will try to get its corresponding SPT from the perspective of the electric charge. To do this, we first introduce an auxiliary trivial time reversal symmetric insulator made of fermions that are Kramers doublets, where these fermions are denoted by $c$ (the lower left system). Next we condense the bound state of $E$, the electric charge of the $U(1)$ spin liquid, and $c^\dagger$, the holes in the auxiliary trivial insulator. This bound state is a boson and a Kramers singlet, so this condensation will preserve the time reversal symmetry. The dynamical $U(1)$ gauge field in the $U(1)$ gauge theory will be confined, and the resulting state is precisely the Fu-Kane-Mele topological insulator, which, viewed from the perspective of the electric charge, is the corresponding SPT of $(E_{fT} M_f)_\theta$ (the right system).

Figure 7.5: That the $G$ symmetric $U(1)$ quantum spin liquid is anomaly-free is equivalent to that it has a corresponding SPT, which is in turn equivalent to that this SPT can have a consistent (but possibly anomalous) 2D surface state.

However, for some more subtle cases, the anomalies are examined by considering the corresponding SPTs. Some examples are given in Sec. 7.7, where $Z_2 \times T$ symmetric states are discussed.
7.6.3 **Couple the Spin Liquids with SPTs**

The above two steps classify symmetry enriched $U(1)$ quantum spin liquids in terms of the properties of the bulk excitations. To complete the classification of the symmetry enriched $U(1)$ quantum spin liquids, one has to consider coupling these spin liquids and 3D bosonic SPTs with the same symmetry. In general, when an SPT is coupled with a $U(1)$ spin liquid, the result is still a $U(1)$ spin liquid with the bulk fractional excitations carrying the same symmetry fractionalization pattern, but the new state can have a different type of surface compared to the original one, due to the nontrivial surface of the SPT. Therefore, one has to check if this SPT can be “absorbed” by the $U(1)$ spin liquid. Physically, this amounts to checking if the nontrivial surface of the SPT remains nontrivial if it is coupled with the bulk excitations in the spin liquid. Examples of such exercises are given in Sec. 7.5.

7.6.4 **A Formal Framework**

We would like to close this section by briefly discussing a more formal approach to classify symmetry enriched $U(1)$ quantum spin liquids. In this formal approach, the problem amounts to classifying the action, or more precisely, the universal part of the partition function, of the $U(1)$ gauge theories. To encode the information about symmetries, in this action the $U(1)$ gauge field should be coupled to a background gauge field corresponding to the symmetry and a background spacetime metric. If the global symmetry includes time reversal the equivalent of coupling a background gauge field is to place the theory on an unorientable space-time manifold.

Note that we are considering spin liquids that arise in a UV system made out of bosons. To impose this restriction directly in the low energy continuum theory we demand that the low energy theory can be consistently formulated on an arbitrary non-spin space-time manifold. On an orientable manifold, this is achieved by requiring that the emergent gauge field be either an ordinary $U(1)$ gauge field (when the emergent electric charge $E$ is a boson) or that it is a Spin$_c$ connection\(^7\) (when $E$ is a fermion). On an unorientable manifold, there is a generalization of a

\(^7\)A Spin$_c$ connection differs from an ordinary $U(1)$ gauge field through a modification of its flux quantization condition: the curvature $F$ of a Spin$_c$ connection satisfies $\int \frac{F}{2\pi} = \int \frac{w_2^{TM}}{2} \pmod{Z}$ on oriented 2-cycles where $w_2^{TM}$ is the second Stieffel-Whitney class of the tangle bundle of the manifold. For more
Spin\(_c\) connection known as Pin\(_{c,\pm}\) connections - the ± sign correspond to the two possibilities that \(E\) is non-Kramers or Kramers under time reversal (more detail is in Ref. [29]). We will not make an explicit distinction in the schematic discussion below between these different kinds of \(U(1)\) connections.

Denote the \(U(1)\) gauge field by \(a\), the gauge field corresponding to the global symmetry by \(b\), and the background metric by \(g\). In general, the action can be written in a form

\[
S[a; b, g] = S_{U(1)}[a] + S_{\text{SPT}}[b, g] + S_{\text{mixed}}[a; b, g]
\]  

(7.20)

The first term contains the Maxwell action and the \(\theta\) term of a \(U(1)\) gauge field, and it is present in general for a \(U(1)\) quantum spin liquid and are independent of symmetries. The third term, \(S_{\text{SPT}}[b, g]\), depends only on \(b\) and \(g\). This term physically describes 3D bosonic SPTs with the same symmetry as the \(U(1)\) spin liquid, and adding it into the action means coupling a \(U(1)\) spin liquid and a bosonic SPT with the same symmetry. As discussed before, this will potentially change the system into a different \(U(1)\) spin liquid. In order to see if such an SPT can be “absorbed” into a \(U(1)\) spin liquid, one needs to check if the universal part of the partition function will change due to the presence of this term. The last term, \(S_{\text{mixed}}[a; b, g]\), only involves terms that couple \(a\) with \(b\) and/or \(g\). This term encodes the information about symmetry fractionalization on the bulk excitations.

In general, such an action is constrained by gauge invariance. In addition, certain constraints on these fields may apply analogous to the modification of the flux quantization condition for Spin\(_c\) connections when \(E\) is a fermion. For example, for fractional topological paramagnets, there is a constraint on such fields given by (7.17). To classify symmetry enriched \(U(1)\) quantum spin liquids, one can first write down all possible such actions and then classify the resulting universal part of the partition function. We leave this for future work.

detail see Refs. [29, 245] and references therein.
7.7 \textit{U(1) Quantum Spin Liquids Enriched by $Z_2 \times T$ Symmetry}

In this section we apply the above general framework to classify $U(1)$ quantum spin liquids enriched by $Z_2 \times T$ symmetry. This symmetry can be relevant for experimental candidates of quantum spin liquids made of non-Kramers quantum spins, \textit{i.e.} for example, two-level systems made of $m_z = \pm 1$ states of a spin-1 atom, where time reversal flips $S_z$ and $Z_2$ acts as a $\pi$ spin rotation around the $x$ axis. Below we will first list all putative states, including the anomalous ones. Then we will examine the anomalies of these states. We will leave the problem of coupling these spin liquids with SPTs for future work.

It turns out there are two types of $Z_2$ actions that deserve separate discussions. In the first type, the $Z_2$ symmetry does not change one type of fractional excitation to another. More precisely, the electric charge and magnetic monopole will both retain their characters under this type of $Z_2$ action. In the second type, the $Z_2$ symmetry changes the fractional excitations. In particular, it can change the electric charge into the anti-electric charge, and at the same time change the magnetic monopole into the anti-magnetic monopole. This type of $Z_2$ action is physically a charge conjugation. One may wonder whether it is possible to change an electric charge into a magnetic monopole, but Ref. [353] pointed out this is impossible in a strictly 3D system.

Below we will discuss these two types of $Z_2$ actions in turn.

7.7.1 \textit{$Z_2$ Not Acting as a Charge Conjugation}

We start from the case where $Z_2$ does not act as a charge conjugation, that is, it does not change a type of fractional excitation to another type.

We will begin with the simpler case that has $\theta = 0$. In this case, to classify the quantum numbers of the electric charge, it is appropriate to look at the projective representations of $Z_2 \times T$, which are classified by $\mathbb{Z}_2^2$, where the nontrivial projective representations can be viewed as being a Kramers doublet under the original time reversal and/or under a new anti-unitary symmetry $T'$, whose generator is the product of the generator of $Z_2$ and the generator of $T$. Although it is not meaningful to talk about whether the magnetic monopole is a Kramers singlet or doublet under $T$ or $T'$, there are still two types of quantum numbers of the magnetic monopole under
Table 7.6: List of non-anomalous $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ not acting as a charge conjugation. All these states are anomaly-free. $T^2_E = 1$ ($T^2_E = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T^{T'}_E = 1$ ($T^{T'}_E = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T'$. $[T, Z_2]_M = +$ ($[T, Z_2]_M = -$) represents the case where $Z_2$ and $T$ commute (anti-commute) on $M$.

<table>
<thead>
<tr>
<th>$T^2_E$</th>
<th>$T^{T'}_E$</th>
<th>$[T, Z_2]_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_b M_b$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b T M_b$</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b T' M_b$</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$E_b M_{b-}$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f M_{b}$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f T M_{b}$</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f T' M_{b}$</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$E_f M_{b-}$</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$E_b M_{f}$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b T M_{f}$</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b T' M_{f}$</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$E_b M_{f-}$</td>
<td>-1</td>
<td>-1</td>
</tr>
</tbody>
</table>

$Z_2 \times T$: on the monopole the $Z_2$ and $T$ can either commute or anti-commute.\(^8\) This relation between $Z_2$ and $T$ is gauge invariant for the monopole, but not gauge invariant for the charge.

Therefore, we can make a list of putative $U(1)$ quantum spin liquids with this type of symmetry, and there are $3 \times 2^2 \times 2 = 24$ of them, as listed in Table 7.6 and Table 7.7. It turns out the 15 states in Table 7.6 are anomaly-free, and the 9 states in Table 7.7 are anomalous, which can be grouped into three anomaly classes. We will give the construction of the non-anomalous states in Appendix G.10. Later in Sec. 7.7.3, we will give the strategy to show the anomalies of the states in Table 7.7, and we will finish the arguments for this anomaly-detection in Appendix G.10.

Before moving to the case with $\theta = \pi$, we note that one point deserves immediate clarification. That is, one may wonder, for example, whether $E_{bT} M_b$ and $E_{bT'} M_b$ are truly distinct, since they are related to each other by relabelling $T \leftrightarrow T'$ and $E \leftrightarrow M$. At the first glance, these two

\(^8\)If $T$ and $Z_2$ commute (anti-commute), $T'$ defined above will also commute (anti-commute) with $Z_2$. 205
Table 7.7: List of anomalous $Z_2 \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ not acting as a charge conjugation. All these states are anomaly-free. $T_E^2 = 1$ ($T'_E^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $\mathcal{T}$. $T_E^2 = 1$ ($T'_E^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $\mathcal{T}'$. $[\mathcal{T}, Z_2]_M = +$ ($[\mathcal{T}, Z_2]_M = -$) represents the case where $Z_2$ and $\mathcal{T}$ commute (anti-commute) on $M$. The last column indicates the anomaly classes.

<table>
<thead>
<tr>
<th>State</th>
<th>$T_E^2$</th>
<th>$T'_E^2$</th>
<th>$[\mathcal{T}, Z_2]_M$</th>
<th>Anomaly class</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bT} M_{f-}$</td>
<td>1</td>
<td>-1</td>
<td>-</td>
<td>class a</td>
</tr>
<tr>
<td>$E_{fT} M_{b-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class a</td>
</tr>
<tr>
<td>$E_{bT} M_{b-}$</td>
<td>1</td>
<td>-1</td>
<td>-</td>
<td>class a</td>
</tr>
<tr>
<td>$E_{fT} M_{b-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
<tr>
<td>$E_{fT} M_{f-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
<tr>
<td>$E_{bT} M_{f-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
<tr>
<td>$E_{bT} M_{f-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
</tbody>
</table>

states indeed seem to have identical physical properties when examined on their own. However, once the definitions of $\mathcal{T}$ and $\mathcal{T}'$ are fixed, these states are distinct. One physical way to see this is to consider the two states at the same time, clearly without breaking either $\mathcal{T}$ or $\mathcal{T}'$, one state cannot be connected to another without encountering a phase transition. Therefore, all these 24 states are truly distinct.

Now we turn to the states with $\theta = \pi$. In this case, the quantum number of the $(\frac{1}{2}, 1)$ dyon determines the quantum numbers of all other dyons. However, the $(\frac{1}{2}, 1)$ dyon does not have any projective representation of the $Z_2 \times \mathcal{T}$ symmetry, so there is only one state: $(E_{fTT'} M_f)_{\theta}$, as described in Table 7.8. The electric charge has to be Kramers doublet under both $\mathcal{T}$ and $\mathcal{T}'$, because it is a bound state of the $(\frac{1}{2}, 1)$ and $(\frac{1}{2}, -1)$ dyons, which have $\pi$ mutual braiding and are exchanged under both $\mathcal{T}$ and $\mathcal{T}'$. Naively, the $M$ particle (the $(0, 2)$ dyon in this context) can either have $Z_2$ and $\mathcal{T}$ commuting or anti-commuting. But it turns out the latter possibility can be ruled out, as shown in Appendix G.11. This $(E_{fTT'} M_f)_{\theta}$ state can be viewed as a descendant of the $SO(3) \times \mathcal{T}$ symmetric $(E_{fT} M_f)_{\theta}$, so it is anomaly-free.

In summary, if $Z_2$ does not change one type of fractional excitation into another type, there are 16 distinct anomaly-free $Z_2 \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids.
Table 7.8: The $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids that have $\theta = \pi$ and have $Z_2$ not acting as a charge conjugation. This state is anomaly-free. $T_E^2 = 1$ ($T_E^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T'_E^2 = 1$ ($T'_E^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T'$. $[T, Z_2]_M = +$ ($[T, Z_2]_M = -$) represents the case where $Z_2$ and $T$ commute (anticommute) on $M$.

<table>
<thead>
<tr>
<th>$(E_{fT}M_f)_{\theta}$</th>
<th>$T_E^2$</th>
<th>$T'_E^2$</th>
<th>$[T, Z_2]_M$</th>
</tr>
</thead>
</table>
| $Z_2$ acting as a charge conjugation

Now we turn to the more complicated case where the $Z_2$ symmetry acts as a charge conjugation. Let us first pause to lay out the principle of organizing these states. Let us focus on the case with $\theta = 0$ for the moment. In this case, it is meaningful to discuss whether $E$ is a Kramers doublet under the original time reversal $T$, and whether $M$ is a Kramers doublet under $T'$. Also, notice now it is also meaningful to ask whether $Z_2$ squares to $+1$ or $-1$ for both $E$ and $M$ (see Appendix G.1 for more details). We will use $(\cdots)_-$ to indicate that $Z_2$ acts as a charge conjugation, and a subscript $Z$ to represent that certain excitation has $Z_2$ squaring to $-1$. For example, $(E_{fT}M_{bT'}Z)_-$ means $Z_2$ flips both the electric charge and magnetic charge, and $E$ is a fermionic Kramers doublet under $T$, while $M$ is a boson where $Z_2$ squares to $-1$, and $M$ is also a Kramers doublet under $T'$.

With this notation, we can list all $3 \times 2^2 \times 2^2 = 48$ possible distinct states with $\theta = 0$ and $Z_2$ acting as a charge conjugation, and they are shown in Table 7.9 and Table 7.10.

Similarly, for states with $\theta = \pi$ and $Z_2$ acting as a charge conjugation, there are only two states: $(E_{fT}M_{fT'})_{\theta-}$ and $(E_{fT}M_{fT'})_{\theta-Z}$. In both states, $Z_2$ takes the $(\frac{1}{2}, 1)$ dyon into the $(-\frac{1}{2}, -1)$ dyon. Because time reversal takes $(\frac{1}{2}, 1)$ into $(\frac{1}{2}, -1)$, then we know $T'$ takes $(\frac{1}{2}, 1)$ to $(-\frac{1}{2}, 1)$. This implies that $M$, the bound state of $(\frac{1}{2}, 1)$ and $(-\frac{1}{2}, 1)$, is a Kramers doublet under $T'$. [241, 339] The difference in these two states is that $Z_2$ squares to $+1$ ($-1$) on the $(\frac{1}{2}, 1)$ dyon in the former (latter). In fact, the former state is just the time reversal symmetric $(E_{fT}M_f)_{\theta}$ further equipped with a charge conjugation symmetry, so it must be anomaly-free.

So without examining anomalies, there are in total 50 possible distinct $Z_2 \times T$ symmetric $U(1)$
quantum spin liquids where $Z_2$ acts as a charge conjugation. It turns out that, together with the anomaly-free $(E_{JT} M_{JT'})_{\theta=0}$, 22 of these states are free of anomaly. The other 28 states are all anomalous, and there are 6 anomaly classes. The strategy to show the anomalies will be given in Sec. 7.7.3, and the arguments for this anomaly-detection will be completed in Appendix G.10.

Therefore, combined with the 16 states where $Z_2$ does not permute any excitation, there are in total 38 distinct non-anomalous $Z_2 \times \mathcal{T}$ symmetric $U(1)$ spin liquid states, and they can be found in Table 7.6, Table 7.8, Table 7.9 and Table 7.11.

We note that models that discuss $Z_2 \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids have been proposed in the literature, and a prototypical set includes but is not limited to Ref. [17–19]. In

Table 7.9: List of anomaly-free $Z_2 \times \mathcal{T}$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ acting as a charge conjugation. $T_E^2 = 1$ ($T_E^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $\mathcal{T}$. $T_M^2 = 1$ ($T_M^2 = -1$) represents the case where $M$ is a Kramers singlet (doublet) under $\mathcal{T}'$. $Z_{E,M}^2$ represents the result of acting the charge conjugation twice on $E$ and $M$, respectively.

<table>
<thead>
<tr>
<th></th>
<th>$T_E^2$</th>
<th>$Z_E^2$</th>
<th>$T_M^2$</th>
<th>$Z_M^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(E_b M_b)_-$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_b M_{bZ})_-$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_b M_{bT})_-$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_b M_{bTZ})_-$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_{fT} M_{bZ})_-$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_{fT} M_{b})_-$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_{fT} M_{b})_-$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_{fTZ} M_{bZ})_-$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_{fT} M_{bTZ})_-$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$(E_{fT} M_{bTZ})_-$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 7.10: List of anomalous $Z_2 \times T$ symmetric $U(1)$ quantum spin spin liquids that have $\theta = 0$ and have $Z_2$ acting as a charge conjugation at $\theta = \pi$. $T_E^2 = 1$ ($T_M^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T_M^2 = 1$ ($T_M^2 = -1$) represents the case where $M$ is a Kramers singlet (doublet) under $T'$. $Z_{E,M}^2$ represents the result of acting the charge conjugation twice on $E$ and $M$, respectively. The last column lists the anomaly classes.

<table>
<thead>
<tr>
<th>$(E_b M_b)_-$</th>
<th>$T_E^2$</th>
<th>$Z_E^2$</th>
<th>$T_M^2$</th>
<th>$Z_M^2$</th>
<th>anomaly class</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(E_b T M_b Z)_-$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_b T M_b T Z)_-$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_b M_f Z)_-$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_b M_f T Z)_-$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$1$</td>
<td>class 2</td>
</tr>
<tr>
<td>$(E_b T M_f Z)_-$</td>
<td>$1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>class 3</td>
</tr>
<tr>
<td>$(E_b M_f T'<em>Z)</em>-$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_f T M_b T Z)_-$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_f T M_b T'<em>Z)</em>-$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_f T M_f Z)_-$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_f T M_f T Z)_-$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_f T M_f T'<em>Z)</em>-$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$1$</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_f M_b Z)_-$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>class 6</td>
</tr>
<tr>
<td>$(E_f M_b T Z)_-$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$1$</td>
<td>class 6</td>
</tr>
<tr>
<td>$(E_f M_b T'<em>Z)</em>-$</td>
<td>$1$</td>
<td>$1$</td>
<td>$-1$</td>
<td>$1$</td>
<td>class 6</td>
</tr>
</tbody>
</table>

These models, the $Z_2 \times T$ symmetric $U(1)$ quantum spin liquid states are $(E_b M_b)_-$.
Table 7.11: List of \(Z_2 \times \mathcal{T}\) symmetric \(U(1)\) quantum spin liquids that have \(\theta = \pi\) and have \(Z_2\) acting as a charge conjugation. \(T_E^2 = 1\) \((T_E^2 = -1)\) represents the case where \(E\) is a Kramers singlet (doublet) under \(\mathcal{T}\). \(T_M^2 = 1\) \((T_M^2 = -1)\) represents the case where \(M\) is a Kramers singlet (doublet) under \(\mathcal{T}'\). \(Z_{E,M,D}^2\) represents the result of acting the charge conjugation twice on \(E\), \(M\) and the \((\frac{1}{2},1)\) dyon, respectively.

<table>
<thead>
<tr>
<th>((E_{JT}M_{JT'})_{\theta-})</th>
<th>(T_E^2)</th>
<th>(Z_E^2)</th>
<th>(T_M^2)</th>
<th>(Z_M^2)</th>
<th>(Z_D^2)</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>((E_{JT}M_{JT'})_{\theta-Z})</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>anomalous, class 1</td>
</tr>
</tbody>
</table>

7.7.3 STRATEGY OF ANOMALY-DETECTION

In this subsection we lay out the strategy to show the anomaly of the other 36 states. It turns out to be easier to first show that \((E_{bZ}M_{bZ})_{-}\) is anomalous with the \(Z_2\) symmetry (independent of time reversal), and this will be done later in this section. This immediately implies that \((E_f M_{bZ})_{-}\) and \((E_{bZ}M_f)_{-}\) are also anomalous with \(Z_2\) symmetry, because these states can be related to \((E_{bZ}M_{bZ})_{-}\) by tuning \(\theta\) by \(2\pi\). It also immediately implies that \((E_{bTZ}M_{bZ})_{-}\), \((E_{bZ}M_{bT'}Z)_{-}\), \((E_{JT}M_{bZ})_{-}\), \((E_{bZ}M_{JT'})_{-}\) are anomalous with \(Z_2 \times \mathcal{T}\) symmetry, because breaking \(\mathcal{T}\) will make them one of \((E_{bZ}M_{bZ})_{-}\), \((E_f M_{bZ})_{-}\) and \((E_{bZ}M_f)_{-}\). Furthermore, this means \((E_{JT}M_{JT'})_{\theta-Z}\) is anomalous, because even if the time reversal symmetry is broken, this state is smoothly connected to the anomalous \((E_{bZ}M_{bZ})_{-}\).

Next, using a generalization of the method for showing the anomaly of \((E_{bZ}M_{bZ})_{-}\), we show \((E_{bT}M_{bT'})_{-}\) and \(E_{bTT'}M_{b-}\) are anomalous with \(Z_2 \times \mathcal{T}\) symmetry in Appendix G.10. It turns out this is enough to show the remaining states are all anomalous. More precisely,

1. showing that \((E_{bT}M_{bT'})_{-}\) is anomalous is sufficient to show that the other entries in Table 7.10 are anomalous.

2. showing that \(E_{bTT'}M_{b-}\) is anomalous is sufficient to show that the rest entries in Table 7.7 are anomalous.

To see the first claim, let us consider \((E_{bT}M_{bZ})_{-}\) and \((E_{bZ}M_{bT'})_{-}\). These two states must be simultaneously anomalous or anomaly-free, because they are related to each other by the relabelling \(\mathcal{T} \leftrightarrow \mathcal{T}'\) and \(E \leftrightarrow M\). Suppose they are anomaly-free, then by combining them with
the states that will be constructed in Appendix G.10.1, we will get \((E_{bT}M_{bT'})_\sim\), which is in contradiction with that \((E_{bT}M_{bT'})_\sim\) is anomalous. This means if \((E_{bT}M_{bT'})_\sim\) is anomalous, then \((E_{bT}M_{bZ})_\sim\) and \((E_{bZ}M_{bT'})_\sim\) will also be anomalous. Combining these three anomalous states with the anomaly-free states constructed in Appendix G.10.1, one can show all other entries in Table 7.10 are also anomalous.

To see the second claim, consider \(E_{fT}M_{b-}\) and \(E_{fT'}M_{b-}\). These two states must also simultaneously be anomalous or anomaly-free, because they are related to each other by the relabelling \(T \leftrightarrow T'\) and \(E \leftrightarrow M\). Suppose they are anomaly-free. By combining them and the anomaly-free states constructed in Appendix G.10.1, we can get \(E_{bT}M_{b-}\), which contradicts that \(E_{bT}M_{b-}\) is anomalous. This means if \(E_{bT}M_{b-}\) is anomalous, then \(E_{fT}M_{b-}\) and \(E_{fT'}M_{b-}\) must also be anomalous. Combining these anomalous states with the anomaly-free states constructed in Appendix G.10.1, one can show all other entries in Table 7.7 are also anomalous.

ANOMALY OF \((E_{bZ}M_{bZ})_\sim\)

In the spirit of Sec. 7.6.2, here we will show that \((E_{bZ}M_{bZ})_\sim\) is anomalous with a \(Z_2\) charge-conjugation symmetry (independent of time reversal), by showing its corresponding SPT has an inconsistent surface.

To show the anomaly of \((E_{bZ}M_{bZ})_\sim\), we will consider from the perspective of \(E_{bZ}\), and suppose there is an SPT made of \(E_{bZ}\) that after gauging becomes \((E_{bZ}M_{bZ})_\sim\). On the surface of this SPT, we first condense the bound state of two \(E_{bZ}\), and this makes the surface a superfluid with a \(Z_4\) symmetry. There will be various vortices, and the \(4\pi\) vortex is the minimal trivial boson. So we can then condense the \(4\pi\) vortices to restore the \(U(1)\) symmetry, and this gives a symmetric gapped surface state where the \(U(1)\) charge of the excitations is quantized in units of \(1/2\).

The particle contents of the surface can be written as \(\{1,M_{bZ},X,N_I\} \times \{1,E_{bZ}\}\), where \(M_{bZ}\) is the remnant of the strength-1 monopole, \(X\) carries half charge under \(U(1)\), and \(N_I\)'s are neutral. In general, there can be many flavors of \(N_I\), but only the case with a single \(X\) needs to be considered, because other \(X\)'s can be related to a single one by attaching certain \(N_I\). Notice in this notation the inverses and bound states of these excitations are understood to be implicitly displayed. We would like to check whether such a surface is consistent. That is, it has consistent
braiding, fusion and symmetry transformation rules.

As for braiding, we know $E_{bZ}$ is local, $M_{bZ}$ and $X$ have mutual $\pi$ statistics, and $M_{bZ}$ has no mutual statistics with $N_I$. $X$ and $N_I$ can have complicated braiding though, and it can even be non-Abelian.

As for fusion, we have

$$M_{bZ} \times M_{bZ} = 1$$  \hfill (7.21)

$$X \times X = E_{bZ} + E_{bZ}N_1 + E_{bZ}N_2M_{bZ}$$  \hfill (7.22)

and

$$N_I \times N_J = N_k + M_{bZ}N_k$$  \hfill (7.23)

(7.21) comes from that this surface is obtained by condensing $M_{bZ}^2$, and (7.22) and (7.23) are obtained under the constraint due to charge conservation. Notice in (7.22) the fusion product cannot be $E_{bZ}M_{bZ}$, because this will be inconsistent with the general condition that a particle and its anti-particle should have the same topological spin. Also, all potential fusion multiplicities are suppressed, and they turn out to be unimportant for our discussion.

Now if we are willing to break the $U(1)$ symmetry on the surface by condensing $E_{bZ}M_{bZ}$,\textsuperscript{9} $X$ will be confined and $N_I$’s will remain, and we will be left with $\{1, M_{bZ}, N_I\}$ that has a $Z_4$ symmetry. Notice that $\{1, M_{bZ}, N_I\}$ is closed under fusion and braiding, and it is known that in three dimensions there is no bosonic SPT protected by $Z_4$ symmetry. This means $N_I$ can be further confined (without breaking the $Z_4$ symmetry), and we are left with $\{1, M_{bZ}\}$. In other words, $N_I$’s can be viewed as emergent particles of a system made of $M_{bZ}$ in the presence of the $Z_4$ symmetry but in the absence of the $U(1)$ symmetry. However, because neither $M_{bZ}$ nor $N_I$ carries a $U(1)$ charge, even in the presence of $U(1)$ symmetry $N_I$ can still be viewed as emergent particles of a system made of $M_{bZ}$.

\textsuperscript{9}This is a boson in this surface topological order.
So we can get rid of \( N_I \) and be left with \( \{1, M_{bZ}, X\} \times \{1, E_{bZ}\} \). Now the fusion of \( X \) must be

\[
X \times X = E_{bZ}
\]  

(7.24)

The only possible consistent topological order of this state is a \( Z_2 \) topological order (or its twisted version, the double-semion theory).

Let us turn to symmetry assignment, and we will particularly consider how charge-conjugation acts on \( X \). Notice when defining the charge-conjugation action on \( X \), there is an ambiguity due to our freedom to multiply it by a gauge transformation. But because this topological order is a \( Z_2 \) gauge theory, the action of the global charge-conjugation symmetry twice should have an unambiguous result on \( X \). So in order to be consistent with the above fusion rule, \( X \) must go to \( \pm iX \) upon acting with charge-conjugation twice. Below we show this is impossible.

Suppose the action of charge-conjugation on \( X \) is implemented by a generic matrix \( C \)

\[
X_i \rightarrow C_{ij} X_j^\dagger, \quad X_i^\dagger \rightarrow C_{ij}^* X_j
\]

(7.25)

Notice the indices label different components of \( X \) that differ by some local operations. This implies that acting charge-conjugation twice on \( X \) gives

\[
X_i \rightarrow (CC^*)_{ij} X_j
\]

(7.26)

Now consider the operator \( X_i M_{ij} X_j \) with an arbitrary matrix \( M \), which is a charge-1 boson, so the charge-conjugation acting on it twice gives \(-1\). This requires

\[
(CC^*)^T M(CC^*) = -M
\]

(7.27)

Because \( M \) is arbitrary, this is possible only if \( CC^* = \pm i \), which confirms the previous statement that \( X \rightarrow \pm iX \) upon acted by \( Z_2 \) twice. However, no matrix \( C \) can possibly satisfy \( CC^* = \pm i \). To see this, suppose \( CC^* = \pm i \), then \( (CC^*)^2 = -1 \). On the other hand, \( C^*C = \mp i \) and \( (CC^*)^2 = CC^*CC^* = C(C^*C)C^* = 1 \), which contradicts with the previous result.
The above contradiction shows that there cannot even be any $X$. So the surface is just $\{1, E_bZ, M_bZ\}$, where everything is local. This means that there is a charge neutral excitation that has $Z_2$ squaring to $-1$, which contradicts the original assumption. Therefore, this SPT cannot exist, and furthermore, $(E_bZ M_bZ)_-$ is anomalous with a $Z_2$ charge-conjugation symmetry. Notice from the above argument we see the anomaly of $(E_bZ M_bZ)_-$ is independent of time reversal.

Note this argument can be easily modified to show that $E_b^{1/2} M_b^{1/2}$ is anomalous with $SO(3)$ symmetry, by changing every $Z_2$ symmetry by $SO(3)$ symmetry, and changing every excitation with charge-conjugation squaring to $-1$ by an excitation with spin-1/2. This is of course consistent with our conclusion from Sec. 7.3, where we have used the $SO(3)$ monopole to show the anomaly.

7.7.4 ANOMALY CLASSES

Before finishing this section, we make some brief remarks on the anomaly classes of these anomalous spin liquid states. Here by an anomaly class, we mean a group of anomalous states which can be turned into each other by coupling it with a state that is anomaly-free. Because the anomalous spin liquid states can in principle be realized on the surface of some 4+1-d $Z_2 \times T$ symmetric bosonic SPTs, analysing the anomaly classes of these 37 anomalous spin liquid states gives some information on the properties of these SPTs.

We first discuss the anomalous spin liquid states with $Z_2$ acting as a charge conjugation. It is straightforward to check that within each of the following 6 groups of anomalous states, all states have the same anomaly:

1. $(E_bZ M_bZ)_-, (E_bT Z M_bT' Z)_-, (E_{fT} M_{fT} Z)_-, (E_{fT} M_{fT'} Z)_-, (E_{fT} Z M_{fT'} Z)_-, (E_{fT} Z M_{fT} Z)_-$.

2. $(E_{fT} Z M_bZ)_-, (E_f M_bZ)_-, (E_{bT} Z M_f)_-$.

3. $(E_bZ M_{bT'} Z)_-, (E_bZ M_f)_-, (E_f M_{bT'} Z)_-$.

4. $(E_bZ M_{bT'} Z)_-, (E_{fT} Z M_{bT'} Z)_-, (E_{fT} Z M_{bT} Z)_-, (E_bT M_{fT} Z)_-, (E_bZ M_{fT} Z)_-$.

5. $(E_{bT} M_bZ)_-, (E_{bT} M_{fT} Z)_-, (E_{bT} Z M_{fT} Z)_-, (E_{bT} Z M_{fT'} Z)_-, (E_{fT} Z M_{bT'} Z)_-, (E_{fT} Z M_{bT} Z)_-$.

6. $(E_{bT} M_{bT'} Z)_-, (E_f M_{bT'} Z)_-, (E_{bT} M_f)_-$.
It is clear that combining group 1 and group 4 results in group 3, combining group 1 and group 5 results in group 2, and combining group 4 and group 5 results in group 6. This implies that the 4D bosonic SPTs with $Z_2 \times T$ symmetry at least have a classification of $Z_3^2$. Group cohomology gives precisely the same classification, [344] and our results suggest that the surface states of these SPTs can be the above anomalous $U(1)$ gauge theories.

Notice there is another SPT that is beyond group cohomology, and that SPT is protected purely by $Z_2$ symmetry. [354, 355] One physical realization of the bulk of this SPT is to consider a decorated domain wall construction, where on each $Z_2$ domain wall we place an $efmf$ state. [354] The surface properties of this beyond-group-cohomology SPT is unclear at this point, and it may be interesting to work it out.

Taking all these together, we propose that the complete classification of 4D bosonic SPTs with $Z_2 \times T$ symmetry is $Z_4^2$. This agrees with the classification of 4D bosonic SPTs with $Z_2 \times Z_2^P$ symmetry, where $Z_2^P$ is a reflection symmetry that results in a trivial action when acted twice. [356]

Next we discuss the anomalous spin liquid states with $Z_2$ not acting as a charge conjugation, whose anomaly classes can be organized as

a. $E_{bT} M_f$, $E_{fT} M_b$, $E_{bT} M_b$.

b. $E_{bT} M_f$, $E_{fT} M_b$, $E_{bT} M_b$.

c. $E_{bT} M_b$, $E_{bT} M_f$, $E_{f} M_b$.

First notice all these states are anomalous with the full $Z_2 \times T$ symmetry. If $T$ is broken, all these states should be non-anomalous $Z_2$ symmetric states. Second, class a and class b differ by the relabelling $T \leftrightarrow T'$ and $E \leftrightarrow M$, and class c can be obtained by combining states in class a and class b. Notice before that class 2 and class 3 differ by this relabelling, so do class 4 and class 5, and that class 6 can be obtained by combining class 2 and class 3, or by combining class 4 and class 5. This suggests states in class c here are of the same anomaly class as states in class 6 as above. We do not attempt to completely settle down the relation among these anomaly classes in this paper.
7.8 $U(1)$ QUANTUM SPIN LIQUIDS ENRICHED BY SOME OTHER SYMMETRIES

In the spirit of the general framework in Sec. 7.6, in this section we briefly discuss $U(1)$ quantum spin liquids with some other symmetries. Part of the motivation comes from the existing lattice models that realize $U(1)$ quantum spin liquids with $O(2) \times T = (U(1) \times Z_2) \times T$ symmetry. [13–15, 357] In all these models, the improper $Z_2$ rotation of the $O(2)$ symmetry acts as a charge conjugation. Ref. [13–15] studied a couple of different lattice models that realize $(E_bM_b^1)_-$, where $\theta = 0$ and both $E$ and $M$ are bosons, and $M$ carries half charge under the $U(1)$ subgroup of $O(2)$. Ref. [357] constructed two other models of $O(2) \times T$ symmetric $U(1)$ quantum spin liquids, where one of them has a bosonic monopole and the other has a fermionic monopole. These two states are $(E_bM_b)_-$ and $(E_bM_f)_-$, respectively. Notice, for simplicity, in this section we will not consider the refined classification that considers coupling these spin liquids with SPTs.

7.8.1 $SO(N) \times T$ SYMMETRY

We can generalize our classification of $SO(3) \times T$ symmetric $U(1)$ quantum spin liquids to $SO(N) \times T$ symmetric $U(1)$ quantum spin liquids, with the integer $N > 3$. The projective representations of $SO(N) \times T$ have the same classification as those of $SO(3) \times T$: there is a spinor representation of $SO(N)$ and a Kramers doublet representation of time reversal. Therefore, the enumeration of states with this symmetry goes in a parallel way as those with $SO(3) \times T$ symmetry, and all non-anomalous states with $SO(3) \times T$ can be generalized to their $SO(N) \times T$ analogs. Furthermore, $\pi_1(SO(N)) = Z_2$ and the monopole structure of an $SO(N)$ gauge field is similar to that of an $SO(3)$ gauge field, so the generalization of the anomalous states in the $SO(3) \times T$ case will still be anomalous. Therefore, we conclude that with $SO(N) \times T$ symmetry there will also be 15 distinct non-anomalous $U(1)$ quantum spin liquids, and they have similar properties as those with only $SO(3) \times T$ symmetry in terms of the bulk fractional excitations.

For the special case of $SO(2) \times T \cong U(1) \times T$, its projective representations on $E$ are classified by $Z_2^2$. One of the nontrivial root projective representation corresponds to Kramers doublet of time reversal, while the other corresponds to half-charge of $SO(2)$, which is protected by time re-
versal here. Therefore, although $SO(2)$ has no projective representation on its own, the projective representations of $SO(2) \times T$ on $E$ can still be viewed as descendants of those of $SO(3) \times T$.

If $\theta = 0$, there is no projective representation on $M$. So there are $3 \times 2^2 = 12$ putative states with $\theta = 0$, which can all be viewed as descendant states of $SO(3) \times T$ symmetric states when the symmetry is reduced to $SO(2) \times T$. Notice some distinct $SO(3) \times T$ symmetric states have the same $SO(2) \times T$ symmetric descendant, because there is no fractional quantum number on monopoles anymore. The descendants of the 15 anomaly-free states of course remain anomaly-free. By inspecting the anomaly classes of anomalous $SO(3) \times T$ symmetric states listed in Sec. 7.3, we see in each anomaly class there is at least one state that has a trivial monopole when the symmetry is reduced to $U(1) \times T$, which means all these anomalous states will become anomaly-free. So all these 12 $SO(2) \times T$ symmetric states with $\theta = 0$ are anomaly-free.

If $\theta = \pi$, the properties of the $(\frac{1}{2}, \pm 1)$ dyons will determine the phase, which does not have any projective representation of this symmetry. So it contributes one anomaly-free state, which can be viewed as a descendant of the $SO(3) \times T$ symmetric $(E_{IT} M_f)_{\theta}$ state when the symmetry is broken down to $U(1) \times T$.

Therefore, there are in total 13 distinct anomaly-free $U(1)$ quantum spin liquids with $U(1) \times T$ symmetry.

### 7.8.2 $SO(N)$ Symmetry

In the following few subsections we will consider the case in the absence of time reversal symmetry. In this subsection we start by discussing $U(1)$ quantum spin liquids with only $SO(3)$ spin rotation symmetry, which can be realized in systems with a spin chirality term in the Hamiltonian.

As for the elementary excitations, we focus on $(1, 0)$ and $(q_e, 1)$, with $q_e$ a real number. It is not hard to see there must be bosons for some $q_e$, and we will consider such a bosonic $(q_e, 1)$ excitation. Due to the absence of time reversal symmetry, $\theta$ in (7.1) can be tuned continuously, so that the $(q_e, 1)$ particle chosen above can be tuned to $(0, 1)$ by Witten effect. That is to say, now we fix $(0, 1)$ to be a boson. Similarly, we can make $(1, 0)$ also a boson. Therefore, the statistics of the elementary excitations is irrelevant here due to the absence of time reversal symmetry.
Next we turn to their quantum numbers under symmetry. Due to the absence of time reversal symmetry, the distinction between electric charge and magnetic monopole also becomes irrelevant. By enumerating the quantum numbers of $E = (1, 0)$ and $M = (0, 1)$ under $SO(3)$, we can have $E_b M_b$, $E_{b1/2} M_b$ and $E_{b1/2} M_{b1/2}$, and these exhaust all (including anomalous) $SO(3)$ symmetric $U(1)$ spin liquids. $E_b M_b$ and $E_{b1/2} M_b$ are clearly not anomalous, and their description as a gauged SPT is similar to their cousins with $SO(3) \times T$ symmetry. Also, as argued in Sec. 7.3, $E_{b1/2} M_{b1/2}$ is anomalous even with only $SO(3)$ symmetry.

Therefore, with only $SO(3)$ symmetry, there are only two (anomaly-free) distinct possible symmetry realizations in the $U(1)$ quantum spin liquids: $E_b M_b$ and $E_{b1/2} M_b$. This concludes the classification of $U(1)$ quantum spin liquids with $SO(3)$ symmetry.

Similar reasoning as above can be applied to $U(1)$ quantum spin liquids with $SO(N)$ symmetry with $N > 3$. First, all these $SO(N)$ groups have one nontrivial projective representation, the spinor representation. Also, as mentioned above, the monopole properties of an $SO(N)$ gauge field is similar to that of an $SO(3)$ gauge field. Therefore, the arguments for the enumeration of all the states, construction of the non-anomalous states and examination of the anomalous states are parallel to that of $SO(3)$, and gives only 2 distinct $SO(N)$ symmetric $U(1)$ quantum spin liquids.

For the special case of $SO(2) \cong U(1)$, because of the absence of any nontrivial projective representation of this symmetry, there will be only a single type of symmetric $U(1)$ quantum spin liquid.

7.8.3 $Z_2$ SYMMETRY

For the case with $Z_2$ symmetry, as above, all states can be symmetrically tuned so that it has the $\theta = 0$-type of charge-monopole lattice with both $E$ and $M$ bosonic. Also notice there is no projective representation of $Z_2$, so nontrivial states must have $Z_2$ acting as charge conjugation. Then there can be $E_b M_b$, $(E_b M_b)_-$, $(E_{bZ} M_b)_-$ and $(E_{bZ} M_{bZ})_-$. The first three states can clearly be realized, but as shown before, $(E_{bZ} M_{bZ})_-$ is anomalous with a $Z_2$ symmetry. So there are 3 distinct $Z_2$ symmetric $U(1)$ quantum spin liquids: $E_b M_b$, $(E_b M_b)_-$ and $(E_{bZ} M_b)_-$. 

218
7.8.4 $O(2)$ symmetry

Similar considerations can be applied to the case with $O(2)$ symmetry. There will still be two spin liquid states where the improper $Z_2$ component does not act as a charge conjugation, and these are the descendants of $E_b M_b$ and $E_{b\frac{1}{2}} M_b$ with $SO(3) \times T$ symmetry. As shown in Sec. 7.3, the descendant of $E_{b\frac{1}{2}} M_{b\frac{1}{2}}$ is still anomalous with $O(2)$ symmetry. Again, for a complete classification, states where $Z_2$ acts as a charge conjugation need to be taken into account. Unlike the case with $O(2) \times T$ symmetry, the fractional excitations always have integer charges under the $U(1)$ subgroup of $O(2)$, if the improper $Z_2$ component acts as a charge conjugation. So these states include $(E_b M_b)_-$, $(E_{bZ} M_b)_-$, $(E_{bZ} M_{bZ})_-$, and the first two are anomaly-free, while the last one is anomalous. Therefore, there are in total 4 distinct non-anomalous $O(2)$ symmetric $U(1)$ quantum spin liquids: $E_b M_b$, $E_{b\frac{1}{2}} M_b$, $(E_b M_b)_-$ and $(E_{bZ} M_b)_-$.  

7.9 Discussion

In this paper we have classified and characterized 3D symmetry enriched $U(1)$ quantum spin liquids. One of our focuses is on such spin liquids with time reversal and $SO(3)$ spin rotational symmetries. 26 states were enumerated based on the properties of the bulk fractional excitations, among which only 15 can be realized in 3D lattice systems. We explain in details how to view these quantum spin liquids as gauged version of some SPTs. The other 11 are shown to be anomalous, i.e. they cannot be realized in a 3D bosonic system with these symmetries. In Appendix G.3, they are constructed on the surface of some 4D bosonic short-range entangled states.

The anomalies of the anomalous states become clear when the properties of the $SO(3)$ monopoles are examined. Although checking the topological defects of the gauge field that corresponds to certain symmetry has been widely applied to detect anomalies, to the best of our knowledge, the properties of $SO(3)$ monopoles have not been investigated in previous studies. We expect it to be helpful in studying other problems that involves $SO(3)$ symmetries.

When combined with bosonic SPTs with time reversal and $SO(3)$ spin rotational symmetry, we find a further refined classification which shows there are 168 different $U(1)$ quantum spin liquids.
After warming up with the example of $SO(3) \times T$ symmetric $U(1)$ quantum spin liquids, we have described a general framework to classify such spin liquid states with a general symmetry. This approach is again physics-based, and it has the advantage of providing us with intuition both on the classification and the physical characterization. However, it is not always easy to implement this framework, and so it is desirable to find a simpler systematic way to do the classification. The field theoretic formal approach discussed in Sec. 7.6 may be potentially helpful. Another possibly helpful formal approach is to generalize the categorical theory that is used to study 2d SETs to $U(1)$ quantum spin liquids. [317, 319] This may be possible because in both cases the excitations are all particle like, although there are infinitely many types of fractional excitations in a $U(1)$ quantum spin liquid.

In the spirit of this general framework, we have also discussed $U(1)$ quantum spin liquids with some other symmetries, and found some very rich structures. In particular, we discussed $U(1)$ spin liquids with $Z_2 \times T$ symmetry in great detail. Based on the properties of the bulk fractional excitations, there are 38 such $Z_2 \times T$ symmetric states that are free of anomaly. The anomalies of the other 37 such $Z_2 \times T$ symmetric states are detected based on the method in the general framework. The study of $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids have some implications on some SPTs, as discussed in Appendix G.10.

Besides looking for a simpler systematic classification of these symmetry enriched $U(1)$ quantum spin liquids, the other most important open questions are of course which microscopic models and experimental systems realizing these different symmetry enriched $U(1)$ quantum spin liquids, and how to detect and distinguish them numerically or experimentally. One particular interesting theoretical aspect of this question is how lattice symmetries interplay with these quantum spin liquids. These are beyond the scope of the current paper and are worth further investigating in the future, and we note some recent progress in this aspect. [358–361]

Another interesting theoretical challenge is to classify and characterize symmetry enriched gapped quantum spin liquids, some of which can be obtained by condensing some excitations in the $U(1)$ quantum spin liquids. One complication in this problem is that there usually exists loop-like excitations, whose properties are not completely understood to date. Because some of these gapped quantum spin liquids are descendants of the $U(1)$ quantum spin liquids, relating
the properties of loop-like excitations in the former to the properties of the particle-like excitations in the latter may shed light on this problem.
8

Bulk characterization of topological crystalline insulators

8.1 Introduction

The notion of symmetry protected distinction of quantum phases of matter is by now well appreciated: some quantum phases are smoothly connected to each other in the absence of symmetry, but when the relevant symmetries are present, these phases are sharply distinguished, i.e. it is not possible to go from one phase into the other without crossing a phase transition (see Figure 8.1). The most well-known example may be the distinction between topological insulators and trivial insulators: in the presence of charge conservation and time reversal symmetries, these two
types of insulators are separated by a phase transition. However, once these symmetries are allowed to be broken, they can be smoothly connected. [4, 5, 362]

**Figure 8.1:** The notion of symmetry protected distinction of quantum phases of matter. As long as the relevant symmetries are preserved, the two phases cannot be connected without crossing a phase transition. However, if symmetry-breaking perturbations are allowed, the phase transition can be avoided.

Topological crystalline insulators (TCIs) are another class of topological materials that have been under intense recent studies. These are short-range entangled phases that are protected by crystalline symmetries: they cannot (can) be connected to a trivial insulator in the presence (absence) of the relevant crystalline (and other) symmetries. [363–367] These states are not only of conceptual interests, but the state-of-the-arts in synthesizing materials also offer great opportunities to realizing them experimentally. [151, 241, 368–376]

The majority of the theoretical studies of TCIs are based on band theories, which are valid for systems made of non-interacting or weakly-interacting fermions. An important question then is whether the nontrivial TCIs predicted by band theories are stable under interactions, or, more precisely, whether these band-theory-based nontrivial TCIs can be smoothly connected to a trivial insulator once strong interactions are switched on (still in the presence of the relevant symmetries). Very recently, the stability of various such TCIs under interactions has been considered and some of them are shown to be unstable, by studying the anomalous surface properties of the TCIs, studying certain dimensionally reduced versions of the TCIs, or studying the formal field theories of the TCIs. [74, 356, 377–386] However, an important and interesting question has remained unanswered: how to characterize nontrivial interacting TCIs directly by their bulk properties in a physical way?

In this paper we consider interacting three dimensional fermionic TCIs protected by crystalline symmetries together with a $U(1)$ charge conservation symmetry, and in some cases other symme-
tries (such as time reversal or $SU(2)$) are also present. Because all these insulators have a $U(1)$ symmetry, it is interesting and helpful to ask what happens if such TCIs are coupled to a (compact) dynamical $U(1)$ gauge field, and in particular, whether the resulting magnetic monopoles in the TCIs will be nontrivial. As a theoretical tool, this idea has provided great insights in other subjects where a $U(1)$ symmetry plays an important role, [10, 29, 316, 327, 339, 351, 361, 387–391] although a magnetic monopole of the real electromagnetism has not been detected.

Therefore, we will discuss a particular type of bulk characterization of these TCIs: the properties of their magnetic monopoles once these TCIs are coupled to a dynamical $U(1)$ gauge field. We will focus on the cases where the crystalline symmetry responsible for the symmetry protection is either a translation symmetry or a mirror symmetry, and we will show several examples of band-theory-based nontrivial TCIs that have monopoles with nontrivial quantum numbers. Because these nontrivial quantum numbers of monopoles are rigid characterizations of the nontrivial TCIs, this implies that these TCIs will remain nontrivial even in the presence of strong interactions.

We will employ two complementary methods to derive the quantum numbers of monopoles, and both methods involve studying symmetric surface states of the TCIs. In general, there can be two types of symmetric surface states of a TCI: it can either be gapless, or it can be gapped at the cost of developing a surface topological order (STO). The two methods that we will be using are based on studying these two types of surface states. In the first method, we start from a symmetric gapless surface of the TCIs that can be described by free Dirac cones. Because the free Dirac theory is a conformal field theory (CFT), we can utilize state-operator-correspondence of a CFT to read off the quantum numbers the monopoles. In the second method, we start from a symmetric STO of the TCIs, then among the surface anyons of the STO we will identify an avatar of the bulk monopole, from which we can also extract the quantum numbers of the monopoles. Notice some TCIs may not have any symmetric STO, and we will discuss two such examples. That consistent results can be derived from different types of surface states is on the one hand assuring, and on the other hand indicating that it is usually a more direct and unified way to characterize a nontrivial TCI by its bulk physical properties.

The specific TCIs that we discuss are conceptually interesting and/or experimentally feasible,
which include:

1. Weak topological insulator (WTI).

A WTI is protected by charge conservation, time reversal and translation symmetries, and it can be viewed as many decoupled layers of two dimensional topological insulators. [4, 5, 362]

2. TCI protected by a mirror symmetry (MTCI).

This TCI has the same symmetries as the experimentally discovered TCI SnTe, and we will consider the ones that can have a symmetric surface with two Dirac cones. We also discuss two time reversal symmetric generalizations of this state (TMTCI-1 and TMTCI-2) in Appendix H.1.

3. Doubled topological insulators protected by a mirror symmetry (DTI).

This example consists of two copies of three dimensional topological insulators further equipped with a mirror symmetry. It is well known that a single topological insulator is protected by charge conservation and time reversal, but two copies of them are smoothly connected to a trivial insulator if there is no other symmetry. [4, 5, 362] However, with a further mirror symmetry, the DTI can still be nontrivial. [367]

4. TCIs with symmetry-enforced-gapless surfaces.

As the last examples, we construct two TCIs (SEG-1 and SEG-2) whose surfaces exhibit symmetry enforced gaplessness: their surfaces have to be gapless as long as the relevant symmetries are preserved, and a symmetric STO is disallowed even with strong interactions. [339]

The studies of these TCIs are related to another intriguing subject. Recently there has been great interest in symmetry enriched long-range entangled phases: phases which cannot be smoothly connected to a trivial state even in the absence of any symmetry, and which acquire symmetry protected distinctions among themselves. In the last few years tremendous and systematic
progress has been made in understanding two dimensional symmetry enriched long-range entangled states, [309–319,392] and the studies of three dimensional symmetry enriched long-range entangled states just began. [320–322] Very recently, in Ref. [361] the author (together with Chong Wang and T. Senthil) discussed, as one of the first systematic studies of three dimensional symmetry enriched long-range entangled phases, the classification and characterization of three dimensional symmetry enriched U(1) quantum spin liquids (QSLs), which are long-range entangled phases with an emergent gapless photon at low energies (see also Ref. [327]). When the TCIs discussed in this paper is coupled to a (compact) dynamical U(1) gauge field, the resulting state is precisely a U(1) QSL enriched by a crystalline symmetry (together with some other symmetries in many cases). We discuss the properties of these crystalline symmetric U(1) QSLs. Moreover, we demonstrate an interesting and potentially useful correspondence between crystalline-symmetry-enriched U(1) QSLs discussed here and some internal-symmetry-enriched U(1) QSLs discussed in Ref. [361]. In particular, the two TCIs with symmetry-enforced-gapless surfaces are found to be the crystalline-symmetric realizations of “fractional topological paramagnets” introduced in Ref. [361].

The rest of the paper is organized as follows. We first describe the two methods to derive the properties of monopoles in Sec. 8.2. Then in Sec. 8.3 we apply these methods to study the monopoles of the TCIs mentioned above, and we find that they all have nontrivial monopoles, so they are all stable under interactions. Next in Sec. 8.4 we discuss the relations between the TCIs discussed in this paper and symmetry enriched U(1) QSLs. Finally, we conclude with some discussions in Sec. 8.5. The appendices contain some supplementary information.

8.2 PROPERTIES OF BULK MONOPOLES FROM A SYMMETRIC SURFACE STATE

In this section we describe two complementary methods that allow us to read off the properties of monopoles of TCIs, and both methods utilize bulk-boundary correspondence. These methods can in principle be generalized and applied to study the properties of monopoles in other problems.
8.2.1 A symmetric gapless surface with Dirac cones

We start with the case where a symmetric gapless surface of a TCI is known. This is usually the case if the TCI under interest can be realized by free fermions, where the symmetric gapless surface can be described by free Dirac cones. From the perspective of the free Dirac fermions on the surface, a monopole in the bulk corresponds to an instanton. Because free Dirac fermions are described by CFTs, the quantum numbers of their instantons can be read off by state-operator-correspondence. More precisely, the instanton operators correspond to the states obtained by threading a $2\pi$ flux through the surface theory. Guaranteed by the index theorem, each Dirac fermion will contribute a zero mode in this flux background. By studying how such states transform under certain symmetries, we can know how the instanton operators transform, and thus extract the quantum numbers of the monopoles. This procedure will be carried out for the examples, and here we just make some remarks.

First of all, we note for fermionic insulators, the monopoles must be bosonic. Therefore, we only need to pay attention to the possible nontrivial (projective) quantum numbers of the monopole under the relevant symmetries. This is not the case for bosonic TCIs, where the monopoles can in principle be fermionic, and one has to pay attention to both the statistics and quantum numbers of the monopole. In that case, as long as either the statistics or the quantum number is nontrivial, the monopole is nontrivial.

Second, if the symmetric surface of a band-theory-based nontrivial TCI is characterized by an odd number of Dirac cones, it is known that the monopole will carry half integer electric charge under the $U(1)$ symmetry (we will always measure the charge in units of the charge of electrons). This is always a nontrivial projective quantum number, so such TCIs will always be stable under interactions.

Therefore in this paper we focus on TCIs whose surfaces host two Dirac cones, where the monopole carries integer charge, which can then always be neutralized by attaching certain charge-1 fermions to it. If this neutral monopole carries nontrivial projective quantum numbers, this TCI is necessarily stable under interactions. Notice the inverse statement is not true. For example, some nontrivial band-theory-based TCIs with four surface Dirac cones have trivial monopoles,
but they are still stable under interactions. These TCIs have bulk characterizations different than nontrivial monopoles, and they are beyond the scope of this paper.

8.2.2 A symmetric surface topological order

Now we describe how to derive the properties of the monopoles from a symmetric STO. [10, 29, 316, 327, 339, 361, 387–391] A topological order is characterized by a set of anyon types and their topological properties, such as fusion and braiding. If this topological order respects a $U(1)$ symmetry, the anyons will have definite $U(1)$ charges. Denote the $U(1)$ charge of an anyon $a$ by $q_a$.

Now suppose dragging a monopole from the outside of a TCI, which is assumed to be a trivial vacuum, to its inside. A $2\pi$ flux will be left on the surface after this event. Because this is a local process, no anyon far away from the flux should be able to tell the difference before and after the monopole goes through the surface. However, for an anyon with charge $q_a$, no matter how far it is from the $2\pi$ flux, it will experience an Aharonov-Bohm phase factor when it moves around the flux:

$$\exp(2\pi i q_a)$$

Therefore, when the monopole goes through, the $2\pi$ flux must nucleate an anyon, $M$, which has braiding that precisely cancels this Aharonov-Bohm phase factor for any anyon $a$. That is, there must be an anyon $M$ such that

$$\theta_{M,a} = \exp(-2\pi i q_a) \quad (8.1)$$

for all anyon $a$, where $\theta_{M,a}$ denotes the braiding phase factor between $M$ and $a$. Because $M$ has Abelian braiding with all anyons, it must be an Abelian anyon. [317]

Therefore, this anyon $M$ is a surface avatar of the bulk monopole, and from how $M$ transforms under the relevant symmetries we can read off the quantum numbers of the monopole under these symmetries.

Before finishing this section and applying these methods to study some specific TCIs, let us make some remarks. All the TCIs studied in this paper have a free-fermion realization, so the first method based on symmetric gapless surface is applicable to all of them. For TCIs that can-
not be realized by free fermions (see, for example, Ref. [74]), it will be difficult to apply the first method. If such states have a symmetric STO, then the second method also applies. Notice we will describe examples where the surface exhibits the phenomenon of symmetry enforced gaplessness first discussed in Ref. [339]: as long as the surface preserves all the relevant symmetries, it must be gapless and a symmetric STO is disallowed. Our examples are the first crystalline-symmetric realizations of this phenomenon. Although we do not know any example so far, there can in principle be intrinsically interacting TCIs that exhibit symmetry-enforced-gapless surfaces. Such TCIs will be very interesting, and it will be difficult to use either method to read off the quantum numbers of their monopoles.

8.3 Examples

In this section we study several examples of TCIs, and we will use the methods described above to show that they all have nontrivial monopoles, so they are all stable under interactions.

8.3.1 Weak topological insulators

We start from a weak topological insulator (WTI), which is a TCI protected by a $U(1)$ charge conservation, a time reversal, $\mathcal{T}$, and a translation, $T_y$. A WTI can be smoothly connected to layers of decoupled two dimensional topological insulators. [4, 5, 362] Both a symmetric gapless surface and a symmetric STO of a WTI have been discussed in Ref. [394]. Below we will derive the quantum numbers of the monopole of a WTI based on these surfaces.

Let us start from the symmetric gapless surface on the $z = 0$ plane, which can be described by two Dirac cones at low energies:

$$H = \psi^\dagger (-i \partial_x \sigma_x - i \partial_y \sigma_y) \psi \quad (8.2)$$

where $\psi = (\psi_1, \psi_2)^T$ collectively denotes the two Dirac cones, $\psi_1$ and $\psi_2$. The $\sigma$ matrices are Pauli matrices acting within each Dirac cone, and we will use $\tau$ matrices to denote Pauli matrices acting between $\psi_1$ and $\psi_2$. These will be the notations throughout this paper.
In order to determine the quantum numbers of a monopole of a WTI, let us imagine threading a $2\pi$ flux through the surface Dirac cones described by (8.2). In this flux background, the Dirac cones $\psi_1$ and $\psi_2$ will contribute a zero mode $f_1$ and $f_2$, respectively. Let us denote $|0\rangle$ to be the state under this $2\pi$ flux background with no zero modes being occupied, so $f_{1,2}^\dagger|0\rangle$ is the state under the flux background with $f_{1,2}$ being occupied, and $f_1^\dagger f_2^\dagger|0\rangle$ is the state under the flux background with both zero modes being occupied. To proceed, let us momentarily assume the system also has an extra anti-unitary particle-hole symmetry that flips the charge but keeps the flux. Under this assumption, the states $f_1^\dagger|0\rangle$ and $f_2^\dagger|0\rangle$ correspond to two components of the neutral monopole operator. [393] By abuse of notations, we will denote these two components of the monopole operators by $M_{1,2}$ such that the correspondence is

$$M_1 \sim f_1^\dagger|0\rangle, \quad M_2 \sim f_2^\dagger|0\rangle$$

(8.3)

We will determine the quantum numbers of the monopole in the presence of this extra symmetry, and then break this extra symmetry. Due to the discrete nature of the nontrivial quantum numbers, they are rigid characterizations of the monopoles which should not change when the extra symmetry is broken (at least weakly).

Suppose the translation symmetry responsible for the symmetry protection is $T_y$, translation along the $y$ direction, then the two Dirac cones have distinct momenta along the $y$ direction that differ by $\pi$ (the translation unit in the $y$ direction is taken to be 1). [394] Therefore, the symmetry actions on the low energy theory can be written as

$$U(1) : \psi \rightarrow e^{i\theta}\psi$$
$$T : \psi \rightarrow i\sigma_y\psi$$
$$T_y : \psi \rightarrow \tau_z\psi$$

(8.4)

Given these symmetries, there are two types of quantum numbers that the monopoles can carry: $T$ and $T_y$ can either commute or anti-commute. The former is a set of trivial quantum numbers, and the latter corresponds to a set of nontrivial projective quantum numbers.

Now we just need to examine how $T$ and $T_y$ act on $f_1^\dagger|0\rangle$ and $f_2^\dagger|0\rangle$. It is straightforward to
examine the action of $T_y$:

\[
M_1 \sim f_1^\dagger |0\rangle \rightarrow f_1^\dagger |0\rangle \sim M_1
\]
\[
M_2 \sim f_2^\dagger |0\rangle \rightarrow -f_2^\dagger |0\rangle \sim -M_2
\]

(8.5)

Extra care is needed when examining the action of $\mathcal{T}$, because this operation will change the $2\pi$ flux background into a $-2\pi$ flux background, which has two other zero modes $\tilde{f}_{1,2}$ contributed by $\psi_{1,2}$. Denote $|\bar{0}\rangle$ to be the state under a $-2\pi$ flux background with no zero mode being occupied, then $\tilde{f}_{1,2}^\dagger |\bar{0}\rangle$ is the state under the $-2\pi$ flux background with one of the zero modes being occupied, and $\tilde{f}_1^\dagger \tilde{f}_2^\dagger |\bar{0}\rangle$ is the state under the $-2\pi$ flux background with both zero modes being occupied. Then under $\mathcal{T}$

\[
M_1 \sim f_1^\dagger |0\rangle \rightarrow \tilde{f}_1^\dagger |\bar{0}\rangle
\]
\[
M_2 \sim f_2^\dagger |0\rangle \rightarrow \tilde{f}_2^\dagger |\bar{0}\rangle
\]

(8.6)

Ref. [361] (see Appendix F 4 therein) showed that under the state-operator-correspondence (8.3),

\[
M_1^\dagger \sim \tilde{f}_2^\dagger |\bar{0}\rangle, \quad M_2^\dagger \sim -\tilde{f}_1^\dagger |\bar{0}\rangle
\]

(8.7)

up to an unimportant phase factor. So the above equation means under $\mathcal{T}$

\[
M_1 \rightarrow -M_2^\dagger, \quad M_2 \rightarrow M_1^\dagger
\]

(8.8)

Now combining (8.5) and (8.8), we see the actions of $T_y$ and $\mathcal{T}$ anti-commute on the monopoles, which represents a nontrivial projective representation of the original symmetry described in (8.4). Therefore, we conclude that the monopoles in a WTI carry nontrivial quantum numbers, and a WTI is stable under interactions as long as the relevant symmetry are preserved.

As a consistency check, let us rederive the quantum numbers of the monopole by studying a symmetric STO of a WTI. As discussed in Ref. [394], a WTI allows a symmetric $Z_4$ STO, where
the anyon contents can be written as

\[ \{1, e, m, e^\alpha m^\beta\} \times \{1, \psi\} \]  

(8.9)

where \( e \) can be viewed as the elementary \( Z_4 \) charge, \( m \) can be viewed as the elementary \( Z_4 \) flux, such that \( \theta_{e,m} = i \), and \( \psi \) is a local fermion. \( \alpha \) and \( \beta \) takes integer values (mod 4).

In this \( Z_4 \) STO, \( e \) carries charge-1/2 and \( m \) is neutral. Based on the principle in Sec. 8.2.2, \( m^2 \) is the surface avatar of the bulk monopole. According to the symmetry assignments of this STO in Ref. [394], it is straightforward to check that \( T \) and \( T_y \) anti-commute on \( m^2 \), which means these two symmetries indeed anti-commute on the monopole. This is consistent with the above result obtained from the symmetric gapless surface, and it confirms that a WTI is stable under interactions as long as the symmetry in (8.4) is preserved.

8.3.2 Topological crystalline insulators protected by a mirror symmetry

Next we discuss TCIs protected by a \( U(1) \) charge conservation and a mirror symmetry \( M \) with respect to the \( x = 0 \) plane, to which class the experimentally discovered TCI SnTe belongs. We denote such a TCI by MTCI.

We will focus on the case where the surface can host two Dirac fermions, whose low energy Hamiltonian is still given by (8.2). In this case, the symmetry assignment is

\[
U(1) : \psi \rightarrow e^{i\theta} \psi \\
M : \psi(x,y) \rightarrow \sigma_x \psi(-x,y)
\]  

(8.10)

where we have assumed this surface is on the \( z = 0 \) plane. For these symmetries, \( M^2 \) can be \( \pm 1 \) on the monopole, and that \( M^2 = 1 \) corresponds to trivial quantum numbers and \( M^2 = -1 \) corresponds to nontrivial projective quantum numbers.

To determine how \( M \) acts on the monopoles, we only need to examine the action of \( M \) on \( f_{1,2}^\dagger |0\rangle \). Notice under the reflection a \( 2\pi \) flux also needs to be converted into a \( -2\pi \) flux, and it is
straightforward to check that under $\mathcal{M}$

\[
M_1 \sim \tilde{f}_1^\dagger |0\rangle \rightarrow \tilde{f}_1^\dagger |\tilde{0}\rangle \sim -M_2^\dagger \\
M_2 \sim \tilde{f}_2^\dagger |0\rangle \rightarrow \tilde{f}_2^\dagger |\tilde{0}\rangle \sim M_1^\dagger
\]  

(8.11)

Therefore, $\mathcal{M}$ squares to $-1$ on the monopole. This is a nontrivial projective representation of the symmetry (8.10), so this TCI is stable under interactions as long as the symmetry (8.10) is preserved.

As a consistency check, we can also derive the quantum numbers of the monopole from a symmetric STO of this TCI. This TCI can have a symmetric $Z_4$ STO similar to the one in (8.9) [356, 386]. In particular, in this case $e$ carries charge-$1/2$ and $m$ is neutral, too. Then based on the principle in Sec. 8.2.2, $m^2$ is the surface avatar of the bulk monopole. As shown in Ref. [356, 386] (see also Appendix H.2), indeed, $\mathcal{M}^2 = -1$ on $m^2$. This is consistent with the above result obtained by studying the symmetric gapless surface.

This TCI is compatible with a further time reversal symmetry, and we discuss two time reversal symmetric generalizations of this state in Appendix H.1. These two states are denoted as TMTCI-1 and TMTCI-2.

### 8.3.3 Doubled topological insulators protected by a mirror symmetry

As our third example, we consider two copies of topological insulators further equipped with a mirror symmetry (DTI). It is well known that a single topological insulator is protected by $U(1)$ charge conservation and time reversal, but two copies of them are smoothly connected to a trivial insulator if there is no other symmetry. [4, 5, 362] However, with a further mirror symmetry $\mathcal{M}$, two copies of topological insulators can still be nontrivial at the free-fermion level. Such an insulator still has a gapless surface described by (8.2), and the symmetries are assigned as

\[
U(1) : \psi \rightarrow e^{i\theta} \psi \\
\mathcal{T} : \psi \rightarrow i\sigma_y \psi \\
\mathcal{M} : \psi(x, y) \rightarrow \sigma_z \tau_y \psi(-x, y)
\]  

(8.12)
It is straightforward to check that no fermion bilinear term can be written down to fully gap out (8.2) while preserving all symmetries.

To understand whether this state is stable under interactions, let us derive the quantum numbers of the monopole from the above surface state. For the above symmetries, the nontrivial projective quantum numbers on monopoles correspond to \( \mathcal{M}^2 = -1 \) or \( Z^2 \equiv (\mathcal{T}\mathcal{M})^2 = -1 \).

Because the \( \mathcal{T} \) action in this state is the same as in a WTI, under \( \mathcal{T} \) we again have (8.8). It is straightforward to check that under \( \mathcal{M} \)

\[
M_1 \sim f_1^\dagger |0\rangle \rightarrow -i f_2^\dagger |\bar{0}\rangle \sim -i M_1^\dagger \\
M_2 \sim f_2^\dagger |0\rangle \rightarrow i f_1^\dagger |\bar{0}\rangle \sim -i M_2^\dagger
\]

(8.13)

So under \( Z \equiv \mathcal{T}\mathcal{M} \), the product of \( \mathcal{T} \) and \( \mathcal{M} \),

\[
M_1 \rightarrow i M_2, \quad M_2 \rightarrow -i M_1
\]

(8.14)

Therefore, \( \mathcal{M} \) squares to 1 and \( Z \equiv \mathcal{T}\mathcal{M} \) squares to \(-1\) on the monopole. So the monopole indeed carries nontrivial quantum numbers, and thus this insulator is stable under interactions.

We can also derive these quantum numbers of the monopoles by examining the symmetric STO of this TCI. As shown in Appendix H.2, this TCI can have a symmetric \( Z_4 \) STO similar to the ones discussed earlier, and we find consistent results for the quantum numbers of monopoles with the above.

Notice both \( \mathcal{T} \) and \( \mathcal{M} \) are important for the symmetry protection of this TCI, and as long as one of them is broken, a symmetric fermion bilinear term can be written down to fully gap the surface. In fact, in Appendix H.2 we argue that a DTI can be obtained from a WTI by suitably breaking the protecting translation symmetry of the latter. In Appendix H.1, we will discuss TMTCI-2, another example which can also be viewed as two copies of topological insulators protected by a further mirror symmetry. However, in TMTCI-2 the mirror symmetry alone is sufficient for the symmetry protection. In fact, that state is just a time reversal symmetric version of the state MTCI described in Sec. 8.3.2.
8.3.4 Topological crystalline insulators with symmetry-enforced-gapless surfaces

As our final examples, we present two TCIs that exhibit symmetry-enforced-gapless surfaces. Both TCIs have a symmetric gapless surface state described by (8.2) at low energies, and they differ by the symmetries.

The first such TCI, SEG-1, has a $U(1)$ charge conservation, a mirror symmetry $M$ and an $SU(2)$ symmetry. These symmetries are assigned as

\begin{align}
U(1) : \psi &\rightarrow e^{i\theta} \psi \\
M : \psi(x, y) &\rightarrow \sigma_z \psi(-x, y) \\
SU(2) : \psi &\rightarrow U_\tau \psi
\end{align}

where $U_\tau$ means an $SU(2)$ matrix generated by $\tau$’s. For these symmetries, there are three types of nontrivial projective quantum numbers on monopoles: having $M^2 = -1$ and/or having spin-$1/2$.

It is easy to see this is just an $SU(2)$ symmetric version of MTCI discussed in Sec. 8.3.2, so $M$ squares to $-1$ on the monopole. Furthermore, it is straightforward to check that the monopole also transforms as a spin-$1/2$ under the $SU(2)$ symmetry. Therefore, the monopole is nontrivial, and this TCI is stable under interactions.

Now we show by contradiction that this TCI does not allow a symmetric STO. Suppose this TCI can have a symmetric STO, then the anyons of this STO have definite spins under the $SU(2)$ symmetry. Because the local fermion carries spin-$1/2$ under the $SU(2)$, it is always possible to make the anyon a spin-singlet by attaching certain local fermions to it. Then the STO can be written as

\begin{align}
\{1, a, \cdots\} \times \{1, \psi\}
\end{align}

where the sector $\{1, a, \cdots\}$ includes all topologically nontrivial spinless anyons, and the sector $\{1, \psi\}$ includes the spin-$1/2$ local fermion. The first sector is already closed under fusion and
braiding due to the $SU(2)$ symmetry. Moreover, these two sectors will not be mixed under the symmetry actions, because the symmetries cannot change a spin-singlet into a spin-1/2, or vice versa. Therefore, the sector $\{1, a, \cdots \}$ alone must be able to emerge from a system made of local spin-singlets. In this system, all local spin-singlets must be bosons that carry even $U(1)$ charges. Notice the symmetries of these bosons include a $U(1)$ charge conservation and a mirror symmetry $\mathcal{M}$.

Now consider the possible properties of the elementary monopoles of a system made of these bosons. Because the charges of such bosons are even, the elementary monopoles of these bosons only emit $\pi$ flux. That is to say, two of such monopoles should be the $2\pi$ monopole discussed earlier, which carries spin-1/2 and $\mathcal{M}^2 = -1$. Based on the understanding of a bosonic TCI in general, these “half monopoles” can be either a boson or a fermion, either a spin-singlet or a spin-1/2, and $\mathcal{M}$ can square to either 1 or $-1$. However, in none of these cases the $2\pi$ monopole will be a spin-1/2 boson with $\mathcal{M}^2 = -1$. This contradiction shows that this TCI cannot have a symmetric STO, and its surface is enforced to be gapless as long as the symmetries $(8.15)$ are preserved. This phenomenon of the surface is dubbed “symmetry enforced gaplessness”, $[339]$ and our example provides a crystalline-symmetric realization of it.

Now we turn to SEG-2, the second TCI that has symmetry-enforced-gapless surface. This TCI has a $U(1)$ charge conservation, an $SU(2)$ symmetry and a symmetry $\mathcal{P}$ that can be viewed as a combination of a mirror symmetry and a unitary charge conjugation. The low energy Hamiltonian of the surface is still described by $(8.2)$, while the symmetries are assigned as

$$
U(1) : \psi \rightarrow e^{i\theta} \psi \\
SU(2) : \psi \rightarrow U_\tau \psi \\
\mathcal{P} : \psi(x, y) \rightarrow \sigma_z \tau_y \psi^*(-x, y)
$$

(8.17)

It is easy to check that no fermion bilinear term can be written down to fully gap $(8.2)$ without breaking these symmetries, so this is a nontrivial TCI at the level of free fermions. In addition, it is straightforward to see that the monopole of this TCI carries nontrivial projective quantum number, \textit{i.e.} spin-1/2 under the $SU(2)$ symmetry, so this nontrivial TCI is stable under interac-
tions. Furthermore, a similar argument as above shows that this TCI also allows no symmetric STO and its surface exhibits symmetry enforced gaplessness.

8.4 Relation to symmetry enriched $U(1)$ quantum spin liquids

After understanding the properties of the monopoles of these TCIs in details, in this section we discuss the relations between these TCIs and symmetry enriched $U(1)$ QSLs. To this end, let us first briefly review the physics of symmetry enriched $U(1)$ QSLs. [361]

A three dimensional $U(1)$ QSL is a three dimensional spin system that has emergent gapless photons at low energies. In a condensed matter system, the appearance of an emergent gapless photon is necessarily associated with the existence of emergent electric charges and magnetic monopoles (and their bound states, dyons). These fractional excitations can in principle also be gapless, but we assume only the photons are gapless for simplicity. The existence of fractional excitations implies long-range entanglement in the ground state, and this nontrivial nature is independent of any symmetry.

In the absence of any symmetry, all $U(1)$ QSLs can be smoothly connected to each other. However, when there are global symmetries in $U(1)$ QSLs, there can be different symmetry enriched $U(1)$ QSLs due to symmetry protected distinctions. In many cases, their differences are reflected in their different spectra. Namely, in different symmetry enriched $U(1)$ QSLs, their electric charges and magnetic monopoles can have different statistics and/or quantum numbers.

It turns out very helpful to view $U(1)$ QSLs as gauged insulators, i.e. an insulator coupled with a dynamical $U(1)$ gauge field. The properties of the insulator will then determine the statistics and quantum numbers of both the electric charge and the magnetic monopole, thus determine the property of the symmetric $U(1)$ QSL.

Therefore, we will consider coupling the TCIs discussed above to a dynamical $U(1)$ gauge field. This gauging procedure turns the TCIs into a $U(1)$ QSLs with some global symmetries. Below we will discuss our examples in turn.

Let us start with a WTI. After gauging it becomes a $U(1)$ QSL with time reversal $\mathcal{T}$ and translation $T_y$. In this $U(1)$ QSL, the electric charge is a fermion. This fermionic charge is a Kramers doublet under the original time reversal $\mathcal{T}$, and also a “Kramers doublet” under a new
anti-unitary symmetry $T'$, which is generated by the product of the generators of $\mathcal{T}$ and $T_y$. As discussed earlier, the magnetic monopole is a boson, and $\mathcal{T}$ and $T_y$ anti-commute on the monopole. For these reasons, we denote this $U(1)$ QSL by $E_{fTT'}M_{b-}$.

Next we turn to MTCI discussed in Sec. 8.3.2. After gauging, the symmetry of the resulting $U(1)$ QSL is just a mirror symmetry $\mathcal{M}$. For later purposes, it is more convenient to twist our notations here. That is, we will regard the fermion of the TCI as the magnetic monopole of the resulting $U(1)$ QSL, which then identifies the monopole of the TCI discussed in Sec. 8.3.2 as the electric charge of this $U(1)$ QSL. So the magnetic monopole is a fermion, and the electric charge is a boson that has $\mathcal{M}^2 = -1$. For this reason, we denote this $U(1)$ QSL by $E_{bM}M_f$.

Now we turn to the DTI discussed in Sec. 8.3.3. After gauging the symmetry of the resulting $U(1)$ QSL is $T \times \mathcal{M}$, where both $T$ and $\mathcal{M}$ square to 1 for all local excitations. The electric charge of the $U(1)$ QSL will be taken as the fermion that is a Kramers doublet under $T$. It is also easy to see $Z^2 \equiv (T\mathcal{M})^2 = -1$ for the charge. The magnetic monopole has $Z^2 = -1$. For this reason, we denote this $U(1)$ QSL by $(E_{fTZ}M_{bZ})_-$, with the parenthesis and minus sign implying that the system has both a time reversal and a mirror symmetry, and the mirror symmetry action preserves the electric charge.

Finally we consider the two TCIs that exhibit symmetry-enforced-gapless surfaces discussed in Sec. 8.3.4. For SEG-1, after gauging the symmetry becomes $SO(3) \times \mathcal{M}$, which means all local excitations carry integer spins and $\mathcal{M}^2 = 1$. In this case, it is again more convenient to identify the fermion of the TCI as the magnetic monopole of the resulting $U(1)$ QSL, and this fermion is a spin-1/2. This lets us identify the bosonic magnetic monopole of the TCI as the electric charge of the $U(1)$ QSL, which is a spin-1/2 boson that has $\mathcal{M}^2 = -1$. For this reason, we denote it by $E_{bM}^{1/2}M_{f1}^{1}$. As for SEG-2, the symmetry of the resulting $U(1)$ QSL is $SO(3) \times \mathcal{P}$, which means all local excitations have integer spins and $\mathcal{P}^2 = 1$. Notice after gauging $\mathcal{P}$ should be interpreted as a mirror symmetry in terms of the local excitations, so from now on we will denote $\mathcal{P}$ as $\mathcal{M}$ for notational consistency. We will identify the electric charge of the resulting $U(1)$ QSL to be the fermion of the TCI, which carries spin-1/2 and has $\mathcal{M}^2 = -1$. Then the magnetic monopole of the $U(1)$ QSL is identified as the magnetic monopole of the TCI, which is a spin-1/2 boson. For this reason, we denote this $U(1)$ QSL by $E_{fM}^{1/2}M_{b1}^{1}$.
There is actually another level of relation between the TCIs and the symmetry enriched $U(1)$ QSLs: the $U(1)$ QSLs enriched by crystalline symmetries have their analogs with internal symmetries discussed in Ref. [327, 361]. Such a correspondence was first noted in Ref. [74] and later elaborated in Ref. [384] as the “crystalline equivalence principle”. Below we illustrate this correspondence.\footnote{The $U(1)$ QSLs with internal symmetries can also be viewed as gauged insulators, and there is also a correspondence between the TCIs studied in this paper and those insulators. However, in this paper we focus on the correspondence between the $U(1)$ QSLs with crystalline symmetries and those with internal symmetries.}

Let us start with the WTI. Notice the translation symmetry $T_y$ of a WTI acts as an internal $Z_2$ symmetry at low energies, so we can ask if there is a $T \times Z_2$ symmetric $U(1)$ QSL that has analogous properties as $E^{fT}M_{b-}$, the gauged WTI. As shown in Ref. [361], there is indeed a $T \times Z_2$ symmetric $U(1)$ QSL also dubbed $E^{fT}M_{b-}$. The properties of these two states are listed in Table 8.1, and the correspondence is clear.

For other crystalline-symmetric $U(1)$ QSLs that involve a mirror symmetry $M$, if we replace the mirror symmetry $M$ by a time reversal symmetry $T'$, we can find their corresponding states in Ref. [361] as well. However, we need to pay extra care when carrying out this procedure: if the charge (monopole) of the $U(1)$ QSL is a fermion and $M$ acts on the charge as a mirror reflection combined with a unitary charge conjugation, then on the charge (monopole) $M^2 = \pm 1$ corresponds to $T'^2 = \mp 1$. [74]

Now we demonstrate this more explicitly. The $M$ symmetric $E_{bM}M_f$ corresponds to the $T$ symmetric $E_{bT}M_f$, where the electric charge has $T^2 = -1$ for the electric charge (see Table 8.2). The $T \times M$ symmetric $(E_{fTZ}M_{bZ})_-$, $(E_{fZ}M_{bMZ})_-$ and $(E_{fTM_{bM}})_-$ correspond to the $T \times T'$ symmetric $(E_{fTZ}M_{bZ})_-$, $(E_{fZ}M_{bTZ})_-$ and $(E_{fTM_{bT}})_-$, respectively. The symmetries in the three latter states can be more conveniently phrased as a time reversal $T$ and a unitary $Z_2$ charge conjugation $C$, which is generated by the product of the generators of $T$ and $T'$. The properties of these states are listed in Table 8.3, and the correspondence is clear. The $SO(3) \times M$ symmetric $E_{bM_{1/2}^\uparrow}M_{1/2}^\downarrow$ and $E_{fM_{1/2}^\uparrow}M_{1/2}^\downarrow$ correspond to the $SO(3) \times T$ symmetric $E_{bT_{1/2}^\uparrow}M_{1/2}^\downarrow$ and $E_{fT_{1/2}^\uparrow}M_{1/2}^\downarrow$, respectively. The properties of these states are summarized in Table 8.4, and one can again see the clear correspondence. Ref. [361] showed that the latter two states possess fractional
Table 8.1: Upper: properties of the $T \times T_y$ symmetric $E_{fTT} M_{b-}$. $T_E^2$ stands for the value of $T^2$ on the electric charge, $T_{T_y}^2$ stands for the value of $T^{T_y^2} \equiv (TT_y)^2$ on the charge, and $[T, T_y]_M = +$ and $[T, T_y]_M = -$ mean that $T$ and $T_y$ commute or anti-commute on the monopole, respectively. Lower: properties of the $T \times Z_2$ symmetric $E_{fTT} M_{b-}$. $T_E^2$ stands for the value of $T^2$ on the electric charge, $T_{Z_2}^2 \equiv (T Z_2)^2$ on the charge, and $[T, Z_2]_M = +$ and $[T, Z_2]_M = -$ mean that $T$ and $Z_2$ commute or anti-commute on the monopole, respectively.

<table>
<thead>
<tr>
<th>$T \times T_y$</th>
<th>$T_E^2$</th>
<th>$T_{T_y}^2$</th>
<th>$[T, T_y]_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{fTT} M_{b-}$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T \times Z_2$</th>
<th>$T_E^2$</th>
<th>$T_{Z_2}^2$</th>
<th>$[T, Z_2]_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{fTT} M_{b-}$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-$</td>
</tr>
</tbody>
</table>

Table 8.2: Upper: properties of the $M$ symmetric $E_{bM} M_f$, where $M_E^2$ stands for the value of $M^2$ on the electric charge. Lower: properties of the $T$ $E_{bT} M_f$, where $T_E^2$ stands for the value of $T^2$ on the electric charge.

<table>
<thead>
<tr>
<th>$M$</th>
<th>$M_E^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bM} M_f$</td>
<td>$-1$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$T$</th>
<th>$T_E^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bT} M_f$</td>
<td>$-1$</td>
</tr>
</tbody>
</table>

topological responses to an external $SO(3)$ gauge field, so they are dubbed “fractional topological paramagnets”. $E_{bM_{1/2}} M_{f_{1/2}}$ and $E_{fM_{1/2}} M_{b_{1/2}}$ are the crystalline-symmetric realizations of fractional topological paramagnets.

The above relation may provide simpler ways to study difficult problems on one side of the correspondence, by translating it to a possibly simpler problem on the other side. For example, if the above correspondence is generally true, we expect the classification and characterization of some crystalline symmetric $U(1)$ QSLs can be directly read off from Ref. [361] (and Ref. [327]). More specifically, the existence of 7 different $T$ symmetric $U(1)$ QSLs leads to the expectation of 7 different $M$ symmetric $U(1)$ QSLs, the existence of 15 different $SO(3) \times T$ symmetric $U(1)$ QSLs leads to the expectation of 15 different $SO(3) \times M$ symmetric $U(1)$ QSLs, and the existence of 38 different $T \times T'$ symmetric $U(1)$ QSLs leads to the expectation of 38 different $T \times M$ symmetric $U(1)$ QSLs. Similar expectations hold for other symmetries. The physical characterizations of the crystalline symmetric $U(1)$ QSLs can also be read off from their analogs with inter-
Table 8.3: Upper: properties of the $\mathcal{T} \times \mathcal{M}$ symmetric $(E_{fTZ} M_{bZ})_-$, $(E_{fZ} M_{bMZ})_-$ and $(E_{fT} M_{bM})_-$. $\mathcal{T}_E^2$ stands for the value of $\mathcal{T}^2$ on the electric charge, $\mathcal{M}_E^2$ stands for the value of $\mathcal{M}^2$ on the magnetic monopole, and $Z_E^2$ and $Z_M^2$ stand for the value of $Z^2 \equiv (\mathcal{T} \mathcal{M})^2$ on the charge and monopole, respectively. Lower: properties of the $\mathcal{T} \times \mathcal{T}'$ symmetric $(E_{fTZ} M_{bZ})_-$, $(E_{fZ} M_{bT'} Z)_{-}$ and $(E_{fT} M_{bT'})_-$. $\mathcal{T}'_E^2$ stands for the value of $\mathcal{T}'^2$ on the electric charge, $\mathcal{M}'_E^2$ stands for the value of $\mathcal{T}'^2$ on the magnetic monopole, and $C_E^2$ and $C_M^2$ stand for the value of $C^2$ on the charge and the monopole, respectively.

<table>
<thead>
<tr>
<th>$\mathcal{T} \times \mathcal{M}$</th>
<th>$\mathcal{T}_E^2$</th>
<th>$Z_E^2$</th>
<th>$\mathcal{M}_E^2$</th>
<th>$Z_M^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(E_{fTZ} M_{bZ})_-$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$(E_{fZ} M_{bMZ})_-$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$(E_{fT} M_{bM})_-$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\mathcal{T} \times \mathcal{T}'$</th>
<th>$\mathcal{T}'_E^2$</th>
<th>$C_E^2$</th>
<th>$\mathcal{T}'_M^2$</th>
<th>$C_M^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(E_{fTZ} M_{bZ})_-$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$(E_{fZ} M_{bT'} Z)_-$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$(E_{fT} M_{bT'})_-$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 8.4: Upper: properties of the $SO(3) \times \mathcal{M}$ symmetric $U(1)$ QSLs $E_{bM\frac{1}{2}} M_{f\frac{1}{2}}$ and $E_{fM\frac{1}{2}} M_{b\frac{1}{2}}$. $\mathcal{M}_E^2$ stands for the value of $\mathcal{M}^2$ on the electric charge, and $S_E$ and $S_M$ stand for the spin of the charge and the monopole, respectively. Lower: properties of the $SO(3) \times \mathcal{T}$ symmetric $U(1)$ QSLs $E_{bT\frac{1}{2}} M_{f\frac{1}{2}}$ and $E_{f\frac{1}{2}} M_{b\frac{1}{2}}$. $\mathcal{T}_E^2$ stands for the value of $\mathcal{T}^2$ on the electric charge, and $S_E$ and $S_M$ stand for the spin of the charge and the monopole, respectively.

<table>
<thead>
<tr>
<th>$SO(3) \times \mathcal{M}$</th>
<th>$\mathcal{M}_E^2$</th>
<th>$S_E$</th>
<th>$S_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bM\frac{1}{2}} M_{f\frac{1}{2}}$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{fM\frac{1}{2}} M_{b\frac{1}{2}}$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$SO(3) \times \mathcal{T}$</th>
<th>$\mathcal{T}_E^2$</th>
<th>$S_E$</th>
<th>$S_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bT\frac{1}{2}} M_{f\frac{1}{2}}$</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>$E_{f\frac{1}{2}} M_{b\frac{1}{2}}$</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
</tr>
</tbody>
</table>

By directly studying the bulk properties of several band-theory-based nontrivial TCIs, we show all of them are stable under interactions, because they all have nontrivial bulk monopoles once they are coupled to a $U(1)$ gauge field. The properties of the monopoles are derived based on two complementary methods, which involve studying different types of surface states of the TCIs.

8.5 Discussion

By directly studying the bulk properties of several band-theory-based nontrivial TCIs, we show all of them are stable under interactions, because they all have nontrivial bulk monopoles once they are coupled to a $U(1)$ gauge field. The properties of the monopoles are derived based on two complementary methods, which involve studying different types of surface states of the TCIs.
That these methods give consistent answers is on the one hand assuring, and on the other hand implying that in many cases it is a more direct and unified way of to characterize nontrivial TCIs in terms of their bulk properties. However, we note that sometimes the boundary can carry subtle information of a topological phase that is not directly visible from the bulk. [9]

All the TCIs we have discussed can be realized by free fermions, and two of them have symmetry-enforced-gapless surfaces. That is, even under strong interactions, their surface states must be gapless as long as the relevant symmetries are preserved. It is intriguing to find a TCI that on the one hand cannot be realized by free fermions, and on the other hand have symmetry-enforced-gapless surfaces. It will be an interesting challenge to study the properties of the monopoles of such TCIs.

The TCIs that we discuss are not only of conceptual importance, but are also of experimental significance. We expect they will either be synthesized in the near future, or have already been synthesized but remain to be further investigated.

We have discussed the relations between the TCIs and symmetry enriched $U(1)$ QSLs at two levels. First, when the TCIs are coupled to a dynamical $U(1)$ gauge field, they become $U(1)$ QSLs enriched by crystalline symmetries. We discuss the properties of these $U(1)$ QSLs in details. In particular, the two TCIs with symmetry-enforced-gapless surfaces are shown to be related to crystalline-symmetric realizations of “fractional topological paramagnets” introduced in Ref. [361]. Second, we demonstrate an interesting correspondence between the $U(1)$ QSLs with crystalline symmetries and $U(1)$ QSLs with internal symmetries discussed in Ref. [361]. This is a manifestation of the crystalline equivalence principle first noted in Ref. [74] and further elaborated in Ref. [384]. Such a correspondence may potentially provide a simpler way of solving a difficult problem, by relating a problem with internal symmetry to a problem with crystalline symmetry, or vice versa.

In this paper we focus on TCIs with a translation symmetry or a mirror symmetry, and it will be interesting to obtain physical bulk characterizations for TCIs with other types of crystalline symmetries.

Besides the magnetic monopoles, there is another type of bulk characterizations that is perhaps more ubiquitous to TCIs: properties of defects of the crystalline symmetries, such as dis-
locations and disclinations. There has already been some work in this direction, and intriguing results have been found in many cases. [384, 394–398] It will be interesting and important to obtain systematic physical characterizations of these defects, as well as their interplay with other defects such as monopoles. We leave this for future work.
In the previous chapters I have presented my work in quantum many-body physics during the past few years. In these works, various topics were covered, different flavors of ideas were applied, and distinct types of techniques were employed. There are still many important open problems in the directions that I have worked on, as detailed at the end of each chapter, and I would like to continue these lines of research in a productive manner in the future.

Finally, I would like to finish this thesis by commenting on some points that I wish I could have done better. Of course, there are infinite interesting problems to study, and it is impossible to list all of them. The following is just an incomplete list of such points which I think are the most prominent, and some of them may be good future research directions.

First of all, I have mainly worked on quantum many-body physics in or close to thermal equi-
librium, and I did not touch the fascinating subject of real-time dynamics of quantum systems far from equilibrium. During the past decade or so, there has been great theoretical, numerical and experimental progress in understanding thermalization, quantum quench dynamics, Floquet dynamics, entanglement growth and operator spreading. In spite of this progress, the universal aspects of real-time dynamics of quantum many-body systems are still rather mysterious, and there are opportunities to contribute in these realms.

Second, within physics in and close to thermal equilibrium, there are some classes of interesting quantum phases that I tried to understand but failed to make any real progress. These include three-dimensional topological orders and fractonic phases. Despite the recent progress in these problems, rich physics of these systems still awaits to be uncovered.

Third, although I have explored many aspects of many types of quantum phase transitions, including constructing interesting critical theories and analyzing dynamical properties of some phase transitions, I was not able to answer some long-standing theoretical questions in this field. Such questions include (but are not limited to):

1. A critical theory between a Fermi liquid with constant Fermi surface volume and an insulator with no (electric or neutral) Fermi surface. Various attempts have been tried, but no fruitful result has been produced so far.

2. IR dynamics of some commonly encountered critical theories, notably the dynamics of some strongly coupled gauge theories.

3. Effects of disorder at various phase transitions, where the most notable example may be the quantum Hall plateau transitions.

Last but not the least, most of my work is analytic, and I did not develop very strong numerical skills. This skill is very important for testing conjectures and providing insights, and it may be good to develop it.
A.1 Review of the conventional regimes

In this appendix, we review the conductivities in the conventional regimes predicted by Chambers’ formula in some detail:

\[
\sigma_{ij} = \frac{e^2}{2\pi^2h} \int \frac{dk}{|v(k)|} v_j(k) \cdot \int_0^\infty dt v_i(q(t)) e^{-\int_0^t \frac{dt'}{\tau(q(t'))}}
\]  

(A.1)

As in the main text, we will focus on the two dimensional case. To define the weak and strong field regimes, we introduce local cyclotron frequency \( \omega(q) \equiv eB/m(q) \), where \( q \) lies on the Fermi surface and \( m(q) \) is the effective mass at that point. For generality, we also introduce \( \tau(q) \).
which is the local relaxation time of electrons at point $q$.

We will see that in the strong field regime, the Hall coefficient is indeed related to the carrier type and carrier density. In the weak field regime, although they are not related in general, in the special case of isotropic Fermi surfaces, the Hall coefficient happens to be related to the carrier type and carrier density in the same way as in the strong field regime.

### A.1.1 Weak field regime

We start by reviewing the weak field regime, where $\omega(q)\tau(q) \ll 1$ for all $q$ on the Fermi surface. In this regime, (A.1) can be expanded in powers of $B$:

$$\begin{align*}
v_i(q(t)) &= v_i(k) - \nabla_kv_i(k) \cdot \frac{ev(k) \times B}{\hbar} t + \mathcal{O}(B^2) \\
\tau(q(t)) &= \tau(k) - \nabla_k\tau(k) \cdot \frac{ev(k) \times B}{\hbar} t + \mathcal{O}(B^2)
\end{align*}$$

which yields

$$\int_0^\infty dt v_i(q(t)) e^{-\frac{t}{\tau(q(t))}}$$

$$= v_i(k)\tau(k) - \tau(k) \frac{ev \times B}{\hbar} \nabla_k(v_i(k)\tau(k)) + \mathcal{O}(B^2)$$

So the weak-field limit of (A.1) is

$$\sigma_{ij} = \frac{e^2}{2\pi^2\hbar} \int dq \frac{v_j(q)v_i(q)\tau(q)}{|v(q)|}$$

$$- \frac{e^3 B}{2\pi^2\hbar} \int dq l_j(q)\hat{t} \cdot \nabla q l_i(q)$$

where $\hat{t}$ is the tangential direction of the Fermi surface at $q$ and $l_i(q) = v_i(q)\tau(q)$.

If $i = j$, the second term vanishes for a closed Fermi surface, so we get

$$\sigma_{xx} = \frac{e^2}{2\pi^2\hbar} \int dq \frac{v_x^2(q)\tau(q)}{|v(q)|} + \mathcal{O}(B^2)$$

With time reversal symmetry, the first term vanishes if $i \neq j$, and we recover Ong’s formula for
Hall conductivity

\[ \sigma_{xy} = -\frac{e^2 B}{2\pi^2 h} \int dq l_y(q) \hat{l} \cdot \nabla q l_x(q) \]
\[ = \frac{e^3}{2\pi^2 h} B \cdot \oint dl \times l/2 = 2 \frac{e^2}{h} \phi \]

(A.6)

where \( \phi = B \cdot \oint dl \times l/2 \) and \( \phi_0 = h/e \).

In general in this regime the Hall number is not related to the carrier type and carrier density, unless the Fermi surface happens to be isotropic. To see this, notice for circular isotropic Fermi surfaces, we have

\[ \sigma_{xx} = \frac{e^2 k_F v_F \tau}{2\pi h} \]

(A.7)

where \( k_F \) and \( v_F \) are the Fermi momentum and Fermi velocity, respectively, and \( \tau \) is the scattering rate on the Fermi surface. Assuming a quadratic band with effective mass \( m \), we have \( k_F^2 = 2\pi n \) and \( \hbar k_F = mv_F \), where \( n \) is the carrier density. Then we get

\[ \sigma_{xx} = \frac{ne^2 \tau}{m} \]

(A.8)

which is Drude’s formula for conductivity for \( \sigma_{xx} \).

As for \( \sigma_{xy} \), in this case \( \phi = -BA \frac{k_F^2 \tau^2}{m^2} \), where \( A \) is the area of the Fermi surface, which is related to the carrier density by Luttinger’s theorem: \( A = 2\pi^2 n \). Plugging these in, we find

\[ \sigma_{xy} = -\frac{e^3}{m^2} Bn \]

(A.9)

So the Hall number is

\[ R_H = \frac{\rho_{yx}}{B} = \frac{1}{B} \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2} \approx \frac{1}{B} \frac{\sigma_{xy}}{\sigma_{xx}^2} \approx -\frac{1}{ne} \]

(A.10)

Therefore, in the case of isotropic Fermi surface the Hall number can still indicate the carrier type and carrier density even in the weak field regime.
A.1.2 Strong field regime

Now we turn to the strong field regime, where \( \omega(q) \tau(q) \gg 1 \) for all \( q \) on the Fermi surface. To study the conductivities in the strong field regime, we first write (A.1) in another form by using that \( \frac{h \, dq}{dt} = -e \mathbf{v} \times \mathbf{B} \):

\[
\sigma_{ij} = \frac{e^2}{2\pi^2 \hbar} \int \frac{dk}{|v(k)|} v_j(k) \cdot \int \frac{hdq}{eB|v(q)|} v_i(q)e^{-\int_{0}^{t} \frac{dt'}{\tau(q(t'))}}
\]

(A.11)

The second integral over momentum goes over the Fermi surface repeatedly. Therefore, we can limit the second integral to be the first complete cyclotron motion and represent the following cyclotron motion as a geometric series:

\[
\sigma_{ij} = \frac{e}{2\pi^2 B} \int \frac{dk}{|v(k)|} v_j(k) \cdot \int \frac{dq}{|v(q)|} v_i(q)e^{-\int_{0}^{t} \frac{dt'}{\tau(q(t'))}} \cdot \left[ 1 - e^{-\frac{T}{\bar{\tau}}} \right]
\]

(A.12)

where \( T \) is the cyclotron period and \( \bar{\tau} \) is defined such that \( T/\bar{\tau} = \int_0^T dt'/\tau(q(t')) \).

In the strong field regime, the electrons travel so fast that \( T \sim 1/B \ll \bar{\tau} \), we can expand the above in terms of \( 1/B \):

\[
\sigma_{ij} \approx \frac{e}{2\pi^2 B} \int \frac{dk}{|v(k)|} v_j(k) \cdot \int \frac{dq}{|v(q)|} v_i(q) \left( 1 - \int_0^{\tau(q)} \frac{dt'}{\tau(q(t'))} \right) \frac{\bar{\tau}}{T}
\]

(A.13)

The integral involving the first term vanishes for closed Fermi surfaces, because the result of the integral over \( q \) will be independent of \( k \), and \( \int dk_l \) vanishes for a closed Fermi surface. For the
second term, we integrate by parts

\[- \int \epsilon_{im} dq_m \int_0^{t(q)} \frac{dt'}{\tau(q(t))} \frac{\bar{\tau}}{T} = -\epsilon_{im} k_m + \int_0^T \frac{dt}{\tau(q(t))} \frac{\epsilon_{im} q_m \bar{\tau}}{T} \]

(A.14)

Again, the second term does not depend on \(k\) so it does not contribute after the integration over \(k\). So we get

\[\sigma_{ij} = -\frac{e}{2 \pi^2 B} \epsilon_{ij} \epsilon_{im} \int dk_l k_m = -\frac{e}{2 \pi^2 B} \epsilon_{ij} A \]

(A.15)

This result tells us that for a closed Fermi surface in the strong field regime, the leading nonzero contribution of the off-diagonal elements of the conductivity tensor is at the order \(1/B\), but leading contribution to the diagonal elements will be at order \(1/B^2\) or higher.

In this regime, the Hall number is given by

\[R_H = \rho_{yx} B \approx \frac{1}{B \sigma_{xy}} = -\frac{1}{ne} \]

(A.16)

where Luttinger’s theorem, \(A = 2 \pi^2 n\), is applied in the last step. So in the strong field regime the Hall number is always related to the carrier type and carrier density, as long as the Fermi surface is closed.

A.2 CORRECTIONS TO THE CONDUCTIVITIES AT THE ORDER OF \((\omega_l \tau)^{-1}\) AND \(\omega_h \tau\)

In the main text the conductivities of the specific model we consider have been calculated to the leading order in \((\omega_l \tau)^{-1}\) and \(\omega_h \tau\). In this appendix we give the leading corrections to those results.

The leading corrections come from two sources. First, it comes from the higher order terms in the expansion of the exponential, and this will give a correction at the order of \((\omega_l \tau)^{-1}\). Second, it comes from the leakage among fast segments and slow segments, and this will give a correction at the order of \(\omega_h \tau\).
To this order, we get conductivities

\[ \sigma_{xx} = \frac{2e}{B} \left[ \frac{k_l^2}{2\pi^2} + \frac{k_l^2}{\pi^2} \left( \frac{\pi}{4} - 1 \right) \frac{1}{\omega_l \tau} \right. \\
\left. + \left( \frac{k_l k_h}{\pi^2} + \frac{k_h^2}{2\pi} \right) \omega_h \tau \right] \]  

(A.17)

\[ \sigma_{xy} = \frac{2e}{B} \left[ \frac{k_l^2}{2\pi^2} \left( \frac{\pi}{2} - 1 \right) - \frac{k_l^2}{2\pi^2} \frac{4 - \pi}{2} \frac{1}{\omega_l \tau} \\
\left. + \frac{k_l k_h}{\pi^2} \omega_h \tau \right] \]  

(A.18)

which yields resistivities

\[ \rho_{xx} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2} \]
\[ = \frac{B}{2ek_l^2} \left( C_1 + C_2 \frac{1}{\omega_l \tau} + C_3 \omega_h \tau \right) \]  

(A.19)

\[ \rho_{yx} = \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2} \]
\[ = \frac{B}{2ek_l^2} \left( D_1 + D_2 \frac{1}{\omega_l \tau} + D_3 \omega_h \tau \right) \]

with

\[ C_1 = \frac{8\pi^2}{(\pi - 2)^2 + 4} \]
\[ C_2 = \frac{4(4 - \pi)\pi^2(\pi^2 - 8)}{((\pi - 2)^2 + 4)^2} \]  

(A.20)
\[ C_3 = \frac{8\pi^2 \left( (\pi - 4)\pi^2 \alpha^2 + 2(\pi(\pi - 8) + 8)\alpha \right)}{((\pi - 2)^2 + 4)^2} \]

and

\[ D_1 = \frac{4(\pi - 2)\pi^2}{(\pi - 2)^2 + 4} \]
\[ D_2 = \frac{4(4 - \pi)\pi^2(\pi^2 - 8)}{((\pi - 2)^2 + 4)^2} \]  

(A.21)
\[ D_3 = \frac{-16\pi^2(2\pi(\pi - 2)\alpha^2 + (\pi^2 - 8)\alpha)}{((\pi - 2)^2 + 4)^2} \]
where $\alpha = k_h/k_l$. 

Appendices of chapter 2

B.1 LATTICE AND SYMMETRIES

In this appendix, we document some details on the conventions and the symmetry transformations.

Consider a monolayer of graphene. We let the primitive lattice vectors $\mathbf{A}$ and reciprocal lattice vectors $\mathbf{B}$ be

$$
\mathbf{A}_1 = a\hat{x}, \quad \mathbf{A}_2 = a\left(-\frac{1}{2}\hat{x} + \frac{\sqrt{3}}{2}\hat{y}\right); \quad \mathbf{B}_1 = \frac{4\pi}{\sqrt{3}a}\left(\frac{\sqrt{3}}{2}\hat{x} + \frac{1}{2}\hat{y}\right), \quad \mathbf{B}_2 = \frac{4\pi}{\sqrt{3}a}\hat{y}, \quad (B.1)
$$

where $a = 2.46$ is the lattice constant (some authors use $a$ to denote the C-C bond length, which is a factor of $\sqrt{3}$ smaller than the lattice constant we are using here). In this choice, we
can choose the basis of the honeycomb lattice sites to be

\[ r_1 = \frac{1}{3} A_1 + \frac{2}{3} A_2; \quad r_2 = \frac{2}{3} A_1 + \frac{1}{3} A_2. \]  

(B.2)

In momentum space, the \( K, K' \) points are given by \( \pm (B_1 + B_2)/3 \), or, for the equivalent ones lying on the \( x \)-axis, \( \pm (2B_1 - B_2)/3 \). Note that \( |K| = 4\pi/(3a) \), as is well known. Furthermore, we take the Dirac speed \( v_F \) to be \( 10^6 \text{ ms}^{-1} \). Besides, we choose the moiré lattice vectors to be

\[ a_1 = \frac{a}{2 \sin(\theta/2)} \left( \frac{\sqrt{3}}{2}, \frac{1}{2} \right); \quad a_2 = \frac{a}{2 \sin(\theta/2)} \left( -\frac{\sqrt{3}}{2}, \frac{1}{2} \right). \]  

(B.3)

In the main text, we have listed all the symmetries of the continuum theory (Table 2.1). Here, we tabulate explicitly the symmetry transformations of the electron operators, which follow from that of the Dirac points in the monolayer problem.

\[ \hat{t}_\rho \hat{\psi}_{\pm \mu; k} \hat{t}_\rho^{-1} \propto e^{ik\cdot \rho} \hat{\psi}_{\pm \mu; k}; \]

\[ \hat{C}_6 \hat{\psi}_{\pm \mu; k} \hat{C}_6^{-1} = \sigma_1 e^{\mp i \frac{2\pi}{3} \sigma_3} \hat{\psi}_{\pm \mu; C_6 k}; \]

\[ \hat{M}_y \hat{\psi}_{\pm \mu; k} \hat{M}_y^{-1} = \sigma_1 \hat{\psi}_{\pm \mu; M_y[k]; M_y k}; \]

\[ \hat{T} \hat{\psi}_{\pm \mu; k} \hat{T}^{-1} = \hat{\psi}_{\pm \mu; -k}, \]  

(B.4)

where \( \mu = t, b \). Note that \( M_y \) is the only symmetry which flips the two layers, i.e., \( M_y[t] = b \) and vice versa.

The symmetries listed in Eq. (B.4) generate all the spatial symmetries of the continuum theory of the TBG [75, 80, 87] (in wallpaper group 17). In particular, we see that \( t_\rho \) and \( M_y \) preserves the valley index (K vs. -K) whereas \( C_6 \) and \( T \) do not. However, their (pair-wise) nontrivial products will leave valley invariant, and it is helpful to also document their symmetry action explicitly (which are fixed by the above):

\[ \hat{C}_3 \hat{\psi}_{\pm \mu; k} \hat{C}_3^{-1} = e^{\mp i \frac{2\pi}{3} \sigma_3} \hat{\psi}_{\pm \mu; C_3 k}; \]

\[ (\hat{C}_6 \hat{T}) \hat{\psi}_{\pm \mu; k} (\hat{C}_6 \hat{T})^{-1} = \sigma_1 e^{\pm i \frac{2\pi}{3} \sigma_3} \hat{\psi}_{\pm \mu; -C_6 k}; \]

\[ (\hat{C}_2 \hat{T}) \hat{\psi}_{\pm \mu; k} (\hat{C}_2 \hat{T})^{-1} = \sigma_1 \hat{\psi}_{\pm \mu; k}. \]  

(B.5)
Here, we make two remarks regarding the subtleties in the symmetry representation documented here: first, the momentum $k$ appearing above are defined as the deviation from the original Dirac points in the monolayer problem. Generally, they correspond to different momenta in the moiré BZ. For instance, the Dirac point labeled by $(+, t)$, i.e., that of the $K$ point in the top layer, is mapped to $K_M$, whereas $(-, t)$ is mapped to $K'_M$. Similarly, $(+, b)$ and $(-, b)$ are respectively mapped to $K'_M$ and $K_M$. If desired, one can also rewrite Eqs. (B.4) and (B.5) using a common set of momentum coordinates defined with respect to the origin of the moiré BZ.

Second, the representation of the translation symmetry, $\hat{t}_\rho$, has a subtlety in its definition. This is because the microscopic translation effectively becomes an internal symmetry for the slowly varying fields appearing in the continuum theory. As such, for a single layer one can only deduce its representation up to an undetermined phase, and hence the appearance of $\propto$ in Eq. (B.4). However, the relative momentum across the different slowly varying fields, say the operators corresponding to the + valley of the top and bottom layers, is a physical quantity. Consequently, there is really only one common ambiguity across all the degrees of freedom appearing in the continuum theory.

### B.2 Valley Symmetry and Wannier Obstruction

We argue here that the valley symmetry resolved band structure does not admit a Wannier representation. Note, since we will ignore spin, this is a two-band model which will be crucial for what follows. If one were to include other bands the arguments below would fail, although precisely what selection of bands would lead to localized Wannier functions (LWFs) remains to be determined. In some ways, it is not very surprising that a valley resolved band structure does not admit a Wannier description, and a simple example is a single valley of monolayer graphene, which is just an isolated Dirac node. But in those cases the band structure does not terminate on raising the energy, and hence does not form an band. In contrast, in our present problem for TBG there is an isolated band, and so one may expect to capture the physics with LWFs. Nonetheless, we will argue there is an obstruction, as can be seen as follows.

We begin with three ingredients: (i) a two band model; (ii) $C_2^T$ symmetry; and a third ingredient which will be specified shortly. The two ingredients above enforce the following form on the
momentum space Hamiltonian:

\[ H = \epsilon_0(p) + \epsilon_1(p)\sigma_1 + \epsilon_2(p)\sigma_2, \quad (B.6) \]

where there is no condition on the function \( \epsilon_i(p) \). Similar to the main text, we implement \( C_2T \) by \( \sigma_1\mathcal{K} \), where \( \mathcal{K} \) denotes complex conjugation. Check that \( C_2T \) leaves the Hamiltonian above invariant. Now, if we are interested in the band wave-functions, they are independent of the first term in the hamiltonian, and we could pass to the following one by imposing a constraint:

\[ H' = \epsilon_1(p)\sigma_1 + \epsilon_2(p)\sigma_2 \quad (B.7) \]

The obvious constraint is to demand:

\[ \sigma_3H'\sigma_3 = -H'. \quad (B.8) \]

This is nothing but the chiral condition that specifies class AIII, and one can show that this can be implemented as an on-site symmetry for a 2-band system. Now, we introduce the third ingredient: (iii) the two Dirac points at the middle of this band structure have the same chirality. This allows us to write down the following effective Hamiltonian close to neutrality:

\[ H'_{\text{low}} = -iv_F(\partial_x\sigma_1 + \partial_y\sigma_2) \quad (B.9) \]

Where we now have a four component structure to include the two Dirac nodes. Note, there is no mass term that will gap out these nodes and also preserve the chiral condition B.8, hence this corresponds to the surface of a three dimensional topological phase in class AIII, with index \( \nu = 2 \), corresponding to the two Dirac nodes. Since this is the surface state of a nontrivial 3D topological phase, it does not admit a Wannier representation. However, when combined with the opposite valley band structure, together the pair of band structures does admit LWFs, but at the price of losing valley conservation symmetry.

Finally let us address a conundrum that the careful reader may be puzzled by. The valley
resolved bands are stated to be the anomalous surface states of a 3D topological phase, nevertheless they appear as isolated bands which seems to contradict the usual expectation that such anomalous bands cannot be separated in energy. The way this is resolved in the present case is through the two band condition, which further allows us to map the problem to one with particle-hole symmetry (class AIII). The later problem can have anomalous surface states that are disconnected from the bulk bands because they are forced to stick at zero chemical potential, and hence cannot be deformed into the bulk. This mapping to class AIII only holds for the two band model, hence if we add bands or fold the Brillouin zone from translation symmetry breaking, the presented arguments no longer hold.

B.3 Mirror-eigenvalue obstruction and chirality-obstruction

In this appendix we present the detailed arguments that relate the mirror-eigenvalue obstruction and the chirality-obstruction, as stated in Sec. 2.4.

To start, let us first record the symmetry algebra:

\[ M_y^2 = 1, \quad (C_2 T)^2 = 1, \quad M_y (C_2 T) = (C_2 T) M_y \tag{B.10} \]

In the following, we will show that having mirror eigenvalues ±1 at the \( M \) point is equivalent to having the same chirality for the two Dirac points at \( K \) and \( K' \). To show this, we take three steps:

1. On an open region of the mBZ that covers the \( K, K' \) and \( M \) points, the action of \( C_2 T \) can always be chosen to be

\[ \psi(k) \rightarrow \sigma_x \mathcal{K} \psi(k) \tag{B.11} \]

for all \( k \) in this region, where \( \mathcal{K} \) denotes complex conjugation. This choice can be made while having a smooth basis.

2. When the above choice of the \( C_2 T \) action is made, if the mirror eigenvalues at \( M \) are
$\pm 1$, we can choose the action of $M_y$ to be

$$\psi(k) \rightarrow \sigma_x \psi(k')$$

(B.12)

for all $k$ in this region, where $k'$ is the $M_y$ partner of $k$. If the mirror eigenvalues at $M$ are the same, then we can choose the action of $M_y$ to be

$$\psi(k) \rightarrow \eta_M \psi(k')$$

(B.13)

with $\eta_M = \pm 1$. For either case, the above choice can be made while having the basis smooth.

3. The above two symmetry actions guarantee that the chiralities of the two Dirac points at $K$ and $K'$ are the same (opposite) if the mirror eigenvalues at $M$ are opposite (same).

Below we prove these statements one by one.

B.3.1 ACTION OF $C_2T$

First assume a generic action of $C_2T$ under a smooth basis:

$$\psi(k) \rightarrow U(k) \psi(k)$$

(B.14)

where $U(k)$ is a $2 \times 2$ unitary matrix that satisfies $U^*U = 1$, because $(C_2T)^2 = 1$. The generic form of such $U(k)$ is

$$U(k) = e^{i\theta_0(k)} (a_0(k) + ia_1(k)\sigma_x + ia_3(k)\sigma_z)$$

(B.15)

The meaning of a smooth basis is that $a_{0,1,2,3}(k)$ and $\theta_0(k)$ are smooth functions of $k$ in this open region. Notice due to the lack of a term proportional to $\sigma_y$ in $U(k)$, $U(k)$ can be diagonalized by an orthogonal transformation. Furthermore, the orthogonal matrix corresponding to this transformation can be chosen to be a smooth function of $k$ because $U(k)$ is smooth.
In order to find a basis in which the action of $C_2\mathcal{T}$ is given by (B.11), we need to find a unitary $V(k)$ such that

$$V(k)\psi(k) \rightarrow \sigma_x K V(k) \psi(k)$$  \hspace{1cm} (B.16)

under $C_2\mathcal{T}$. If this unitary can be found, then the action of $C_2\mathcal{T}$ is given by (B.11) on the basis $V(k)\psi(k)$. It is not hard to see that this is equivalent to finding a unitary $\tilde{V}(k)$, such that under $C_2\mathcal{T}$

$$\tilde{V}(k)\psi(k) \rightarrow K \tilde{V}(k) \psi(k)$$  \hspace{1cm} (B.17)

So below we will show that this latter $\tilde{V}(k)$ exists.

Combining (B.14) and (B.17) yields

$$U(k) = \tilde{V}(k)^T \tilde{V}(k)$$  \hspace{1cm} (B.18)

Because $U(k)$ can be diagonalized by an orthogonal transformation, there must be a solution of $\tilde{V}(k)$ to the above equation, and the solution can be made smooth as a function of $k$, because $U(k)$ is a smooth function of $k$ itself.

This means that, within this open region, the action of $C_2\mathcal{T}$ can always be chosen as

$$\psi(k) \rightarrow \sigma_x K \psi(k)$$  \hspace{1cm} (B.19)

while preserving the smoothness of the basis. This concludes the first step listed above.

**B.3.2 Action of $M_y$**

Now we go to the smooth basis under which the action of $C_2\mathcal{T}$ is given by (B.11), and assume a generic $M_y$ action under this basis

$$\psi(k) \rightarrow M(k)\psi(k')$$  \hspace{1cm} (B.20)
with a unitary $M(k)$. The symmetry algebra (B.10) implies that

\[ M(k)M(k') = M(k')M(k) = 1, \]
\[ M(k)^* = \sigma_x M(k)\sigma_x \]  

(B.21)

This means there are two possible types of $M(k)$:

\[ M(k) = \eta(k) e^{i\sigma_z \theta(k)} \]  

(B.22)

where $\eta(k) = \pm 1$, or

\[ M(k) = \cos \theta(k)\sigma_x + \sin \theta(k)\sigma_y \]  

(B.23)

Furthermore, the type of the mirror actions at $k$ and $k'$ must be the same. If both at $k$ and $k'$ the mirror action is of the first type, then

\[ \eta(k)e^{i\sigma_z \theta(k)}\eta(k')e^{i\sigma_z \theta(k')} = 1 \]  

(B.24)

If both at $k$ and $k'$ the mirror action is of the second type, then $\theta(k) = \theta(k')$.

Now we consider the case where the mirror eigenvalue at $M$ is $\pm 1$. Then the mirror action at $M$ must be of the second type, because the first type will not give rise to two different eigenvalues at $M$. Without loss of generality, we can take the action of $M_y$ at $M$ to be

\[ \psi(k = M) \rightarrow \sigma_x \psi(k = M) \]  

(B.25)

Next we look for a unitary $W(k)$ such that under $M_y$

\[ W(k)\psi(k) \rightarrow \sigma_x W(k')\psi(k') \]  

(B.26)

If such a $W(k)$ can be found, the action of $M_y$ can be chosen to be (B.12). This requires that

\[ W(k)M(k) = \sigma_x W(k') \]  

(B.27)
Notice this requirement automatically implies $W(k')M(k') = \sigma_x W(k)$ by noting that $M(k)M(k') = 1$.  We need to choose $W(k)$ such that the $C_2T$ action is still given by (B.11), which requires

$$W(k)^* = \sigma_x W(k) \sigma_x$$

(B.28)

The smoothness of the basis means that in this region $M(k)$ is always of the second type, given by (B.23).  Then to satisfy the requirements (B.27) and (B.28), we can choose

$$W(k) = W(k') = e^{-i\frac{\pi}{2} \theta(k)}$$

(B.29)

This is indeed a smooth transformation given that $\theta(k) = \theta(k')$ in this case, as discussed earlier.

This tells us that as long as the mirror eigenvalues at $M$ are $\pm 1$, we can always choose the mirror action in this region to be given by (B.12), while preserving the smoothness of the basis.

In contrast, if both mirror eigenvalues at $M$ are the same, the mirror action at $M$ must be of the first type. That is, we can write the action of $M_y$ at $M$ as

$$\psi(k = M) \rightarrow \eta_M \psi(k = M)$$

(B.30)

where $\eta_M = \pm 1$ is the mirror eigenvalue at $M$. The smoothness of the basis implies that $M(k)$ in the entire open region is of the first type, given by (B.22). In addition, in this entire open region $\eta(k) = \eta_M$. Then we can find a unitary $W(k)$ transformation, such that

$$W(k)M(k) = \eta_M W(k')$$

(B.31)

so that the mirror action is given by (B.13) in the entire open region, and the basis is still smooth while the action of $C_2T$ is still given by (B.11). More explicitly, we can choose

$$W(k) = W(k')^\dagger = e^{-i\frac{\pi}{2} \theta(k)}$$

(B.32)

This concludes the second step listed above.
B.3.3 Relative chiralities

Now we go to a basis where the action of $C_2\mathcal{T}$ is given by (B.11), and the action of $M_y$ is given by either (B.12) or (B.13), depending on whether the mirror eigenvalues at $M$ are different or identical. This $C_2\mathcal{T}$ action constrains the first-quantized Hamiltonian to be

$$H(k) = n_0(k) + n_1(k)\sigma_x + n_2(k)\sigma_y$$

(B.33)

And the winding of $(n_1(k), n_2(k))^T$ along a closed path defines the chirality along this closed path. From this definition, we see that only gapless points in the Brillouin zone contribute to the net chirality.

If the mirror eigenvalues are opposite at $M$, that is, the action of $M_y$ is given by (B.12), then $n_1(k) = n_1(k')$ and $n_2(k) = -n_2(k')$, where $k'$ is the $M_y$-partner of $k$. Now consider a pair of gapless points in the Brillouin zone that are related by $M_y$, and draw a small closed loop around each of them. It is straightforward to see that the windings around these two loops are the same. In contrast, if the mirror eigenvalues are identical at $M$, that is, the action of $M_y$ is given by (B.13), then $n_{1,2}(k) = n_{1,2}(k')$. It is straightforward to see that the windings around a pair of gapless points related by $M_y$ are opposite. This concludes that third step listed above.

In short, when $M_y$ is preserved, having opposite (same) mirror eigenvalues at $M$ is equivalent to having same (opposite) chiralities at a pair of gapless points related by $M_y$. Given that both $K$ and $K'$ are host Dirac points and they are related by $M_y$, the net chirality of the entire Brillouin zone is nonzero (zero) if the mirror eigenvalues at $M$ are opposite (identical). Upon breaking $M_y$, the net chirality cannot change due to its discrete nature. This enables us to check the net chirality by simply looking at the eigenvalues of $M_y$ at its high symmetry points.
Appendices of chapter 3

C.1 Symmetry representations

In this appendix, we providing some details regarding the symmetry representations in the problem. In Table C.1 we provide the symmetry properties of the fermion operators. Note that we have included a time-reversal-like operation $\tilde{T}$, which is not a symmetry of the problem, but is considered for simplifying our discussion.

In Table C.2, we provide the momentum-space representations arising from the full-filling of orbitals in real-space, i.e., atomic insulators. The real-space orbitals we will consider include: $(\tau, s)$, $(\tau, p_z)$, $(\tau, p_{\pm})$, $(\eta, s)$, $(\eta, p_z)$, $(\eta, p_{\pm})$, $(\kappa, s)$ and $(\kappa, p_z)$. Notice that when a pair of states at $\Gamma$ have $C_3$ representation of $(\omega, \omega^*)$, these two states must have $M_y$ eigenvalues $(+1, -1)$. In
Table C.1: Action of symmetries on the real-space orbitals. Given a symmetry $g$ and the fermion creation operator $\hat{c}^\dagger$ for an orbital, we tabulate the outcome of $\hat{g}\hat{c}^\dagger\hat{g}^{-1}$. Note that the action of anti-unitary operators is ambiguous up to an arbitrarily choice on $U(1)$ phases. Also, the time-reversal-like symmetry $\tilde{T}$ is not a symmetry of our problem, as the actual time-reversal symmetry of twisted bilayer graphene exchanges the two valleys. We include $\tilde{T}$ here merely to simplify the discussion, and it would be broken explicitly. We let $\omega \equiv e^{i2\pi/3}$.

\[
\begin{array}{cccc}
  g & \hat{s}^\dagger & \hat{p}_z^\dagger & \hat{p}_+^\dagger & \hat{p}_-^\dagger \\
  C_3 & \hat{s}^\dagger & \hat{p}_z^\dagger & \omega^* \hat{p}_+^\dagger & \omega \hat{p}_-^\dagger \\
  M_y & \hat{s}^\dagger & -\hat{p}_z^\dagger & \hat{p}_+^\dagger & \hat{p}_-^\dagger \\
  C_2T & \hat{s}^\dagger & \hat{p}_z^\dagger & \hat{p}_-^\dagger & \hat{p}_+^\dagger \\
  \tilde{T} & \hat{s}^\dagger & \hat{p}_z^\dagger & -\hat{p}_-^\dagger & -\hat{p}_+^\dagger \\
\end{array}
\]

other words, these two states form the two dimensional representation of symmetry $D_3 \cong \langle C_3 \rangle \times \langle M_y \rangle$, where $\langle g \rangle$ indicates the subgroup generated by the element $g$.

From Table C.2, one can find all possible representation-matching equations that can be used to resolve the Wannier obstruction in a “fragile” manner. (Alternatively, one could have also resolved it by appending topological bands, as we discussed in Ref. [128].) For example, there are two representation-matching equations involving only three bands:

\[
(\tau, s) \oplus (\eta, p_z) \overset{\text{rep.}}{=} (\tau, p_z) \oplus (\text{target});
(\tau, p_z) \oplus (\eta, s) \overset{\text{rep.}}{=} (\tau, s) \oplus (\text{target}).
\]  

(C.1)

This is the minimal number of bands that are needed to resolve the obstruction in terms of representations. However, as we will see in Appendix C.5 this resolution cannot correctly capture the band topology of the two nearly flat bands in TBG. Building from these two representation-matching equations, one can add trivial bands on both sides of the equation and get a new equation. We will call the latter a derived equation from the former.

When the number of bands involved is four, the representation-matching equations that can potentially lead to a fragile resolution of the obstruction are all derived equations of the above
Table C.2: The resulting symmetry representations at high-symmetry points from various real-space orbitals. Eigenvalues from degenerate bands are grouped by parenthesis. Note that $M_y$ is not a symmetry at $K$.

<table>
<thead>
<tr>
<th>$(\tau, s)$</th>
<th>$\Gamma$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_3$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$M_y$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\tau, p_\pm)$</th>
<th>$\Gamma$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_3$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$M_y$</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\eta, s)$</th>
<th>$\Gamma$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_3$</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$M_y$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\eta, p_\pm)$</th>
<th>$\Gamma$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_3$</td>
<td></td>
<td>$(\omega, \omega^*)$</td>
</tr>
<tr>
<td>$M_y$</td>
<td>-1, -1</td>
<td>$(\omega, \omega^*)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\kappa, s)$</th>
<th>$\Gamma$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_3$</td>
<td>1</td>
<td>$(\omega, \omega^*)$</td>
</tr>
<tr>
<td>$M_y$</td>
<td>1, 1, 1</td>
<td>$(\omega, \omega^*)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\kappa, p_z)$</th>
<th>$\Gamma$</th>
<th>$K$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_3$</td>
<td>1</td>
<td>$(\omega, \omega^*)$</td>
</tr>
<tr>
<td>$M_y$</td>
<td>-1, -1</td>
<td>$(\omega, \omega^*)$</td>
</tr>
</tbody>
</table>

ones. With five bands, however, there are new representation-matching equations:

$$(\eta, p_z) \oplus (\kappa, s) \overset{\text{rep.}}{=} (\kappa, p_z) \oplus \text{(target)}; \quad (C.2)$$

$$(\eta, s) \oplus (\kappa, p_z) \overset{\text{rep.}}{=} (\kappa, s) \oplus \text{(target)}.$$  

From these equations and using

$$(\tau, s) \oplus (\tau, p_\pm) \overset{\text{rep.}}{=} (\kappa, s); \quad (C.3)$$

one obtains

$$(\eta, p_z) \oplus (\tau, s) \oplus (\tau, p_\pm) \overset{\text{rep.}}{=} (\kappa, p_z) \oplus \text{(target)} \overset{\text{rep.}}{=} (\tau, p_z) \oplus (\tau, p_\pm) \oplus \text{(target)}; \quad (C.4)$$

$$(\eta, s) \oplus (\tau, p_z) \oplus (\tau, p_\pm) \overset{\text{rep.}}{=} (\kappa, s) \oplus \text{(target)} \overset{\text{rep.}}{=} (\tau, s) \oplus (\tau, p_\pm) \oplus \text{(target)}.$$
derived ones of equations with fewer bands:

\[(\tau, p_z) \oplus (\tau, p_{\pm}) \oplus (\kappa, s) \overset{\text{rep.}}{=} (\eta, p_{\pm}) \oplus \text{(target)}; \tag{C.5}\]

\[(\tau, s) \oplus (\tau, p_{\pm}) \oplus (\kappa, p_z) \overset{\text{rep.}}{=} (\eta, p_{\pm}) \oplus \text{(target)}, \]

where the first one is precisely Eq. (3.1) used in the main text. From these equations and using
Eq. (C.3), one can obtain other equations:

\[(\tau, s) \oplus (\tau, p_z) \oplus (\tau, p_{\pm}) \oplus (\tau, p_{\pm}) \overset{\text{rep.}}{=} (\eta, p_{\pm}) \oplus \text{(target)}; \tag{C.6}\]

\[(\kappa, p_z) \oplus (\kappa, s) \overset{\text{rep.}}{=} (\eta, p_{\pm}) \oplus \text{(target)}.

Up to six bands, it is straightforward to check that the above are all the independent representation-
matching equations that can potentially lead to a fragile resolution of the Wannier obstructions.
Here, by “independent” we mean that these equations cannot be viewed as a derived one from
another.

**C.2 DETAILS OF THE BLOCH HAMILTONIAN**

In this appendix, we provide further details on the Bloch Hamiltonians constructed in this work.
In particular, we provide the explicit expressions for the Hamiltonians. Unlike the presentation in
the main text, we find it more natural to first discuss the six-band model, and then move on to
the ten-band one.

**C.2.1 SIX-BAND MODEL**

Here, we document explicitly the symmetry-allowed terms entering into the six-band model we
constructed. In the following, all the coupling parameters \( t \) are real numbers.

Recall our six-band Hilbert space arises from \((\tau, p_z), (\tau, p_{\pm}), \text{and} (\kappa, s)\). Let us write the fermion
operator for orbitals in the unit cell \( r \) as

\[\mathcal{C}\hat{c}_{r}^{\dagger} \equiv \begin{pmatrix} \hat{c}_{p_{\pm},r}^{\dagger} & \hat{c}_{\tau_{p_{\pm},r}}^{\dagger} & \hat{c}_{\tau_{p_{\pm},r}}^{\dagger} & \hat{c}_{\kappa s_{r},r}^{(1)} & \hat{c}_{\kappa s_{r},r}^{(2)} & \hat{c}_{\kappa s_{r},r}^{(3)} \end{pmatrix} \tag{C.7}\]
which fixes our basis choice of the Bloch Hamiltonian.

As discussed in the main text, the terms in our six-band Hamiltonian will mostly be conventional nearest-neighbor bonds, with the sole exception of a second nearest-neighbor bond included for $(\kappa, s)$. Let us discuss these terms one-by-one. First, the nearest-neighbor bond between the $(\tau, p_z)$ orbitals takes the standard form on the triangular lattice:

$$H_{p_z} = t_{p_z} (\phi_{01} + \phi_{11} + \phi_{10} + \text{h.c.}), \quad (C.8)$$

where we let $\phi_{lm} = e^{-ik \cdot (la_1 + ma_2)}$, and we denote negative numbers by $\bar{l} \equiv -l$.

The nearest-neighbor couplings for the $(\tau, p_{\pm})$ orbitals are slightly more complicated due to the two-orbital structure. First, the intra-orbital piece is identical to $H_{p_z}$, but with $t_{p_z} \mapsto t_{p_{\pm}}$ and multiplied by the $2 \times 2$ identity matrix. Second, there is also an inter-orbital coupling, which in momentum-space is given by

$$C_{p_{\pm}p_{\pm}} = t_{p_{\pm}p_{\pm}}^+ (\phi_{01} + \phi_{11} \omega + \phi_{10} \omega^*) + t_{p_{\pm}p_{\pm}}^- (\phi_{01} + \phi_{11} \omega + \phi_{10} \omega^*), \quad (C.9)$$

where $\omega = e^{i2\pi/3}$. This gives the Bloch Hamiltonian

$$H_{p_{\pm}} = t_{p_{\pm}} (\phi_{01} + \phi_{11} + \phi_{10} + \text{h.c.}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & C_{p_{\pm}p_{\pm}}^t \\ C_{p_{\pm}p_{\pm}} & 0 \end{pmatrix}. \quad (C.10)$$

Next, we consider the kagome lattice. Aside from the standard nearest-neighbor bond $t_{\kappa}$, we find it natural to also incorporate the second nearest-neighbor bond $t_{\kappa}'$, for otherwise there would be an artificial (almost) flat band in the spectrum. This gives

$$H_{\kappa} = t_{\kappa} \begin{pmatrix} 0 & \phi_{10} & 1 \\ 1 & 0 & \phi_{01} \\ \phi_{11} & 1 & 0 \end{pmatrix} + t_{\kappa}' \begin{pmatrix} 0 & \phi_{11} & \phi_{10} \\ \phi_{01} & 0 & \phi_{10} \\ \phi_{01} & \phi_{11} & 0 \end{pmatrix} + \text{h.c.} \quad (C.11)$$

Lastly, we consider the nearest-neighbor coupling between the different lattices. In momentum
space they are characterized by

\[
\begin{align*}
C_{p\pm p_z} &= it^+_{p\pm p_z} \begin{pmatrix}
\phi_{01} + \phi_{11} \omega + \phi_{10} \omega^* \\
- (\phi_{01} + \phi_{11} \omega^* + \phi_{10} \omega)
\end{pmatrix} - it^-_{p\pm p_z} \begin{pmatrix}
\phi_{01} + \phi_{11} \omega + \phi_{10} \omega^* \\
- (\phi_{01} + \phi_{11} \omega^* + \phi_{10} \omega)
\end{pmatrix}; \\
C_{\kappa p\pm} &= t^+_{\kappa p\pm} \begin{pmatrix}
\phi_{10} & \phi_{\bar{1}1} \\
\phi_{\bar{1}1} \omega^* & \omega \\
\omega & \phi_{10} \omega^*
\end{pmatrix} - t^-_{\kappa p\pm} \begin{pmatrix}
\phi_{\bar{1}1} & \phi_{10} \\
\omega^* & \phi_{\bar{1}1} \omega \\
\phi_{10} \omega & \omega^*
\end{pmatrix}.
\end{align*}
\]

(C.12)

Here, the subscript \(p\pm p_z\) indicates coupling between the \((\tau, p_{\pm})\) and \((\tau, p_z)\) orbitals, and \(\kappa p\pm\) indicates that between \((\kappa, s)\) and \((\tau, p_{\pm})\). The real-space form of all the nontrivial couplings above is represented diagrammatically in Fig. C.1.

Altogether, the full Bloch Hamiltonian of the six-band model is given by

\[
H^{(6)}_\kappa = \begin{pmatrix}
H_{p_z} + \mu_{p_z} & C^\dagger_{p\pm p_z} & 0 \\
C_{p\pm p_z} & H_{p_{\pm}} + \mu_{p_{\pm}} & C^\dagger_{\kappa p\pm} \\
0 & C_{\kappa p\pm} & H_{\kappa} + \mu_{\kappa}
\end{pmatrix}
\]

(C.13)

where we have added relative chemical potentials \(\mu_i\) across the different lattices. For reasons that will become clearer later, we find it convenient to reparameterize them as

\[
\mu_{p_z} \equiv -6t_{p_z} + \delta_{p_z}; \quad \mu_{p_{\pm}} \equiv 3t_{p_{\pm}} + \delta_{p_{\pm}}; \quad \mu_{\kappa} \equiv -4(t_{\kappa} + t'_{\kappa}) + \delta_{\kappa}.
\]

(C.14)

Before we proceed, we make two remarks regarding the model parameters. First, note that we have ignored the coupling between the \((\tau, p_z)\) and \((\kappa, s)\) orbitals, as we find its inclusion to be unnecessary for reproducing the key energetic features of the dispersion. In practice, such terms are symmetry-allowed and would never be exactly zero, but since we only address symmetry-robust features in the problem their presence has little physical implications. As such, we choose to keep it at 0 to simplify the discussion.

Second, we have parameterized the coupling strengths such that when \(t^+_{p\pm p\pm} - t^-_{p\pm p\pm} = t^+_p - t^-_{p\pm p_z} = 0\), \(H^{(6)}\) will be \(\tilde{T}\)-invariant. Since the time-reversal-like operation \(\tilde{T}\) is not an actual symmetry of the problem, this relation will not hold in our model parameters. This
parameterization is nonetheless adopted as it provides a simple way to control the degree of $\tilde{T}$-breaking in the spectrum.

Figure C.1: Coupling terms in the six-band model. The full coupling terms consist of the indicated hopping together with their Hermitian conjugates. We always take the center site to be a triangular site in the “home” unit cell, and indicate the unit cell coordinates $l a_1 + m a_2$ of connected sites by $(lm)$ with $\bar{l} \equiv -l$. For the kagome sites in (e,f), we further specify their sublattice indices. The strength of the terms in panels (a-f) are respectively denoted by $t_{p\pm p\pm}^\pm$, $t_{p\pm p\pm}^{-\pm}$, $t_{p\pm p\pm}^\pm$, $t_{p\pm p\pm}^{-\pm}$, $t_{\kappa p\pm}^\pm$, and $t_{\kappa p\pm}^{-\pm}$.

Next, we expand on the discussion in the main text concerning a physical picture for our model parameters. Recall that the electronic weights of TBG sit mostly at the AA spots, which form a triangular lattice at the moiré scale [76, 77, 79, 82, 88, 93, 123]. In addition, the Dirac points at K and K’ are naturally explained by the symmetry characters of the $(\tau, p_{\pm})$ orbitals. The main nontrivial feature of the nearly flat bands, therefore, stems from the fact that at $\Gamma$ points the non-degenerate bands cannot come from the $(\tau, p_{\pm})$ orbitals due to a mismatch in the symmetry representations. Rather, in our Hilbert space they can only arise from the $(\tau, p_{\pm})$ and $(\kappa, s)$ orbitals. Our goal is to choose parameters such that the orbital characters of the nearly flat bands behave as expected across the entire Brillouin zone (BZ).

First, let us ignore all the nontrivial coupling terms by setting $t_{p\pm p\pm}^\pm = t_{p\pm p\pm}^{-\pm} = t_{p\pm p\pm}^\pm = t_{p\pm p\pm}^{-\pm} = 0$. 

269
As discussed, the nearly flat bands will be formed by the \((\tau, p_z)\) and \((\kappa, s)\) orbitals at \(\Gamma\), and the \((\tau, p_{\pm})\) Dirac points at \(K\) and \(K'\). In our parameterization, we can arrange all these orbitals to be at zero-energy by setting \(\delta_{p_z} = \delta_{p_{\pm}} = \delta_{\kappa} = 0\). Furthermore, by adjusting the values of \(\delta_{p_z}\) and \(\delta_{\kappa}\) we can set the bandwidth of the nearly flat bands.

As the wave functions of the nearly flat bands at \(\Gamma\) must arise from the \((\tau, p_z)\) and \((\kappa, s)\) orbitals, we should bring down the energy of the \((\tau, p_{\pm})\) bands at \(\Gamma\). This is achieved by choosing \(t_{p_{\pm}} < 0\). This sets up the required orbital characters for the bands near zero energy: \((\tau, p_z)\) and \((\kappa, s)\) at \(\Gamma\) but \((\tau, p_{\pm})\) in the rest of the BZ. We are then left with the \((\tau, p_z)\) and \((\kappa, s)\) bands at most of the BZ, as well as the \((\tau, p_{\pm})\)-Dirac point at \(\Gamma\). Observe that the bandwidth of the four connected bands below the nearly flat bands set the dominant energy scale in the band manifold we intend to describe. This bandwidth can be reconciled with the size of \(t_{\kappa}\), and therefore we choose it to be the reference energy scale and measure all the other terms with respect to it. We also choose the values of \(t_{p_z}\) and \(t'_{\kappa}\) to reproduce the broad energetic features of the bands.

Fig. 3.3b in the main text shows the band dispersion we obtained following the discussion above. Observe that the broad energetic features are already in place. It remains to turn on the nontrivial couplings between the orbitals, and open band gaps to isolate the two nearly flat bands at the top. Our chosen parameters are tabulated in Table C.3, and the corresponding band structure is shown in Fig. 3.3c in the main text. Note that some parameters are roughly an order of magnitude smaller than the others; they are adjusted to capture the fine energetic features in the nearly flat bands.

### C.2.2 Ten-band model

Next, we discuss the terms in the ten-band model. Here, the dominant energetic features are imprinted by the choice of the “quasi-orbital” fermion operator \(\hat{h}^{(l)}_{p_\pm, r} = \sum_x c_x^{\dagger} h^{(l)}_{p_\pm, r}(x)\). We have already provided the explicit form of the localized wave functions \(h^{(l)}_{p_\pm, r}(x)\) in real space in Fig. 3.1b of the main text. However, it will be convenient to also write down explicitly the Fourier
Table C.3: Hopping parameters in the six-band model. We abbreviate “nearest neighbor” to “nnbr.” We set the dominant energy scale to be $t_\kappa = 27$ meV, and measure all the other terms relative to that.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Ratio to $t_\kappa$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{p_z} \equiv \mu_{p_z} + 6t_{p_z}$</td>
<td>$(\tau, p_z)$ chemical potential</td>
<td>0</td>
</tr>
<tr>
<td>$\delta_{p_\pm} \equiv \mu_{p_\pm} - 3t_{p_\pm}$</td>
<td>$(\tau, p_\pm)$ chemical potential</td>
<td>-0.23</td>
</tr>
<tr>
<td>$\delta_\kappa \equiv \mu_\kappa + 4(t_\kappa + t_{\kappa}')$</td>
<td>$(\kappa, s)$ chemical potential</td>
<td>0.25</td>
</tr>
<tr>
<td>$t_{p_z}$</td>
<td>$(\tau, p_z)$ nnbr</td>
<td>0.17</td>
</tr>
<tr>
<td>$t_{p_\pm}$</td>
<td>$(\tau, p_\pm)$ intra-orbital nnbr</td>
<td>-0.017</td>
</tr>
<tr>
<td>$t_{p_\pm p_\pm}$</td>
<td>$(\tau, p_\pm)$ inter-orbital nnbr</td>
<td>-0.065</td>
</tr>
<tr>
<td>$t_{p_\pm p_\pm}$</td>
<td>$(\tau, p_\pm)$ inter-orbital nnbr</td>
<td>-0.055</td>
</tr>
<tr>
<td>$t_\kappa$</td>
<td>$(\kappa, s)$ nnbr</td>
<td>1</td>
</tr>
<tr>
<td>$t_\kappa'$</td>
<td>$(\kappa, s)$ second-nnbr</td>
<td>0.25</td>
</tr>
<tr>
<td>$t_{p_\pm p_\pm}$</td>
<td>$(\tau, p_\pm)$-(\tau, p_z) nnbr</td>
<td>0.095</td>
</tr>
<tr>
<td>$t_{p_\pm p_\pm}$</td>
<td>$(\tau, p_\pm)$-(\tau, p_z) nnbr</td>
<td>0.055</td>
</tr>
<tr>
<td>$t_{\kappa p_\pm}$</td>
<td>$(\kappa, s)$-(\tau, p_\pm) nnbr</td>
<td>0.6</td>
</tr>
<tr>
<td>$t_{\kappa p_\pm}$</td>
<td>$(\kappa, s)$-(\tau, p_\pm) nnbr</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Transform of these wave function in momentum space:

$$h^{(A)}_{p_+:k} = \begin{pmatrix} -\omega + \phi_{10}\omega^* + \phi_{01}\omega \zeta^* a \\ \omega^* + \phi_{11}\omega + \phi_{01}\omega \zeta b \\ 1 + \phi_{11} + \phi_{01} c \\ -i\phi_{10} d \\ -i\omega d \\ -i\phi_{01} \omega^* d \end{pmatrix}; \quad h^{(A)}_{p_-:k} = \begin{pmatrix} 1 + \phi_{11}\omega^* + \phi_{01}\omega \zeta^* a \\ 1 + \phi_{11} + \phi_{01} c \\ \omega^* + \phi_{11}\omega + \phi_{01}\omega \zeta b \\ -i\phi_{10} d \\ -i\omega^* d \\ -i\phi_{01} \omega d \end{pmatrix};$$

$$h^{(B)}_{p_+:k} = \begin{pmatrix} \omega + \phi_{10}\omega^* + \phi_{11}\omega \zeta^* a \\ \omega^* + \phi_{10}\omega + \phi_{11}\omega \zeta b \\ 1 + \phi_{10} + \phi_{11} c \\ id \\ i\omega d \\ i\omega^* d \end{pmatrix}; \quad h^{(B)}_{p_-:k} = \begin{pmatrix} -\omega + \phi_{10} + \phi_{11}\omega \zeta^* a \\ \omega^* + \phi_{10} + \phi_{11}\omega \zeta b \\ 1 + \phi_{10} + \phi_{11} c \\ id \\ i\omega^* d \\ i\omega d \end{pmatrix}. \tag{C.15}$$
Furthermore, it is natural to group these four column vectors into a single $6 \times 4$ matrix:

$$h_{k} = \begin{pmatrix} h^{(A)}_{p_+; k} & h^{(A)}_{p_-; k} & h^{(B)}_{p_+; k} & h^{(B)}_{p_-; k} \end{pmatrix}.$$  \hspace{1cm} (C.16)

The Bloch Hamiltonian corresponding to Eq. (3.3) with $\delta = 0$ can then be written as

$$H_k(t, 0) = t \begin{pmatrix} 0_{6 \times 6} & h_k \\ h_k^\dagger & 0_{4 \times 4} \end{pmatrix}.$$  \hspace{1cm} (C.17)

Note the block structure of $H(t, 0)$; we see immediately that $H(t, 0)$ anti-commutes with $6_{\times 6} \oplus (-4_{\times 4})$, implying $H_k(t, 0)$ will be particle-hole symmetric at every $k$. Also, as $h_k^\dagger$ is a $4_{\times 6}$ rectangular matrix, the equation

$$h_k^\dagger \phi = 0$$  \hspace{1cm} (C.18)

must have at least two non-trivial solutions at every $k$. In other words, there will be (at least) two exactly flat bands at zero-energy in the spectrum of $H(t, 0)$ (Fig. 3.2a and d in the main text). For generic choices of the parameters $a$–$d$, these will be the only states at zero-energy, and they serve as the precursor for the nearly flat bands in the final model.

Our next step is to add generic perturbations to reproduce the actual energetic features of the TBG spectrum. There are two main features which we wish to capture: (i) the dispersion of the nearly flat bands; and (ii) the absence of $\tilde{T}$ in the higher-energy bands. For (i), we simply add a subset of the terms we used in constructing the six-band model. For (ii), we consider an additional nearest-neighbor coupling between the $(\eta, p_\pm)$ orbitals, which are present only in the ten-band Hilbert space. As in the previous discussion, we represent the term diagrammatically in Fig. C.2, and provide the explicit expression:

$$H_\eta = t_\eta \begin{pmatrix} 0 & e^{i\phi_\eta(1 + \phi_{01} + \phi_{10})} \\ e^{-i\phi_\eta(1 + \phi_{01} + \phi_{10})} & 0 \end{pmatrix} \otimes_{2 \times 2},$$  \hspace{1cm} (C.19)

where $t_\eta, \phi_\eta$ are real parameters. In order to break the undesirable $\tilde{T}$-invariance, we will choose
Table C.4: Hopping parameters in the perturbation $\hat{V}$ to the ten-band model. We abbreviate “nearest neighbor” to “nnbr.” We measure the strengths of the various terms relative to the dominate one, $t_\eta = 32.5$ meV. All terms present in Table C.3 but not here are set to 0.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Meaning</th>
<th>Ratio to $t_\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta_{p_z} \equiv \mu_{p_z} + 6t_{p_z}$</td>
<td>$(\tau, p_z)$ chemical potential</td>
<td>-0.100</td>
</tr>
<tr>
<td>$\delta_\kappa \equiv \mu_\kappa + 4(t_\kappa + t'_\kappa)$</td>
<td>$(\kappa, s)$ chemical potential</td>
<td>0.110</td>
</tr>
<tr>
<td>$t_{p_z}$</td>
<td>$(\tau, p_z)$ nnbr</td>
<td>0</td>
</tr>
<tr>
<td>$t_{p_\pm}$</td>
<td>$(\tau, p_\pm)$ intra-orbital nnbr</td>
<td>0.003</td>
</tr>
<tr>
<td>$t_{p_\pm p_\pm}$</td>
<td>$(\tau, p_\pm)$ inter-orbital nnbr</td>
<td>0.004</td>
</tr>
<tr>
<td>$t'_\kappa$</td>
<td>$(\kappa, s)$ nnbr</td>
<td>0</td>
</tr>
<tr>
<td>$t_{\kappa}$</td>
<td>$(\kappa, s)$ second-nnbr</td>
<td>0</td>
</tr>
<tr>
<td>$t_{p_\pm p_\pm}$</td>
<td>$(\tau, p_\pm)-(\tau, p_z)$ nnbr</td>
<td>0.016</td>
</tr>
<tr>
<td>$t_{kp_\pm}$</td>
<td>$(\kappa, s)-(\tau, p_\pm)$ nnbr</td>
<td>0.016</td>
</tr>
<tr>
<td>$t_{kp_\pm}$</td>
<td>$(\kappa, s)-(\tau, p_{\pm})$ nnbr</td>
<td>-0.016</td>
</tr>
<tr>
<td>$t_\eta e^{i\phi_\eta}$</td>
<td>$(\eta, p_{\pm})$ nnbr</td>
<td>$i$</td>
</tr>
</tbody>
</table>

$\phi_\eta \neq 0, \pi$.

Figure C.2: The leading perturbation to the ten-band model, which breaks the undesirable $\tilde{T}$-invariance. Each circle denotes a honeycomb site, with the orbital $p_{\pm}$ indicated.

Altogether, the perturbation to the ten-band model is, in a block-matrix form, given by

\[
V_k = \begin{pmatrix}
\mu_{p_z} & C_{p_\pm p_z}^\dagger & 0 & 0 \\
C_{p_\pm p_z} & H_{p_\pm} + \mu_{p_\pm} & C_{kp_\pm}^\dagger & 0 \\
0 & C_{kp_\pm} & \mu_\kappa & 0 \\
0 & 0 & 0 & H_\eta
\end{pmatrix},
\]

where the four blocks correspond respectively to the $(\tau, p_z)$, $(\tau, p_\pm)$, $(\kappa, s)$, and $(\eta, p_{\pm})$ orbitals.

The chosen parameters for the relative strengths of the terms are provided in Table C.4. Note
that, if we switch off all the perturbations other than $H_\eta$, the two bands at charge neutrality will remain exactly flat.

### C.3 Deformation to Explicit Atomic Limits for the Six-Band Model

For completeness, we demonstrate the triviality of the complementary bands in the six-band model in this appendix. Unlike the ten-band model, our six-band model was defined using only conventional hopping terms without resorting to the notion of quasi-orbitals. Consequently, the four complementary bands are not automatically trivial. This can be settled by a similar deformation to an explicit atomic limit, achieved by first augmenting the Hilbert space to include the $(\eta, p_{\pm})$ bands, and then writing down a deformation Hamiltonian akin to that in Eq. (3.4) in the main text:

$$
H'_\mu^{(6)} = t_0 \cos \left( \frac{\pi \mu}{2\mu_0} \right) \begin{pmatrix} 0 & \tilde{h}_k \\ \tilde{h}_k^\dagger & 0 \end{pmatrix} + \mu \begin{pmatrix} \frac{1}{\mu_0} H_k^{(6)} & 0 \\ 0 & \frac{1}{10} \end{pmatrix}. 
$$

(C.21)

Here, $\tilde{h}_k$ is identical to that defined in Appendix C.2.2, but with a different set of wave-function parameters $(a, b, c, d) = (0.48 - 0.24i, 0.13 + 0.42i, 0.04 + 0.30i, 0.24 - 0.29i)$. As before, we set $t_0 = \mu_0 = 130$ meV.

When $\mu = \mu_0$, the lowest four bands coincide with that of $H^{(6)}$; when $\mu = -\mu_0$, the lowest four bands correspond to the strict atomic insulator arising from the $(\eta, p_{\pm})$ orbitals. In Fig. C.3, we plot the evolution of the band gaps above and below the two nearly flat bands at charge neutrality. The gaps never close, and hence $H'_\mu^{(6)}$ represents a smooth deformation of the lowest four bands of $H^{(6)}$ to a strict atomic insulator.

### C.4 Five-Band Model

In the main text, we have introduced two (closely related) tight-binding models, with six or ten bands, which captures the key properties of both the two nearly flat bands as well as the set(s) of four bands further away from charge neutrality. Yet, since our ultimate goal is to provide a real-space description of the two nearly flat bands, it might be beneficial to consider models with...
Figure C.3: Evolution of the band gaps above and below the two nearly flat bands from Eq. (C.21), which establishes an adiabatic deformation of the lowest four bands of our six-band model to a strict atomic insulator. More than $4 \times 10^4$ momenta are sampled in the BZ in determining $\Delta$. 

In this appendix, we provide a five-band model constructed in the spirit of compromising the accuracy of the high-energy states in order to reduce the number of bands involved. Unlike that in Ref. [128], however, our model here faithfully reproduces the features of the TBG bands in the energy window from roughly $-100$ meV to the top of the two nearly flat bands. As is evident below, this five-band model will still belong to the same class as those already constructed, and it will capture all the energetic, symmetry and topology features of the two nearly flat bands in TBG. In addition, the three complementary bands in the model will again be atomic in nature. In fact, by combining the discussions in Appendices C.1 and C.5, one can show that our five-band model is the minimal (in the sense of minimum number of bands) fragile resolution of the band topology in TBG.

In parallel with the discussion in the main text, our starting point will be a representation-matching equation which involves five bands:

$$(\eta, s) \oplus (\tau, p_z) \oplus (\tau, p_\pm) \overset{\text{rep.}}{=} (\kappa, s) \oplus \text{(target)},$$  \hspace{1cm} (C.22)

which is a part of Eq. (C.4). We will again construct a set of pseudo-orbitals which have identical
symmetry properties as \((\kappa, s)\) but residing in the Hilbert space defined by the five bands on the left. Due to a shortage of alphabets we will label these orbitals as \(\rho_{s;1}^{(l)}\), where \(l = 1, 2, 3\) labels the three kagome sites in each unit cell. The site convention and a real-space description of the \(\rho_{s;1}^{(1)}\) is provided in Fig. C.4.

\[\begin{align*}
\rho_{s;1}^{(1)} &= \begin{pmatrix}
 i\tilde{a}(\phi_{11} - \phi_{10}) \\
 \tilde{b}\phi_{11} + \tilde{c}\phi_{10} \\
 \tilde{c}\phi_{11} + \tilde{b}\phi_{10} \\
 \tilde{d}^*\phi_{10} \\
 \tilde{d}
\end{pmatrix}; &
\rho_{s;1}^{(2)} &= \begin{pmatrix}
 i\tilde{a}(1 - \phi_{11}) \\
 \omega(\tilde{b} + \tilde{c}\phi_{11}) \\
 \omega^*(\tilde{c} + \tilde{b}\phi_{11}) \\
 \tilde{d}^* \\
 \tilde{d}
\end{pmatrix}; &
\rho_{s;1}^{(3)} &= \begin{pmatrix}
 i\tilde{a}(\phi_{10} - 1) \\
 \omega^*(\tilde{b}\phi_{10} + \tilde{c}) \\
 \omega(\tilde{c}\phi_{10} + \tilde{b}) \\
 \tilde{d}^*\phi_{01} \\
 \tilde{d}
\end{pmatrix},
\end{align*}\]

where \(\tilde{a}, \tilde{b}, \tilde{c}\) are real and \(\tilde{d}\) is complex. Going from top to bottom, the entries correspond to the fermion operators \(\tilde{\tau}_{p_z;1;\mathbf{k}}, \tilde{\tau}_{p_+;1;\mathbf{k}}, \tilde{\tau}_{p_-;1;\mathbf{k}}, \tilde{\eta}_{s;1;\mathbf{k}}^{(A)}\), and \(\tilde{\eta}_{s;1;\mathbf{k}}^{(B)}\). Furthermore, we again aggregate the three
column vectors into a $5 \times 3$ matrix $\rho_k \equiv \begin{pmatrix} \rho^{(1)}_{s; k} & \rho^{(2)}_{s; k} & \rho^{(3)}_{s; k} \end{pmatrix}$. Then we can write the Bloch Hamiltonian as

$$H^{(5)}_k = -t'_0 \rho_k \rho_k^\dagger + \text{diag}(\mu'_{p_z}, \mu'_{p_\pm}, \mu'_{p_\pm}, \mu'_{\eta}, \mu'_{\eta}),$$

(C.24)

where the $\mu'_i$ are again chemical potential. Again, if we set $\mu'_i = 0$, there will be two exactly flat bands at zero energy. Hence, we will adjust $\mu'_i$ to reproduce the energetic features of the two nearly flat bands.

In Fig. C.5a,b we show the band structures for the following set of parameters:

$$\tilde{a} = 0.25; \quad \tilde{b} = 0.2; \quad \tilde{c} = 0.1; \quad \tilde{d} = 0.67;$$

$$t'_0 = 80 \text{ meV}; \quad \mu'_{p_z} = -0.043 t'_0; \quad \mu'_{p_\pm} = 0; \quad \mu'_{\eta} = 0.05 t'_0.$$  

(C.25)

Comparing with the other band structures shown in the main text, we see that this five-band model faithfully captures the energetics from the top of the nearly flat bands down to $\sim -100$ meV. In addition, we again find exactly two Dirac points in the nearly flat bands, pinned respectively to K and K’. By design, all the symmetry representations of the nearly flat bands here are identical to those in TBG, and, therefore, based on the results in Ref. [128] we again conclude these two nearly flat bands showcase the known band topology of TBG.

Lastly, let us show that the three complementary bands in our five-band model are atomic in nature. This is anticipated, as the bands are constructed through the notion of quasi-orbitals transforming in the same way as $(\kappa, s)$. Unsurprisingly, we will show it using the same technique in Ref. [136], as is discussed in Appendix C.3: We first augment the Hilbert space to introduce explicitly a set of $(\kappa, s)$ orbitals, and then consider the deformation Hamiltonian

$$H'_{\mu}^{(5)} = t'_D \cos \left( \frac{\pi \mu}{2 \mu_0} \right) \begin{pmatrix} 0 & \rho_k \\ \rho_k^\dagger & 0 \end{pmatrix} + \mu \left( \begin{pmatrix} \frac{1}{\mu_0} H^{(5)}_k & 0 \\ 0 & \frac{1}{4} \end{pmatrix} \right),$$

(C.26)

here, all the parameters are the same as those listed in Eq. (C.25), and we set $\mu_0 = t'_D = t'_0 = 80$ meV. The evolution of the band gaps above and below the two nearly flat bands are shown in
Figure C.5: Band structure from the five-band model. (a) The full spectrum. (b) A zoom-in for the two nearly flat bands. (c) The band gaps above and below the two nearly flat bands at charge neutrality never close as we deform the bands into strict atomic limits. More than $4 \times 10^4$ momenta are sampled in the BZ in determining $\Delta$.

Fig. C.5c. As with the other models, the gaps never close, which establishes an adiabatic deformation of the three complementary bands to a strict atomic limit.

C.5 “Wilson loop” analysis

In Ref. [399], it was proposed that the band topology of the two nearly flat bands of TBG could be inferred from a certain topological property of the Wilson loop, and the authors further suggested that this band topology is “stable,” in that it will survive even in the presence of additional trivial degrees of freedom.

In essence, “Wilson loops” are the multi-band generalization of the Berry phase. To be self-contained, we define it as follows: Consider a set of $n$ bands whose eigenvectors are collected into a matrix $\Psi_k$ for every $k$. Let $C$ be a closed path in the BZ, and let $\{k_i\}$ be a discretization of $C$ into $N$ momenta such that $|k_{i+1} - k_i| \to 0$ as $N \to \infty$. We further label the momenta such that $k_N = k_1 + b$, where $b$ (possibly $= 0$) is a reciprocal lattice vector encoding the topological
property of $C$ (as a loop over the BZ). Then we define the Wilson loop to be the $n \times n$ matrix

\[ W(C) \equiv \lim_{N \to \infty} \Psi_{k_1}^\dagger \Psi_{k_2}^\dagger \Psi_{k_3}^\dagger \cdots \Psi_{k_{N-1}}^\dagger \Psi_{k_N}^\dagger, \quad (C.27) \]

where care must be taken to relate $\Psi_{k_N} = U_{k_1,b} \Psi_{k_1}$ for some unitary matrix $U_{k_1,b}$. $W(C)$ is unitary when $N \to \infty$, and in the presence of $C_2 T$ symmetry it can be further shown to be orthogonal [149–152, 399].

Following the recipe in Ref. [399], we compute $W(C)$ for a particular set of contours: write any $k$ in the BZ as $k = \frac{k_1}{2\pi} b_\parallel + \frac{k_1}{2\pi} b_\perp$, where $b_\parallel$ and $b_\perp$ are distinct reciprocal lattice vectors. Then pick $C$ to be “straight lines” running along $b_\parallel$ wrapping around the BZ once. The different contours are then labeled by the remaining coordinate $k_\perp$, and one studies the properties of the family of Wilson loops $\{W(k_\perp) : k_\perp \in [0, 2\pi)\}$.

We compute this Wilson loop spectrum for our ten-band model $H(t_0, \delta_0)$ (Eq. (3.3) in the main text), choosing $b_\parallel = b_2$ and $b_\perp = b_1$. The results are shown in Fig. C.6, which can be studied using the extensive results derived earlier in Ref. [152] concerning the $C_2 T$-protected topological properties of the Wilson loops.

The Wilson loop spectrum for the two nearly flat bands is shown in Fig. C.6a. We find the same nontrivial spectral flow for the nearly flat bands as in Ref. [399]. This is expected, as it is quite likely that the $Z$-valued Wilson loop invariant suggested in Ref. [399] (arising from $\pi_1(O(2)) = Z$, as was noted earlier in Refs. [149, 152]), is equivalent to the chirality obstruction we identified in Ref. [98]. From the spectrum, one can utilize the characterization in Ref. [152] to infer that the set of two nearly flat bands has trivial weak invariants (which are simply 1D Berry phases quantized to 0 vs. $\pi$), but a nontrivial second Stiefel-Whitney (SW) class [400] of $w_2 = 1 \in Z_2$. The meaning of having $w_2 = 1 \in Z_2$ is that there is an $O(n)$ monopole inside the BZ torus [150, 152]. For two-band problems, i.e., $n = 2$, our results imply the total monopole charge is half of the net chirality [33, 92, 97, 98, 122, 128]. Also, we note that $w_2$ is an additive invariant with respect to band stacking, subjected to the constraint that the weak 1D invariants of the bands are trivial [152]. This constraint is satisfied by all the sets of bands we consider here.

Fig. C.6b shows the spectrum obtained for the lowest four bands, which, as we have shown,
are adiabatically connected to an explicit atomic insulator. Note the existence of continuous spectral gaps separating each band from the rest. Based on the characterization in Ref. [152], all topological invariants are trivial. This is consistent with the atomic nature of the four complementary bands.

Fig. C.6c shows the spectrum for the lowest six bands, i.e., the composite bands formed by those in (a) and (b). Importantly, a continuous spectral gap is also found, similar to that observed in Ref. [149], which is consistent with the fact that these six bands together form an atomic insulator. Despite its atomic nature, the characterization in Ref. [152] indicates that the SW invariant is \( w_2 = 1 \). In fact, this follows simply from the additive nature of \( w_2 \). In any case, as we have demonstrated in the main text, these six bands are smoothly deformable into a strict atomic limit. This implies the proposed SW insulator in Ref. [152] is actually atomic in nature.

Figure C.6: Wilson loop spectra for the following set of bands in the ten-band model: (a) the two nearly flat bands at charge neutrality; (b) the lowest four bands; and (c) the six bands of (a) and (b) combined. Note that a nontrivial spectral flow, which forbids any atomic description, is found only in (a). This is consistent with the fragile nature of the band topology.

More generally, we remark that our results provide a concrete physical interpretation of the 2D \( C_2 T \)-protected SW invariant [152]. Recall the lowest six bands in our ten-band model corresponds to the atomic insulator \((\tau, p_z) \oplus (\tau, p_{\pm}) \oplus (\kappa, s)\). We have shown that

\[
    w_2 \left[ (\tau, p_z) \oplus (\tau, p_{\pm}) \oplus (\kappa, s) \right] = 1. \tag{C.28}
\]

Since the Wilson loop of the atomic insulator \((\tau, p_z) \oplus (\tau, p_{\pm})\) is identity in the strict atomic
limit, we may conclude $w_2 [\tau, p_z] = 0$. More carefully, this can be argued as follows: First, notice that $w_2$ is well-defined so long as $C_2 T$ and lattice translation symmetries are retained. Imagine breaking $C_3$ and $M_y$, such that there is no symmetry distinction between the orbitals which we originally labeled as $s$, $p_z$, $p_+$, and $p_-$. This implies $w_2 [\tau, p_z] = 3w_2 [(\tau, s)] = w_2 [(\tau, s)]$. Then our claim follows as the single band problem $(\tau, s)$ is in the trivial SW class $w_2 = 0$ [152].

Using the additive nature of $w_2$ with respect to band stacking [150,152], we can conclude

$$1 = w_2 [\tau, p_z] = w_2 [(\tau, p_z) \oplus (\tau, p_\pm)] + w_2 [(\kappa, s)] = w_2 [(\kappa, s)] . \quad (C.29)$$

By definition, $(\kappa, s)$ is manifestly atomic, and hence these bands are regarded as trivial in our context. Therefore, $w_2$ only indicates (stable) mutual distinction between atomic insulators, but does not imply nontrivial band topology which forbids any atomic (i.e., product-state) description. Lastly, we note that $w_2 [(\eta, \ell)] = 0$ for $\ell = s, p_z, p_\pm$. This is because the two honeycomb sites in each unit cell are related by $C_2 T$ and, upon the breaking of $C_3$ and $M_y$, one can smoothly collapse the two honeycomb sites at the point-group origin while respecting the protecting symmetries for $w_2$.

![](image.png)

**Figure C.7**: Wilson loop spectra for the following set of bands in $H_{\mu=0}^{(5)}$: (a) the two nearly flat bands at charge neutrality; (b) the lowest three bands; and (c) the five bands of (a) and (b) combined.

With these observations in mind, one may inspect the representation-matching equations in Appendix C.1 again and demand the equality of $w_2$ on the two sides. This narrows down the minimal “fragile resolution” of the band topology to involve 5 bands (if the complementary bands
are allowed to be topological, four-band models are possible, as we showed in Ref. [128]). We have already constructed a five-band fragile resolution in Appendix C.4. For completeness, we also compute the Wilson loop for that model. As the Wilson loop spectrum will be completely flat in a strict atomic limit, instead of studying $H^{(5)}$ directly we consider $H^{(5)}_{\mu=0}$ in Eq. (C.26), whose lowest five bands have the same band topology as $H^{(5)}$ due to the persistence of band gaps as $\mu$ is varied. The results are shown in Fig. C.7, which verify our earlier discussions on the relation between the $w_2$ invariant [152] and atomic insulators. In particular, the Wilson loop invariant defined in Ref. [399] is trivial in Fig. C.7c, which is obtained by simply appending a set of atomic bands, corresponding to $(\kappa, s)$, to the two nearly flat bands.
D.1  **An example that does not display dimensional decoupling: non-Fermi liquid near an Ising-nematic critical point**

In this appendix we provide an example of non-Fermi liquid that does not display dimensional decoupling: the non-Fermi liquid near an Ising-nematic critical point.

Consider an electronic system that has an instability towards the formation of an Ising ferromagnet, which breaks the $Z_2$ symmetry of spin flips. Such a phase transition can viewed as a realization of the “Ising-nematic” transition [401]. Theoretically, the universal physics of the Ising-Nematic transition in 2+1 dimension is believed to be described by focusing on a pair of patches on the Fermi surface that are parallel to each other. The low-energy effective Lagrangian
is

\[ \mathcal{L}_{\text{IN}} = \sum_{s=\pm} \psi_s^\dagger \left( \eta \frac{\partial}{\partial \tau} - is \frac{\partial}{\partial x} - \frac{\partial^2}{\partial y^2} - \lambda \phi \right) \psi_s \]

\[ + (\partial_y \phi)^2 + r\phi^2 \]

(D.1)

where \( \psi_s \) with \( s = \pm \) are the fermion operators on the two parallel patches, and \( \phi \) is the Ising-nematic order parameter that flips sign under the \( Z_2 \) transformation [168, 401].

It is found in the QC regime of this transition, the system is very non-Fermi liquid like, and the electron spectral function near this critical point also has the form of (4.2). As calculated in the framework of a controlled expansion, in this case \( \alpha \approx 0.7 \) and \( z = 3 \) [168].

Now consider we have a stack of many layers of such systems and examine the effects of interlayer interactions on the decoupled fixed point. First consider the interlayer electric conductivity \( \sigma \) induced by interlayer electron tunneling. Because \( \alpha \approx 0.7 \), the arguments in the main text indicates it is relevant. Furthermore, according to (4.32), the electrons are not incoherent enough and \( \sigma \) actually diverges as the temperature goes to zero.

We can also consider the coupling between the order parameters between different layers. This coupling is symmetric under a global \( Z_2 \) transformation for all layers and is thus an allowed perturbation, and it does not involve interlayer electron tunneling. It has the following form

\[ \sum_{\alpha \beta} g_{\alpha \beta} \int d\tau dx dy \phi_\alpha(x, y, \tau) \phi_\beta(x, y, \tau) \]

(D.2)

Following the RG analysis in Ref. [401], the scaling dimensions of various quantities of interests are \( [y] = -1 \), \( [x] = -2 \), \( [\tau] = -2 \) and \( [\phi] = 2 \), from which we can deduce that \( [g_{\alpha \beta}] = 2 \). This means this coupling is strongly relevant.

These results means that the decoupled fixed point is unstable to interlayer interactions in this example, so it does not display the phenomenon of dimensional decoupling.

In fact, under reasonable assumptions, it can be argued that for any non-Fermi liquid obtained near a conventional quantum critical point associated with a phase transition that involves internal symmetry breaking, the coupling between the order parameters on different layers is relevant. To see this, consider a single 2D layer and denote the order parameter associated with this tran-
sition by $O$. It is reasonable to assume that at the critical point the susceptibility of the order parameter diverges, which implies the scaling form,

$$\langle O(k, i\omega)O(k, i\omega) \rangle \sim \frac{1}{k^\delta} \cdot h \left( \frac{\omega}{k^s} \right)$$

has $\delta > 0$, where $h$ is a universal function. In terms of $\delta$, the scaling dimension of the order parameter is $[O] = D - \frac{\delta}{2}$, where $D$ is the total scaling dimension of the space-time.

Now consider a stack of many such 2D layers, the interlayer coupling of order parameters $\int d\tau dx dy O_i O_j$ has scaling dimension $-\delta < 0$, so it is relevant and prevents dimensional decoupling. Notice to get this result we have assumed the susceptibility of the order parameter diverges at the critical point, but we cannot prove this must hold for all symmetry-breaking transition in a metallic environment, neither can we find any counterexample.

Before ending this appendix, we note this statement does not imply the theoretic works that treat these systems as a single 2D layer are all incorrect. In fact, as long as the interlayer interactions are much weaker than the intralayer interactions at the lattice level, this treatment is valid unless one goes to extremely low energies.

### D.2 RG EQUATION OF PERTURBATION $\delta L_1$

In this appendix we derive the RG equation of the perturbation $\delta L_1$, (4.12).

**Figure D.1:** Feynman diagrams that contribute to the RG equation of $\delta L_1$ to the one-loop order.

At tree-level, this perturbation is marginal in two spatial dimensions. Up to one-loop order, there are three distinct Feynman diagrams that contribute to the RG equation of $\delta L_1$, as shown in Fig. D.1. To obtain the RG equation, we will carry out a Wilsonian procedure by integrating
out the fast modes with momenta between the old and new cutoff scales, $\Lambda$ and $\Lambda e^{-l}$, respectively, where $e^{-l}$ is a scaling factor. We do not put any cutoff for the frequency. The first two diagrams vanish after the frequency integral is carried out, which, physically speaking, is because for non-relativistic bosons there need to be at least two particles to have any interaction at all and the ground state at the unperturbed critical point is a vacuum with no particle. The third diagram does not vanish, and different $g_{\alpha\beta}$’s do not mix from this diagram. For the purpose of obtaining the RG equation, we can set the external momenta and frequencies to be zero, then the third diagram gives

$$
- \frac{g_{1\alpha\beta}^2}{(2\pi)^3} \int_{\Lambda e^{-l}}^{\Lambda} d^2 k \int_{-\infty}^{\infty} d\omega \left( \frac{1}{i\omega - \frac{k^2}{2mb}} - \frac{1}{i\omega - \frac{k^2}{2mb}} \right)
= -C g_{1\alpha\beta}^2 l
$$

with $C = \frac{m_b}{2\pi}$. Therefore, the one-loop RG equation of $g_{\alpha\beta}$ is

$$
\frac{dg_{1\alpha\beta}}{dl} = -C g_{1\alpha\beta}^2
$$

In fact, this RG equation is valid up to arbitrary order of perturbation, which is again related to the fact that in this model there need to be at least two particles so that there is interaction at all and the ground state at the critical point is the vacuum with no particle [174,176].

### D.3 Electron spectral function at criticality of the chemical potential tuned Mott transition

In this appendix we calculate the electron spectral function at the critical point of the chemical potential tuned Mott transition.

Because the electron vertex is not singularly enhanced [157], we can get the electron spectral function by convolving the boson spectral function

$$
\mathcal{A}_b(k, \omega) \sim \delta \left( \omega - \frac{q^2}{2m_b} \right)
$$
and the spinon spectral function

\[ A_f(k, \omega) \sim \frac{\omega^2}{\left( \lambda \sin \frac{\pi}{3} \omega^2 - \epsilon_k^2 \right)^2 + \lambda^2 \cos^2 \frac{\pi}{3} \omega^2} \]  

(D.7)

where spinon dispersion is taken to be in the form of (4.41).

The electron spectral function is then

\[ A(k, \omega) = \int_0^\omega d\Omega \int_q A_b(q, \omega - \Omega) A_f(k - q, \Omega) \]  

(D.8)

From the expressions of spectral functions of boson and spinon, we see in the low-energy and low-momentum regime the important region of integral is \( \omega \sim \Omega \sim q_\parallel^2 \sim q_\perp^2 \sim k_\parallel^2 \). Then in the spinon spectral function, \( k_\parallel - q_\parallel \) is much larger than all other terms, and we can approximate it by a delta function

\[ A_f(k - q, \Omega) \sim \delta(k_\parallel - q_\parallel) \]  

(D.9)

Now plug (D.6) and (D.9) into (D.8), we can do the integral and get

\[ A(k, \omega) \sim |\omega|^{\frac{1}{2}} f \left( \frac{2m_b\omega}{k_\parallel^2} \right) \]  

(D.10)

with

\[ f(x) = \left( 1 - \frac{1}{x} \right)^{\frac{1}{2}} \theta(x - 1) \]  

(D.11)

Fitting into the scaling form (4.2), we have \( \alpha = -1 \) and \( z = 2 \). Notice the electron spectral function (D.10) has the same form as (4.9), the electron spectral function at the critical point of a \( Z_4 \) orthogonal metal transition.
D.4 Effects on the interlayer pairing of (4.20)

In this appendix we discuss the effects on the interlayer pairing of the coupling (4.20). First consider a circular Fermi surface in a single layer, then the dominate pairing occurs between opposite patches on the Fermi surface. Because these opposite patches carry opposite electric currents, the Amperean force between them is repulsive, which suppresses pairing [168, 169]. Now suppose we have many layers coupled by (4.20), the dominate interlayer pairing is among patches on different layers that have parallel norms. If \( g_{3\alpha\beta} \) is positive (negative), the magnetic fields on the \( \alpha \)-th layer and on the \( \beta \)-th layer will favor to have opposite (same) signs. When they have opposite (same) signs, the Amperean force between patches on the two layers with parallel norms will be attractive (repulsive). Therefore, (4.20) will enhance interlayer pairing if \( g_{3\alpha\beta} > 0 \) and will suppress it if \( g_{3\alpha\beta} < 0 \).  

In the rest of this appendix, we will apply an RG analysis to justify the above physical picture.

It is convenient to write (4.20) in the frequency-momentum space

\[
\delta L_3 = \sum_{\alpha\beta} \int_{\mathbf{q},\omega} g_{3\alpha\beta} \mathbf{q}^2 \cdot a_{\alpha}(\mathbf{q}, i\omega) a_{\beta}(\mathbf{-q}, -i\omega)
\]

and introduce the momentum modes along the \( z \) direction for the gauge fields:

\[
a_{qz}(\mathbf{q}, i\omega) = \frac{1}{\sqrt{N}} \sum_{\alpha} a_{\alpha}(\mathbf{q}, i\omega) e^{-iqz\alpha}
\]

Then at the critical point of the chemical potential tuned Mott transition or in the spinon Fermi surface phase, the gauge field propagator can be diagonalized in the \( q_z \) basis as

\[
D_{qz}(\mathbf{q}, i\omega) = \frac{1}{k_0 |\omega| + q^2 (\chi_d + \mathcal{E}(q_z))}
\]

with the dispersion along the \( z \) direction \( \mathcal{E}(q_z) = \sum_{\alpha\beta} g_{3\alpha\beta} e^{iq_z(z_{\alpha} - z_{\beta})} \).

The coupling between the gauge field and the spinons on the \( \alpha \)-th layer should also be prop-

---

\(^1\)Similar line of thinking implies that for multiple layers of composite Fermi liquids interacting with repulsive interlayer density-density interactions, the system is unstable against interlayer pairing. This is consistent with the result in Ref. [402].
erly written in the $q_z$ basis as

$$a_\alpha \tilde{f}_\alpha f_\alpha = \left( \frac{1}{\sqrt{N}} \sum_{q_z} a_{q_z}(q, i\omega) e^{iq_z z_\alpha} \right) \tilde{f}_\alpha f_\alpha$$  \hspace{1cm} (D.15)

As we can see, each $q_z$ mode of the gauge field and the spinons on each layer are all coupled, but there is a factor $e^{iq_z z_\alpha}$ in the coupling constant. Due to the oscillating nature of this factor, different layers may appear to have opposite “charges” for a given $q_z$ mode, which can potentially induce attractive interlayer four-fermion interaction and enhance the interlayer pairing instability. We note that similar phenomenon has been studied in the context of bilayer composite Fermi liquids [402].

An RG approach that is suitable to study the effect of pairing in the presence of a gauge field in 2D has been recently developed in Ref. [169]. It is pointed out, the renormalization of the BCS channel scattering amplitude, $V$, comes from two aspects. The first is the conventional contribution arising from renormalizing the thickness of the momentum shell around the Fermi surface being considered to zero, which contributes to the beta function of $V$ by an amount $-V^2$. The second is from the interactions among different patches of the Fermi surface generated by integrating the high momentum modes of the gauge field as the sizes of the patches are renormalized to zero, and this contributes a positive constant to the beta function.

This RG procedure can be straightforwardly generalized into our multilayered system. Denote the interlayer BCS scattering amplitude of the fermions on the $\alpha$-th and the $\beta$-th layers by $V_{\alpha\beta}$, the aforementioned first type of contribution remains to be $-V_{\alpha\beta}^2$, but for the second type of contribution we need to add all $q_z$ modes with each of them multiplied by the oscillating exponential factor $e^{iq_z (z_\alpha - z_\beta)}$. Borrowing the result from Ref. [169], the beta function of $V$ in the weak interlayer coupling regime can be calculated to be

$$\frac{dV_{\alpha\beta}}{dl} = -\epsilon \frac{g_{3\alpha\beta}}{\chi_d} - V_{\alpha\beta}^2$$  \hspace{1cm} (D.16)

where $\epsilon$ is a positive constant.

From this equation we see if $g_{3\alpha\beta} < 0$, it suppresses the interlayer pairing instability, while it enhances this instability if $g_{3\alpha\beta} > 0$. This confirms the physical picture at the beginning of this
But since $g_{3\alpha\beta}$ is assumed to be very small, the potential pairing instability will only occur at extremely low temperatures and hence we ignore it.

### D.5 Derivation of the formula for interlayer electric conductivity

In this appendix we sketch the derivation of the formula for interlayer electric conductivity (4.30).

Starting from the Hamiltonian (4.28), we replace $t$ by $te^{ieAd}$ where $A$ is the external field. The partition function of the system in the presence of the external field is

$$Z[A] = \int [Dc]e^{-\sum a S_a + \int d\tau \sum \alpha t(e^{ieAd}c_{i,\alpha+1}c_{i,\alpha} + h.c.)}$$

where $S_a$ is the Euclidean action of the $a$th layer and the second term in the exponent is the action corresponding to interlayer electron tunneling.

Suppose this external field induces current $j$, going to the frequency-momentum space, the conductivity in the linear response regime is

$$\sigma(\omega) = \frac{1}{i\omega} j(\omega) = \frac{1}{i\omega} \frac{\delta^2}{A(\omega)^2} \ln \left( \frac{Z[A]}{Z[0]} \right)$$

Expanding $\ln \left( \frac{Z[A]}{Z[0]} \right)$ in terms of $A$, the quadratic term is

$$- \frac{t(ed)^2}{2} \int d\tau A(\tau)^2 \sum_{i\alpha} \langle c_{i,\alpha+1}c_{i,\alpha} + h.c. \rangle - \frac{(ted)^2}{2} \int d\tau \cdot$$

$$\int d\tau' A(\tau)A(\tau') \sum_{i\alpha,i'\alpha'} \left\langle c_{i,\alpha+1}(\tau)c_{i,\alpha}(\tau) - h.c. \right\rangle \cdot$$

$$\left\langle c_{i',\alpha'+1}(\tau')c_{i',\alpha'}(\tau') - h.c. \right\rangle$$

Assuming different layers are decoupled and expanding to the leading order of $t$, the above expression in frequency-momentum space becomes

$$\frac{N(ted)^2}{\beta} \sum_{k,\omega_1,\omega_2} |A(i\omega_1)|^2$$

$$\cdot G(k, i\omega_2) (G(k, i\omega_2) - G(k, i(\omega_1 + \omega_2)))$$

290
with the electron Green’s function on each layer

\[ G(k, i\omega) = \int d^2x \, d\tau \langle T \{ c_i(\tau)c_0(0) \} \rangle e^{-ik \cdot r_i + i\omega \tau} \]  

(D.21)

By (D.18),

\[ \sigma(\omega) = \frac{\left( i\omega_n \rightarrow \omega + i\eta \right)}{i\omega} \]  

(D.22)

with

\[ K(i\omega_n) = \frac{N(ted)^2}{\beta}. \]

\[ \sum_{k, \omega_m} G(k, i\omega_m) \left( G(k, i\omega_n) - G(k, i(\omega_n + \omega_m)) \right) \]  

(D.23)

Plugging spectral decomposition

\[ G(k, i\omega) = \int d\Omega \frac{A(k, \Omega)}{i\omega - \Omega} \]  

(D.24)

into (D.22) and taking its real part, we get the interlayer electric conductivity at frequency \( \omega \):

\[ \sigma(\omega) = -\pi N(ted)^2. \]

\[ \sum_k \int d\Omega A(k, \Omega) A(k, \omega + \Omega) \frac{f(\omega + \Omega) - f(\Omega)}{\omega} \]  

(D.25)

Taking the limit \( \omega \rightarrow 0 \), we get (4.30) as the DC interlayer electric conductivity.

### D.6 Calculations of the imaginary parts of the spinon self-energies in the IFL and MFL regimes, and of the in-plane conductivity in the MFL regime

In this appendix we first provide more details of the calculations of the imaginary parts of the spinon self-energies in the IFL and MFL regimes, namely, (4.45) and (4.47), respectively. Then we also sketch the calculation of the in-plane conductivity in the MFL regime.
Both self-energies are calculated from an integral of the form

\[
\Sigma_f(k, i\omega) = v_F^2 T \sum_{\Omega} \int \frac{D(q, i\Omega) G_f^{(0)}(k - q, i(\omega - \Omega))}{q^2} \]

Whether the result is for chemical potential tuned Mott transition or bandwidth controlled Mott transition depends on whether the gauge field propagator is of the form (4.18) or (4.24).

We first use spectral decomposition to write the above integral as

\[
\Sigma_f(k, i\omega) = -\frac{v_F^2 T}{\pi} \sum_{\Omega} \int q D(q, i\Omega) A_f^{(0)}(k - q, \Omega_1) A_f^{(0)}(k - q, \Omega_2)
\]

Now we can carry out the Mastubara frequency summation and get

\[
\Sigma_f(k, i\omega) = \frac{v_F^2}{\pi} \int d\Omega_1 d\Omega_2 \left[ n(\Omega_1) + 1 - f(\Omega_2) \right]
\]

Performing the analytic continuation to real frequency \(i\omega \rightarrow \omega + i\eta\) and taking the imaginary part, we get

\[
\Sigma''(k, \omega) = -v_F^2 \int d\Omega [n(\Omega) + 1 - f(\omega - \Omega)]
\]

Now we need to carry out the integral over \(q\). To do that, we need to specify the gauge field propagator.
D.6.1 Calculation of the imaginary part of the spinon self-energy in the IFL regime

In the IFL regime in the chemical potential tuned Mott transition, the gauge field propagator is given by (4.18). As noted before, the important region of integral involves $q_\parallel \sim q^2_\perp \sim \Omega \sim \omega \ll q^2_\parallel$, so we can ignore the $q_\parallel$ dependence of the gauge field propagator. Now the integral over $q_\parallel$ can be carried out and we get

$$\Sigma''_f(k, \omega) = -\frac{v_F}{4\pi^2} \int d\Omega \frac{\cosh (\frac{\beta \omega}{2})}{\sinh (\beta (\Omega - \frac{\omega}{2})) + \sinh (\frac{\beta \omega}{2})} \cdot \int dq_\perp \frac{k_0 \Omega_{q_\perp}}{(k_0 \Omega_{q_\perp})^2 + (\chi a q^2_\perp + \rho_s)^2} \tag{D.30}$$

The integral over $q_\perp$ is

$$\int dq_\perp \frac{k_0 \Omega_{q_\perp}}{(k_0 \Omega_{q_\perp})^2 + (\chi a q^2_\perp + \rho_s)^2} = \begin{cases} \frac{2k_0 \Omega}{\rho^2} \ln \left( \frac{\Omega_0}{a|\Omega|} \right), & |\Omega| \lesssim \Omega_c \\ \frac{2k_0 \Omega}{\chi^2 (k_0 \Omega)^{\frac{3}{2}}}, & |\Omega| \gtrsim \Omega_c \end{cases} \tag{D.31}$$

where the characteristic frequency scale $\Omega_c = a^{(c)} \rho^2 k_0 \sqrt{\chi_d}$ with a constant $a^{(c)}$ that is on the order of unity. Plugging this result in (D.30), performing the integral over $\Omega$ and using that in the IFL regime $T \gg \Omega_c$, we get (4.45).

D.6.2 Calculation of the imaginary part of the spinon self-energy in the MFL regime

In the MFL regime of the bandwidth controlled Mott transition, the gauge field propagator is given by (4.24). Again, the important region of integral involves $q_\parallel \sim q^2_\perp \sim \Omega \sim \omega \ll q^2_\parallel$, so we can ignore the $q_\parallel$ dependence of the gauge field propagator. After carrying out the integral over
\[ q_{\parallel} \text{ and } q_{\perp}, \text{ we get} \]

\[ \Sigma''_f(k, \omega) = \]

\[ -\frac{v_F k_0}{4\pi^2 \sigma_0^2 \rho_s^2} \int d\Omega \frac{\cosh \left( \frac{\beta \omega}{2} \right)}{\sinh \left( \beta \left( \Omega - \frac{\omega}{2} \right) \right) + \sinh \left( \beta \frac{\omega}{2} \right)} \]

\[ \cdot \left[ \ln \left( 1 + \left( \frac{\Omega_b}{\Omega} \right)^2 \right) + \frac{\Omega_b}{\Omega} \left( \frac{\pi}{2} - \arctan \left( \frac{\Omega_b}{\Omega} \right) \right) \right] \]

(D.32)

where the characteristic frequency scale \( \Omega_b = \frac{a^{(b)} \sigma_0 \rho_s^2}{k_0} \) with another constant \( a^{(b)} \) on the order of unity. Finally, performing the integral over \( \Omega \) and using that in the MFL regime \( T \gg \Omega_b \), we get (4.47).

**D.6.3 Calculation of the in-plane conductivity in the MFL regime**

In order to calculate the in-plane conductivity, we need to calculate the transport scattering rate of spinons, which is the imaginary part of the spinon self-energy multiplied by an additional factor \( q_{\perp}^2 / k_0^2 \). So from (D.29), we see the transport scattering rate of spinons is

\[ \gamma_{tr} = \frac{v_F^2}{k_0^2} \int d\Omega n(\Omega) + f(\Omega) \]

\[ \cdot \int \frac{q_{\perp}^2 \text{Im} D(q, \Omega) A_f^{(0)}(k - q, \omega - \Omega)}{q} \]

(D.33)

As before, because the important region of integral involves \( q_{\parallel} \sim q_{\perp}^2 \Omega \sim \omega \ll q_{\parallel}^2 \), we will ignore the \( q_{\parallel} \) dependence of the gauge field propagator (4.24). Then the integral can be done and we find the dominant contribution is

\[ \gamma_{tr} \sim T^2 \]

(D.34)

This implies the in-plane conductivity behaves as

\[ T^{-2} \]

(D.35)

294
D.7 Validity of the perturbative calculations of the interlayer electric conductivity

In this appendix we discuss the validity of the perturbative calculation of the interlayer electric conductivity, which only keeps the leading order terms in $t$, the interlayer electron tunneling amplitude. Notice in general one needs to consider all types of interlayer interactions discussed in the main text, but here, for simplicity, we only consider interlayer electron tunneling specified by the model (4.28).

According to (D.19), schematically, each higher order term in the perturbative expansion picks up one more $t^2 |\langle b \rangle|^4 G_f^2$ factor compared to the previous term. This implies, due to the structure of (4.42), the $n$th term in the perturbative expansion of $\sigma$ has the structure

$$\frac{t^{2n} |\langle b \rangle|^{4n}}{\Sigma_f^{2n-1}}.$$  \hfill (D.36)

So the ratio of the $n+1$th term to the $n$th term can be schematically written as

$$\frac{t^2 |\langle b \rangle|^4}{\Sigma_f^2}.$$  \hfill (D.37)

For the IFL regime of the chemical potential tuned Mott transition, we expect (D.37) is of the order of $t^2 |\langle b \rangle|^4 / T^4 \sim t^2 (g - g_c)^{4\beta / 3}$. Because $T \gg \rho_\Delta^3 \sim (g - g_c)^{3\nu}$ in the IFL regime, we have $t^2 (g - g_c)^{4\beta / 3} / T^4 \ll t^2 (g - g_c)^{4(\beta - \nu)}$. In the universality class of 2D dilute bose gas $\beta = \nu = \frac{1}{2}$, so this ratio is much smaller than unity as long as the system is close to the quantum critical point and $t$ is small, which means the higher order terms in the perturbative expansion is much smaller than the leading order contribution. Moreover, as shown in Ref. [161], the in-plane electric conductivity scales as $T^{-\frac{3}{2}}$, so the interlayer electric conductivity in (4.46) is much smaller than the in-plane one as long as the temperature is low enough. Therefore, it is expected that the perturbative series in the IFL regime of the chemical potential tuned Mott transition will converge.

The issue is a little more subtle in the MFL regime of the bandwidth controlled Mott transition, which lies in the temperature regime $(g - g_c)^{\frac{4\beta}{1 + \eta}} < T < (g - g_c)^{\frac{2\beta}{1 + \eta}}$, where we have...
used \( \rho_s \sim (g - g_c)^\nu \) and the scaling relation \( \beta = \frac{\nu(1+\eta)}{2} \) in the universality class of 3D XY model. (D.37) in this case is expected to be of the order of \( t^2|\langle b \rangle|^4 / (T \ln \frac{T}{\rho_s^2})^2 \). Ignoring the logarithmic correction, the condition under which the perturbative calculation is valid is that \( T \gtrsim t|\langle b \rangle|^2 \sim t(g - g_c)^{2\beta} \). So long as the interlayer electron tunneling amplitude \( t \) is very small, we can still have \( (g - g_c)^{\frac{4\beta}{(1+\eta)}} > t(g - g_c)^{2\beta} \), then the higher order terms in the perturbative series are small compared to the leading order one. Moreover, according to Appendix D.6.3, the in-plane conductivity in this regime scales as \( T^{-2} \), which is much larger than the leading order contribution to the interlayer conductivity in the low energy regime. Therefore, the perturbative calculation is expected to be valid in the regime where \( g - g_c \gtrsim t^{\frac{1+\eta}{2(1-\eta)}} \approx t^{\frac{1}{2}} \).

On the other hand, if the parameters of the system are in a narrow window where \( g - g_c \) is extremely small so that \( g - g_c \lesssim t^{\frac{1+\eta}{2(1-\eta)}} \) and the temperature is so low that \( T \lesssim t(g - g_c)^{2\beta} \), the higher order terms in the perturbative series can be larger than the leading order term, so the perturbative calculation breaks down. In this case, we expect the temperature dependence will be more singular than \( T^{-1} \) but no more singular than \( T^{-2} \), which is the Fermi liquid form. But as long as the interlayer electron tunneling amplitude \( t \) is small, this only occurs in an extremely narrow window of the parameter space, so it may be more difficult to detect experimentally.
E.1 Symmetry actions of the Kitaev model and the neutral Fermi surface state

In this appendix we review the symmetry actions of the Kitaev model, and present a gauge transformation that converts them into the symmetry actions in Table 5.2.

The symmetry actions of the Kitaev model are discussed in details in Ref. [236]. For completeness, let us first review the relevant results therein. Ref. [236] also starts from the parton construction given by (5.6), but in a way where the $SU(2)$ gauge structure is manifest. More pre-
cisely, on the site $i$ define

$$F_i = \begin{pmatrix} f_{i1} & -f_{i2}^\dagger \\ f_{i2} & f_{i1}^\dagger \end{pmatrix}$$  \hspace{1cm} (E.1)$$

so that the physical spin operators on this site can be written in terms of $F_i$ as

$$S_i = \frac{1}{4} \text{Tr} \left( F_i^\dagger \sigma F_i \right)$$  \hspace{1cm} (E.2)$$

Because (E.2) is invariant under a local $SU(2)$ transformation

$$F_i \rightarrow F_i W_i$$  \hspace{1cm} (E.3)$$

where $W$ is an $SU(2)$ matrix, this parton construction has an $SU(2)$ gauge redundancy.

Generically, a symmetry action on $F_i$ is

$$F_i \rightarrow U_g(i)^\dagger F_{g(i)} G_g(i)$$  \hspace{1cm} (E.4)$$

where $i$ labels the site and $g$ labels a symmetry action. Ref. [236] gives the expressions for $U$ and $G$’s for $T_{1,2}$, $C_6^*$ and $\sigma^*$:

$$U_{T_{1,2}} = 1,$$

$$U_{C_6^*}(A) = U_{C_6^*}(B) = \sigma_{C_6^*}$$  \hspace{1cm} (E.5)$$

$$U_{T_{\sigma^*}}(A) = U_{T_{\sigma^*}}(B) = \sigma_{T_{\sigma^*}}$$

and

$$G_{T_{1,2}} = 1$$

$$G_{C_6^*}(A) = -G_{C_6^*}(B) = \sigma_{C_6^*}$$  \hspace{1cm} (E.6)$$

$$G_{T_{\sigma^*}}(A) = G_{T_{\sigma^*}}(B) = \sigma_{T_{\sigma^*}}$$

where $A$ and $B$ label the sublattice, and $\sigma_{C_6^*} \equiv \frac{1+i(\sigma_1+\sigma_2+\sigma_3)}{2}$ and $\sigma_{T_{\sigma^*}} \equiv e^{-i\sigma_3\frac{\pi}{4}}$. 298
Because of the $SU(2)$ gauge redundancy (E.3), it is equally valid to define another set of partons that are related to (E.2) by a gauge transformation

$$\tilde{F}_i = \begin{pmatrix} \tilde{f}_{i1} & -\tilde{f}_{i2}^\dagger \\ \tilde{f}_{i2} & \tilde{f}_{i1}^\dagger \end{pmatrix} = F_i W_i$$

(E.7)

In terms of $\tilde{F}$, the symmetry actions become

$$\tilde{F}_i \rightarrow U_g(i)^\dagger \tilde{F}_g(i) \tilde{G}_g(i)$$

(E.8)

where

$$\tilde{G}_g(i) = W_g(i)^\dagger G_g(i) W_i$$

(E.9)

when $g$ is unitary, or

$$\tilde{G}_g(i) = W_g(i)^\dagger G_g(i) W_i^*$$

(E.10)

when $g$ is anti-unitary.

Now take $W$ to be $W_{A(B)}$ on the $A(B)$ sublattice, with

$$W_A = a + i(b\sigma_1 + a\sigma_3)$$

$$W_B = e^{-i\sigma_3 \bar{\tau}} W_A e^{i\sigma_3 \bar{\tau}}$$

(E.11)

where

$$a = \sqrt{\frac{1}{6 - 2\sqrt{3}}} \ 	ext{,} \ b = (\sqrt{3} - 1)a$$

(E.12)
one can verify that

\[
\begin{align*}
\tilde{G}_{T_{1,2}} &= 1 \\
\tilde{G}_{C_6}(A) &= \tilde{G}_{C_6}(B) = e^{i\sigma_3 \frac{\pi}{4}} \\
\tilde{G}_{T_{\sigma^*}}(A) &= \tilde{G}_{T_{\sigma^*}}(B) = e^{-i\sigma_3 \frac{\pi}{4}}
\end{align*}
\] (E.13)

In terms of \( \tilde{f}(r_i) = \left( \tilde{f}_1(r_i), \tilde{f}_2(r_i) \right)^T \), the symmetry actions using \( \tilde{G}_i \) are

\[
\begin{align*}
T_{1,2} : \tilde{f}(r_i) &\to \tilde{f}(r_i + n_{1,2}) \\
C_6^* : \tilde{f}(r_i) &\to e^{i\frac{\pi}{4}} \sigma_{C_6}^\dagger \tilde{f}(C_6 r_i) \\
T_{\sigma^*} : \tilde{f}(r_i) &\to e^{-i\frac{\pi}{4}} \sigma_{T_{\sigma^*}}^\dagger \tilde{f}(\sigma r_i)
\end{align*}
\] (E.14)

which is recorded in Table 5.2.

E.2 KITAEV MODEL IN TERMS OF SU(2) Partons

In this appendix we give the explicit expressions of (5.8).

First recall that at the mean field level the original Kitaev model described in [9] becomes a quadratic Hamiltonian of Majorana fermions \( \chi_{A,B}^a \), with \( a = 0, 1, 2, 3 \) and \( A(B) \) represents the sublattice index (we adopt the same notation as in Ref. [236]). It contains three types of terms:

\[
H_K = H_1 + H_2 + H_3
\] (E.15)

where \( H_1 \) is a nearest-neighbor coupling

\[
H_1 = 2iJ_1 \sum_{r_i} \left( \chi_B^0(r_i)\chi_A^0(r_i) + \chi_B^0(r_i + n_2)\chi_A^0(r_i) + \chi_B^0(r_i - n_1)\chi_A^0(r_i) \right),
\] (E.16)
$H_2$ is a next-nearest-neighbor coupling

$$H_2 = 2iJ_2 \sum_{r_i} \left( \chi^0_A(r_i + n_1)\chi^0_A(r_i) + \chi^0_A(r_i + n_2)\chi^0_A(r_i) + \chi^0_A(r_i + n_3)\chi^0_A(r_i) - \chi^0_B(r_i + n_1)\chi^0_B(r_i) - \chi^0_B(r_i + n_2)\chi^0_B(r_i) - \chi^0_B(r_i + n_3)\chi^0_B(r_i) \right)$$

(E.17)

and $H_3$ is a nearest-neighbor coupling

$$H_3 = 2iJ'_1 \sum_{r_i} \left( \chi^1_A(r_i)\chi^2_B(r_i) + \chi^1_A(r_i - n_2)\chi^1_B(r_i) + \chi^2_A(r_i + n_1)\chi^2_B(r_i) \right)$$

(E.18)

Notice for any values of $J_1$, $J_2$ and $J'_1$, this theory describes the same quantum phase (or the time reversal partner, depending on the sign of $J_2$) as in the Kitaev model. Since we are only concerned with universal properties, we will not worry the precise values of $J_1$, $J_2$ and $J'_1$.

In Ref. [236], the parton $f$ in (5.6) or (E.2) is related to the Majorana fermions via

$$f_1 = \frac{1}{\sqrt{2}} (\chi^0 + i\chi^3), \quad f_2 = \frac{1}{\sqrt{2}} (i\chi^1 - \chi^2)$$

(E.19)

We will instead relate these Majorana fermions to the transformed partons that is related to the above $f$’s via (E.7), with the gauge transformation given by (E.11). In terms of these partons, the Kitaev model (E.15) can be written as (5.8).

Below we spell out the terms in (5.8). For notational simplicity, we will drop the tildes from now on, with the understanding that now we are using the transformed partons.

All terms in $H_1$ are of the form $2iJ_1 \chi^0_B(r_j)\chi^0_A(r_i)$, which in terms of the transformed spinons reads

$$2iJ_1 \chi^0_B(r_j)\chi^0_A(r_i) = J_1 \left\{ f^T_B(r_j) \left[ \left( a^2 + \frac{b^2}{2} \right) + (a^2 - \frac{b^2}{2})\sigma_3 + ab\sigma_1 + ab\sigma_2 \right] f_A(r_i) + \text{h.c.} \right\}$$

(E.20)

where the first line goes into $H_{\text{hopping}}$, and the second line goes into $H_{\text{pairing}}$.

All terms in $H_2$ are of the form $2iJ_2 \tau_z \chi^0(r_j)\chi^0(r_i)$, with $\tau_z = 1$ ($\tau_z = -1$) in the A (B) sublat-
The terms on the $y$-bonds (bonds along $a_2$) are of the form $2iJ'_1\chi_A^2(r_j)\chi_B^2(r_i)$, which in terms of the transformed spinons reads

$$2iJ'_1\chi_A^2(r_j)\chi_B^2(r_i) = J'_1\left\{\begin{array}{l}
  f_A^\dagger(r_j) \left[ \begin{array}{c}
  -a^2 - \frac{b^2}{2} \\
  a^2 - \frac{b^2}{2}
  \end{array} \right] \sigma_3 - ab\sigma_1 - ab\sigma_2 f_B(r_i) + \text{h.c.} \\
  + f_A^T(r_j) \left[ \begin{array}{c}
  ia^2 + \frac{b^2}{2} \\
  -ia^2 + \frac{b^2}{2}
  \end{array} \right] \sigma_3 - ab(1 + i)\sigma_1 f_B(r_i) + \text{h.c.}
\end{array} \right\} \tag{E.23}$$

The terms on the $z$-bonds (bonds along $a_3$) are of the form $2iJ'_1\chi_A^3(r_j)\chi_B^3(r_i)$, which in terms of the transformed spinons reads

$$2iJ'_1\chi_A^3(r_j)\chi_B^3(r_i) = J'_1\left\{\begin{array}{l}
  f_A^\dagger(r_j) \left[ \begin{array}{c}
  -a^2 - \frac{b^2}{2} \\
  a^2 - \frac{b^2}{2}
  \end{array} \right] \sigma_3 + ab\sigma_1 + ab\sigma_2 f_B(r_i) + \text{h.c.} \\
  + f_A^T(r_j) \left[ \begin{array}{c}
  ia^2 + \frac{b^2}{2} \\
  -ia^2 + \frac{b^2}{2}
  \end{array} \right] \sigma_3 + ab(1 + i)\sigma_1 f_B(r_i) + \text{h.c.}
\end{array} \right\} \tag{E.24}$$

The first lines in the right-hand-side always go into $H_{\text{hopping}}$, and the second lines always go into $H_{\text{pairing}}$. 

$$302$$
E.3 Derivation of duality of critical theories

We can use the level-rank duality in Ref. [218] to show that the bosonic critical theory Eq. (5.15) is dual to the fermionic critical theory Eq. (5.18). We begin with a level-rank duality, namely, that the $U(2)^2$ theory with $N_f$ fundamental bosons is dual to the $SU(2)^{-2}N_f/2$ theory with $N_f$ fundamental fermions. The duality only holds for $N_f = 1, 2$. The bosonic theory is,

$$
\mathcal{L} = \sum_{I=1}^{N_f} |(\partial_\mu - i b_\mu)\Phi_I|^2 - m \sum |\Phi_I|^2 - V(|\Phi|) - \frac{2}{4\pi} \text{Tr}(bd\bar{b} - \frac{2i}{3}b^3) - \frac{1}{2\pi} B'd(Tr b).
$$

(E.25)

And the fermionic dual is,

$$
\mathcal{L} = \sum_{I=1}^{N_f} \bar{\Psi}_I(i\partial + a + \frac{B'}{2}\mathbf{1}_2 + m)\Psi_I + \frac{2-N_f/2}{4\pi} \text{Tr} \left[ (a + \frac{B'}{2}\mathbf{1}_2)d(a + \frac{B'}{2}\mathbf{1}_2) - \frac{2i}{3}(a + \frac{B'}{2}\mathbf{1}_2)^3 \right]
+ (4-N_f)CS_g
$$

(E.26)

Here $b$ is a $U(2)$ gauge field, $a$ is a $SU(2)$ gauge field and $B'$ is a $U(1)$ probe field. $V(\Phi)$ is the $SU(N_f)$ invariant quartic term.

Next we add a TQFT $U(1)^{-2}$ to both theories, yielding two new theories that are dual to each other. The bosonic theory changes to

$$
\mathcal{L} = \sum_{I=1}^{N_f} |(\partial_\mu - i b_\mu)\Phi_I|^2 - m \sum |\Phi_I|^2 - V(|\Phi|) - \frac{2}{4\pi} \text{Tr}(bd\bar{b} - \frac{2i}{3}b^3) - \frac{1}{2\pi} B'd(Tr b)
$$

+ $\frac{2}{4\pi} \beta d\beta - \frac{1}{2\pi} \beta d(B - B')$

(E.27)

And the fermionic theory changes to,

$$
\mathcal{L} = \sum_{I=1}^{N_f} \bar{\Psi}_I(i\partial + a + \frac{B'}{2}\mathbf{1}_2 + m)\Psi_I + \frac{2-N_f/2}{4\pi} \text{Tr} \left[ (a + \frac{B'}{2}\mathbf{1}_2)d(a + \frac{B'}{2}\mathbf{1}_2) - \frac{2i}{3}(a + \frac{B'}{2}\mathbf{1}_2)^3 \right]
+ (4-N_f)CS_g + \frac{2}{4\pi} \beta d\beta - \frac{1}{2\pi} \beta d(B - B')$

At last, we gauge the $U(1)$ probe field $B' \rightarrow \alpha$. In the bosonic theory, we can simply integrate
out α, yielding β = Tr b, and the theory exactly reduces to the bosonic critical theory Eq. (5.15) we introduced in the main text:

\[ L = \sum_{I=1}^{N_f} |(\partial_\mu - ib_\mu)\Phi_I|^2 - m \sum |\Phi_I|^2 - V(|\Phi|) - \frac{2}{4\pi} \text{Tr}(\bar{b}b - \frac{2i}{3} b^3) + \frac{2}{4\pi} \text{Tr}(b)d(Tr b) - \frac{1}{2\pi} B d(Tr b), \]  

(E.28)

In the fermionic theory, gauging B' will promote a + B'212 to a U(2) gauge field a,

\[ L = \sum_{I=1}^{N_f} \bar{\Psi}_I (i\partial + a + m)\Psi_I + \frac{2 - N_f/2}{4\pi} \text{Tr} \left[ ada - \frac{2i}{3} a^3 \right] + (4-N_f) \text{CS}_g + \frac{2}{4\pi} \beta d\beta - \frac{1}{2\pi} \beta d(B - (Tr a)). \]  

(E.29)

This theory is exactly the fermionic critical theory Eq. (5.18) we introduced in the main text.

In the rest of this appendix we will derive the topological nature of the resulting phases for different signs of m and for N_f = 1, 2. For this purpose, it is sufficient to switch off the probe gauge field B, i.e., we will set B = 0. The method presented below can be straightforwardly adopted to determine the topological nature of similar theories.

\[ m \ll -1 \text{ with } N_f = 1, 2: \text{ ITO} \]

Let us start with m \ll -1. In this case, integrating out the fermions results in the following effective Lagrangian (for both N_f = 1 and N_f = 2):

\[ L = \frac{2}{4\pi} \text{Tr} \left[ a d a - \frac{2i}{3} a^3 \right] + 4\text{CS}_g + \frac{2}{4\pi} \beta d\beta + \frac{1}{2\pi} \beta d(Tr a) \]  

\[ = \frac{2}{4\pi} \text{Tr} \left[ a d a - \frac{2i}{3} a^3 \right] + \frac{4}{4\pi} \tilde{a} d\tilde{a} + \frac{2}{4\pi} \beta d\beta + \frac{2}{2\pi} \beta d\tilde{a} + 4\text{CS}_g \]  

(E.30)

where a = a + \tilde{a}1, with a an SU(2) gauge field and \tilde{a} a U(1) gauge field.

Now we would like to understand why this Chern-Simons-matter theory describes the ITO state, i.e., a topological order with anyon contents \{1, σ, ε\}, where σ is a non-Abelian anyon with topological spin \(\theta_\sigma = e^{i\pi}\) and ε is a Majorana fermion. To this end, let us first understand different sectors of this theory. Denote the Lagrangian of the first sector by \(L_1\):

\[ L_1 = \frac{2}{4\pi} \text{Tr} \left[ a d a - \frac{2i}{3} a^3 \right] \]  

(E.31)
and the Lagrangian of the second sector by $\mathcal{L}_2$:

$$\mathcal{L}_2 = \frac{4}{4\pi} \tilde{a} \tilde{d} \tilde{a} + \frac{2}{4\pi} \beta d \beta + \frac{2}{2\pi} \beta \tilde{d} \tilde{a} \quad (E.32)$$

If $\mathcal{L}_1$ described a Chern-Simons field coupled to bosonic matter fields, it was precisely $SU(2)_{-2}$, i.e., it described a topological order with anyon content $\{1, \sigma_{-3}, \epsilon\}$, where $\sigma_{-3}$ is a non-Abelian anyon with topological spin $\theta_{\sigma_{-3}} = e^{-i\frac{3\pi}{8}}$, and $\epsilon$ is a Majorana fermion. This is also Kitaev’s $\nu = -3$ state in the 16-fold way [9]. In terms of the Chern-Simons-matter field theory, the $\sigma_{-3}$ excitation is obtained by exciting a matter field in the spinor representation of the $SU(2)$ gauge group, then the Chern-Simons term will associate some $SU(2)$ flux to this excitation and convert it into the non-Abelian anyon $\sigma_{-3}$. Importantly, here our Chern-Simons gauge field $a$ is coupled to a fermionic matter in the fundamental representation, and the fermionic nature of the matter field will change the topological spin of this excitation from $e^{-i\frac{3\pi}{8}}$ to $e^{-i\frac{3\pi}{8}} \times (-1) = e^{i\frac{3\pi}{8}}$. Let us suggestively denote this excitation as $\sigma_5$. The $\epsilon$ excitation is obtained by exciting a matter field in the integer-spin representation of the $SU(2)$ gauge field. Since matter fields in such representations are bosonic, the topological spin of this excitation will not be modified. One can further check the fusion and braiding, and verify that $\mathcal{L}_1$ coupled to fermionic matter in the fundamental representation is actually the topological order with $\nu = 5$ in Kitaev’s 16-fold way (up to the chiral central charge on the edge), with anyon content $\{1, \sigma_5, \epsilon\}$ [9].

Next we examine the property of $\mathcal{L}_2$. Using the standard K-matrix formalism [1], the topological nature of this theory can be determined by first rewriting $\mathcal{L}_2$ as

$$\mathcal{L}_2 = \frac{K_{IJ} a_I d a_J}{4\pi} \quad (E.33)$$

where $a_I = (\tilde{a}, \beta)^T$ and

$$K_{IJ} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \quad (E.34)$$
To read off the topological properties of this state, we need to invert the matrix $K$ and get

$$K^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{pmatrix}$$ (E.35)

The excitations of this theory can be labelled by an excitation vector $l$, and the elementary ones are $l_1 = (1, 0)^T$ and $l_2 = (0, 1)^T$. If the Chern-Simons gauge fields in $L_2$ are coupled to bosonic matter fields, the excitation labelled by $l_1$ has topological spin $\theta_1 = e^{-i\frac{\pi}{2}}$, and the excitation labelled by $l_2$ has topological spin $\theta_2 = -1$. These two excitations, $l_1$ and $l_2$, have mutual braiding $\theta_{l_1l_2} = -1$.

Notice $l_1$ carries charge-1 under both $\tilde{a}$ and $a$, so the excitation associated with $l_1$ is actually bound with $\sigma_5$ in the sector of $L_1$. This composite excitation has topological spin $e^{i\frac{\pi}{8}} \times e^{-i\frac{\pi}{2}} = e^{i\frac{\pi}{8}}$, and it will be identified as the $\sigma$ excitation in the ITO. At this point, there seem to be three nontrivial topological excitations: $\sigma$, $\epsilon$ and excitation $l_2$ in the sector $L_2$. As argued before, $l_2$ is a fermion. In fact, $l_2$ should be identified with $\epsilon$. To see this, consider the bosonic bound state $\epsilon \cdot l_2$. It is straightforward to check that this bound state has no nontrivial braiding with all other excitations. Therefore, this excitation must be local. In other words, $\epsilon$ and $l_2$ are in fact in the same topological sector.

In summary, the final anyon content is $\{1, \sigma, \epsilon\}$, which is precisely the same anyon content as the ITO state. Also, the fusion and braiding properties of these excitations also match with ITO. Furthermore, in Sec. 5.4 we have verified that the chiral central charge of the edge states of this theory matches with that of the ITO. Therefore, we conclude that the theory described by $m \ll -1$ is precisely the ITO state, for both $N_f = 1$ and $N_f = 2$.

$m \gg 1$ with $N_f = 1$: A SHORT-RANGE ENTANGLED STATE

Next, let us move to the case with $m \gg 1$ and $N_f = 1$. In this case, integrating out the fermions leads to the following effective Lagrangian:

$$\mathcal{L} = \frac{1}{4\pi} \text{Tr} \left[ ada - \frac{2i}{3} a^3 \right] + \frac{2}{4\pi} \tilde{a} d \tilde{a} + \frac{2}{4\pi} \beta d \beta + \frac{2}{2\pi} \beta d \tilde{a} + 2\text{CS}_g$$ (E.36)
Again, let us look at the two sectors separately:

\[
\mathcal{L}_1 = \frac{1}{4\pi} \text{Tr} \left[ \alpha d\alpha - \frac{2i}{3} \dot{\alpha}^3 \right]
\]

\[
\mathcal{L}_2 = \frac{2}{4\pi} \dot{a}d\dot{a} + \frac{2}{4\pi} \beta d\beta + \frac{2}{2\pi} \beta \dot{d} \dot{a}
\]

In the sector described by \(\mathcal{L}_1\), if the Chern-Simons gauge field is coupled to a bosonic matter, it is precisely the \(SU(2)_{-1}\) theory, which has a topological order with only one nontrivial excitation, an anti-semon \(\bar{s}\) with topological spin \(\theta_\bar{s} = e^{-i\frac{\pi}{2}}\). Again, this excitation comes from exciting a matter field in the spinor representation of the \(SU(2)\) gauge field. Because here the spinor representations are all fermionic, this anti-semon will be converted into a semion \(s\) with topological spin \(\theta_s = e^{i\frac{\pi}{2}}\).

In the sector described by \(\mathcal{L}_2\), let us first define \(a_{\pm} = \frac{1}{2}(\dot{a} \pm \beta)\). In terms of \(a_{\pm}\), \(\mathcal{L}_2\) can be written as

\[
\mathcal{L}_2 = \frac{8}{4\pi} a_+ da_+
\]

Notice the absence of a Chern-Simons term for \(a_-\) here, which means this gauge field should be confined due to monopole proliferation. That is to say, the deconfined excitation in the sector described by \(\mathcal{L}_2\) should have zero charge under \(a_-\). It is straightforward to verify that these excitations all have even charges under \(a_+\). So the elementary nontrivial excitation in this sector is given by having charge-2 under \(a_+\), and this excitation has topological spin \(\theta = e^{-i\frac{\pi + 2\pi x^2}{8}} = e^{-i\frac{\pi}{2}}\).

It is also easy to see that any excitation with charge-2 under \(a_+\) also carries a spinor representation of the \(SU(2)\) gauge field \(a\) in the sector of \(\mathcal{L}_1\), so this excitation is always bound with the \(s\) excitation from \(\mathcal{L}_1\), and the resulting composite excitation is a boson. One can verify there is no other nontrivial excitation in this theory, which means the topological order is actually trivial.

Furthermore, integrating out the gauge field in \(\mathcal{L}_1\) of Eq. (E.37) generates \(-CS_g\), and integrating out the gauge field in \(\mathcal{L}_2\) of Eq. (E.37) also generates \(-CS_g\). Adding them together cancels the background term \(2CS_g\), so the resulting state has vanishing chiral central charge on the edge.

In summary, the case with \(m \gg 1\) and \(N_f = 1\) is a short-range entangled state, \(i.e.,\) it has no nontrivial topological excitation or nontrivial edge mode.
Finally, let us turn to the case with \( m \gg 1 \) and \( N_f = 2 \). In this case, integrating out the fermions gives rise to the following effective Lagrangian:

\[
\mathcal{L} = \frac{2}{4\pi} \beta d\beta + \frac{2}{2\pi} \beta d\tilde{a} \tag{E.39}
\]

There is no Chern-Simons term for \( a \), which means the fermionic matter field that carries a fundamental representation of the \( SU(2) \) gauge field \( a \) is confined. The possible elementary deconfined topological excitations should carry charge-1 under \( \beta \) or charge-2 under \( \tilde{a} \). Using the K-matrix formalism it is easy to verify that these excitation are all bosons and they have no mutual braiding. Therefore, the resulting state actually contains no nontrivial anyon. Furthermore, one can check that integrating out the gauge fields in Eq. (E.39) generates no gravitational Chern-Simons term, which means that this theory has a zero chiral central charge on its edge.

Therefore, the case with \( m \gg 1 \) and \( N_f = 2 \) is also a short-range entangled state, i.e., it has no nontrivial topological excitation or nontrivial edge mode.

**E.4 Parton mean field of the fermionic \( U(2) \) critical theory**

In this appendix, we discuss the parton mean-field ansatz for the ITO and its confinement transitions. As discussed in the main text, the \( U(2) \) parton construction is \( \tilde{S}^+ = \phi^f \hat{f}^+_a \hat{f}^+_b \). There is a \( U(2) \) gauge redundancy, and the \((f_a, f_b)\) is the \( U(2) \) fundamental. We further rewrite \( \phi^f = c_1^f c_2^f \).

The mean field Hamiltonian of the fermionic partons \((c, f)\) generally has the first, second and third nearest-neighbor hoppings, which should be consistent with the symmetries: translation symmetry, inversion \( C_2 \) and \( \mathcal{T} \sigma^* \). We note that the symmetry actions of translation and inver-
sion are simple on \((c, f)\), while the \(T\sigma^*\) symmetry transformation is implemented as,

\[
\mathcal{T}\sigma^* : \quad i \rightarrow -i, \quad (E.40)
\]

\[
\mathcal{S}^{x,y}_r \rightarrow \mathcal{S}^{x,y}_{\sigma r}, \quad (E.41)
\]

\[
\mathcal{S}^z_r \rightarrow -\mathcal{S}^z_{\sigma r}, \quad (E.42)
\]

\[
c^\dagger_r \rightarrow c_{\sigma r}, \quad (E.43)
\]

\[
f^\dagger_r \rightarrow f_{\sigma r}, \quad (E.44)
\]

The mean-field Hamiltonian for the partons (both \(c\) and \(f\)) takes a generic form with the first-, second- and third-nearest-neighbor hoppings, \(H = -\sum_{ij} t_{ij} d_i^d d_j\), where \(d\) can represent either \(f\) or \(c\). Specifically, we consider a symmetry preserving hopping pattern, which has parameters \(t_{1x} = t_{1y}, t_{1z}, t_{2z}, t_{3x} = t_{3y}\) and \(t_{3z}\), as shown in Fig. E.1.

![Figure E.1: The hopping terms of the parton mean-field ansatz.](image)

The parton \(c_{1,2}\) is always in a \(C = -1\) band, it corresponds to \(\phi\) realizes \(\nu = -1/2\) bosonic Laughlin state. Specifically we take the hopping amplitude as \(t_{1x}^{c_1} = t_{1z}^{c_1} = 1, t_{1z}^{c_1} = 1, t_{2z}^{c_1} = 0.5e^{i\pi/2}\), and \(t_{1x}^{c_2} = t_{1z}^{c_2} = 1, t_{1z}^{c_2} = -1, t_{2z}^{c_2} = 0.5e^{i\pi/2}\).

ITO is realized by putting \(U(2)\) \(f\)-partons into a \(C = 2\) band, which for example can be realized with hopping amplitude \(t_{1x}^f = t_{1y}^f = 1, t_{1z}^f = 1, t_{2z}^f = 0.5e^{i\pi/2}, t_{3x}^f = t_{3y}^f = 0.3, t_{3z}^f = 1\).

To realize the zigzag magnetic order, we need to tune the Chern number of \(f\)-partons to \(C = 1\). It can be triggered by tuning \(t_{1x}^f = t_{1y}^f\), and the transition happens at \(t_{1x}^f = t_{1y}^f = 1.3\). Using this mean-field ansatz, we work out the symmetry quantum numbers of the relevant operators as summarized in Table 5.1. We note that the quantum numbers of \(d\) (\(\text{Tr a}\)) turn out to be identical to that of \(\bar{\Psi}\gamma^\mu \Psi\) in all our fermionic dual theories, so they are not displayed in the tables.
Table E.1: Symmetries of operators in the $N_f = 2$ critical theory with Dirac nodes at the $M_1$ and $M_2$ points. $s = 0, 1$ cannot be determined using our current method. There is one symmetry allowed relevant operator, $\bar{\Psi} \gamma^0 \tau^z \Psi$, which will destroy the quantum critical point.

<table>
<thead>
<tr>
<th>$\bar{\Psi} T^x \Psi$</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$C_2$</th>
<th>$T \sigma^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\Psi} T^y \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi} T^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^0 \Psi$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^1 \Psi$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^2 \Psi$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^0 \tau^x \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^1 \tau^x \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^2 \tau^x \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^0 \tau^y \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^1 \tau^y \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^2 \tau^y \Psi$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^0 \tau^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^1 \tau^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$\bar{\Psi} \gamma^2 \tau^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\mathcal{M}_1$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>$(-1)^* \mathcal{M}_2^*$</td>
</tr>
<tr>
<td>$\mathcal{M}_2$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>$(-1)^* \mathcal{M}_1^*$</td>
</tr>
<tr>
<td>$\mathcal{M}_3$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$-(-1)^* \mathcal{M}_3^*$</td>
</tr>
</tbody>
</table>
Table E.2: Symmetries of operators in the $N_f = 2$ critical theory with two Dirac cones on the high symmetry line $K - K'$. $s = 0, 1$ cannot be determined using our current method. There is one symmetry allowed relevant operator, $\bar{\Psi} \gamma^1 \tau^z \Psi$, which however only moves the location of Dirac points without destroying the quantum critical point.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$C_2$</th>
<th>$\mathcal{T} \sigma^s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\Psi}_1 \Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$-\bar{\Psi}_2 \Psi_1$</td>
<td>$-\bar{\Psi}_2 \Psi_1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2 \Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{2ik}$</td>
<td>$-\bar{\Psi}_1 \Psi_2$</td>
<td>$-\bar{\Psi}_1 \Psi_2$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1 \gamma^0 \Psi$</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1 \gamma^1 \Psi$</td>
<td>1</td>
<td>1</td>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2 \gamma^2 \Psi$</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1 \gamma^0 \Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$-\bar{\Psi}_2 \gamma^0 \Psi_1$</td>
<td>$\bar{\Psi}_2 \gamma^0 \Psi_1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1 \gamma^1 \Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$\bar{\Psi}_2 \gamma^1 \Psi_1$</td>
<td>$-\bar{\Psi}_2 \gamma^1 \Psi_1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1 \gamma^2 \Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$\bar{\Psi}_2 \gamma^2 \Psi_1$</td>
<td>$\bar{\Psi}_2 \gamma^2 \Psi_1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2 \gamma^0 \Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{2ik}$</td>
<td>$-\bar{\Psi}_1 \gamma^0 \Psi_2$</td>
<td>$\bar{\Psi}_1 \gamma^0 \Psi_2$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2 \gamma^1 \Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{2ik}$</td>
<td>$\bar{\Psi}_1 \gamma^1 \Psi_2$</td>
<td>$-\bar{\Psi}_1 \gamma^1 \Psi_2$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2 \gamma^2 \Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{2ik}$</td>
<td>$\bar{\Psi}_1 \gamma^2 \Psi_2$</td>
<td>$\bar{\Psi}_1 \gamma^2 \Psi_2$</td>
</tr>
<tr>
<td>$\bar{\Psi}_0 \tau^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1 \tau^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>$1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2 \tau^z \Psi$</td>
<td>1</td>
<td>1</td>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\mathcal{M}_1$</td>
<td>$-e^{2ik}$</td>
<td>$-e^{2ik}$</td>
<td>$-\mathcal{M}_2$</td>
<td>$(-1)^s \mathcal{M}_1^\dagger$</td>
</tr>
<tr>
<td>$\mathcal{M}_2$</td>
<td>$-e^{-2ik}$</td>
<td>$-e^{-2ik}$</td>
<td>$-\mathcal{M}_1$</td>
<td>$(-1)^s \mathcal{M}_2^\dagger$</td>
</tr>
<tr>
<td>$\mathcal{M}_3$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$(-1)^s \mathcal{M}_3^\dagger$</td>
</tr>
</tbody>
</table>
Table E.3: Symmetries of operators in the $N_f = 2$ critical theory with two Dirac cones on the high symmetry line $M_3 - M_3$. $s = 0, 1$ cannot be determined using our current method. There is one symmetry allowed relevant operator, $\bar{\Psi}\gamma^2\tau^z\Psi$, which however only moves the location of Dirac points without destroying the quantum critical point.

<table>
<thead>
<tr>
<th>Operator</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$C_2$</th>
<th>$T\sigma^s$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\Psi}\Psi$</td>
<td>$e^{-2ik}$</td>
<td>$e^{2ik}$</td>
<td>$-\bar{\Psi}\Psi_1$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\Psi_2$</td>
<td>$e^{2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$-\bar{\Psi}_1\Psi_2$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\gamma^0\Psi_2$</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\gamma^1\Psi_2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\gamma^2\Psi_2$</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\gamma^0\Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{2ik}$</td>
<td>$-\bar{\Psi}_1\gamma^0\Psi_2$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\gamma^1\Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{2ik}$</td>
<td>$\bar{\Psi}_1\gamma^1\Psi_2$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_1\gamma^2\Psi_2$</td>
<td>$e^{-2ik}$</td>
<td>$e^{2ik}$</td>
<td>$\bar{\Psi}_1\gamma^2\Psi_2$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2\gamma^0\Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$-\bar{\Psi}_2\gamma^0\Psi_1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$\bar{\Psi}_2\gamma^1\Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$\bar{\Psi}_2\gamma^1\Psi_1$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_2\gamma^2\Psi_1$</td>
<td>$e^{2ik}$</td>
<td>$e^{-2ik}$</td>
<td>$\bar{\Psi}_2\gamma^2\Psi_1$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_2\gamma^0\tau^z\Psi$</td>
<td>1</td>
<td>1</td>
<td>$-1$</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_2\gamma^1\tau^z\Psi$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\bar{\Psi}_2\gamma^2\tau^z\Psi$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$M_1$</td>
<td>$-e^{2ik}$</td>
<td>$-e^{-2ik}$</td>
<td>$-M_2$</td>
<td>$-(-1)^sM_2^\dagger$</td>
</tr>
<tr>
<td>$M_2$</td>
<td>$-e^{-2ik}$</td>
<td>$e^{2ik}$</td>
<td>$-M_1$</td>
<td>$-(-1)^sM_1^\dagger$</td>
</tr>
<tr>
<td>$M_3$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$-1$</td>
<td>$(-1)^sM_3^\dagger$</td>
</tr>
</tbody>
</table>
To realize the transition from the ITO to the trivially polarized state, we need to tune the Chern number directly from $C = 2$ to $C = 0$. There are three different types of ways to realize this transition:

1. Tuning $t_{1z}^f$, and the transition happens at $t_{1z}^f = 1.6$. The two Dirac cones are at the $M_1$ and $M_2$ points. In this case, the symmetry actions on $\Psi$ are given by

   \begin{align}
   T_1 : \Psi &\rightarrow -\tau^z \Psi \\
   T_2 : \Psi &\rightarrow \tau^z \Psi \\
   C_2 : \Psi(r) &\rightarrow \gamma^0 \tau^z \Psi(-r) \\
   T \sigma^* : \Psi(x,y) &\rightarrow i\gamma^1 \tau^x \Psi^\dagger(x,-y)
   \end{align}

   (E.45)

   The quantum numbers of the gauge invariant relevant operators are summarized in Table E.1.

2. Tuning $t_{3x}^f = t_{3y}^f$, and the transition happens at $t_{3x}^f = t_{3y}^f = 1$. The two Dirac cones are at $(k_1, k_2) = (k, k), (-k, -k)$, which are on the high symmetry line $K - K'$. In this case, the symmetry actions on $\Psi$ are given by

   \begin{align}
   T_1 : \Psi &\rightarrow e^{ik\tau^x} \Psi \\
   T_2 : \Psi &\rightarrow e^{ik\tau^y} \Psi \\
   C_2 : \Psi(r) &\rightarrow \gamma^0 \tau^y \Psi(-r) \\
   T \sigma^* : \Psi(x,y) &\rightarrow i\gamma^1 \tau^y \Psi^\dagger(x,-y)
   \end{align}

   (E.46)

   The quantum numbers of the gauge invariant relevant operators are summarized in Table E.2.

3. Tuning $t_{3z}^f$, and the transition happens at $t_{3z}^f = 1.6$. The two Dirac cones are at $(k_1, k_2) = (k, -k), (-k, k)$, which are on the high symmetry line $M_3 - M_3$. In this case, the symmetry
actions on $\Psi$ are given by

\begin{equation}
\begin{align*}
T_1 : & \Psi \rightarrow e^{ikr^z} \Psi \\
T_2 : & \Psi \rightarrow e^{-ikr^z} \Psi \\
C_2 : & \Psi(r) \rightarrow \gamma^0 \tau^b \Psi(-r) \\
T\sigma^* : & \Psi(x, y) \rightarrow i\gamma^1 \tau^x \Psi^\dagger(x, -y)
\end{align*}
\end{equation}

(E.47)

The quantum numbers of the gauge invariant relevant operators are summarized in Table E.3.

At last, we make a few comments on the monopole operators. Technically, we follow the method in Ref. [238, 403, 404] to calculate the quantum numbers of the monopoles. Namely, we explicitly construct the monopole states on a torus, and then extract the quantum number of the monopole states. Specifically, we put the system on a $2 \times L \times L$ lattice, and spread a uniform $2\pi$ flux for each parton $c, f$. Each Dirac fermion will form Landau levels with one exact zero mode. When $N_f = 1$, the gauge invariant monopole corresponds to a state with all negative-energy Fermi sea filled. In contrast, in the $N_f = 2$ critical theory, the gauge invariant monopoles should have two zero modes filled (each from one $U(2)$ color) together with the filled negative-energy Fermi sea. There are three gauge invariant ways to fill the zero modes,

\begin{align*}
M_1 = & \bar{\mathcal{M}}\psi_{1a}\psi_{1b}, \\
M_2 = & \bar{\mathcal{M}}\psi_{2b}\psi_{2a}, \\
M_3 = & \frac{1}{\sqrt{2}}\bar{\mathcal{M}}(\psi_{1a}\psi_{2b} - \psi_{1b}\psi_{2a}).
\end{align*}

(E.48)

Here $\bar{\mathcal{M}}$ is the bare monopole with $2\pi$ flux and filled negative-energy Fermi sea. $\psi$ represents the zero mode, and $1, 2$ are the flavor indices and $a, b$ are the color indices. The three monopoles are in the adjoint representation of the $SU(2)$ flavor symmetry. Using our current method, we are not able to determine the quantum number of the monopoles under $T\sigma^*$, for which there is an undetermined sign $\mathcal{M} \rightarrow \pm \mathcal{M}^\dagger$. In the $N_f = 1$ theory, we speculate the sign is $-1$, hence it matches the quantum number of the zigzag order. In the $N_f = 2$ theory, we leave this sign undetermined, and it has no influence on our discussion on the nature of the confined state.
More details on the quantum numbers of monopoles

Before finishing this appendix, we discuss in more details the symmetry actions on the monopoles of the fermionic critical theory with $N_f = 2$, using state-operator correspondence. Including both colors and spins, there will be four zero modes in the presence of $\pm 2\pi$ background flux. In this case there is no Chern-Simons term for $\text{Tr}(a)$, so two of the zero modes need to be occupied to form a gauge invariant state (operator). In terms of states, there are three different ways to occupy these zero modes and make a color singlet:

$$f_{1a}^\dagger f_{1b}^\dagger |0\rangle, \quad f_{2b}^\dagger f_{2a}^\dagger |0\rangle, \quad \frac{1}{\sqrt{2}}(f_{1a}^\dagger f_{2b}^\dagger - f_{1b}^\dagger f_{2a}^\dagger) |0\rangle$$

(E.49)

where the $f$’s are the operators of the zero modes, and $|0\rangle$ is the ground state under a $2\pi$ background flux with no zero mode occupied. We use 1 and 2 to label the two different flavors, and $a$ and $b$ to label the two different colors. These states correspond to the operators $M_{1,2,3}$ in Eq (E.48), respectively.

The actions of $T_{1,2}$ and $C_2$ are simpler because they do not take the monopole operators to their hermitian conjugates. To determine the action of $\mathcal{T}\sigma^*$, which takes the monopoles to their hermitian conjugates, it will be important to first identify the corresponding states of the hermitian conjugates of these operators. This can be worked out using the methods in Refs. [242, 243].

More precisely, let us write the three states in Eq. (E.49) in a more suggestive form

$$\mathcal{M}_1 \sim f_{1a}^\dagger f_{1b}^\dagger |0\rangle = \left( f^T \frac{(1 + \tau^z) \epsilon f}{4} \right)^* |0\rangle = \left( f^T \tau^y \frac{(\tau^y + i \tau^x) \epsilon f}{4} \right)^* |0\rangle$$

$$\mathcal{M}_2 \sim f_{2b}^\dagger f_{2a}^\dagger |0\rangle = \left( - f^T \frac{(1 - \tau^z) \epsilon f}{4} \right)^* |0\rangle = \left( - f^T \tau^y \frac{(\tau^y - i \tau^x) \epsilon f}{4} \right)^* |0\rangle$$

$$\mathcal{M}_3 \sim \frac{1}{\sqrt{2}}(f_{1a}^\dagger f_{2b}^\dagger - f_{1b}^\dagger f_{2a}^\dagger) |0\rangle = \frac{1}{\sqrt{2}} \left( f^T \frac{\tau^x \epsilon f}{2} \right)^* |0\rangle = \frac{1}{\sqrt{2}} \left( f^T \frac{\tau^y - i \tau^z \epsilon f}{2} \right)^* |0\rangle$$

(E.50)
where $\tau$ acts on the flavor space and $\epsilon$ acts on the color space. From these we get

$$i(M_1 + M_2) \sim \left( f^T \tau y \tau x \frac{\epsilon}{2} f \right)^* |0\rangle
$$

$$M_1 - M_2 \sim \left( f^T \tau y \tau x \frac{\epsilon}{2} f \right)^* |0\rangle
$$

$$-iM_3 \sim \frac{1}{\sqrt{2}} \left( f^T \tau y \tau z \frac{\epsilon}{2} f \right)^* |0\rangle$$

(E.51)

Therefore, $(i(M_1 + M_2), M_1 - M_2, -iM_3)$ transforms as a vector under the $SU(2)$ flavor symmetry. Because this representation of the $SU(2)$ transformation is real, $(-i(M_1^\dagger + M_2^\dagger), M_1^\dagger - M_2^\dagger, iM_3^\dagger)$ also transforms in the same representation under the $SU(2)$ flavor symmetry. This observation tells us what the corresponding states of these hermitian conjugates are (up to an undetermined phase factor):

$$-i(M_1^\dagger + M_2^\dagger) \sim \left( \tilde{f}^T \tau y \tau x \frac{\epsilon}{2} \tilde{f} \right)^* |\tilde{0}\rangle
$$

$$M_1^\dagger - M_2^\dagger \sim \left( \tilde{f}^T \tau y \tau y \frac{\epsilon}{2} \tilde{f} \right)^* |\tilde{0}\rangle
$$

$$iM_3^\dagger \sim \frac{1}{\sqrt{2}} \left( \tilde{f}^T \tau y \tau z \frac{\epsilon}{2} \tilde{f} \right)^* |\tilde{0}\rangle$$

(E.52)

Now we can check the action of $T\sigma^*$ on $M_{1,2,3}$. We have two types of actions of $T\sigma^*$ on the fermions. For the first type:

$$T\sigma^* : \Psi(x,y) \to i\gamma^1 \tau^x \Psi(x,-y)^\dagger$$

(E.54)
we have

\[ M_1 \sim f_{1a}^\dagger f_{1b}^\dagger |0\rangle \rightarrow f_{1a}^* \bar{f}_{1b} f_{1a}^\dagger f_{1b}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle = -f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle \sim M_2^\dagger \]

\[ M_2 \sim f_{2b}^\dagger f_{2a}^\dagger |0\rangle \rightarrow \bar{f}_{1b} f_{1a} f_{1b}^\dagger f_{1a}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle = \bar{f}_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle \sim M_1^\dagger \]

\[ M_3 \sim \frac{1}{\sqrt{2}} \left( f_{1a}^\dagger f_{2b}^\dagger - f_{1b}^\dagger f_{2a}^\dagger \right) |0\rangle \rightarrow \frac{1}{\sqrt{2}} \left( \bar{f}_{2a} - \bar{f}_{2b} \right) f_{1a}^\dagger f_{1b}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle \]

\[ = \frac{1}{\sqrt{2}} \left( f_{1b}^\dagger f_{2a}^\dagger - f_{1a}^\dagger f_{2b}^\dagger \right) |\tilde{0}\rangle \sim -M_3^\dagger \]  \hspace{1cm} (E.55)

In the above we have taken the convention that, under \( T \sigma^* \), \( |0\rangle \rightarrow f_{1a}^\dagger f_{1b}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle \). Notice these transformation rules have a common undetermined phase factor for \( M_{1,2,3} \).

For the second type of \( T \sigma^* \):

\[ T \sigma^* : \Psi(x, y) \rightarrow i\gamma^1 r^z \Psi(x, -y)^\dagger \]  \hspace{1cm} (E.56)

we have

\[ M_1 \sim f_{1a}^\dagger f_{1b}^\dagger |0\rangle \rightarrow f_{1a}^* \bar{f}_{1b} f_{1a}^\dagger f_{1b}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle = -f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle \sim -M_1^\dagger \]

\[ M_2 \sim f_{2b}^\dagger f_{2a}^\dagger |0\rangle \rightarrow \bar{f}_{1b} f_{1a} f_{1b}^\dagger f_{1a}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle = f_{1a}^\dagger f_{1b}^\dagger |\tilde{0}\rangle \sim -M_2^\dagger \]

\[ M_3 \sim \frac{1}{\sqrt{2}} \left( f_{1a}^\dagger f_{2b}^\dagger - f_{1b}^\dagger f_{2a}^\dagger \right) |0\rangle \rightarrow \frac{1}{\sqrt{2}} \left( \bar{f}_{1a} - \bar{f}_{1b} \right) f_{1a}^\dagger f_{1b}^\dagger f_{2a}^\dagger f_{2b}^\dagger |\tilde{0}\rangle \]

\[ = \frac{1}{\sqrt{2}} \left( f_{1b}^\dagger f_{2a}^\dagger - f_{1a}^\dagger f_{2b}^\dagger \right) |\tilde{0}\rangle \sim M_3^\dagger \]  \hspace{1cm} (E.57)

Again, there is a common undetermined phase factor for the transformation rules of \( M_{1,2,3} \).

To summarize, this discussion tells us about the action of \( T \sigma^* \) on the monopole operators in the fermionic critical theory with \( N_f = 2 \). Similar arguments can establish the actions of other symmetries on these monopoles. However, this method leaves an undetermined common phase factor in each transformation of \( M_{1,2,3} \). For unitary symmetries \( T_{1,2} \) and \( C_2 \), we determine this phase factor numerically. For the anti-unitary symmetry \( T \sigma^* \), the current numerical method is insufficient to pin down this phase factor, and we leave it open. The results are listed in Tables E.1 E.2 and E.3. Notice for the fermionic critical theory with \( N_f = 2 \), the quantum numbers of monopoles will not affect the nature of the confined phase at all. Furthermore, in the cases with
the Dirac points located at two generic momenta on the $K - K'$ line or the $M_3 - M_3$ line, where no symmetry allowed fermion bilinear perturbation can destroy the critical point, the unitary symmetries already forbid single monopole operators, while two-fold monopole operators are always symmetry allowed, regardless of what the undetermined phase factors in the actions of $T\sigma^*$ are.
Appendices of chapter 6

F.1 Intertwined chains of cluster state

We present examples where the sub-leading term $\gamma$ has a system-size-dependent oscillation. Recall the local tensor for the cluster state Eq. (6.29):

$$M^{(0)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad M^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} \quad (F.1)$$
where the superscripts are the physical indices. Define a local tensor $K^{(ab)}$ with two physical qubits per site and bond dimension 4 as

$$K^{(ab)} = \left( M^{(a)} \otimes M^{(b)} \right) \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}. \tag{F.2}$$

We trace out the physical qubit $a$ and keep $b$. One can verify that the transfer matrix for $\alpha = 2$ has four nonzero eigenvalues $\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{2}$. Thus, the entanglement entropy is

$$S = -\log \left[ 3 \left( \frac{1}{2} \right)^L + \left( -\frac{1}{2} \right)^L \right] \tag{F.3}$$

under the periodic boundary condition. We omitted Rényi index $\alpha$; in fact, this formula is true for any $0 \leq \alpha \leq \infty$ since the distribution of nonzero Schmidt coefficients (entanglement spectrum) is flat.

This local tensor represents two independent cluster states that are intertwined. When the chain length $L$ is even, there are two symmetries supported entirely on $a$ qubits. But, these symmetries cannot be defined separately when $L$ is odd. The even-odd behavior may be attributed to the extra symmetry when $L$ is even, which are not uniformly on-site.

More generally, one can define an MPS such that the subleading term of the entanglement entropy has periodicity $n$ as a function of chain length $L$ for any positive integer $n$. Eq. (F.2) will corresponds to $n = 2$. The construction is to imagine a helix of $n$ strands, project it into a plane, put a qubit to each outer vertex, and interpret the line between vertices as the bonds of the cluster state. Under periodic boundary condition with length $L$, the number of distinct strands is given by $\gcd(n, L)$. Still, the translation-invariance of the state is observed. The local tensor can be given as

$$K^{(ab)} = \left( M^{(a)} \otimes I^{\otimes (n-2)} \otimes M^{(b)} \right) \mathcal{C} \tag{F.4}$$

$\mathcal{C} = \text{cyclic rotation of tensor factors.}$
The entanglement entropy for any Rényi index \(0 \leq \alpha \leq \infty\) is given by

\[
S = -\log \left[ \sum_{k=1}^{n} c_k \left( e^{2\pi i k/n} \right)^L \right] = L \log 2 - \gcd(L, n) \log 2
\]

\[(F.5)\]

where \(c_k\) are multiplicities of the transfer matrix’ eigenvalues. We do not compute \(c_k\) from the transfer matrix, but they have to be determined by this formula since \(L \mapsto \gcd(L, n)\) is a periodic function. The appearance of \(\gcd\) is because there are \(\gcd(L, n)\) rings of cluster states. This proves an interesting statement that

\[
c_k = \frac{1}{n} \sum_{j=1}^{n} e^{2\pi i jk/n} 2^{\gcd(j, n)}
\]

\[(F.6)\]

are nonnegative integers for all \(k = 1, 2, \ldots, n\).

The examples of this section pose a challenge to unconditionally define the subleading term \(\gamma\). For example, the following limit

\[
\lim_{L \to \infty} S(L) - L(S(L + 1) - S(L))
\]

\[(F.7)\]

does not exist for the MPS state in Eq. (F.2). Even a Kitaev-Preskill-like combination, canceling off the length (area) contribution, will not converge. One might inevitably have to introduce some perturbation to define \(\gamma\), as our examples are not generic in the absence of any symmetry.

### F.2 Time reversal and lattice symmetries

In this appendix, we raise a question whether other symmetries on a 1D chain reduced from a 2D chain can give rise to a robust finite sub-leading term of the entanglement entropy. We consider three kinds of symmetries: the time reversal \(\mathcal{T}\), lattice reflection \(\mathcal{R}\), and lattice inversion \(\mathcal{I}\). For 1D systems the lattice reflection and the inversion may coincide. The reason we are distinguishing the two is that the 1D chain is divided into two parts, upper and lower, across the entanglement cut. Upon the lattice reflection, the upper part and the lower part are not exchanged;
however, by the lattice inversion, which amounts to $\pi$-rotation about a point, the two parts are exchanged.

The scope of this appendix is restricted to situations where the symmetry group $G$ is $G = T$, $G = R$, $G = I$, $G = T \times R$, or $G = T \times I$. The argument below does not apply to a situation where $G = R \times I$ or $G = T \times R \times I$.

We will find symmetry-respecting deformations of states, after which the entanglement entropy becomes

$$S_{\text{new}} = \alpha L \quad \text{(F.8)}$$

for some $\alpha$, or

$$S_{\text{new}} = \text{const.} \geq 0. \quad \text{(F.9)}$$

This will prove that if the non-positive subleading term ($-\gamma$) is robust under those symmetries, it must be zero. This is in contrast to the situation where the 1D state form a nontrivial SPT under a product group of internal symmetry, where a strictly negative subleading term ($-\gamma$) is stabilized by the internal symmetry.

### F.2.1 Lattice Reflection

Recall that our 1D chain has two physical qudits, $a_i$ and $b_i$, at each site $i$. The entanglement cut separates the physical qudits so that all $a_i$ qudits are in one partition and all $b_i$ are in another.

The lattice reflection is realized as

$$\mathcal{R} : a_i \leftrightarrow a_{-i}, \quad b_i \leftrightarrow b_{-i} \quad \text{(F.10)}$$

for all $i = \ldots, -2, -1, 0, 1, 2, \ldots$.

Consider inserting auxiliary qudits $c_i$ in the product state into the chain. Each site is now consisting of $a_i, b_i, c_i$, and we assume that $b_i, c_i$ belong to the same partition with respect to the entanglement cut. Now we introduce the unitary operator $W_i$ on site $i$ that implements the swap between $a_i$ and $c_i$ (see Fig. F.1):

$$W = \sum_{u,v} |v, u\rangle \langle u, v| \quad \text{(F.11)}$$
Figure F.1: Deforming the state while respecting the lattice reflection or inversion symmetry. From the original 1D chain which consists of red ($a_i$) and green ($b_i$) qudits, one can insert an auxiliary yellow ($c_i$) qudit. Then one can apply the swap operator that exchanges the red qudits and yellow qudits circled by the dashed ellipses. This swap operation can be implemented continuously without breaking the lattice reflection symmetry. The numbers in each qudits label the positions of the corresponding qudits before the swap operation. It is understood that we have a 1D chain of qudits, although only a few sites are shown here.

The uniform application $\prod_i W_i$ obviously respects the lattice reflection symmetry $\mathcal{R}$. Since $c_i$ were in the product state, they had no entanglement with the rest. Thus, after the swap operation, the entanglement entropy becomes identically zero across the existing cut. One can implement $W$ smoothly since the unitary group is connected. In this way, we have found a smooth deformation of the state such that the entanglement entropy in the final state is simply zero. In particular, we have smoothly changed the subleading term, if any, to zero.
F.2.2 LATTICE INVERSION WITHOUT TRANSLATION

The lattice inversion is implemented as

\[ \mathcal{I} : a_i \leftrightarrow b_{-i} \]  

(F.12)

for all \( i = \ldots, -1, 0, 1, \ldots \). Similarly as in the previous subsection, by introducing auxiliary qudits in the product state and swap unitary, one can push the physical qudits, expect those at \( i = 0 \), to one side of the entanglement cut, while respecting the lattice inversion symmetry (see Fig.F.1). The deformed state can be viewed as a 1D state where the entanglement cut divides the chain into halves of length \( L/2 \). The entanglement entropy does not depend on the system size (the “area” law of entanglement entropy) and is equal to some positive constant \( h \). If the chain was a nontrivial SPT under this inversion symmetry, such as the Haldane spin-1 chain, then \( h \) cannot be made to become zero; if it was trivial SPT, then a smooth deformation such that \( h \to 0 \) is possible.

If we used Levin-Wen combination to define the subleading term \((-\gamma)\), then

\[ \gamma = S_{AB} + S_{BC} - S_{ABC} - S_B \]  

(F.13)

which is nonnegative by the strong subadditivity. (At this point, we should not use the Rényi entropy, but von Neumann) The deformed state clearly gives \( \gamma = 0 \).

F.2.3 TIME REVERSAL

We assume that the system consists of spin-\( J \)’s, and the time reversal is implemented by

\[ \mathcal{T} = e^{-i\pi J_y} K \]  

(F.14)

for spin systems in the \( J_z \)-basis, where \( K \) is the complex conjugation. We will construct a similar deformation as in the previous cases. We insert auxiliary spins in product states, and swap the spins of the original chain with the auxiliary ones. Complication arises from two sources: The
first one is that a half-integer spin cannot be time-reversal invariant. This is easily resolved by inserting singlets formed by two spins. The second one is that the swap $W$ and its smooth implementation $W(t)$ must commute with $\mathcal{T}$. To resolve the second one, we will shortly prove that there exists $W_2(t)$ for any $J$ such that

$$W_2(t = 0) = I, \quad W_2(t = \pi) = W \otimes W,$$

$$[W_2(t), \mathcal{T}] = 0 \text{ for all } t. \quad (F.15)$$

Equipped with $W_2 = W_2(t)$, we can deform the state so that the final state has entanglement entropy

$$S = \alpha_J L \quad (F.16)$$

exactly without any subleading term, where $\alpha_J = \log(2J + 1)$ is the entanglement entropy of a singlet consisting of two spin-$J$’s. To see this, insert singlets $a'_i$ to the partition where $a_i$ belong, and another set of singlets $b'_i$ to the partition where $b_i$ belong. Note that each of $a'_i$ or $b'_i$ consists of two spins, whereas each of $a_i$ or $b_i$ consists of one spin. Apply $W_2$ such that the pair of $a_i$ and one auxiliary spin from $a'_i$ is exchanged with the whole singlet $b'_i$. See Figure F.2. The original $b_i$ is not moved at all, and $a_i$ is brought to the partition where $b_i$ belongs. The singlet $b'_i$ is moved to the opposite partition, and the singlet $a'_i$ is now shared between the entanglement partitions. Thus, the entanglement entropy of the deformed state entirely comes from the singlets $a'_i$, and Eq. (F.16) holds.

Figure F.2: Deforming state while respecting symmetry. In addition to the red ($a_i$) and green ($b_i$) qudits, a pair of time reversal-invariant spin singlets are inserted in each site (the yellow qudits are $a'_i$ and blue qudits are $b'_i$). The swap unitary is applied to the qudits circled by the dashed ellipses, so that the entanglement across the cut solely arises from the inserted singlet. The swap can be implemented continuously during which the time-reversal symmetry is unbroken. It is understood we have a 1D chain of qudits, although only one site is shown here.
Remark that this time-reversal invariant deformation respects lattice reflection and translation symmetry, if they were present in the original state. The deformation using $W_2$ can also be adapted to a situation where there is a lattice inversion symmetry.

We now construct the promised $W_2(t)$. In the basis where $J_z$ is diagonal, we will show that there exists a real orthogonal matrix $W_2(t)$ such that it commutes with $(e^{-i\pi J_y})^{\otimes 4}$, and it smoothly interpolates between the identity and $W \otimes W$. Observe that $R = e^{-i\pi J_y}$ is a real matrix since $J_y = (J_+ - J_-)/2i$ is purely imaginary. Since $J_y$ is hermitian, we have $R^T = R^\dagger = R^{-1}$. Moreover, $R^2$ is $+1$ for integer spins or $-1$ for half-integer spins. Therefore, $(R \otimes R)^2 = R^2 \otimes R^2 = 1$. It follows that $R^{\otimes 2}$ is real symmetric with eigenvalues $\pm 1$. The swap matrix $W$ is obviously real symmetric and squares to 1. Since $W$ and $R^{\otimes 2}$ commute, they can be simultaneously diagonalized by a real orthogonal matrix. Likewise, $W^{\otimes 2}$ and $R^{\otimes 4}$ can be simultaneously diagonalized, and there exists a real orthonormal basis $|w = \pm 1, r = \pm 1, k\rangle$ consisting of common eigenvectors of $W^{\otimes 2}$ and $R^{\otimes 4}$, where $w, r$ are the eigenvalue of $W^{\otimes 2}$ and $R^{\otimes 4}$, respectively, and the index $k$ runs from 1 to the degeneracy $k_{w,r}$ of the common eigenspace. We claim that both $k_{-1,1}$ and $k_{-1,-1}$ are always even. Given this claim, we can construct $W_2(t)$ by

$$W_2(t)|_{\text{span}(|-1, r, 2m-1, |1, r, 2m\rangle)} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix} \quad (F.17)$$

for $m = 1, \ldots, k_{-1,r}/2$, and the identity on $w = +1$ subspace. The constructed $W_2(t)$ is clearly real orthogonal, and commutes with $R^{\otimes 4}$ since it preserves the eigenspaces of $R^{\otimes 4}$. We have $W_2(0) = I$ by definition, and $W_2(\pi) = +1$ on the $w = +1$ subspace and $W_2(\pi) = -1$ on the $w = -1$ subspace; hence, $W_2(\pi) = W^{\otimes 2}$.

It remains to compute the degeneracy $k_{-1,r}$ to show that it is even. Let $\{|a\rangle : a = 1, \ldots, 2J+1\}$ be a complete orthonormal set of eigenvectors of $R$. These $|a\rangle$ may not be real vectors, but the degeneracy of eigenspaces can be computed with respect to any basis we choose. Then, $(|ab\rangle \pm |ba\rangle)/\sqrt{2}$ are complete common eigenvectors of $R \otimes R$ and the swap $W$. Then, the $(-1)$-eigenvectors of $W^{\otimes 2}$ in an eigenspace of $R^{\otimes 4}$ are

$$\frac{(|ab\rangle - |ba\rangle)(|cd\rangle + |dc\rangle)}{2}, \quad \frac{(|ab\rangle + |ba\rangle)(|cd\rangle - |dc\rangle)}{2} \quad (F.18)$$

326
which always come in pairs. The degeneracy is 
\[ k_{-1,r} = 2^{N(N-1)/2} \frac{N(N+1)}{2} = \frac{1}{2} N^2 (N + 1)(N - 1), \]
an even number for any \( N = 2J + 1 \).

Without introducing \( W_2 \), a continuous real implementation of \( W \) alone from the identity is not possible. For spin-\( \frac{1}{2} \), the swap \( W \) has determinant \(-1\), so \( W \) belongs to the non-identity component of \( O(4) \).

### F.3 Cylinder extrapolation method on the 2D cluster state at nonzero temperature

Here we calculate the mutual information for the cluster state across the circumferential cut of a cylinder. We consider the geometry of Fig. 6.3. The mutual information

\[ I(A : B) = S(A) + S(B) - S(AB) \]  
(F.19)

is preferred to the entropy \( S(A) \) (or \( S(B) \)) because it obeys an area law even at finite temperatures. \([289,304]\)

The Hamiltonian is given by Eq. (6.8). By definition of the entropy, the mutual information is invariant under local unitary in either region \( A, B \), and it is oblivious to any tensor product factor. Hence, by the same argument as for the ground state of the cluster state, the mutual information of the 2D cluster state reduces to that of mutual information of the 1D cluster state with the extensive bipartition. The Gibbs state at inverse temperature \( \beta \) is given by

\[ \rho_{AB} = Z^{-1} \prod_{j}^{2L} \exp (\beta \sigma_j^z \sigma_{j+1}^z) \]  
(F.20)

\[ = Z^{-1} \prod_{j}^{2L} (I \cosh \beta + \sigma_j^z \sigma_{j+1}^z \sinh \beta). \]  
(F.21)

The partition function \( Z \) is equal to that of uncoupled \( 2L \) spins in a magnetic field,

\[ Z = 2^{2L} \cosh^{2L} \beta, \]  
(F.22)

and the spectrum of \( \rho_{AB} \) is the tensor product of \( 2L \) identical spectra \( \{ \frac{1}{2}(1 \pm \tanh \beta) \} \). Hence,
the von Neumann entropy is

$$S(AB) = 2L f \left( \frac{1 + t} {2} \right)$$  \hspace{1cm} (F.23)

where $t = \tanh \beta$ and $f(x) = -x \log x - (1 - x) \log(1 - x)$ is the binary entropy function.

Generalizing the result in Sec. 6.5.1, we get the reduced density matrix for $A$

$$\rho_A = \Tr_B(\rho_{AB}) = \frac{1}{2L} \left( I + t^L \prod_{j \in A} \sigma_j^z \right),$$  \hspace{1cm} (F.24)

whose entropy is

$$S(A) = 2^{L-1} \left( \frac{1 + t^L} {2L} \log \frac{2L} {1 + t^L} + \frac{1 - t^L} {2L} \log \frac{2L} {1 - t^L} \right)$$

$$= L \log 2 - \frac{1}{2} \log(1 - t^{2L}) - \frac{t^L} {2} \log \frac{1 + t^L} {1 - t^L}.$$  \hspace{1cm} (F.25)

Therefore, the mutual information is

$$I(A : B) = 2L (\log 2 - f((1 + t)/2))$$

$$- \log(1 - t^{2L}) - t^L \log \frac{1 + t^L} {1 - t^L}$$

$$= 2\alpha_t L - t^{2L} + O(t^{4L})$$  \hspace{1cm} (F.27)

where $\alpha_t = O(t^2)$ for small $t = \tanh \beta$ and $\alpha_t \sim \log 2$ for $t \sim 1$.

In conclusion, at any finite $\beta$, the subleading term is exponentially small in $L$. Note that for any $0 < \beta \leq \infty$, the usual correlation length of the cluster state is zero since it differ from a product Gibbs state by a quantum circuit of depth 2. So, the length scale of the subleading term is greater than the usual correlation length, and is diverging as $\beta \to \infty$. 

328
G.1 Some remarks on time reversal and charge-conjugation symmetry

In this appendix we discuss some general properties of time reversal and charge-conjugation symmetry. In particular, we would like to check the values of $\mathcal{T}^2$ and $\mathcal{C}^2$ on a fractionalized excitation. By abuse of notation, we will denote this excitation by $E$ in general, which includes but does not limit to the case where $E$ is the electric charge of a $U(1)$ quantum spin liquid. In general, $E$ can be a multicomponent object, and let us denote the $i$th component as $E_i$.

We first consider time-reversal symmetry $\mathcal{T}$. We assume the anti-unitary time reversal symmetry acts on $E_i$ as

$$\mathcal{T} : E_i \rightarrow U_{ij} E_j$$  \hspace{1cm} (G.1)
with $U$ a generic matrix. That is, time reversal changes $E$ to something that differs from it only by a local operator. We also assume $E_i^\dagger M_{ij}E_j$ is always a local object, for any matrix $M$. Notice in the above notation, some components of $E$ can be a bound state of the fractionalized excitation $E$ and some local particles.

Straightforward algebra indicates that acting time reversal twice, the original $E$ becomes

$$E_i \rightarrow (U^*U)_{ij}E_j \quad (G.2)$$

and the local object becomes

$$E_i^\dagger M_{ij}E_j \rightarrow E_i^\dagger((U^*U)^\dagger M(U^*U))_{ij}E_j \quad (G.3)$$

Suppose the system is made of Kramers singlets. That is, $E_i^\dagger M_{ij}E_j$ is invariant upon acted by time reversal twice for any $M$, which is equivalent to that

$$(U^*U)^\dagger M(U^*U) = M \quad (G.4)$$

for any matrix $M$. This is possible only if $U^*U = e^{i\phi}I$, where $I$ is the identity matrix. Because $(U^*U)^2 = e^{2i\phi}I$ and $U^*UU^*U = U^*(UU^*)U = I$. This implies $e^{i\phi} = \pm 1$. That is, $T^2$ acting on such fractionalized excitations must give $\pm 1$. [241]

The above argument also shows even if there is microscopic Kramers doublet in the system, as long as time reversal takes the form $(G.1)$ and $T^2$ is well-defined, $T^2$ can only be $\pm 1$. However, if there are microscopic Kramers doublets in the system, it is possible to have $T^4 = -1$. Notice in the case where $T^4 = -1$, $T^2$ does not have to be well-defined.

Now $T^4 = -1$ (or $T^2 = \pm i$) can happen if time reversal changes the relevant excitation by a nonlocal operation, and a typical example for this case is $(eCmC)^T\epsilon$ mentioned in the main text, where under time reversal $e$ and $m$ are exchanged. It can also happen if time reversal changes this excitation by a local operation. In this case, time reversal has to attach a local Kramers doublet to the relevant nontrivial excitation. To see this, without loss of generality, let us assume the time reversal partner of $E$ is $F$, i.e., $E \rightarrow F$ under time reversal. To have $T^2 = \pm i$ for $E$, we
need $F \to \pm iE$ under time reversal. These also imply $T^2 = \mp i$ for $F$, so $E$ and $F$ defer by a local Kramers doublet.

We now consider charge-conjugation symmetry $C$. In general it is a unitary symmetry acting on $E_i$ as

$$C : E_i \to V_{ij} E_j^\dagger,$$

where $V$ is a matrix. Taking the hermitian conjugate of the above equation, we have

$$C : E_i^\dagger \to V_{ij}^* E_j.$$

So $C$ acting twice should give

$$C^2 : E_i \to (VV^*)_{ij} E_j.$$

Again we require the local operator $E_i^\dagger M_{ij} E_j$ to have $C^2 = 1$, for any matrix $M$. Following the same logic as we did for time-reversal symmetry, we conclude that $C^2 = \pm 1$ on the fractionalized excitation $E$.

Also notice that the value of $C^2$ is invariant under a $U(1)$ gauge transform $U_\theta$, namely $(U_\theta C)^2 = C^2$. This means that the value $C^2 = \pm 1$ is a physically meaningful quantity.

The simple discussion here on $C^2$ should be enough for our purpose. In the context of gapped topological orders (for example in $Z_{2N}$ gauge theories), the mathematically more precise meaning of $C^2$ on fractionalized excitations has been discussed in Ref. [317].

G.2 An SPT: $eCm_{\frac{1}{2}}$

In this appendix we describe a 3D SPT, $eCm_{\frac{1}{2}}$, under symmetry $U(1) \times SO(3)$, $U(1) \times \mathcal{T} \times SO(3)$ or $(U(1) \times \mathcal{T}) \times SO(3)$. The defining property of this SPT is that it can have a symmetric surface $Z_2$ topological order with excitations $\{1, e, m, \epsilon\}$, where $e$ carries charge-$1/2$ under $U(1)$ and $m$ carries spin-$1/2$ under $SO(3)$. We begin by giving field theoretic descriptions of this surface state with the above symmetries.
If the symmetry is simply $U(1) \times SO(3)$, a field theory for this surface can be described by the following Lagrangian:

$$
\mathcal{L} = \sum_{s=\pm} \left| (\partial_\mu - i a_\mu) z_s \right|^2 + V(|z|^2) + \frac{1}{4e^2} (\epsilon^{\mu\nu\lambda} \partial_\nu a_\lambda)^2 - \frac{1}{2\pi} A_d a
$$

(G.8)

where $A_d$ is a shorthand for $\epsilon^{\mu\nu\lambda} A_\mu \partial_\nu a_\lambda$. $z_s$ is a two-component complex field that transforms as a doublet under $SO(3)$, $a_\mu$ is a non-compact $U(1)$ gauge field, and $A$ is a background $U(1)$ gauge field corresponding to the global $U(1)$ symmetry.

Condensing a bound state of two $z$’s in the singlet channel (i.e. letting $\langle z_1 \partial_x z_2 - z_2 \partial_x z_1 \rangle \neq 0$ for example) gives us the above surface topological order, where the uncondensed single $z$ will be identified as $m$ that carries spin-1/2. After this condensation, the flux of $a$ is quantized in unit of $\pi$, and the last term in the above Lagrangian implies that this $\pi$-flux carries charge-1/2 under the global $U(1)$. This $\pi$-flux can be identified as $e$. The resulting state is precisely $eCm^1_2$.

If the symmetry is $(U(1) \ltimes T) \times SO(3)$, the surface theory of $eCm^1_2$ can still be described by the field theory given by (G.8), but now the spin operator is represented as $S \sim z^\dagger \sigma i \partial_z z$ and under time reversal

$$
z \rightarrow z^\dagger, \quad a \rightarrow a, \quad a_0 \rightarrow -a_0
$$

(G.9)

In order to obtain $eCm^1_2$, we need to make $\langle z_1 \partial_x z_2 - z_2 \partial_x z_1 \rangle = \langle z_1^\dagger \partial_x z_2^\dagger - z_2^\dagger \partial_x z_1^\dagger \rangle \neq 0$.

If the symmetry is $U(1) \times T \times SO(3)$, the surface state of $eCm^1_2$ can be described by a field theory similar to (G.8):

$$
\mathcal{L} = \sum_{s,\alpha=\pm} \left| (\partial_\mu - i a_\mu) z_{s\alpha} \right|^2 + V(|z|^2) + \frac{1}{4e^2} (\epsilon^{\mu\nu\lambda} \partial_\nu a_\lambda)^2 - \frac{1}{2\pi} A_d a
$$

(G.10)

Notice now each component of $z_s$ contains two components, $z_{s\alpha}$ with $\alpha = \pm$, and the generators
of spin rotations become \( S = \frac{1}{2} \bar{z}_{s\alpha} \tau_{ss'} z_{s'\alpha} \). Under time reversal

\[
z_{s\alpha} \rightarrow (\sigma_2)_{\alpha \alpha'} (\tau_2)_{ss'} z_{s'\alpha'}, \quad a \rightarrow -a, \quad a_0 \rightarrow a_0
\]

where \( \sigma \) and \( \tau \) are the standard Pauli matrices. The reason to give more components to \( z \) is to make it not a Kramers doublet. To get \( eCm\frac{1}{2} \), we can also condense the bound state of two \( z \)'s in the singlet channel, that is, let \( \langle z^T \sigma_1 \tau_2 \partial_z z \rangle \neq 0 \). Similar argument as above implies the resulting state is \( eCm\frac{1}{2} \).

To see that this state as a strictly 2D system is anomalous, consider tunneling a \( U(1) \) monopole through this 2D system. This will leave a \( 2\pi \) flux on the system. For such a local process, no excitations far away should be able to tell the existence of this \( 2\pi \)-flux. But \( e \) carries half charge under \( U(1) \), it will pick up a nontrivial phase factor upon circling around this \( 2\pi \)-flux, regardless how far it locates away from it. To cancel this phase factor, an \( m \) needs to be present at the \( 2\pi \)-flux. Because \( m \) carries spin-1/2, this then implies tunneling a monopole leaves a spin-1/2 on the surface. This is not possible for a strictly 2D system with symmetry \( U(1) \times SO(3) \). Notice that time reversal symmetry is not involved in the anomaly, so this surface is still anomalous even if the symmetry is \( U(1) \times SO(3) \).

To visualize this SPT, the simplest way is to do a layer construction similar to that used in Ref. Wang2013. Because similar method will be used in Appendix G.3 to construct \( 4 + 1 \)-d systems whose surfaces realize the anomalous quantum spin liquids, we do not explicitly display it for \( eCm\frac{1}{2} \) here.

Notice when the symmetry is \((U(1) \rtimes T) \times SO(3)\), to realize \( eCm\frac{1}{2} \), we have assumed that the microscopic bosons are Kramers singlet. Below we argue that for microscopic Kramers doublet charged bosons, \( eCm\frac{1}{2} \) cannot be realized. This fact is important, because otherwise gauging \( eCm\frac{1}{2} \) in such a system would lead to \( E_{bT} M_{b\frac{1}{2}} \), which is argued to be anomalous in Sec. 7.3.

Suppose \( eCm\frac{1}{2} \) can be realized in a system made of Kramers doublet charged bosons. Fusing two \( e \)'s gives a charge-1 local particle, which must be a Kramers doublet by assumption. Given the excitation content of this theory and that time reversal keeps the \( U(1) \) charge, the time re-
versal action on $e$ can always be represented as

$$e_i \rightarrow U_{ij} e_j \quad (G.12)$$

Now notice $e_i M_{ij} e_j$ is a local charge-1 operator for any matrix $M$, so this operator must be a Kramers doublet, which implies that

$$(U^* U)^T M (U^* U) = -M \quad (G.13)$$

This is possible only if $U^* U = \pm i$. As shown in Appendix G.1, no matrix $U$ can have this property. This implies that $e C m^1_2$ cannot be realized in a system made of Kramers doublet charged bosons.

**G.3 Anomalous spin liquids as surface states of some 4 + 1-D systems**

In this appendix we will show the anomalous spin liquids can be obtained on the surface of some 4 + 1-d systems. The simplest way to construct these 4 + 1-d surface states is the following layer construction, which has been widely used to construct topological states [316, 321, 340, 405].

For example, to construct a 4 + 1-d system whose surface can realize $E_{b T} M_{b \frac{1}{2}}$, one can start by stacking alternating layers of non-anomalous spin liquids $E_{b T} M_b$ and $E_{b \frac{1}{2}} M_b$ (see Fig. G.1). Then on the $i$th, $i + 1$th and $i + 2$th layers, one can condense the bound state of $E_i$, $M_{i+1}$ and $E_{i+2}$ with the subscript indicating the layer index. This bound state, $B_i = E_i M_{i+1} E_{i+2}$, is a trivial boson, and $B_i$’s with different $i$’s commute, so they can be simultaneously condensed without breaking any symmetry. After this condensation, the gauge field in the 4 + 1-d bulk will be confined (or Higgsed) and this bulk becomes short-range entangled, but on the surface some non-trivial excitations still survive. These survivors are $E_1$ and $M_1^T E_2$ on the top surface, and $E_N$ and $M_{N-1}^T M_N^T$ on the bottom surface. Now the top (bottom) surface realizes $E_{b T} M_{b \frac{1}{2}}$, and $E_1 (E_{N-1} M_N^T)$ and $M_1^T E_2 (E_N)$ can be viewed as the electric charge and magnetic monopole, respec-
tively. The 4 + 1-d system constructed here is an SPT under symmetry \(SO(3) \times T\) because its surface, \(E_{bT} M_{b\frac{1}{2}}\), is anomalous.

To obtain 4+1-d systems whose surface realize all other anomalous spin liquids, one only needs to replace each layer by the appropriate non-anomalous spin liquid and condense the proper bound states.

For some spin liquids, there is a more isotropic construction of the corresponding 4 + 1-d systems by using a non-linear Sigma model (NLSM) with appropriate topological terms and anisotropies, similar to that used in Ref. Bi2015. For example, to construct the corresponding 4 + 1-d bulk of \(E_{bT} M_{b\frac{1}{2}}\), consider a 4 + 1-d \(O(6)\) NLSM with a theta-term at \(\theta = 2\pi\). Its surface theory is a 3 + 1-d six-component NLSM with a Wess-Zumino-Witten (WZW) term at level-1, with Lagrangian

\[
\mathcal{L} = \frac{1}{g} (\partial^\mu n_a)^2 + \frac{2\pi i}{\Omega_5} \int_0^1 du \epsilon^{abcdef} n_a \partial u n_b \partial x n_c \partial y n_d \partial z n_e \partial \tau n_f
\]

with \(\Omega_5\) the surface area of a five-dimensional unit sphere. The six-component vector transforms under time reversal as

\[
n_{1,2,3} \rightarrow -n_{1,2,3} \quad n_{4,5,6} \rightarrow n_{4,5,6}
\]

This theory is invariant under \(O(6) \times T\).

To see how \(E_{bT} M_{b\frac{1}{2}}\) can be accessed by the above theory, let us first add some \(SO(3) \times SO(3)\) anisotropy, such that the first (second) three components transform as a vector under the first (second) \(SO(3)\). Consider the weak coupling limit of the theory where both the first and the second three components are ordered. Now disorder the second three components by proliferating its hedgehog defects. In this way, the second three components themselves form a trivial state that preserves the second \(SO(3)\) symmetry. Due to the WZW term, the hedgehog defects of the first three components carry spin-1/2 under the second \(SO(3)\), and it will be identified as the magnetic monopole of \(E_{bT} M_{b\frac{1}{2}}\) later.

Now disorder the first three components by proliferating the spin wave excitations while keep-
ing its hedgehog defects gapped. In this way we will get a $U(1)$ spin liquid. One way to see this is to write the first three components in the $CP^1$ representation,

$$n_a = z_\alpha^\dagger \sigma_\alpha^{a\beta} z_\beta, \quad \text{for } a = 1, 2, 3 \quad (G.16)$$

where $z$ is a two-component complex spinon field with $|z_1|^2 + |z_2|^2 = 1$, and $\sigma$’s are the standard Pauli matrices. Under the previous defined time reversal and the first $SO(3)$, the spinon field transforms as a Kramers doublet and $SU(2)$ doublet. This spinon will be identified as the electric charge of $E_{bT_1}M_{b_1}^{1/2}$ later.

The Lagrangian that only involves the first three components can now be written as the following gauge theory

$$\mathcal{L}' = |(\partial_\mu - ia_\mu)z|^2 + V(|z|^2) + \frac{1}{4e^2}(\epsilon^{\mu\nu\lambda}\partial_\nu a_\lambda)^2 \quad (G.17)$$

where $a_\mu$ is an emergent $U(1)$ gauge field due to the $U(1)$ gauge redundancy in (G.16), i.e., $n_a$ is invariant when $z \rightarrow z e^{i\theta}$ for any real $\theta$. The ordered state of the first three components corresponds to the Higgs phase of the $U(1)$ gauge theory, and proliferating spin wave excitations corresponds to making spinons gapped and give rise to a $U(1)$ spin liquid, where the gapped spinons are the electric charge. The monopole of this $U(1)$ spin liquid, which is the source of magnetic flux, should be identified as the un-proliferated hedgehog defect, which is the source of the skyrmions.

Finally, adding a weak anisotropy to collapse the $SO(3) \times SO(3)$ symmetry to its diagonal $SO(3)$ subgroup, we get $E_{bT_1}M_{b_1}^{1/2}$ with $SO(3) \times T$ symmetry, where the electric charge is a Kramers doublet and $SU(2)$ doublet, and the magnetic monopole is an $SU(2)$ doublet.

The construction above gives a 4 + 1-d system whose surface realizes $E_{bT_1}M_{b_1}^{1/2}$. If time reversal is ignored, the above construction gives the 4 + 1-d system whose surface realizes $E_{b_1}M_{b_1}^{1/2}$. The 4 + 1-d system whose surface can realize $E_{bT}M_{b_1}^{1/2}$ can be obtained similarly, where time reversal acts in the same way as before, while under $SO(3)$ only the last three components transform as a vector and the first three components do not transform.

Notice in the construction based on NLSMs, even if $SO(3)$ is broken to $Z_2 \times Z_2$, all components
still transform nontrivially under this symmetry. Then it is believed that the constructed $4 + 1$-d states are still nontrivial SPTs. This motivates us to conjecture that even if the symmetry is broken to $Z_2 \times Z_2 \times T$, the descendants of the anomalous states remain anomalous because they still live on the surface of some SPTs.

G.4 Classification of some SPTs

In this appendix we classify some SPTs which, once gauged, can become some of the $U(1)$ quantum spin liquids studied in the main text.

G.4.1 Bosonic SPT with symmetry $(U(1) \times T) \times SO(3)$

We start with bosonic SPT with symmetry $(U(1) \times T) \times SO(3)$, where the microscopic boson is a Kramers singlet. Without $SO(3)$ symmetry, the classification of this SPT is well established. They are classified by $\mathbb{Z}_3^2$, where the three root states are $eCmC$, $eTmT$ and $efmf$. [337] With $SO(3)$ symmetry, Appendix G.2 shows there is another root state: $eCm_2^1$. In fact, there two more root states: $e_7^1mT$ and $e_7^1m_2^1$. That these two are nontrivial SPTs can be inferred from the classification of bosonic SPT with symmetry $U(1) \times T$. Indeed, once $SO(3)$ is broken to $U(1)$, $e_7^1mT$ and $e_7^1m_2^1$ become $eCmT$ and $eCmC$, respectively.

Therefore, we propose the classification of these SPTs is $\mathbb{Z}_2^6$. Notice that among the six root states, only two of them need protection from the $U(1)$ symmetry: $eCmC$ and $eCm_2^1$.

G.4.2 Bosonic SPT with symmetry $U(1) \times T \times SO(3)$

If the symmetry is $U(1) \times T \times SO(3)$, the understanding of bosonic SPTs with symmetry $U(1) \times T$ implies there is one more root state: $eCmT$. This state is protected by both $U(1)$ and time reversal.

Therefore, we propose the classification of these SPTs is $\mathbb{Z}_2^7$. The properties of the surface $Z_2$ topologically ordered states of the root states are summarized in Table G.1.
Table G.1: Surface $Z_2$ topological ordered states of SPTs under symmetry $U(1) \times T \times SO(3)$. The topological sectors are denoted by $\{1, e, m, \epsilon\}$. $q_e$ and $q_m$ represents the charge of $e$ and $m$ under $U(1)$, $T^2_e$ and $T^2_m$ represents the Kramersness of $e$ and $m$ under time reversal, and $S_e$ and $S_m$ represents the spin of $e$ and $m$, respectively. If the symmetry is $(U(1) \times T) \times SO(3)$, $eCmT$ will be absent and all other six root states remain.

<table>
<thead>
<tr>
<th></th>
<th>$q_e$</th>
<th>$q_m$</th>
<th>$T^2_e$</th>
<th>$T^2_m$</th>
<th>$S_e$</th>
<th>$S_m$</th>
<th>comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$eCmC$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$eTmT$</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$efmf$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>e and m are fermions</td>
</tr>
<tr>
<td>$eCm\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td></td>
</tr>
<tr>
<td>$e\frac{1}{2}m\frac{1}{2}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td></td>
</tr>
<tr>
<td>$e\frac{1}{2}mT$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>$\frac{1}{2}$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$eCmI$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

G.4.3 SPT with symmetry $((U(1) \times SU(2))/Z_2) \times T$ of fermions

For fermionic SPT with symmetry $((U(1) \times SU(2))/Z_2) \times T$, free fermion band theory gives a $\mathbb{Z}$ classification, and each state can be labelled by an integer $k$, which is basically the number of pairs of massless Dirac fermions on the surface.

The root state can have a surface state with two massless Dirac fermions with the following Hamiltonian

$$H = \psi^\dagger (-i \partial_x \sigma_x - i \partial_y \sigma_y) \otimes \tau_0 \psi$$  \hspace{1cm} (G.18)

where $\sigma$ and $\tau$ are the standard Pauli matrices with $\sigma_0 = \tau_0 = I$, and $\sigma$ acts on the internal indices of the Dirac fermions and $\tau$ acts on the spin indices. Under $U(1)$,

$$\psi \rightarrow \psi e^{i\theta}$$  \hspace{1cm} (G.19)

Under $SU(2)$,

$$\psi \rightarrow \sigma_0 \otimes U \psi$$  \hspace{1cm} (G.20)
with $U$ an $SU(2)$ matrix. And under time reversal

$$\psi \rightarrow i\sigma_y \otimes \tau_0 \psi^\dagger$$  \hspace{1cm} (G.21)

Notice the inverse of this root state, i.e. the state that can trivialize the root state when coupled together, can have the same surface Hamiltonian as this root state except that time reversal acts as $\psi \rightarrow -i\sigma_2 \otimes \tau_0 \psi^\dagger$. This means the state labelled by $k$ and that labelled by $-k$ are identical after gauging the $U(1)$, because the aforementioned sign difference in the time reversal action can be eliminated by a $U(1)$ gauge transformation. \cite{241}

When the $U(1)$ symmetry is gauged, the monopole of the corresponding $U(1)$ gauge field is a Kramers doublet that carries spin-1/2. \cite{393,406} Because the method that leads to this result will be used extensively later, it is helpful to review it here.

Since the surface is described by two free Dirac fermions, which is a conformal field theory, one can use state-operator correspondence to study the properties of monopoles, by imagining putting the surface on a sphere with $2\pi$ flux threading out. Guaranteed by the index theorem, each Dirac fermion will contribute a zero mode in the background of the $2\pi$ flux, in our case denoted by $f_1$ and $f_2$, respectively. We also denote the flux background with both zero modes empty by $|0\rangle$. Because the time reversal symmetry flips the $U(1)$ charge here, the physical gauge invariant states must have one of the zero modes being occupied. That is, it should be $f_1^\dagger |0\rangle$ or $f_2^\dagger |0\rangle$, which are bosonic. In light of state-operator correspondence, these two states correspond to two different charge-neutral monopole operators, denoted by $M_1$ and $M_2$, respectively. Also, $|0\rangle$ corresponds to the operator of the $(-1,1)$ dyon, and $f_1^\dagger f_2^\dagger |0\rangle$ corresponds to the operator of $(1,1)$ dyon. Then the quantum numbers of the monopole can be read off from the properties of these states.

For example, for the surface theory described above, because the two Dirac fermions transform as spin-1/2 under $SU(2)$, the monopoles $M_1 \sim f_1^\dagger |0\rangle$ and $M_2 \sim f_2^\dagger |0\rangle$ also transform as spin-1/2.
As for time reversal, these two states transform as

\[ M_1 \sim f_1^\dagger |0\rangle \rightarrow f_1 f_1^\dagger f_2^\dagger |0\rangle = f_2^\dagger |0\rangle \sim M_2 \]
\[ M_2 \sim f_2^\dagger |0\rangle \rightarrow f_2 f_2^\dagger f_2^\dagger |0\rangle = -f_1^\dagger |0\rangle \sim -M_1 \]

(G.22)

This means the monopoles are Kramers doublet under time reversal. Therefore, after gauging the \( U(1) \) symmetry, this state becomes \( E_{BT_1^2} M f_1^\dagger \).

Now we turn to the classification of such fermionic SPTs. Upon adding interactions, the free fermion classification collapses to \( \mathbb{Z}_4 \). \[406\] It can be shown that the state with 8 massless Dirac fermions on the surface is trivial, and the state with 4 massless Dirac fermions on the surface is equivalent to \( eTmT_1 \), a bosonic SPT with symmetry \( SO(3) \times T \). There can be interacting SPTs beyond band theory, which can be viewed as bosonic SPTs with symmetry \( SO(3) \times T \). They are classified by \( \mathbb{Z}_2^4 \). One of the root states of these bosonic SPTs coincide with a free fermion SPT that can have 4 massless Dirac fermions on the surface, so we propose the complete classification is \( \mathbb{Z}_4 \times \mathbb{Z}_2^3 \).

G.4.4 SPT with symmetry \(((U(1) \times T) \times SU(2))/\mathbb{Z}_2)\) of Kramers singlet fermions

Consider fermionic SPT with symmetry \(((U(1) \times T) \times SU(2))/\mathbb{Z}_2\) and assume \( T^2 = 1 \) for these fermions. Free fermion band theory gives a \( \mathbb{Z}_2 \) classification. The root state can have a surface state with two massless Dirac fermions, described by the same Hamiltonian as (G.18), with the only difference that under time reversal

\[ \psi \rightarrow \sigma_y \otimes \tau_y \psi \]

(G.23)

When the \( U(1) \) symmetry of these Dirac fermions is gauged, the monopole of the corresponding \( U(1) \) gauge field carries spin-1/2. \[331\] This can also be seen by using the method of state-operator correspondence reviewed above. Notice in this case the time reversal symmetry does not flip the \( U(1) \) charge, so it is convenient to momentarily suppose equipping the system with a further charge conjugation symmetry. We will determine the properties of the monopoles in the presence of this further symmetry first, and then break this symmetry. Because the properties of
the monopoles are described by some discrete data, breaking this symmetry will not change any of them.

Again, each Dirac fermion will contribute a zero mode to the $2\pi$ flux background, and the neutral bosonic monopoles correspond to the two states with one zero mode occupied: $M_{1,2} \sim f_{1,2}^\dagger |0\rangle$. Because the two Dirac fermions carry spin-1/2 under $SU(2)$, as above, the monopoles must also carry spin-1/2. It is not meaningful to discuss whether monopoles are Kramers doublet or not under time reversal, so this finishes determining the properties of the monopoles. From this discussion, we see that after gauging the $U(1)$ symmetry this state becomes $E_{f_1^\dagger M_{b_2}^\dagger}$.

Although it is not meaningful to discuss whether monopoles are Kramers doublet or not under time reversal, it is interesting and helpful to determine how monopoles transform under time reversal. To be consistent with that time reversal and $SU(2)$ commute on the monopole, under time reversal the monopole operators must transform as $M_{1,2} \rightarrow M_{1,2}^\dagger$ (with a possible phase ambiguity).

How do we understand this time reversal action on monopoles from the point of view of state-operator correspondence? This is a little tricky because in this case the $2\pi$ flux will be turned into a $-2\pi$ flux under time reversal, which also has two zero modes, denoted by $\tilde{f}_{1,2}$, where $\tilde{f}_{1,2}$ is contributed by $\psi_{1,2}$, respectively. In particular, denote $|\tilde{0}\rangle$ as the state with $-2\pi$ flux background and neither zero mode occupied. This is the time reversal partner of $|0\rangle$, so it corresponds to the $(-1,-1)$ dyon. Similarly, $\tilde{f}_1^\dagger \tilde{f}_2^\dagger |\tilde{0}\rangle$ is the time reversal partner of $f_1^\dagger f_2^\dagger |0\rangle$, so it corresponds to the $(1,-1)$ dyon.

Under time reversal,

\begin{align}
  f_1^\dagger |0\rangle &\rightarrow \tilde{f}_2^\dagger |\tilde{0}\rangle \\
  f_2^\dagger |0\rangle &\rightarrow -\tilde{f}_1^\dagger |\tilde{0}\rangle
\end{align}

(G.24)

where an unimportant $U(1)$ phase factor has been suppressed. For this to be compatible with that $M_{1,2} \rightarrow M_{1,2}^\dagger$ under time reversal, we must identify (with an unimportant phase factor)

\begin{align}
  M_1^\dagger \sim \tilde{f}_2^\dagger |\tilde{0}\rangle, \quad M_2^\dagger \sim -\tilde{f}_1^\dagger |\tilde{0}\rangle
\end{align}

(G.25)
To the best of our knowledge, this identification of the Hermitian conjugate of the monopoles in the context of state-operator correspondence has not been given before, and it will be used later. We remark that this identification is true as long as the long-distance conformal field theory is described by two massless Dirac fermions, and it should be independent of the microscopic symmetries of the system, although we obtained it by considering a system with a particular symmetry.

Now we return to the classification of these fermionic SPTs. Upon adding interaction, the nontrivial state is stable. There are also SPTs beyond band theory with root state that can be viewed as bosonic SPTs with symmetry $SO(3) \times T$, and they can be classified by $\mathbb{Z}_2^4$. One of the four root states becomes trivial in the presence of fermions with this symmetry (see Table 7.5), therefore, we propose that the total classification is $\mathbb{Z}_2^4$.

G.5 Relations between the classification of $SO(3) \times T$ symmetric $U(1)$ quantum spin liquids and the classification of some relevant SPTs

In the main text $SO(3) \times T$ symmetric $U(1)$ quantum spin liquids can be classified into 15 phases, as summarized in Table 7.1 and Table 7.2. In particular, how they can be viewed as gauged SPTs are also discussed. It is helpful to understand the relation between the classification of the $U(1)$ quantum spin liquids and the classification of the relevant SPTs. Below we give some examples.

From the point of view of $E$, $E_b M_b$, $E_b M_f$, $E_b M_{b \frac{1}{2}}$ and $E_b M_{f \frac{1}{2}}$ can all be regarded as the gauged insulators of Kramers singlet bosons with symmetry $(U(1) \times T) \times SO(3)$. In Appendix G.4 we propose that the bosonic SPTs under this symmetry are classified by $\mathbb{Z}_2^6$, where only two of the six root states requires protection from the $U(1)$ symmetry. It is not hard to see, after gauging this $U(1)$ symmetry, the $\mathbb{Z}_2^2$ subset of SPTs coming from these two root states become precisely the above four quantum spin liquids.

With the same symmetry as above, if the bosons are Kramers doublet, Appendix G.2 shows that only one of the two root states survives. Gauging the trivial insulator and the nontrivial SPT from the other root state leads to $E_{bT} M_b$ and $E_{bT} M_f$, respectively.

From the point of view of $M$, $E_b M_b$, $E_{bT} M_b$, $E_{b \frac{1}{2}} M_b$, $E_{bT \frac{1}{2}} M_b$, $E_f M_b$, $E_{fT} M_b$, $E_{f \frac{1}{2}} M_b$ and
$E_{f\tau\frac{1}{2}}M_b$ can be thought of as the gauged bosonic insulators with symmetry $U(1) \times \mathcal{T} \times SO(3)$. In Appendix G.4 we propose that the bosonic SPTs under this symmetry are classified by $\mathbb{Z}_2^4$, where only three of the root states requires protection from the $U(1)$ symmetry. Gauging the $\mathbb{Z}_2^4$ subset of the SPTs generated by these three root states gives precisely the above eight quantum spin liquids.

From the point of view of $M$, both $E_bM_{f\frac{1}{2}}$ and $E_{\theta\tau\frac{1}{2}}M_{f\frac{1}{2}}$ can be viewed as a gauged topological superconductor of fermions with symmetry $((U(1) \times SU(2))/\mathbb{Z}_2) \times \mathcal{T}$. In Appendix G.4 we propose that the topological superconductors with this symmetry are classified by $\mathbb{Z}_4 \times \mathbb{Z}_2^3$, where the first $\mathbb{Z}_4$ factor represents those can be realized with free fermions, and the last $\mathbb{Z}_2^3$ factor corresponds to interacting topological superconductors beyond band theory. For states that can be realized by band theory, the nontrivial topological superconductors can have $2k$ massless Dirac fermions on the surface, where $k = 0, 1, 2, 3$ (mod 4). Gauging states with even $k$ leads to $E_bM_{f\frac{1}{2}}$ (up to a bosonic SPT $e\theta\tau\tau\frac{1}{2}$) and gauging states with odd $k$ leads to $E_{\theta\tau\frac{1}{2}}M_{f\frac{1}{2}}$. For states beyond band theory, gauging them results in the same quantum spin liquid as their corresponding state within band theory up to a bosonic SPT with symmetry $SO(3) \times \mathcal{T}$.

From the point of view of $E$, both $E_{f\frac{1}{2}}M_b$ and $E_{f\frac{1}{2}}M_{b\frac{1}{2}}$ can be viewed as gauged topological insulators of fermions with symmetry $((U(1) \times \mathcal{T}) \times SU(2))/\mathbb{Z}_2$, where the microscopic fermions are Kramers singlets. In Appendix G.4 we propose that these topological insulators are classified by $\mathbb{Z}_2^5$, where the first $\mathbb{Z}_2$ factor corresponds to those realizable by free fermions, and the nontrivial state can have 2 massless Dirac fermions on the surface. Gauging the trivial state leads to $E_{f\frac{1}{2}}M_b$ and gauging the nontrivial state leads to $E_{f\frac{1}{2}}M_{b\frac{1}{2}}$. Gauging the states beyond band theory gives the same quantum spin liquids as their corresponding free fermion cousins up to a bosonic SPT under symmetry $SO(3) \times \mathcal{T}$.

The above examples show that gauging different SPTs may results in the same quantum spin liquid, and no one single class of SPTs will lead to all quantum spin liquids after gauging, so the classification of these quantum spin liquids does not form a simple group structure, while the classification of SPTs does. As mentioned in the introduction, viewing a single quantum spin liquid as two different gauged SPTs leads to some helpful dualities on the surface theories of these SPTs, which can be inferred from the above discussion.
Also, if two different quantum spin liquids can be viewed as two different SPTs with the same microscopic constitutes and symmetry, the quantum phase transition between them can also be viewed as the gauged version of the quantum phase transition between the two corresponding SPTs. For example, the quantum phase transition between $E_{f_\frac{1}{2}}M_b$ and $E_{f_\frac{1}{2}}M_b \frac{1}{2}$ can be viewed as the gauged version of the quantum phase transition between the trivial and nontrivial fermionic insulators with symmetry $((U(1) \times T) \times SU(2))/Z_2$. We will not go into the details in this paper.

G.6 Bosonic SPTs with $SO(3) \times T$ symmetry

In this appendix we discuss bosonic SPTs with symmetry $SO(3) \times T$, with the assumption that the microscopic degrees of freedom are non-Kramers bosons with spin-1. Group cohomology gives classification $\mathbb{Z}_2^3$, [344] but it misses one root state [337], and the complete classification should be $\mathbb{Z}_2^4$. The four root states all have anomalous surface $Z_2$ topological orders, where symmetries are realized in a way impossible in a purely two dimensional system (see Table 7.3). Among the four root states, $eTmT$ and $efmf$ are protected by time reversal alone. Below we review the anomalies of $e\frac{1}{2}m\frac{1}{2}$ and $e\frac{1}{4}mT$.

Suppose $e\frac{1}{2}m\frac{1}{2}$ is realizable in a purely two dimensional system, then tunneling an $SO(3)$ monopole through it leaves a $\pi$-flux seen by both $e$ and $m$. Because such a local process should not have nonlocal observable effect, an $\epsilon$, the fermionic bound state of $e$ and $m$, must be trapped at this $\pi$-flux. Due to time reversal symmetry, there is no polarization spin in this process and this flux is just a fermion. Therefore, a local process generates a fermion in this 2D system, which is impossible. Notice this anomaly is just the $SO(3)$ version of the anomaly of $eCmC$ under symmetry $U(1) \times T$. In fact, when the symmetry $SO(3) \times T$ is broken down to $U(1) \times T$, the descendant state of $e\frac{1}{2}m\frac{1}{2}$ is precisely $eCmC$, which has a $U(1)$ theta angle to be $2\pi$. This also implies the $SO(3)$ $\Theta = 2\pi$ for $e\frac{1}{2}m\frac{1}{2}$.

As for $e\frac{1}{4}mT$, tunneling an $SO(3)$ monopole leaves a $\pi$-flux seen by $e$ and $\epsilon$, so this process must trap an $m$. Because $SO(3)$ commutes with $T$, the $SO(3)$ flux is invariant under time reversal, and a Kramers doublet is generated by this local process. This contradicts the assumption that there is no local Kramers doublet particles. Again, this anomaly is the $SO(3)$ version of the anomaly of $eCmT$ under symmetry $U(1) \times T$, and the descendant state of $e\frac{1}{4}mT$ is just $eCmT$.
when the symmetry $SO(3) \times T$ is broken down to $U(1) \times T$.

G.7 Constraints on the Hall conductances due to time reversal and spin rotational symmetries

Suppose in addition to a $U(1)_c$ charge conservation symmetry, a two dimensional gapped system also has time reversal symmetry and spin rotational symmetry. One can also consider the spin quantum Hall conductance, $\sigma_{xy}^s$, which relates the spin current due to a gradient Zeeman field. [407] This can be formally viewed as the response of the system to a probe gauge field, $A_s$, which corresponds to the $S_z$ rotation symmetry, $U(1)_s$. There can also be quantum spin Hall conductance, $\sigma_{xy}^{cs}$, which relates the spin current and the electric field of the gauge field $A_c$, the gauge field corresponding to symmetry $U(1)_c$. [408] This appendix discusses the constraints on these Hall conductances due to time reversal and spin rotational symmetries. The results are useful in determining what polarization charge or spin will be generated when flux is inserted in the system, or equivalently, when a monopole tunnels through the system.

To this end, we first reorganize the charge conservation and $S_z$ rotation symmetries in terms of two other $U(1)$ symmetries, denoted by $U(1)_\uparrow$ and $U(1)_\downarrow$. These two $U(1)$ symmetries can be viewed as separate conservations of spin-up and spin-down particles. The corresponding gauge fields and charges of these two symmetries are related to $A_c$ and $A_s$ by

$$A_\uparrow = A_c + A_s, \quad A_\downarrow = A_c - A_s$$
$$Q_\uparrow = \frac{Q_c + Q_s}{2}, \quad Q_\downarrow = \frac{Q_c - Q_s}{2}$$

(G.26)

Notice under the charge quantization condition that $Q_1$ and $Q_2$ can independently take any integers, $Q_c$ and $Q_s$ have to either be both even or both odd.

Now we can discuss the Hall conductances due to coupling to $A_\uparrow$ and $A_\downarrow$. The general Hall response theory reads

$$\mathcal{L} = \frac{1}{4\pi}(\sigma_{xy}^\uparrow A_\uparrow dA_\uparrow + \sigma_{xy}^\downarrow A_\downarrow dA_\downarrow + 2\sigma_{xy}^{\uparrow\downarrow} A_\uparrow dA_\downarrow)$$

(G.27)
where $AdB$ is a shorthand for $\epsilon_{\mu\nu\lambda} A_\mu \partial_\nu B_\lambda$. Using (G.26), we get

\[
\begin{align*}
\sigma_{xy}^c &= \sigma_{xy}^\uparrow + \sigma_{xy}^\downarrow + 2\sigma_{xy}^{\uparrow\downarrow} \\
\sigma_{xy}^s &= \sigma_{xy}^\uparrow + \sigma_{xy}^\downarrow - 2\sigma_{xy}^{\uparrow\downarrow} \\
\sigma_{xy}^{cs} &= \sigma_{xy}^\uparrow - \sigma_{xy}^\downarrow
\end{align*}
\]  

(G.28)

Clearly any element in the spin rotational symmetry that takes spin-up to spin-down (such as rotation around $x$-axis by $\pi$) requires $\sigma_{xy}^\uparrow = \sigma_{xy}^\downarrow$. Below we study the constraints from time reversal symmetry. Notice that the $U(1)_s$ charge is always odd under time reversal, but the $U(1)_c$ charge can either be time reversal even or odd, and we discuss these two cases separately.

We start from the case where the $U(1)_c$ charge is even under time reversal, which means

\[
A_{1,2}^0 \to A_{2,1}^0, \quad A_{\uparrow,\downarrow}^0 \to -A_{\downarrow,\uparrow}^0
\]  

(G.29)

For the response theory to be time reversal symmetric, we need

\[
\sigma_{xy}^\uparrow = -\sigma_{xy}^\downarrow, \quad \sigma_{xy}^{\uparrow\downarrow} = 0
\]  

(G.30)

For the case where the $U(1)_c$ is odd under time reversal, time reversal transformation takes

\[
A_{\uparrow,\downarrow}^0 \to -A_{\downarrow,\uparrow}^0, \quad A_{\uparrow,\downarrow}^0 \to A_{\downarrow,\uparrow}^0
\]  

(G.31)

For the response theory to be time reversal symmetric, we need

\[
\sigma_{xy}^\uparrow = \sigma_{xy}^\downarrow = \sigma_{xy}^{\uparrow\downarrow} = 0
\]  

(G.32)

From these constraints and (G.28) one can easily obtain the constraints on $\sigma_{xy}^c$, $\sigma_{xy}^s$ and $\sigma_{xy}^{cs}$. We note all these constraints can also be obtained simply by applying Laughlin’s flux insertion argument.

We notice that all these Hall conductance vanishes if the system has both time reversal symmetry and spin rotational symmetry that contains at least $O(2) \simeq U(1)_s \rtimes Z_2$, where $U(1)_s$ is a
rotational symmetry around one axis and $Z_2$ is the $\pi$-rotation around another axis perpendicular to the previous one. This implies that inserting flux or tunneling a monopole through such two dimensional systems will not lead to any polarization charge or spin.

G.8 Non-edgeability of some $Z_2$ topological orders in the presence of non-trivial particles

In the Sec. 7.5.1 we claimed some $Z_2$ topological orders are not edgeable even in the presence of some nontrivial particles, i.e. they do not allow for a physical edge separating it and the trivial vacuum. In this appendix, we will justify this claim by showing that these $Z_2$ topological orders allow no $K$-matrix theory to describe them.

G.8.1 Brief review of the $K$-matrix theory

We begin with a brief review of some general aspects of the $K$-matrix theory. For more details, see Ref. Wen2004Book and Ref. Levin2012,Lu2012,LuVishwanath2013.

The Lagrangian of a $K$-matrix theory of a system that couples to an external $U(1)$ gauge field $A_c$ is given by

$$\mathcal{L} = \frac{K_{IJ}}{4\pi} a_I d a_J - \frac{q^{I}_c}{2\pi} A_c d a_I$$

(G.33)

with $K$ a symmetric invertible matrix with all entries integers, and $q_c$ a vector with all entries integers.

An excitation of this theory can be represented by an integral excitation vector, $l$. The charge of this excitation under the external gauge field $A_c$ is

$$l^T K^{-1} q_c$$

(G.34)

And this excitation has self-statistics angle

$$\pi l^T K^{-1} l$$

(G.35)
For two excitations represented by excitation vectors $l_1$ and $l_2$, respectively, the mutual braiding angle between them is

$$2\pi l_1^T K^{-1} l_2$$  \hspace{1cm} (G.36)

A simple example is that

$$K = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$$  \hspace{1cm} (G.37)

which represents $Z_2$ topological order, where $e$ can be taken to be represented by excitation vector $(1, 0)^T$ and $m$ can be taken to be represented by $(0, 1)^T$.

The 2+1-d bulk theory (G.33) allows the following boundary theory:

$$\mathcal{L} = \frac{1}{4\pi} (K_{IJ} \partial_t \phi_I \partial_x \phi_J - V_{IJ} \partial_x \phi_I \partial_x \phi_J)$$  \hspace{1cm} (G.38)

where $\phi_I$ are bosonic fields such that $e^{i\phi_I}$ is the annihilation operator of excitation $l$ on the boundary. These bosonic fields satisfy Kac-Moody algebra

$$[\phi_I(x), \partial_y \phi_J(y)] = 2\pi i(K^{-1})_{IJ} \delta(x - y)$$  \hspace{1cm} (G.39)

$V_{IJ}$ is called the velocity matrix that gives the velocities of these bosonic fields.

The above summarizes the topological properties of the $K$-matrix theory, (G.33). Below we review the symmetry actions on this theory.

In general symmetries act on the gauge fields $a_I$ as a matrix. For example, we denote the time reversal action as

$$a_I \rightarrow T_{IJ} a_J$$  \hspace{1cm} (G.40)

with $T$ an integral matrix. Notice the above equation only gives the transformation of the spatial components of the gauge fields, and the temporal components should have a minus sign in front.
due to the anti-unitary nature of time reversal symmetry.

It is important to notice the above does not fully specify the symmetry action, and to that end, one needs to specify how the boundary bosonic fields transform. In general, they transform as

\[ \phi_I \rightarrow T_{IJ} \phi_J + t_I \]  

(G.41)

with \( t_I \) a real vector. [309]

To make the bulk theory (G.33) invariant under anti-unitary time reversal symmetry, we need

\[ K \rightarrow T^K T = -K \]  

(G.42)

If the \( U(1) \) charge is even under time reversal, we further require

\[ q_c \rightarrow T^K q_c = q_c \]  

(G.43)

while if the \( U(1) \) charge is odd under time reversal we require

\[ q_c \rightarrow T^K q_c = -q_c \]  

(G.44)

G.8.2 Non-edgeability of some \( Z_2 \) topological orders in the presence of nontrivial particles

Now by showing some \( Z_2 \) topological orders even in the presence of nontrivial particles do not allow a \( K \)-matrix theory description, we show their non-edgeability because \( K \)-matrix theories are supposed to capture all two dimensional Abelian states.

Here we list the \( Z_2 \) topological orders of interests. We denote \( eTmT \) in the presence of bosons with quantum numbers \( C^2 T^{1/2} \) by \((eTmT, bC^2 T^{1/2})\), and \( eTmT \) in the presence of fermions with quantum number \( \tilde{C}^2 \) by \((eTmT, f\tilde{C}^2)\). Besides these two, we will also consider \((e^{1/2}_T mT, bC^2 T^{1/2})\), \((e^{1/2}_T mT, fC^2 T^{1/2})\), \((eTmT, fC^{21})\), \((eTmT, f\tilde{C}^{21})\), \((eTmT, fC^2 T^{1/2})\), \((e^{1/2}_T mT, fC^{21})\) and \((e^{1/2}_T mT, fC^2 T^{1/2})\), with similar notations as before.
We immediately have two main difficulties in showing their non-edgeability. First, in some cases the nontrivial particles carry spin-1/2 and we need to incorporate $SU(2)$ symmetry in the $K$-matrix theory, but continuous non-Abelian symmetries are usually not manifest in a $K$-matrix theory and dealing with them directly is generally difficult. To resolve this difficulty, we will instead just show that the descendants of the relevant states are still not edgeable when the $SU(2)$ symmetry is broken down to $U(1)$, which is sufficient to show the original states are non-edgeable with the full $SU(2)$ symmetry. To distinguish this $U(1)$ from the original charge $U(1)$, we will denote the charge $U(1)$ by $U(1)_c$, and this $U(1)$ by $U(1)_s$. Unit charge under $U(1)_s$ will be denoted by $C^2$. Therefore, for example, we will consider $(eTmT, fC^2C'^2)$ instead of $(eTmT, fC^2_{1/2})$.

Now (G.33) needs to be modified to include the coupling to $A_s$, the external gauge field corresponding to $U(1)_s$

$$\mathcal{L} = \frac{K_{IJ}}{4\pi} a_I da_J - \frac{q_c}{2\pi} A_c da_I - \frac{q_s}{2\pi} A_s da_I$$

(G.45)

Notice the charge of $U(1)_s$ is always odd under time reversal, so time reversal symmetry requires that

$$q_s \rightarrow T^T q_s = -q_s$$

(G.46)

The second difficulty is that in general the state that we are interested in may be described by a $K$-matrix with a large dimension, but dealing with a large-dimensional $K$-matrix is daunting. However, the following observation suggests we actually only need to consider a $2 \times 2$ $K$-matrix.

Notice all these $Z_2$ topological orders come from $eTmT$, $e\frac{1}{2}mT$ and $eT\frac{1}{2}mT$. For both cases, the nontrivial topological quasiparticles only need a single component to describe them. This is because whenever there are two components of them, one can condense some bound states of them that are singlets under all symmetries. This will not change the topological order or the symmetry of the system, but only one component will be left over. [316] More concretely, this means to describe the putative $Z_2$ topological orders that we are interested in, we should always
be able to write the $K$-matrix as

$$K = 2\sigma_x \oplus L$$  \hspace{1cm} (G.47)

where $L$ is an invertible symmetric integral matrix that can be large in dimension, and $L$ describes only local excitations. For bosons, $L$ can be written as $\sigma_x \oplus \sigma_x \oplus \cdots \oplus \sigma_x$, while for fermions, $L$ can be written as $\sigma_z \oplus \sigma_z \oplus \sigma_x \oplus \cdots$. In this form, any excitation with an excitation vector of the form $(1, 0, \cdots)^T$ can be identified as $e$, and all excitations with an excitation vector of the form $(0, 1, \cdots)^T$ can be viewed as $m$, where the “…” can be nonzero. At this moment, the nontrivial topological quasiparticle that we are after, for example, the Kramers doublet $e$ particle in $eTmT$, can still be represented by $(1, 0, \cdots)^T$ with “…” nonzero. But we can always bind proper local excitations to this excitation so that the excitation vector becomes $(1, 0, 0, 0, \cdots)^T$ with “…” all zeros.

The argument above shows that, in order to show the non-edgeability of those $Z_2$ topological orders, it is sufficient to show that no $2 \times 2$ $K$-matrix can describe the topological quasiparticles with the corresponding quantum numbers, up to binding local excitations. Let us demonstrate this via the following example. In $(eTmT, bC^2TC'^2)$, $eTmT$ can be relabelled as, for example, $eC^2C'^2mT$. The above statement means that, in order to show that $(eTmT, bC^2TC'^2)$ is not edgeable, it is sufficient to show that none of $eTmT, eC^2C'^2mT$ and all other states related to these by binding a local excitation made up of $bC^2TC'^2$ can be realized by a $2 \times 2$ $K$-matrix.

Because time reversal should not convert a local excitation into a nonlocal one, we expect that the matrix $T$ can be written in the following form

$$T = \begin{pmatrix} T_0 & T_1 \\ 0 & T_2 \end{pmatrix}$$  \hspace{1cm} (G.48)

where $T_0$ is a $2 \times 2$ integral matrix. Plugging this form of $K$ and $T$ into (G.42), we see time reversal symmetry requires that

$$T_0^T \sigma_x T_0 = -\sigma_x$$  \hspace{1cm} (G.49)
It is easy to show the only solutions are \( T_0 = \pm \sigma_z \) or \( T_0 = \pm \epsilon \), where \( \epsilon = i \sigma_y \).

Notice in all the cases we consider, the quantum numbers of \( e \) and \( m \) are always nontrivial. If \( T_0 = \pm \epsilon \), then \( T_0^2 = -1 \). This does not allow Kramers doublet structure, and it also does not allow nonzero \( q_e \) and \( q_s \) that satisfy (G.43) or (G.44) and (G.46). So this choice of \( T_0 \) can never work.

So we can focus on the case with \( T_0 = \pm \sigma_z \). Without loss of generality, we take \( T_0 = -\sigma_z \). Notice now \( e \), represented by the excitation vector \((1,0)^T\), is always a Kramers singlet independent of \( t \). If the second entry of \( t \) is \( \pi/2 \), \( m \), represented by excitation vector \((0,1)^T\), is a Kramers doublet.

With this choice of \( T_0 \), in order to satisfy (G.43) or (G.44) and (G.46), \( q_e \) and \( q_s \) can only be taken as

\[
q_e = (0, q_1)^T \quad q_s = (q_2, 0)^T \tag{G.50}
\]

when the charge under \( A_c \) is even under time reversal, or

\[
q_e = (q_1, 0)^T \quad q_s = (q_2, 0)^T \tag{G.51}
\]

when the charge under \( A_c \) is odd under time reversal. In the first case, \( e \) carries charge \( q_1/2 \) under \( A_c \) and zero charge under \( A_s \). In the second case, \( e \) carries zero charge under both \( A_c \) and \( A_s \).

\( e T m T \) IS NOT EDGEABLE IN THE PRESENCE OF \( bC^2 TC'^2 \)

The above discussions immediately imply that \( e T m T \) is not edgeable in the presence of \( bC^2 TC'^2 \). This is because \( e \) cannot be a Kramers doublet, an odd number of \( bC^2 TC'^2 \)s have to be attached to it to cancel its Kramersness. Then \( e \) carries nonzero charge under \( A_s \), which is in contradiction with \( e \) always carrying zero charge under \( A_s \). So \((e T m T, bC^2 T \frac{1}{2})\) is not edgable.
\(eC'^2mT\) is not edgeable in the presence of \(bC'^2TC'^2\)

Here \(e\) is not a Kramers doublet but it carries nonzero charge under \(A_s\). To cancel this charge, an odd number of \(bC'^2TC'^2\)s have to be attached to \(e\), which makes it a Kramers doublet. This is again impossible as argued above. One can also try to switch the label between \(e\) and \(m\), then it becomes \(eTmC'^2\). The argument above implies this is inconsistent even in the presence of \(bC'^2TC'^2\). So \((e^\frac{1}{2}mT, eC'^2T^\frac{1}{2})\) is not edgeable.

\(eTC'^2mT\) is not edgeable in the presence of \(bC'^2C'^2\)

Here no matter how many \(bC'^2C'^2\)s are attached, the \(Z_2\) topological order always has both \(e\) and \(m\) being Kramers doublet. This cannot be realized. So \((eT^\frac{1}{2}mT, bC'^2_{\frac{1}{2}})\) is not edgeable.

\(eTmT\) is not edgeable in the presence of \(fC'^2\) or \(fC'^2C'^2\)

Here \(e\) is a Kramers doublet, so an odd number of \(fC'^2\)s or \(fC'^2C'^2\) need to be attached to it, which makes it become \(e\) and the new \(e\) carry \(C'^2\). This is in contradiction with \(e\) carrying zero charge under \(A_c\) in this case. So \((eTmT,fC'^2)\) and \((eTmT, fC'^2_{\frac{1}{2}})\) are not edgeable.

\(eTmT\) and \(eTC'^2mT\) are not edgeable in the presence of \(fC'^2T\) or \(fC'^2TC'^2\)

Here \(e\) and \(m\) are Kramers doublets, and attaching any number of \(fC'^2T\) or \(fC'^2TC'^2\) always leaves both \(e\) and \(m\) Kramers doublets, so \((eTmT, fC'^2T), (eTmT, fC'^2T^\frac{1}{2})\) and \((eT^\frac{1}{2}mT, fC'^2T^\frac{1}{2})\) are not edgeable.

\(eTmT\) and \(eC'^2mT\) are not edgeable in the presence of \(fC'^2C'^2\)

Here an odd number of \(fC'^2C'^2\)s need to be attached to \(e\), which converts it to \(e\) and make the new \(e\) carry nonzero charge under \(A_c\) and \(A_s\). But \(e\) cannot carry nonzero charge under both \(A_c\) and \(A_s\). So \((eTmT, fC'^2_{\frac{1}{2}})\) and \((e^\frac{1}{2}mT, fC'^2_{\frac{1}{2}})\) are not edgeable.

In summary, none of the \(Z_2\) topological orders is edgeable in the presence of the relevant non-trivial excitations. This implies they are all still anomalous.
G.9 PROJECTIVE REPRESENTATIONS: THE ELECTRIC (STANDARD), THE MAGNETIC (TWISTED) AND THE DYONIC (MIXED) ONES

In this appendix we discuss in detail various projective representations of a symmetry group $G$: the electric (standard), the magnetic (twisted) and the dyonic (mixed) ones. We always assume this group $G$ can contain time reversal, but besides, for simplicity, all other elements form a connected unitary group. That is, these elements are all unitary and they can all be continuously connected to the identity element. We will see although all these projective representations are classified by some group cohomologies, but different cases are classified by different group cohomologies.

G.9.1 ELECTRIC (STANDARD) PROJECTIVE REPRESENTATIONS

We begin with the familiar case of standard projective representations. Although our results will be identical to the ones in textbooks, we will use a different formulation that is more appropriate for our purposes and easier to generalize to twisted projective representations.

Suppose there is a symmetry $G$, which can in principle contain anti-unitary element. If all elements of $G$ only change an excitation by a local operation, then it is appropriate to discuss the standard projective representations of $G$ on this excitation.

A prototypical example of this case is that the relevant excitation is the electric charge $E$ of a $U(1)$ quantum spin liquid. In general, the action of an element $g \in G$ on $E$ can be written as

$$E_i \rightarrow U(g)_{ij} E_j \quad (G.52)$$

Here different components of $E_i$ differ from each other by a local operation, and $U(g)$ is a matrix representation of $g$.

Because $E_i^* A_{ij} E_j$ is a local operator for any matrix $A$, this operator is supposed to transform
in the linear representation of $G$. For $g_1, g_2 \in G$, acting $g_1$ followed by $g_2$ on this operator gives

$$E^\dagger \left( U(g_1)^{s(g_2)}U(g_2) \right)^\dagger \cdot A^{s(g_2g_1)} \cdot \left( U(g_1)^{s(g_2)}U(g_2) \right) E$$

where now $E$ represents a column vector with components $E_i$, and, for an arbitrary matrix $M$, $M^{s(g)} = M$ if $g$ is unitary, while $M^{s(g)} = M^*$ if $g$ is anti-unitary. For the special case where $M$ is just a phase factor, $s(g) = 1$ ($s(g) = -1$) if $g$ is unitary (anti-unitary).

The above result should be identical to the one obtained by acting $g_2g_1$ on this local operator directly:

$$E^\dagger \cdot U^\dagger (g_2g_1) \cdot A^{s(g_2g_1)} \cdot U(g_2g_1) \cdot E$$

For these two results to be identical for an arbitrary matrix $A$, we must have

$$U(g_1)^{s(g_2)}U(g_2) = \omega(g_2, g_1)U(g_2g_1)$$

where $\omega(g_2, g_1)$ is a phase factor.

The above equation can be written in the following equivalent way:

$$U(g_2g_1) = \omega(g_2, g_1)^{-1}U(g_1)^{s(g_2)}U(g_2)$$

For any $g_1, g_2, g_3 \in G$, applying this equation to $U(g_1g_2g_3)$ yields

$$U(g_1g_2g_3)$$

$$= \omega(g_1, g_2g_3)^{-1}U(g_2g_3)^{s(g_1)}U(g_1)$$

$$= \omega(g_1, g_2g_3)^{-1}\omega(g_2, g_3)^{-s(g_1)}U(g_3)^{s(g_1g_2)}U(g_2)^{s(g_1)}U(g_1)$$

$$= \omega(g_1g_2, g_3)^{-1}U(g_3)^{s(g_1g_2)}U(g_1g_2)$$

$$= \omega(g_1g_2, g_3)^{-1}\omega(g_1, g_2)^{-1}U(g_3)^{s(g_1g_2)}U(g_2)^{s(g_1)}U(g_1)$$
This implies the following associativity condition for the phase factor $\omega$’s:

$$\omega(g_1g_2, g_3)\omega(g_1, g_2) = \omega(g_1, g_2g_3)\omega(g_2, g_3)^{s(g_1)}$$  \hspace{1cm} (G.57)

There is a gauge freedom for the phase factor $\omega$’s. To see this, notice the symmetry action of $g \in G$ on $E$ can be modified by a gauge transformation:

$$E_i \rightarrow \lambda(g) U(g)_{ij} E_j \equiv \tilde{U}(g)_{ij} E_j$$  \hspace{1cm} (G.58)

where $\lambda(g)$ is a $U(1)$ phase factor. The action of $g \in G$ on any local operator will be the same, which means $\tilde{U}(g)$ is an equally good representation of $g$. Under this transformation, it is straightforward to check that

$$\tilde{U}(g_2)^{s(g_1)}\tilde{U}(g_1) = \tilde{\omega}(g_1, g_2)\tilde{U}(g_1g_2)$$  \hspace{1cm} (G.59)

where

$$\tilde{\omega}(g_1, g_2) = \omega(g_1, g_2) \cdot \frac{\lambda(g_1)\lambda(g_2)^{s(g_1)}}{\lambda(g_1g_2)}$$  \hspace{1cm} (G.60)

The factor systems $\omega$ and $\tilde{\omega}$ related in this way should be regarded to be in the same class, because they give rise to identical results in any local operator. It is straightforward to check that the relation (G.60) is an equivalence relation, that is, it is reflexive, symmetric and transitive. Furthermore, it is clear if $\omega_1$ and $\omega_2$ are the two classes of factor systems corresponding to representations $U_1(g)$ and $U_2(g)$, respectively, $\omega_1 \cdot \omega_2$ will be the factor system of the representation $U_1(g) \cdot U_2(g)$. This defines a multiplication operation among the classes of factor systems. With this multiplication, the classes of factor systems form an Abelian group, where the trivial element is the class of factor systems of a linear representation. In fact, this Abelian group form a structure of group cohomology $H^2(G, U_T(1))$. [344] In this group cohomology, the $n$-cochains
ωn(g1, g2, · · · , gn) take value as a phase factor, the 1-coboundary operation is defined as

\[ d_1\omega_1(g_1, g_2) = \frac{\omega_1(g_1)\omega_1(g_2)^s(g_1)}{\omega_1(g_1g_2)} \]  \hspace{1cm} (G.61)

and the 2-coboundary operation is defined as

\[ d_2\omega_2(g_1, g_2, g_3) = \frac{\omega_2(g_2, g_3)^s(g_1)\omega_2(g_1, g_2g_3)}{\omega_2(g_1g_2, g_3)\omega_2(g_1, g_2)} \]  \hspace{1cm} (G.62)

It is straightforward to check that \( d_2d_1 = 1 \). Also, the solutions to the associativity condition (G.57) are 2-cocycles, and different solutions are identified up to a 1-coboundary. Therefore, the classes of factor systems, or the (standard) projective representations, are indeed classified by this cohomology. Below we will see the twisted and mixed projective representations are also classified by some group cohomologies, which are however different from this \( H^2(G, U_T(1)) \).

**G.9.2 Magnetic (twisted) projective representations**

Next we turn to twisted projective representations. A prototypical example where it is appropriate to consider twisted projective representations is, when the symmetry includes time reversal, to consider the fractional quantum numbers on the magnetic monopole, \( M \).

Suppose \( g \in G \) is unitary, its action on \( M \) can be represented as

\[ M_i \rightarrow U(g)_{ij}M_j \]  \hspace{1cm} (G.63)

Suppose \( g \in G \) is anti-unitary, its action on \( M \) can be represented as

\[ M_i \rightarrow U(g)_{ij}M_j^* \]  \hspace{1cm} (G.64)

Again, because \( M_i^*A_{ij}M_j \) is a local operator for any matrix \( A \), it is supposed to transform in the linear representation of \( G \). Similar to the analysis in the previous case, this implies, for
\( g_1, g_2 \in G, \)

\[
U(g_1)^{s(g_2)}U(g_2)^{s(g_1)} = \omega(g_2, g_1)U(g_2g_1) \tag{G.65}
\]

where \( \omega(g_2, g_1) \) is a phase factor.

Similar as standard projective representations, these phase factors need to satisfy an associativity condition:

\[
\omega(g_1g_2, g_3)\omega(g_1, g_2)^{s(g_3)} = \omega(g_1, g_2g_3)\omega(g_2, g_3)^{s(g_1)} \tag{G.66}
\]

Further, there is also a gauge freedom that leads to the following equivalence relation

\[
\omega(g_1, g_2) \sim \tilde{\omega}(g_1, g_2) = \omega(g_1, g_2) \cdot \frac{\lambda(g_1)^{s(g_2)}\lambda(g_2)^{s(g_1)}}{\lambda(g_1g_2)} \tag{G.67}
\]

where \( \lambda \)'s are phase factors.

Just as in the case of the standard projective representations, the classes of factor systems of a twisted projective representation also form an Abelian group, whose multiplication, trivial element, and inverse element are defined in parallel as in the case of the standard projective representation. This group is also described by a cohomology, denoted by \( H^2(G, U_T^M(1)) \). This cohomology is different from the previous one, \( H^2(G, U_T(1)) \), in the coboundary operations. In this cohomology, the \( n \)-cochains \( \omega_n(g_1, g_2, \cdots, g_n) \) still take values as phase factors, the 1-coboundary operation is defined as

\[
d_1\omega_1(g_1, g_2) = \frac{\omega_1(g_1)^{s(g_2)}\omega_1(g_2)^{s(g_1)}}{\omega_1(g_1g_2)} \tag{G.68}
\]

and the 2-coboundary operation is defined as

\[
d_2\omega_2(g_1, g_2, g_3) = \frac{\omega_2(g_1, g_2g_3)\omega_2(g_2, g_3)^{s(g_1)}}{\omega_2(g_1g_2, g_3)\omega_2(g_1, g_2)^{s(g_3)}} \tag{G.69}
\]
It is straightforward to check $d_2d_1 = 1$. Again, the solutions of the associativity condition are 2-cocycles, and they are identified up to a 1-coboundary. Therefore, twisted projective representations are classified by $H^2(G, U^M_1(1))$.

Interestingly, for the cases with $G = 1$, $G = \mathbb{Z}_2$ and $G = U(1)$, $H^2(G \times \mathcal{T}, U^M_1(1)) = H^2(G \times \mathbb{Z}_2, U(1))$, where the latter is the standard group cohomology with $G \times \mathbb{Z}_2$ acting trivially on the $U(1)$ coefficient. It will be interesting to know if this relation is always true.

G.9.3 Dyonic (mixed) projective representations

In the case of a $G$ symmetric $U(1)$ quantum spin liquid at $\theta = \pi$, the property of the phase is determined by the $\left(\frac{1}{2}, 1\right)$ and $\left(\frac{1}{2}, -1\right)$ dyons. Denote these two dyons by $D^{(+)}$ and $D^{(-)}$, respectively. The symmetry quantum numbers of these two dyons are given by the dyonic (mixed) projective representations.

Again, assume the only part of the symmetry that can change the type of fractional excitations is time reversal, the action of $g \in G$ on $D^{(+)}$ and $D^{(-)}$ can be written as

$$D_i^{(+)} \to U_+(g)_{ij} D_j^{(+)} \quad D_i^{(-)} \to U_-(g)_{ij} D_j^{(-)} \quad (G.70)$$

if $g$ is unitary, and

$$D_i^{(+)} \to U_+(g)_{ij} D_j^{(-)} \quad D_i^{(-)} \to U_-(g)_{ij} D_j^{(+)} \quad (G.71)$$

if $g$ is anti-unitary.

Now using that $D_i^{(+)*} A_{ij} D_j^{(+)}$ and $D_i^{(-)*} A_{ij} D_j^{(-)}$ are local operators for any matrix $A$, we get

$$U_i^{s(g_2)}(g_1) U_{s(g_1)-i}(g_2) = \omega_i(g_2, g_1) U_i(g_2 g_1) \quad (G.72)$$

where $i = \pm$, and $\omega_i(g_2, g_1)$ is a phase factor.
In this case the associativity condition becomes

\[
\omega_i(g_1, g_2, g_3) \omega_i(g_2, g_3) s(g_1) = \omega_i(g_1, g_2, g_3) \omega_i(g_3, i)(g_1, g_2)
\]  \hfill (G.73)

And the equivalence relation becomes

\[
\omega_i(g_1, g_2) \sim \tilde{\omega}_i(g_1, g_2) = \omega_i(g_1, g_2) \frac{\lambda s(g_2) \lambda_i(g_1)\lambda_i(g_2) s(g_1)}{\lambda_i(g_1, g_2)}
\]  \hfill (G.74)

Similar as the twisted projective representations, the mixed projective representations also form an Abelian group and are also classified by a group cohomology, denoted by

\[
H^2(G, U^D(1) \times U^D(1))
\]

Here the \( n \)-cochains \( \omega_{i,n}(g_1, g_2, \cdots, g_n) \) take values as a phase factor (for \( i = \pm \) separately), the 1-coboundary operation is defined as

\[
d_1\omega_{i,1}(g_1, g_2) = \frac{\omega_{s(g_2)+i}(g_1) \omega_i(g_2) s(g_1)}{\omega_{i,1}(g_1, g_2)}
\]  \hfill (G.75)

and the 2-coboundary operation is defined as

\[
d_2\omega_{i,2}(g_1, g_2, g_3) = \frac{\omega_{i,2}(g_1, g_2, g_3) \omega_{i,2}(g_2, g_3) s(g_1)}{\omega_{i,2}(g_1, g_2, g_3) \omega_{i,2}(g_3, i)(g_1, g_2)}
\]  \hfill (G.76)

It is straightforward to check \( d_2d_1 = 1 \). Clearly, the solutions of the associativity condition are 2-cocycles, and different solutions are identified up to a 1-coboundary. So the mixed projective representations are classified by \( H^2(G, U^D(1) \times U^D(1)) \).
Table G.2: List of non-anomalous $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ not acting as a charge conjugation. All these states are anomaly-free. $T_E^2 = 1$ ($T_E^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T_{E'}^2 = 1$ ($T_{E'}^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T'$. $[T, Z_2]_M = +$ ($[T, Z_2]_M = -$) represents the case where $Z_2$ and $T$ commute (anti-commute) on $M$.

<table>
<thead>
<tr>
<th></th>
<th>$T_E^2$</th>
<th>$T_{E'}^2$</th>
<th>$[T, Z_2]_M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_b M_b$</td>
<td>1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>$E_b T M_b$</td>
<td>-1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>$E_b T' M_b$</td>
<td>1</td>
<td>-1</td>
<td>+</td>
</tr>
<tr>
<td>$E_{bTT} M_b$</td>
<td>-1</td>
<td>-1</td>
<td>+</td>
</tr>
<tr>
<td>$E_f M_b$</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>$E_f T M_b$</td>
<td>-1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>$E_{fTT} M_b$</td>
<td>1</td>
<td>-1</td>
<td>+</td>
</tr>
<tr>
<td>$E_{fTT'} M_b$</td>
<td>-1</td>
<td>-1</td>
<td>-</td>
</tr>
<tr>
<td>$E_b M_f$</td>
<td>1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>$E_b T M_f$</td>
<td>-1</td>
<td>1</td>
<td>+</td>
</tr>
<tr>
<td>$E_b T' M_f$</td>
<td>1</td>
<td>-1</td>
<td>+</td>
</tr>
<tr>
<td>$E_{bTT} M_f$</td>
<td>-1</td>
<td>-1</td>
<td>+</td>
</tr>
<tr>
<td>$E_b M_f$</td>
<td>1</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

G.10 EXAMINE THE ANOMALIES OF $Z_2 \times T$ SYMMETRIC $U(1)$ QUANTUM SPIN LIQUIDS WITH $\theta = 0$

In this appendix we will give more details of the anomaly-detection of the 72 different putative $U(1)$ quantum spin liquids with $Z_2 \times T$ symmetry that have $\theta = 0$. Among these states, $Z_2$ does not act as a charge conjugation for 24 of them and acts as a charge conjugation for the other 48. These 72 states are all listed in Sec. 7.7, and they are copied in Table G.2, Table G.3, Table G.4 and Table G.5 for convenience.

We will first show that the 15 states in Table G.2 and the 21 states in Table G.4 are anomaly-free, and give their constructions. Then we will show that the 9 states in Table G.3 and the 27 states in Table G.5 are anomalous.

Among the 36 anomaly-free states mentioned above, 26 of them have at least one of $E$ and $M$ being a trivial boson. These states clearly do not suffer from any anomaly, and they can be
Table G.3: List of anomalous $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ not acting as a charge conjugation. All these states are anomaly-free. $T^2_E = 1$ ($T^2_E = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T_E'^2 = 1$ ($T_E'^2 = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T'$. $[T, Z_2]_M = +$ ($[T, Z_2]_M = -$) represents the case where $Z_2$ and $T$ commute (anti-commute) on $M$. The last column indicates the anomaly classes.

<table>
<thead>
<tr>
<th>State</th>
<th>$T^2_E$</th>
<th>$T^2_E'$</th>
<th>$[T, Z_2]_M$</th>
<th>Anomaly Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{bT}M_{f-}$</td>
<td>1</td>
<td>-1</td>
<td>-</td>
<td>class a</td>
</tr>
<tr>
<td>$E_{fT}M_{b-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class a</td>
</tr>
<tr>
<td>$E_{bT'}M_{b-}$</td>
<td>1</td>
<td>-1</td>
<td>-</td>
<td>class a</td>
</tr>
<tr>
<td>$E_{fT'}M_{b-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
<tr>
<td>$E_{bT}M_{f-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
<tr>
<td>$E_{fT}M_{b-}$</td>
<td>-1</td>
<td>1</td>
<td>-</td>
<td>class b</td>
</tr>
<tr>
<td>$E_{bTT'}M_{f-}$</td>
<td>-1</td>
<td>-1</td>
<td>-</td>
<td>class c</td>
</tr>
<tr>
<td>$E_{fTT'}M_{b-}$</td>
<td>-1</td>
<td>-1</td>
<td>-</td>
<td>class c</td>
</tr>
<tr>
<td>$E_{f}M_{b-}$</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>class c</td>
</tr>
</tbody>
</table>

viewed as some gauged trivial insulators. The other 10 states,

$$E_{bT}M_f, E_{bT'}M_f, E_{bTT'}M_f, E_{fTT'}M_{b-},$$

$$(E_{fT}M_{bT'})_-, (E_{bT}M_{fT'})_-, (E_{fZ}M_{bT'}Z)_-, \quad (G.77)$$

$$(E_{bTZ}M_{fZ})_-, (E_{fTZ}M_{bZ})_-, (E_{bZ}M_{fT'Z})_-,$$

can be viewed as gauged free-fermion SPTs, which will be constructed below. To show that all other states are anomalous, as discussed in Sec. 7.7.3, it is sufficient to show that $(E_{bT}M_{bT'})_-$ and $E_{bTT'}M_{b-}$ are anomalous.

The rest of this appendix is organized as follows. In Appendix G.10.1 we will construct the relevant free-fermion SPTs, which after gauging give rise to states in (G.77). Then we will show that $(E_{bT}M_{bT'})_-$ is anomalous in Appendix G.10.2, and that $E_{bTT'}M_{b-}$ is anomalous in Appendix G.10.3.

G.10.1 Constructions of the relevant free-fermion SPTs

This subsection gives the construction of the free-fermion SPTs corresponding to states in (G.77). All these free-fermion topological insulators have two Dirac cones on the surface, and the surface
Table G.4: List of anomaly-free $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ acting as a charge conjugation. $T_2^E = 1$ ($T_2^E = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T_2^M = 1$ ($T_2^M = -1$) represents the case where $M$ is a Kramers singlet (doublet) under $T'$. $Z_2^{E,M}$ represents the result of acting the charge conjugation twice on $E$ and $M$, respectively.

<table>
<thead>
<tr>
<th>$E_b M_b$</th>
<th>$T_2^E$</th>
<th>$Z_2^E$</th>
<th>$T_2^M$</th>
<th>$Z_2^M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_b M_b$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b M_{bZ}$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b M_{bT}$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_b M_{bTZ}$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f M_f$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f M_{fZ}$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f M_{fT}$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$E_f M_{fTZ}$</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Hamiltonian can be written as

$$H = \sum_{i=1}^{2} \psi_i^\dagger (-i \partial_x \sigma_x - i \partial_y \sigma_z) \psi_i$$  \hspace{1cm} (G.78)

The differences among these states are in the symmetry assignments. Denote $\psi = (\psi_1, \psi_2)^T$, in all cases there is a $U(1)$ symmetry under which $\psi \rightarrow e^{i\theta} \psi$. We will also assign time reversal and $Z_2$ symmetries to these states, such that there is no symmetry-allowed fermion bilinear term that can open a gap on the surface. Then we will show the bosonic monopoles of these topological insulators have the desired nontrivial properties, using the method in Ref. Borokhov2002 (reviewed in Appendix G.4).
Table G.5: List of anomalous $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids that have $\theta = 0$ and have $Z_2$ acting as a charge conjugation at $\theta = \pi$. $T_2^E = 1$ ($T_2^E = -1$) represents the case where $E$ is a Kramers singlet (doublet) under $T$. $T_{M}^{2} = 1$ ($T_{M}^{2} = -1$) represents the case where $M$ is a Kramers singlet (doublet) under $T'$. $Z_{E,M}^{2}$ represents the result of acting the charge conjugation twice on $E$ and $M$, respectively. The last column lists the anomaly classes.

<table>
<thead>
<tr>
<th>Operation</th>
<th>$T_{E}^{2}$</th>
<th>$Z_{E}^{2}$</th>
<th>$T_{M}^{2}$</th>
<th>$Z_{M}^{2}$</th>
<th>Anomaly Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(E_{bZ}M_{bZ})_-$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_{bTZ}M_{bT'Z})_-$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_{fT}M_{bZ})_-$</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_{bZ}M_{fT'})_-$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_{fT}M_{bT'Z})_-$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_{bTZ}M_{fT'})_-$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>class 1</td>
</tr>
<tr>
<td>$(E_{bTZ}M_{f})_-</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>class 2</td>
</tr>
<tr>
<td>$(E_{bZ}M_{bT'Z})_-$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>class 2</td>
</tr>
<tr>
<td>$(E_{bZ}M_{f})_-</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>class 2</td>
</tr>
<tr>
<td>$(E_{bZ}M_{bT'Z})_-$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>class 3</td>
</tr>
<tr>
<td>$(E_{fT}M_{bT'Z})_-$</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>class 3</td>
</tr>
<tr>
<td>$(E_{bZ}M_{fT'})_-$</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>class 3</td>
</tr>
<tr>
<td>$(E_{fT}M_{bT'Z})_-$</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_{fT}M_{bT'Z})_-</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_{fT}M_{fT})_-</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_{bZ}M_{f})_-</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>class 4</td>
</tr>
<tr>
<td>$(E_{bT}M_{bZ})_-</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_{bT}M_{fT')_-$</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_{bTZ}M_{fT'})_-</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_{bTZ}M_{f})_-</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_{fZ}M_{f})_-</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>class 5</td>
</tr>
<tr>
<td>$(E_{fT}M_{f})_-</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>class 6</td>
</tr>
<tr>
<td>$(E_{fT}M_{bT'})_-</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>class 6</td>
</tr>
</tbody>
</table>

Let us start with the case where $Z_2$ does not act as a charge conjugation, and give the construction of free-fermion SPTs corresponding to $E_{bT}M_f$ and $E_{fT'}M_{b-}$. For the corresponding
SPT of $E_{bT} M_f$, let the symmetries be assigned as

\[
\begin{align*}
\mathcal{T} : \psi &\rightarrow \sigma_y \psi^\dagger \\
Z_2 : \psi &\rightarrow \tau_y \psi \\
\mathcal{T}' : \psi &\rightarrow \sigma_y \tau_y \psi^\dagger
\end{align*}
\]

Clearly the action of $\mathcal{T}$ and $Z_2$ commute on the fermion, so after gauging the fermion will become the $M_f$. Now we check the symmetry quantum number of the $E$, which is the monopole of $\psi$. Using state-operator correspondence, this is equivalent to checking the properties of the two zero-energy state in the presence of a $2\pi$ flux background with one of the two zero modes being occupied. Denote these zero modes by $f_1$ and $f_2$, which are related to $\psi_1$ and $\psi_2$, respectively. And denote the state with a $2\pi$ flux background and none of the zero modes being occupied by $|0\rangle$. Because $\mathcal{T}$ will flip the charge but keep the flux, under $\mathcal{T}$,

\[
\begin{align*}
|0\rangle &\rightarrow f_1^\dagger f_1^\dagger f_2^\dagger f_2^\dagger |0\rangle = f_2^\dagger |0\rangle \\
|0\rangle &\rightarrow f_2^\dagger f_1^\dagger f_2^\dagger f_1^\dagger |0\rangle = -f_1^\dagger |0\rangle
\end{align*}
\]

Notice the above transformations can be modified by an unimportant phase factor. This means $E$ will be a Kramers doublet under $\mathcal{T}$. In fact, here the particle-hole-like $\mathcal{T}$ is enough to protect the Dirac cones, and it is shown this is sufficient to show that $E$ is a Kramers doublet under $\mathcal{T}$. [406] Under $\mathcal{T}'$,

\[
\begin{align*}
|0\rangle &\rightarrow f_2 f_1 f_2 f_1 |0\rangle = -f_1 |0\rangle \\
|0\rangle &\rightarrow -f_1 f_1 f_2 f_2 |0\rangle = -f_2 |0\rangle
\end{align*}
\]

This means $E$ will be a Kramers singlet under $\mathcal{T}'$. This is consistent with that $\mathcal{T}'$ is not enough to protect the Dirac cones. Therefore, after gauging this state indeed becomes $E_{bT} M_f$.

From this, $E_{bT'} M_f$ can be constructed similarly, and $E_{bTT'} M_f$ can be obtained by combining
To obtain \( E_{\bar{T}T'M_f} \) and \( E_{\bar{T}'T'M_f} \), let the symmetries be assigned as

\[
\mathcal{T} : \psi \rightarrow \sigma_y \psi \\
Z_2 : \psi \rightarrow \tau_x \psi
\]  

(G.82)

One can show that this state becomes \( E_{\bar{T}T'M_{b-}} \) after gauging by using state-operator correspondence, but an alternative point of view can be obtained by considering this state as a descendant of the corresponding SPT of the \( SO(3) \times \mathcal{T} \) symmetric \( E_{f\frac{1}{2}M_{b\frac{1}{2}}} \), which has been described in details in Appendix G.4. To see it, denote the three generators of \( SO(3) \) by \( S_x, S_y \) and \( S_z \), and denote the generator of \( \mathcal{T} \) by \( t \). Now break the \( SO(3) \times \mathcal{T} \) to \( Z_2 \times ̂\mathcal{T} \), where the \( Z_2 \) is generated by \( \exp (iS_x \pi) \), and \( ̂\mathcal{T} \) is generated by \( \exp (iS_y \pi) \cdot t \). It is straightforward to check that the descendant state is \( E_{f̂\bar{T}'M_{b-}} \).

Now we turn to states with \( Z_2 \) acting as a charge conjugation. Let us start with the example corresponding to \((E_{\bar{T}T'M_{bZ}})_{-}\). This state is actually the gauged version of the free-fermion topological insulator in class CII, which has been discussed in Ref. Potter2016 (but using a different notation as here). The time reversal and charge conjugation symmetries are assigned as

\[
\mathcal{T} : \psi \rightarrow \sigma_y \psi \\
Z_2 : \psi \rightarrow \tau_y \psi^\dagger
\]  

(G.83)

Now let us first examine the \( \mathcal{T}' \) action on the two states corresponding to monopoles, whose action on \( \psi \) is

\[
\mathcal{T}' : \psi \rightarrow \sigma_y \tau_y \psi^\dagger
\]  

(G.84)

This is the same \( \mathcal{T}' \) action as in the corresponding SPT of \( E_{bT'M_f} \), so \( f_1^\dagger |0\rangle \) and \( f_2^\dagger |0\rangle \) correspond to Kramers singlets under \( \mathcal{T}' \).

Next let us examine the \( Z_2 \) action on \( f_1^\dagger |0\rangle \) and \( f_2^\dagger |0\rangle \). Notice the \( 2\pi \) flux background is converted into a \(-2\pi \) flux background, which also has two zero modes. Using the method in Ref. Potter2016, we argue that for such systems with two symmetry-protected Dirac cones, the value of charge-conjugation squared on the neutral monopole is the same as the value of charge-conjugation...
squared on the Dirac fermions. The simplest way to see this is to notice that these states have $\theta = 2\pi$. For the state corresponding to a trivial insulator, which has $\theta = 0$, the monopole has trivial quantum number, that is, the value of charge-conjugation squared is 1. Then one can tune $\theta$ by $2\pi$ to get a state corresponding to the topological insulator. In intermediate process of tuning $\theta$, the time reversal symmetry is generically broken. But the existence of such nontrivial topological insulator implies at the end the system will have time reversal symmetry when $\theta$ becomes $2\pi$. On the other hand, this process will not break the charge conjugation symmetry. Then according to the Witten effect, [409] the original $(1, 1)$ dyon now becomes the $(0, 1)$ monopole, and this new monopole has the value of charge-conjugation squared to be $-1$.

Ref. Potter2016 obtained this result by considering tuning $\theta$ of the $U(1)$ gauge theory. It can actually also be obtained directly by using state-operator correspondence. Recall it has been shown in Appendix G.4 that $M_1^\dagger \sim f_2^\dagger |\tilde{0}\rangle$ and $M_2^\dagger \sim -f_1^\dagger |\tilde{0}\rangle$, where $|\tilde{0}\rangle$ is the state with $-2\pi$ flux background and neither zero mode being occupied. Under charge conjugation, both charge and flux will be occupied. So under a convention of phase factors we can choose $|0\rangle \rightarrow f_1^\dagger f_2^\dagger |\tilde{0}\rangle$ under charge conjugation, then the monopole operators transform as

$$M_1 \sim f_1^\dagger |0\rangle \rightarrow f_2^\dagger f_1^\dagger f_2^\dagger |\tilde{0}\rangle = - f_1^\dagger |\tilde{0}\rangle \sim M_2^\dagger \tag{G.85}$$

$$M_2 \sim f_2^\dagger |0\rangle \rightarrow - f_1^\dagger f_1^\dagger f_2^\dagger |\tilde{0}\rangle = - f_2^\dagger |\tilde{0}\rangle \sim - M_1^\dagger$$

Again, unimportant $U(1)$ phase factors have been suppressed. The above transformation shows that the value of charge-conjugation squared is indeed $-1$ on the monopoles. Therefore, after gauging this state becomes $(E_{FTZ} M_{bZ})_-$.

Next we turn to the free-fermion SPT corresponding to $(E_{FT} M_{bT'})_-$, where the assignments of the time reversal and charge conjugation symmetries are

$$\mathcal{T}: \psi \rightarrow \sigma_y \psi \tag{G.86}$$

$$Z_2: \psi \rightarrow \psi^\dagger$$
Now we check the whether these monopoles are Kramers doublets under $T'$, whose action on $\psi$ is

$$T' : \psi \rightarrow \sigma_y \psi^\dagger \quad \text{(G.87)}$$

This is the same as the $T$ action in the SPT corresponding to $E_{bT}M_f$, so here the monopole must be a Kramers doublet under $T'$.

As for the value of charge-conjugation squared on the monopole,

$$f_1^\dagger |0\rangle \rightarrow f_1^\dagger \tilde{f}_1^\dagger |0\rangle \sim f_1^\dagger |\tilde{0}\rangle \sim M_1^\dagger$$

$$f_2^\dagger |0\rangle \rightarrow f_2^\dagger \tilde{f}_2^\dagger |\tilde{0}\rangle \sim -f_1^\dagger |\tilde{0}\rangle \sim M_2^\dagger$$

so the value of charge-conjugation squared is 1 for the monopoles. Therefore, after gauging this state becomes $(E_{fT}M_{bT'}Z)_-$.

Finally, for the free-fermion SPT corresponding to $(E_{fZ}M_{bT'}Z)_-$, the assignments of the time reversal and charge conjugation symmetries are

$$T : \psi \rightarrow \sigma_y \tau_y \psi$$

$$Z_2 : \psi \rightarrow \tau_y \psi^\dagger$$

This state has the same $Z_2$ action as the one giving rise to $(E_{fT}M_{bZ})_-$, and the same $T'$ action as the one giving rise to $(E_{fT}M_{bT'})_-$. In light of the previous discussion, the monopole should have charge-conjugation squared to be $-1$ and be a Kramers doublet under $T'$. Therefore, after gauging this state becomes $(E_{fZ}M_{bT'}Z)_-$.

By now the constructions of free-fermion SPTs corresponding to states in (G.77) are given. Before we finish this subsection, we make some remarks on free-fermion topological insulators with time reversal and a unitary $Z_2$ symmetry, which may or may not act as a charge conjugation. In all these free-fermion topological insulators, the surface will always have an even number of Dirac cones in order to have $\theta = 0$. If it has 4 Dirac cones on the surface, the corresponding $U(1)$ quantum spin liquid in general has a trivial monopole. In order to get a $U(1)$ quantum spin liquid with nontrivial monopole, the corresponding free-fermion SPT should have only 2 surface Dirac cones. We have actually exhausted all possible free-fermion topological insulators.
with two surface Dirac cones, and they only give the 10 $U(1)$ quantum spin liquids in (G.77) that have nontrivial monopole. On the other hand, if a nontrivial fermionic topological insulator is equivalent to a bosonic SPT, the monopole must also be trivial. [339, 406] Thus, if a state with fermionic charge and nontrivial monopole is anomaly-free and distinct from the above three (such as $(E_fM_{bT'})_-$), it implies the existence of an intrinsically interacting fermionic SPT, which is a nontrivial fermionic SPT that cannot be realized by free-fermions and is not equivalent to a bosonic SPT. [410, 411] These SPTs are very interesting, but in the discussion below we will argue that no other spin liquid state with fermionic charge and nontrivial monopole is anomaly-free, which means no such intrinsically interacting fermionic SPT can be found with $U(1)$, time reversal and $Z_2$ symmetries (even if the fermions are allowed to transform projectively under these symmetries).

G.10.2 Anomaly of $(E_{bT}M_{bT'})_-$

In this subsection, by using the same logic as before, we will examine the anomaly of $(E_{bT}M_{bT'})_-$. That is, we will consider the corresponding SPT from the perspective of $M_{bT'}$, and check whether it is possible to have a consistent surface topological order. However, unlike in the case of $(E_{bZ}M_{bZ})_-$, where we can reach the conclusion by quite general arguments, here we need to examine some rather detailed properties of the surface states.

Again, we will first condense the bound state of two $M_{bT'}$ on the surface, which reduces the surface symmetry to $T' \times Z_2$. We would like to point out that there are two possibilities for the surface at this point: the surface can either be a simple superfluid, or the surface superfluid has to coexist with another anomalous topological order. The latter happens if the bulk is still a non-trivial SPT even if the bulk symmetry is broken down to $T' \times Z_2$, in which case there must be another anomalous surface topological order of an SPT with $T' \times Z_2$ symmetry, if this symmetry is to be preserved.

For the case of a simple superfluid surface, we can show there is inconsistency of the surface topological order. As for the more complicated case where the surface superfluid has to coexist with another anomalous topological order, we need the properties of the surfaces of 3D bosonic SPTs with $T' \times Z_2$ symmetry, which, to the best of our knowledge, are lacking in the literature.

369
So we will first discuss the classification of such SPTs, and then show that there will still be some inconsistency even for the more complicated case. This leads us to concluding that there is no such corresponding SPTs that can become \((E_{bT} M_{bT'})_−\) after gauging, which means \((E_{bT} M_{bT'})_−\) is anomalous.

Before the detailed discussion on this problem, let us first collect a few useful tools that will be applied repeatedly below.

1. In a topological order, a particle always has the same topological spin as its anti-particle.

   That is, denote \(a^{-1}\) as the anti-particle of \(a\), then

   \[
   \theta_a = \theta_{a^{-1}} \quad \text{(G.90)}
   \]

2. Suppose \(a\) and \(b\) are two anyons in a topological order. Suppose \(c\) is a possible fusion product of \(a\) and \(b\), that is, \(a \times b = N_{ab}^c c + \cdots\) with \(N_{ab}^c\) the fusion multiplicity. Then \((R_{ab}^c)^2 = \frac{\theta_c}{\pi_a \pi_b}\), where \((R_{ab}^c)^2\) is the mutual braiding between \(a\) and \(b\) when their fusion product is fixed to be \(c\), and \(\theta_a\) is the topological spin of \(a\). In the case of Abelian topological order, the mutual braiding between \(a\) and \(b\) can be simply denoted as \(\theta_{a,b}\), and the above formula becomes

   \[
   \theta_{a,b} = \frac{\theta_c}{\theta_a \theta_b} \quad \text{(G.91)}
   \]

3. Braiding and fusion commute in a topological order. For example, in an Abelian topological order,

   \[
   \theta_{ab,c} = \theta_{a,c} \theta_{b,c} \quad \text{(G.92)}
   \]

4. If the time reversal partner of an anyon \(a\) is \(b\), and \(c = a \times b\) is the bound state of \(a\) and \(b\), then the Kramers parity of \(c\) is determined by

   \[
   T_c^2 = \theta_c \quad \text{(G.93)}
   \]

370
Simple superfluid

We start our discussion with the case of a simple superfluid surface. This superfluid has vortices with vorticity quantized in units of $\pi$, and the minimal trivial vortex is the $4\pi$ vortex. We will condense these $4\pi$ vortices to restore the full symmetry of the surface and get a symmetric gapped surface topological order, where the $U(1)$ charge is quantized in units of $1/2$. The $\pi$ vortices and the $2\pi$ vortices will remain gapped, and we will denote the $\pi$ vortex by $v$ and the $2\pi$ vortex by $E_{bT}$.

The vortex condensation above will in general generate a charge-$1/2$ boson, which we denote by $X$. Physically, $X$ is the $2\pi$ vortex of the $4\pi$ vortices. This $X$ should be an Abelian boson. To see this, let us go to the energy scale below which we can consider only the $4\pi$ vortices that are to be condensed. Limiting ourselves below this energy scale should not change the topological data of $X$. Although the $\pi$ and $2\pi$ vortices are nontrivial, below this energy scale they do not play any role. Then because the $4\pi$ vortices are trivial bosons, $X$, the $2\pi$ vortex of the $4\pi$ vortices, is expected to be a simple Abelian boson.

The bound state of two $X$’s can be combined with $M_{bT'}$ to generate a charge-neutral bosonic excitation, which will be denoted by $N$. The bound state of two $N$’s have trivial braiding with all other excitations, so this should be viewed as a local excitation. Therefore, the particle contents of the surface theory can be written as

$$
\{1, X, N, X^{-1}, v, E_{bT}, v^{-1}\} \times \{1, M_{bT'}\}
$$

(G.94)

The various bound states of these excitations are understood to be implicitly displaced. Also, $X^{-1}$ represents the excitation that can fuse with $X$ into the trivial vacuum, 1, which does not carry any quantum number, and it should be distinguished from $XM_{bT'}^\dagger$.

Below we will determine the braiding, fusion and symmetry assignments of these excitations. Without loss of generality, we will always take $v$ to be neutral, because its charge can always be cancelled by binding it with certain amount of $X$ and $M_{bT'}$.

We start with braiding. For self-braiding, the only uncertain part is about $v$: it can either be Abelian or non-Abelian. Now we turn to mutual braiding. The mutual braiding within the
charge sector (built up with $X$ and $N$) is always trivial. For the vortex sector (built up with $v$ and $E_{bT}$), the braiding between $v$ and $E_{bT}$ is trivial because $v$ is neutral and $E_{bT}$ is the remnant of the $2\pi$ vortex, and the braiding between $v$ and $v^{-1}$ is to be determined.

The mutual braiding between the charge sector and the vortex sector can be determined in the following way. Because condensing $X$ will make the surface back into the simple superfluid, we can view $X$ as something that is condensed in the superfluid phase. From the Meissner effect we know the vortices come with certain fluxes in the superfluid phase, and this combined object of vortices and fluxes should be local with respect to the $X$ condensate. Therefore, the mutual braiding between the vortices themselves with the $X$ condensate is the conjugate of the charge-flux Aharonov-Bohm phase. This tells us

\[ \theta_{X,v} = -i, \quad \theta_{X,E_{bT}} = \theta_{N,v} = -1, \quad \theta_{N,E_{bT}} = 1 \]  

(G.95)

Notice the third relation comes from the identification $N = X^2 M_{bZ}^4$ and that $X^2$ is condensed in the superfluid ($N$ is not condensed in the superfluid, so we cannot say $\theta_{v,N} = 1$ because $N$ is neutral). The mutual braiding listed here will be used repeatedly below.

Now we turn to fusion. Most fusion rules can be determined by the charge and vorticity assignment:

\[ X \times X = NM_{bZ}, \quad N \times N = 1, \quad E_{bT} \times E_{bT} = 1 \]  

(G.96)

However, there is some flexibility for $v$. For example, even if the $v$ is Abelian, we can have either $v \times v = E_{bT}$ or $v \times v = E_{bT} \times N$. Of course when $v$ is non-Abelian, we must have $v \times v = E_{bT} + E_{bT} N$ (with potential fusion multiplicities suppressed). Because $N$ is a boson that is local with respect to $E_{bT}$, $N$ must have trivial braiding with $v$ in this non-Abelian case, otherwise $v$ and its anti-particle would have opposite topological spins, which violates (G.90) and is thus disallowed. However, $\theta_{v,N} = -1$. This implies $v$ cannot be non-Abelian. \footnote{In contrast, if the local particle is a fermion, $v$ can still be non-Abelian.} Furthermore,
for the same reason, the fusion rule for \( v \) has to be

\[
v \times v = E_{bT}
\]  

(G.97)

Finally we discuss the symmetry assignment. The \( U(1) \) charges of these excitations are clear: \( X \) carries half charge, \( M_{bT} \) carries unit charge, and other excitations are neutral. The assignment of the \( \mathcal{T} \) and \( Z_2 \) symmetries is constrained by some general rules. First, \( \mathcal{T} \) should conjugate the topological spins of the excitations, and \( Z_2 \) should keep their topological spins. Second, the behavior of charge and vorticity under various symmetries is fixed. For example, because the charge flips and the vorticity does not change under \( \mathcal{T} \), \( \mathcal{T} \) will take \( v \) to either \( v \) or \( vN \). Because of the fusion rule, \( v \times v = E_{bT} \), \( v \) cannot become \( v \) under time reversal. That is, \( v \) should go to \( vN \).

Putting all these constraints together, there is actually not too much freedom for this topological order. One choice is the \( Z_4 \) gauge theory listed in Table G.6,\(^2\) and the only thing that one can modify on top of this state is to change the values of \( Z_2 \) for \( X \) and \( v \), and the value of \( T'^2 \) for \( X \). Notice in all these cases, \( N \) is a Kramers doublet under \( \mathcal{T}' \).

The above theory is actually inconsistent. To see this, notice that since the \( \mathcal{T} \) partner of \( v \) is \( vN \) and \( \theta_{v,N} = -1 \), \( v \) must be a semion or anti-semion, so that \( \mathcal{T} \) can conjugate the topological spin of \( v \). This means the bound state of \( v \) and its \( \mathcal{T}' \) partner, \( v^{-1}N \), is a boson, so this bound state should be a Kramers singlet under \( \mathcal{T}' \) according to (G.93). However, as discussed above, this bound state is \( v \times v^{-1}N = N = X^2 M_{bT}^{\dagger} \), which is a Kramers doublet under \( \mathcal{T}' \). This contradiction shows that the simple superfluid surface is inconsistent.

**Superfluid coexisting with another anomalous topological order**

Now we turn to the case where the surface superfluid has to exist with another anomalous topological order. As discussed earlier, this happens if the bulk remains to be a nontrivial SPT when the bulk symmetry is also reduced to \( \mathcal{T}' \times Z_2 \). We will call such SPTs the reduced bulk SPTs. To complete the discussion, we need the properties of 3D bosonic SPTs with this symmetry, which

\(^2\)To fit into the usual notation of a \( Z_4 \) gauge theory, one can take the \( Z_4 \) charge to be \( X \), and take the \( Z_4 \) flux to be \( Xv \).
Table G.6: Symmetry assignments of the surface topological order from the simple superfluid surface of the corresponding SPT of $(E_{bT} M_{bT})^-$. The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the charges of these excitations under $U(1)$. The third row lists the time reversal partners of these excitations. The fourth row lists the values of $T^2$ of these excitations, with empty entries representing that $T^2$ is not well-defined. The fifth row lists the $Z_2$ partners of these excitations. The sixth row lists the values of $Z_2^2$ of these excitations. The seventh row lists the $T'$ partners of these excitations. And the last row lists the values of $T'^2$ of these excitations.

\[
\begin{array}{|c|c|c|c|c|}
\hline
& X & N & v & E_{bT} \\
\hline
U(1) & \frac{1}{2} & 0 & 0 & 0 \\
\hline
T & X^{-1} & N & vN & E_{bT} \\
\hline
T^2 & -1 & -1 & -1 & -1 \\
\hline
Z_2 & X^{-1} & N & v^{-1} & E_{bT} \\
\hline
Z_2^2 & \pm1 & 1 & \pm1 & 1 \\
\hline
T' & X & N & v^{-1}N & E_{bT} \\
\hline
T'^2 & \pm1 & -1 & \pm1 & \pm1 \\
\hline
\end{array}
\]

will be discussed below.

3D bosonic SPTs with $T' \times Z_2$ symmetry

Notice there exists local Kramers doublet under $T'$, so more precisely, the symmetry group of the surface superfluid should be denoted by $Z_4^T \times Z_2$. From group cohomology, the classification of such SPTs is $Z_3^2$, and there should still be another SPT whose surface is $efmf$, and this SPT is beyond group cohomology. So we propose that the complete classifications of such SPTs is $Z_4^2$. This proposal is further supported by the classification of 3D bosonic SPTs with $Z_4^T \times Z_2$ symmetry, where $Z_4^T$ is a reflection symmetry that results in a trivial action when acted four times. SPTs with $Z_2 \times Z_4^T$ are believed to have the same classification of SPTs with $Z_2 \times Z_4^P$, where the latter are classified by $Z_2^4$. [356]

What are the surface topological orders of the other three root states? We show that they can all be $Z_2$ topological orders, and they are denoted by $(e T'_i m T'_i)_{T'}$, $eZmZ$ and $e T'_i mZ$. Below we explain the properties of these states.

The first state, $(e T'_i m T'_i)_{T'}$, is protected by $T'$ alone, and in this state $e$ and $m$ are exchanged under $T'$. Furthermore, $T'$ acting on $e$ or $m$ four times gives $-1$ (the meaning of $T'_i$).
of $Z_2$ is trivial on both $e$ and $m$. In fact, this state is the descendant of $(eCmC)_{T'}e$ when $e^4$ is condensed without breaking time reversal, and $(eCmC)_{T'}e$ is a surface state of the bosonic topological insulator made of Kramers bosons.

To justify that this is a legitimate surface state of an SPT protected by $T'$, we need to show this descendant state is still a nontrivial SPT. That is, the bosonic topological insulator made of Kramers bosons is still a nontrivial SPT when double charge is condensed without breaking time reversal. This can be seen by checking the time reversal domain wall of this state. Consider breaking $T'$ in two opposite ways in the two sides of a 2D domain wall, while keeping a unitary $Z_2$ symmetry intact through the entire system (this unitary $Z_2$ is just the symmetry generated by acting the generator of $T'$ twice). Notice before the double charges are condensed, the time reversal domain wall of this bosonic topological insulator is the elementary bosonic integer quantum Hall state, [349] because it has $\sigma_{xy} = 2e^2/h$. [337] When the double charge is condensed, this bosonic integer quantum Hall state becomes the Levin-Gu state. [412, 413] That is to say, the time reversal domain wall of this descendant state is a Levin-Gu state. But this cannot happen unless the original $T'$ symmetric system is a nontrivial SPT. 3

The above discussion establishes that there is a 3D bosonic SPT protected by $T'$, and its surface can be $(eT'_i mT'_i)_{T'}$. For our purposes, it will be useful to think about what state this SPT becomes if the symmetry is enhanced to the full $((U(1) \times Z_2) \times T)/Z_2 = (O(2) \times T)/Z_2$ symmetry of $M_{bT'}$, which can be obtained by considering what the surface topological order becomes when the symmetry is enhanced. Because the $\pi$ rotation of the $U(1)$ is locked with acting $T'$ twice, $e$ and $m$ should carry half charge under $U(1)$. The entire symmetry assignment of this surface state is shown in Table G.7.

It is straightforward to check that this state can be the surface topological order of the corresponding SPT of $(E_{fT} M_{bT'})_\perp$ (viewed from the perspective of $M_{bT'}$). This observation implies that if the reduced bulk SPT is $(eT'_i mT'_i)_{T'}$, we can reduce the surface state into a simple superfluid by coupling the original SPT to the corresponding SPT of $(E_{fT} M_{bT'})_\perp$. Because coupling $(E_{bT} M_{bT'})_\perp$ and $(E_{fT} M_{bT'})_\perp$ should result in $(E_{f} M_{bT'})_\perp$, if we can show in the scenario of a

\[375\] More precisely, this is because the Levin-Gu state is the root state of 2D $Z_2$ SPTs, which means there is no 2D $Z_2$ symmetric short-range entangled bosonic state that becomes the Levin-Gu state when it is stacked with its time reversal partner.
Table G.7: Symmetry assignments of the surface topological order of the corresponding SPT of $(E_{f T} M_{b T'})_-$. The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the charges of these excitations under $U(1)$. The third row lists the time reversal partners of these excitations. The fourth row lists the values of $T^2$ of these excitations, with empty entries representing that $T^2$ is not well-defined. The fifth row lists the $Z_2$ partners of these excitations. The sixth row lists the $Z_2^2$ of these excitations. The seventh row lists the $T'$ partners of these excitations. And the last row lists the values of $T'^2$ of these excitations, with $\pm i$ standing for that $T'^4 = -1$ on the excitation.

<table>
<thead>
<tr>
<th></th>
<th>$e$</th>
<th>$m$</th>
<th>$\epsilon \equiv \epsilon m^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U(1)$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
</tr>
<tr>
<td>$T$</td>
<td>$m^\dagger$</td>
<td>$e^\dagger$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>$T^2$</td>
<td></td>
<td></td>
<td>$-1$</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>$e^\dagger$</td>
<td>$m^\dagger$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>$Z_2^2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$T'$</td>
<td>$m$</td>
<td>$e$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>$T'^2$</td>
<td>$\pm i$</td>
<td>$\pm i$</td>
<td>1</td>
</tr>
</tbody>
</table>

simple superfluid surface, no SPT made of $M_{b T'}$ can become $(E_{fT} M_{bT'})_-$ after gauging, it is sufficient to show no SPT with $(e T'_i m T'_i)_{T'}$ reduced bulk SPT can become $(E_{bT} M_{bT'})_-$ after gauging.

Next we turn to explaining the properties of $eZmZ$. In this surface $Z_2$ topological order, the $Z_2$ symmetry acting on $e$ or $m$ twice gives a $-1$ phase factor, and $T'$ acts trivially on $e$ and $m$.

Again, we need to show that the bulk with this surface topological order is a nontrivial SPT with $T' \times Z_2$ symmetry, or equivalently, that $eZmZ$ is anomalous with $T' \times Z_2$ symmetry.

The way to understand the anomaly of $eZmZ$ is to relate it to $eCmC$, the surface state of a nontrivial SPT with $T' \times U(1)$ symmetry. Notice this symmetry is not $(U(1) \times T)/Z_2$, the symmetry of charged Kramers bosons. In particular, the $\pi$ rotation of the $U(1)$ here is not locked with acting time reversal twice, as in the latter symmetry. There is a bosonic topological insulator at $\theta = 2\pi$ with $T' \times U(1)$ symmetry, and this is independent of the presence of local bosons that are Kramers doublets under $T'$. The surface state of this bosonic topological insulator is $eCmC$, which means a $Z_2$ topological order with both $e$ and $m$ carrying half charge under $U(1)$, and $T'$ acts trivially on $e$ and $m$. Again, the time reversal domain wall of this bosonic topological insulator will have the character of an elementary bosonic integer quantum Hall state. Break-
ing the $U(1)$ symmetry down to $Z_2$ results in the $eZmZ$ state, and as before, the time reversal domain wall of this descendant 3D state will be a Levin-Gu state, which is disallowed unless the 3D bulk is a nontrivial SPT. This shows that $eZmZ$ is still anomalous with $T' \times Z_2$ symmetry, and we also see $T'$ plays a role in protecting this state, even though it appears to act on $e$ and $m$ trivially.

Again, it will be useful for us to understand what this state becomes when the symmetry is enhanced to the full $(O(2) \times T)/Z_2$ symmetry of $M_{bt'}$. Because only the $Z_2$ acts nontrivially in this state, when the symmetry is enhanced, $U(1)$ and $T'$ should still act trivially. This means there is a corresponding SPT of $(E_b M_{bt'})_-$ that becomes $eZmZ$ when the symmetry is broken down to $T' \times Z_2$. Therefore, when the reduced bulk SPT is $eZmZ$, we can always cancel the anomaly of $eZmZ$ by coupling the original putative SPT to this corresponding SPT of $(E_b M_{bt'})_-$. 

Lastly, we turn to discuss $eT'_i mZ$, which is a $Z_2$ topological order with $e$ having $T'^4 = -1$ and $m$ having $Z_2$ squaring to $-1$. As argued in Appendix G.1, in order for $e$ to have $T'^4 = -1$, $T'$ should attach $M_{bt'}$, a local Kramers doublet under $T'$, to $e$. This surface state is anomalous because when two $Z_2$ fluxes are inserted in the system, both $m$ and $\epsilon$ will see a nontrivial phase factor when moving around it. But this is a local process, so this $-1$ phase factor must be compensated by something nucleated in the $Z_2$ flux when it is inserted, and this nucleated object must be $e$. That is to say, a local process produces an object with $T^4 = -1$, which is absent in the system by assumption. Therefore, this state is anomalous.

What does $eT'_i mZ$ become when the symmetry is enhanced to the full $(O(2) \times T)/Z_2$ symmetry? It turns out the symmetry cannot be enhanced to the full $(O(2) \times T)/Z_2$ symmetry. In other words, there is no bosonic SPT with $(O(2) \times T)/Z_2$ symmetry whose descendant can be $eT'_i mZ$. However, we still know that, because the $\pi$ rotation of the $U(1)$ is locked with acting $T'$ twice, when the $4\pi$ vortices are condensed and the full symmetry is recovered on the surface, $e$ should carry half charge of $U(1)$, and $m$ carries integer charge but has $Z_2$ squaring to $-1$.

For completeness, we also mention that if the reduced bulk SPT is $efmf$, when the full symmetry is recovered, none of $e$ and $m$ get nontrivial action under the symmetry, and its anomaly

---

4We point out that in order to establish that $eZmZ$ is anomalous, it is actually important that $T'$ acts trivially on both $e$ and $m$, so that this $eZmZ$ state can be viewed as the descendant of $eCmC$ after the $U(1)$ is broken down to $Z_2$. 

377
can be cancelled by just coupling it with another effm state with the full symmetry, which after gauging becomes \((E_bM_{bT'})_\cdot\).

The above discussion implies that, in the scenario where the surface superfluid has to coexist with another anomalous topological order, in order to show there is no SPT made of \(M_{bT'}\) that can become \((E_{bT}M_{bT'})_\cdot\) after gauging, it is sufficient to show:

1. In the scenario of a simple superfluid surface, there is no SPT made of \(M_{bT'}\) that can become \((E_fM_{bT'})_\cdot\).

2. In the scenario of \(eT'_imZ\) reduced bulk SPT, there is no SPT made of \(M_{bT'}\) that can become \((E_{bT}M_{bT'})_\cdot\).

3. In the scenario of \(eT'_imZ\) reduced bulk SPT, there is no SPT made of \(M_{bT'}\) that can become \((E_fM_{bT'})_\cdot\).

Below we will show these three statements in turn.

**Simple superfluid surface for \((E_fM_{bT'})_\cdot\)**

We start from the first statement. The similar argument as before implies that the surface topological order in this case will be a \(Z_4\) topological order:

\[
\{1, X, N, X^{-1}, v, E_f, v^{-1}\} \times \{1, M_{bT'}\}
\]  

with the symbols standing for parallel excitations as before, and a difference is that here \(E_f\) is a fermion with no symmetry fractionalization in terms of \(T\) and \(Z_2\).

In this case, most of the topological data (braiding and fusion) will be the same as in the simple superfluid for \((E_{bT}M_{bT'})_\cdot\). The only difference is that, because \(E_f\) is a fermion, the fusion product of two \(v\)'s should be \(E_fN\), otherwise the antiparticle of \(v\) will have a different topological spin from \(v\). That is,

\[
v \times v = E_fN
\]
Because $E_fN$ is still a fermion, $v$ has topological spin $\theta_v = \exp \left( i \left( \frac{\pi}{4} + \frac{n\pi}{2} \right) \right)$ with $n$ an integer. In order for $T$ to conjugate $\theta_v$, the $T$ partner of $v$ has to be $vX$ or $vX^{-1}$. This then changes the charge of $v$, which is not legitimate.

This establishes the first statement: in the scenario of a simple superfluid surface, there is no SPT made of $M_{bT'}$ that can become $(E_fM_{bT'})_-$.

\[ eT'_i mZ \text{ reduced bulk SPT for } (E_bT'M_{bT'})_- \]

Now we turn to the second statement, where the surface superfluid has to coexist with the anomalous topological order, $eT'_i mZ$. In the superfluid phase of the surface, the vortices and anyons of this anomalous topological order are distinct. One of their distinctions is that the vortices carry (logarithmically) expensive energy cost, while the anyons have a finite energy gap. The fusion rules of $e$ and $m$ in the superfluid phase is that

\[ e \times e = M_{bT'}, \ m \times m = 1 \quad (G.100) \]

These seemingly innocuous fusion rules actually deserve further clarification. Purely in terms of topological sectors, there is no difference in $M_{bT'}$ and 1 in the right hand sides of these fusion rules, because they can be turned into each other by binding an $M_{bT'}^\dagger$, a local excitation. However, in terms of symmetry quantum numbers, $M_{bT'}$ and 1 are of course different. It makes sense to talk about fusing two $e$ particles or two $m$ particles that have the same symmetry quantum number (with the difference due to attaching an $M_{bZ}$ resolved), and the above equations should be interpreted as the fusion rules of fusing two identical $e$’s or $m$’s. This distinction is important when we try to determine the fusion rules of $e$ and $m$ after the surface gets into the symmetric topologically ordered phase, where the right hand sides of these fusion rules can potentially be modified by multiplying something condensed in the superfluid phase.

When $4\pi$ vortices are condensed and the full symmetry is restored, $e$ will remain deconfined because it will carry half charge under $U(1)$ ($m$ will of course also remain deconfined because it is neutral). Then the surface topological order can be written as

\[ \{1, X, N, X^{-1}, v, E_{bT}, v^{-1}, e, m\} \times \{1, M_{bT'}\} \quad (G.101) \]
where the symbols have parallel meaning as in the case of a simple superfluid. As before, the various bound states of these excitations are understood to be implicitly displayed.

Again, \( X \) will be an Abelian boson that carries half charge, and \( N \) will be a boson that is a Kramers doublet under \( T' \). Because condensing \( X \) makes the surface back to the superfluid coexisting with \( eT'_mZ \), the superfluid phase can again be viewed as a condensate of \( X \). This has two important consequences.

First, the mutual braiding in (G.95) still holds, and \( \{ e, m \} \) will have trivial braiding with \( \{ X, N, X^{-1} \} \). Then \( N^2 \) is trivial, and (G.96) still holds. However, the mutual braiding between \( v \) and \( \{ e, m \} \) is undetermined at this point. We only know, because \( \theta_{E_{bT},m} = 1 \), we should have \( \theta_{v,m} = \pm 1 \).

Second, no condensate in the superfluid can be multiplied to the right hand sides of the fusion rules in (G.100) without violating charge conservation, which means that (G.100) is also the right fusion rules for \( e \) and \( m \) in the topologically ordered phase.

Now let us determine the fusion rule of two \( v \)'s. Just from charge conservation and vorticity conservation, there seem to be many possible fusion products of two \( v \)'s:

\[
E_{bT}, E_{bT}N, E_{bT}m, E_{bT}mN,
E_{bT}eX^{-1}, E_{bT}eX^{-1}N, E_{bT}eX^{-1}, E_{bT}eX^{-1}N
\]

However, based on some topological arguments, in the following we can rule out all of them but \( E_{bT}m \) and \( E_{bT}mN \).

To see this, first notice that for each Abelian anyon, it has a unique braiding phase factor with all fusion products of two \( v \)'s. Because of this, some of the above cannot simultaneously be the fusion products. For example, because \( E_{bT} \) and \( E_{bT}m \) have different braiding phase factors with \( e \), they cannot both be the fusion products of two \( v \)'s. Using this, we see the fusion products can be one of the four possible pairs

\[
\{ E_{bT}, E_{bT}N \}, \{ E_{bT}m, E_{bT}mN \},
\{ E_{bT}eX^{-1}, E_{bT}eX^{-1}N \}, \{ E_{bT}eX^{-1}, E_{bT}eX^{-1}N \}
\]

380
However, because $N$ is a boson with $\theta_{v,N} = -1$, within each pair at most one of them can be the fusion product of two $v$’s, otherwise the anti-particle of $v$ will have opposite topological spin of $v$, and (G.90) is violated. This means $v$ has to be Abelian again, and the entire topological order is Abelian. Together with $\theta_{E_{bT},v} = 1$, this also implies $E_{bT}N$ cannot be a fusion product of two $v$’s. Furthermore, because $\theta_{X,v} = -i$, the last four excitations cannot be fusion products of two $v$’s, otherwise $v$ would have an antiparticle that has a different topological spin from itself.

In fact, $E_{bT}$ cannot be a fusion product of two $v$’s, either. This is because $\theta_{E_{bT},e} = -1$, which means $\theta_{v,e} = \pm i$ and $\theta_{v,e^2} = -1$, if $v \times v = E_{bT}$. However, $e^2$ is local, so $\theta_{v,e^2} = 1$. This contradiction implies that $E_{bT}$ cannot be the fusion product of two $v$’s.

So we are finally left with two possibilities:

\[ v \times v = E_{bT}m \quad (G.104) \]

or

\[ v \times v = E_{bT}mN \quad (G.105) \]

In the first possibility, $\theta_{v,m} = 1$, while $\theta_{v,m} = -1$ in the second possibility.

Now we discuss the symmetry assignment. Recall that in the superfluid phase, the topological sectors of $e$ and $m$ are transformed as

\[ \mathcal{T}^\prime: \ e \to e_{bT}, \ m \to m \quad (G.106) \]

under $\mathcal{T}'$, and

\[ Z_2: \ e \to e, \ m \to m \quad (G.107) \]

or

\[ Z_2: \ e \to e_{bT}, \ m \to m \quad (G.108) \]
under $Z_2$. Notice that the above expressions only imply the action on the topological sectors, and in this case there are two possibilities for the $Z_2$ transformation.

In the symmetric topologically ordered phase, the right hand sides of these transformation rules can be multiplied by something condensed in the superfluid phase. Also, according to the values of $T'^2$ of $e$ and $m$ and the fact that the $\pi$ rotation of $U(1)$ is locked with acting $T'$ twice, $e$ carries half charge and $m$ carries no charge under $U(1)$ in the topologically ordered phase. In order for $T'$ to maintain the $U(1)$ charge, and for $Z_2$ to flip the $U(1)$ charge, the unique choice for the transformation rules of $e$ and $m$ in the topologically ordered phase are

$$e \rightarrow eN, \quad m \rightarrow m$$

under $T'$, and

$$e \rightarrow eM_{bT'}, N, \quad m \rightarrow m$$

(G.109)

or

$$e \rightarrow eM_{bT'}, \quad m \rightarrow m$$

(G.110)

under $Z_2$, corresponding to the two possible $Z_2$ transformations in the superfluid phase, respectively.

To further determine the symmetry assignments, it is crucial to determine the symmetry actions on $v$. Let us consider the $Z_2$ action on $v$ first, and we will begin with the case where $v \times v = E_{bT}m$ and $e \rightarrow eM_{bT'}N$ and $m \rightarrow m$ under $Z_2$. Notice in this case $\theta_{v,m} = 1$. As the $Z_2$ flips both the charge and the vorticity, the options for the $Z_2$ partner of $v$ are:

$$v^{-1}, v^{-1}N, v^{-1}m, v^{-1}mN,$$

$$v^{-1}eX^{-1}, v^{-1}eX^{-1}N, v^{-1}\epsilon X^{-1}, v^{-1}\epsilon X^{-1}N$$

(G.112)

All of them except $v^{-1}$ and $v^{-1}m$ can be ruled out, because in those cases the $Z_2$ action cannot keep the topological spin of $v$ invariant. If the $Z_2$ partner of $v$ is $v^{-1}$, then under the $Z_2$ transfor-
mation $\theta_{e,v}$ becomes

\[
\theta_{eNM_{\beta T}^\dagger,v^{-1}} = \theta_{eNM_{\beta T}^\dagger,v^3} = \theta_{e,v^3} \theta_{N,v^3} = \theta_{e,v} \theta_{N,v}^2 = -\theta_{e,v}
\]

which is disallowed. In the above we have used (G.91) and (G.92). The above discussion implies that the $Z_2$ partner of $v$ can only be $v^{-1}m = vE_{bT}$.

Just from that $\mathcal{T}$ flips the charge but keeps the vorticity, the options for the $\mathcal{T}$ partner of $v$ are:

\[
v, vN, vm, vmN, veX^{-1}, veX^{-1}N, v\epsilon X^{-1}, v\epsilon X^{-1}N
\]

The last four can be ruled out due to some topological reasons. For example, suppose $v$ becomes $veX^{-1}$ under $\mathcal{T}$. In order for $\mathcal{T}$ to conjugate the topological spin of $v$, using the (G.95) and $\theta_{e,v} = \pm 1$, the topological spin of $v$ must be $\theta_v = \exp\left(i\left(\pm \frac{\pi}{4} + n\pi\right)\right)$, with $n$ an integer. Then the bound state of two $v$’s must be a fermion, contradicting to the fusion rule $v \times v = E_{bT}m$. This means $veX^{-1}$ cannot be the $\mathcal{T}$ partner of $v$. Similar arguments show that $veX^{-1}N, v\epsilon X^{-1}$ and $v\epsilon X^{-1}N$ cannot be the $\mathcal{T}$ partner of $v$, either.

If the $\mathcal{T}$ partner of $v$ is $vm$, then $v$ is either a boson or a fermion, and $\theta_{v,vm} = 1$. But $\theta_{v,vm}$ should be locked to the Kramers parity of $vvm = E_{bT}$, which is $-1$. This contradiction implies that $vm$ cannot be the $\mathcal{T}$ partner in this case. The time reversal partner of $v$ can also not be $vNm$, because in this case under time reversal $\theta_{v,e}$ becomes

\[
\theta_{vNm_{\beta T}^\dagger,e} = -\theta_{v,e} \neq \theta_{v,e}^* \tag{G.115}
\]

which is disallowed. This is actually another reason why $vm$ cannot be the time reversal partner of $v$, because $vm$ cannot conjugate $\theta_{v,e}$, either. If the time reversal partner of $v$ is $v$, then $v$ has a well defined Kramers parity, but its $Z_2$ partner, $v^{-1}m = vE_{bT}$, has an opposite Kramers parity.
This is disallowed, otherwise $Z_2$ and time reversal cannot commute for the spin liquid. To see this, suppose $v$ transforms under time reversal as

$$v_i \rightarrow T_{ij}v_j \quad \text{(G.116)}$$

and under $Z_2$ as

$$v_i \rightarrow C_{ij}\tilde{v}_j \quad \text{(G.117)}$$

Its $Z_2$ partner, $\tilde{v}$, transforms under time reversal as

$$\tilde{v}_i \rightarrow \tilde{T}_{ij}\tilde{v}_j \quad \text{(G.118)}$$

Because $v_i^*M_{ij}v_j$ is a local operator that has no charge or vorticity, its Kramers parity should be 1 and $Z_2$ should commute with time reversal on this operator, for any matrix $M$. This requires $T^*T = \pm 1$, $\tilde{T}^*\tilde{T} = \pm 1$, and $TC = e^{i\phi}C^*\tilde{T}$, with $\phi$ a phase. Taking all these together, we get

$$T^*T = C\tilde{T}^*\tilde{T}C^{-1} = \tilde{T}^*\tilde{T} \quad \text{(G.119)}$$

That is, $v$ and $\tilde{v}$ should have the same Kramers parity. Notice to claim that for $v_i^*M_{ij}v_j$ the Kramers parity is 1 and $Z_2$ commutes with $T$, it is important that this operator is not only local, but also carries no charge or vorticity, otherwise by a gauge transformation its Kramers parity and the commutation relation can be changed.

So the time reversal partner can only be $vN$. Notice in this case $v$ must be a semion or anti-semion, because $\theta_{v,N} = -1$ and time reversal conjugates the topological spin of $v$. Then $\theta_{vvN} = 1$, which means the Kramers parity of $vvN = E_{bT}mN$ is 1, so $N$ has to be a Kramers doublet.

The above discussion implies that if $v \times v = E_{bT}m$ and the $Z_2$ action on $e$ and $m$ is given by (G.110), the $Z_2$ partner of $v$ can only be $vE_{bT}$, and the $T$ partner of $v$ can only be $vN$. Using similar arguments, one can actually check that if $v \times v = E_{bT}mN$ and the $Z_2$ action on $e$ and $m$ is given by (G.110), the $Z_2$ partner of $v$ can still only be $vE_{bT}$, and the $T$ partner of $v$ can only be
vN. In both cases, the entire symmetry assignments are largely determined, as shown in Table G.8.

This surface state is actually problematic. To see this, notice under \( \mathcal{T} \) the topological sector of \( eX^{-1}v \) is invariant, so \( eX^{-1}v \) has a well defined Kramers parity. However, under the \( Z_2 \) this topological sector becomes \( eX^{-1}vE_{bT} \), which carries an opposite Kramers parity. As discussed above, this is disallowed, otherwise \( \mathcal{T} \) and \( Z_2 \) cannot commute for the spin liquid.

**Table G.8:** Symmetry assignments of the surface topological order if the \( Z_2 \) action on \( e \) and \( m \) is given by (G.110), for both the case with \( v \times v = E_{bT}m \) and the case with \( v \times v = E_{bT}mN \). The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the charges of these excitations under \( U(1) \). The third row lists the time reversal partners of these excitations. And the fourth row lists the \( Z_2 \) partners of these excitations.

<table>
<thead>
<tr>
<th>( U(1) )</th>
<th>( Z_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{1}{2} )</td>
<td>( 0 )</td>
</tr>
<tr>
<td>( \mathcal{T} )</td>
<td>( X^{-1} )</td>
</tr>
<tr>
<td>( Z_2 )</td>
<td>( X^{-1} )</td>
</tr>
</tbody>
</table>

The above discussion implies that if the \( Z_2 \) action on \( e \) and \( m \) is given by (G.110), the surface state of this SPT is problematic. Now we are only left with the case where the \( Z_2 \) action on \( e \) and \( m \) is given by (G.111). In this case, one can use similar method to constrain the rest of the symmetry assignments, and the resulting symmetry assignment is shown in Table G.9. There is also a problem of this topological order: the \( Z_2 \) partner of \( v \) is \( v^{-1} \), so it has a well defined value for charge-conjugation squared. However, its \( \mathcal{T} \) partner has an opposite value of charge-conjugation squared, because the values of charge-conjugation squared for \( N \) and \( m \) are 1 and \(-1\), respectively. This contradicts the fact that \( Z_2 \) and \( \mathcal{T} \) should commute for the spin liquid.

To see this, suppose \( \mathcal{T} \) acts as

\[
v_i \rightarrow T_{ij} \tilde{v}_j \tag{G.120}
\]

and \( Z_2 \) acts as

\[
v_i \rightarrow C_{ij}v_j^*, \quad \tilde{v}_i \rightarrow \tilde{C}_{ij} \tilde{v}_j^* \tag{G.121}
\]
Table G.9: Symmetry assignments of the surface topological order if the $Z_2$ action on $e$ and $m$ is given by \((G.111)\). The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the charges of these excitations under $U(1)$. The third row lists the time reversal partners of these excitations. And the fourth row lists the $Z_2$ partners of these excitations.

<table>
<thead>
<tr>
<th></th>
<th>$X$</th>
<th>$N$</th>
<th>$v$</th>
<th>$E_{gT}$</th>
<th>$e$</th>
<th>$m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U(1)$</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
</tr>
<tr>
<td>$\mathcal{T}$</td>
<td>$X^{-1}$</td>
<td>$N$</td>
<td>$vm/vmN$</td>
<td>$E_{gT}$</td>
<td>$eM_{gT}^1N$</td>
<td>$m$</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>$X^{-1}$</td>
<td>$N$</td>
<td>$v^{-1}$</td>
<td>$E_{gT}$</td>
<td>$eM_{gT}^1N$</td>
<td>$m$</td>
</tr>
</tbody>
</table>

where $T$, $C$ and $\tilde{C}$ are three unitary matrices (in fact, being invertible is enough for the following argument).

For an arbitrary matrix $M$, because $v_i^*M_{ij}v_j$ is a neutral local operator that carries no vorticity, $Z_2$ and time reversal should commute on it. Demanding the results of acting time reversal first and $Z_2$ later and the results of acting $Z_2$ first and time reversal later to be the same gives

$$(T\tilde{C})^\dagger M^*(T\tilde{C}) = (C^*T^*)^\dagger M^*C^*T^* \quad (G.122)$$

For this equation to be true for arbitrary $M$, we must have

$$T\tilde{C} = e^{i\phi}C^*T^* \quad (G.123)$$

with $\phi$ a phase. Or equivalently,

$$\tilde{C} = e^{i\phi}T^{-1}C^*T^* \quad (G.124)$$

Taking the complex conjugation on both sides yields $\tilde{C}^* = e^{-i\phi}(T^{-1})^*CT$. So

$$\tilde{C}\tilde{C}^* = T^{-1}C^*T^*(T^{-1})^*CT = T^{-1}C^*CT \quad (G.125)$$

Notice because $v_i^*M_{ij}v_j$ is local and neutral, $Z_2$ squares to 1 for it because the system is made of $M_{bZ}$. Combined with the discussion in Appendix G.1, the above equation implies that $\tilde{C}\tilde{C}^* = $
That is, \( v \) and \( \tilde{v} \) should have the same charge-conjugation squared.

Taking all these arguments together, the second statement is established: in the scenario of \( eT_i'mZ \) reduced bulk SPT, there is no SPT made of \( M_{bT'} \) that can become \((E_{bT}M_{bT'})_-\).

\[
eT_i'mZ \quad \text{reduced bulk SPT for} \quad (E_{f}M_{bT'})_-
\]

Finally, we turn to the last statement. Using similar arguments as before, the surface topological order can be written as

\[
\{1, X, N, X^{-1}, v, E_f, v^{-1}, e, m\} \times \{1, M_{bT'}\}
\]

with the similar symbols representing parallel excitations as before.

As before, in this topological order, the fusion rules for \( e \) and \( m \) are still given by \((G.100)\).

The symmetry assignments for \( e \) and \( m \) are such that \( e \) carries half charge under \( U(1) \), while \( m \) carries no charge, and the other symmetry assignments for \( e \) and \( m \) are given by \((G.109)\) and \((G.110)\), or \((G.109)\) and \((G.111)\).

In this case, most of the topological data will be the same as the case with \( eT_i'mZ \) reduced bulk SPT for \((E_{bT}M_{bT'})_-\). The only difference is in the fusion product of two \( v \)'s. Modifying the arguments before while keeping in mind that now \( E_f \) is a fermion, we find two possible fusion rules for two \( v \)'s

\[
v \times v = E_{f}m
\]

or

\[
v \times v = E_{f}Nm
\]

In the first possibility, \( \theta_{v,m} = -1 \), while \( \theta_{v,m} = 1 \) in the second possibility. In both cases the right hand side of the fusion rules are fermions, so the topological spin of \( v \) must be \( \theta_v = \exp (i \left( \frac{\pi}{2} + \frac{n\pi}{2} \right)) \), with \( n \) an integer.

In this case, in order for \( \mathcal{T} \) to keep the vorticity of \( v \) and conjugate the topological spin of \( v \),
the $\mathcal{T}$ partner of $v$ can be one of the following:

$$veX^{-1}, \, veX^{-1}N, \, veX^{-1}, \, veX^{-1}N$$  \hspace{1cm} (G.129)

Which one of these can conjugate the topological spin of $v$ depends on the values of $\theta_v = \exp\left(i\left(\frac{\pi}{4} + \frac{n\pi}{2}\right)\right)$, $\theta_{e,v} = \pm1$ and $\theta_{e,v} = \pm1$. However, no matter which one of the above four excitations is the $\mathcal{T}$ partner of $v$, $\theta_{v,m}$ becomes $-\theta_{v,m} \neq \theta_{v,m}^*$. This means there is no consistent symmetry assignment for this topological order.

This establishes the third statement: in the scenario of $eT'_i mZ$ reduced bulk SPT, there is no SPT made of $M_{bT'}$ that can become $(E_f M_{bT'})_-$.

Taking all the arguments above together, we have established that $(E_f M_{bT'})_-$ is anomalous with $Z_2 \times \mathcal{T}$ symmetry. Notice unlike $(E_{bZ} M_{bZ})_-$, which is anomalous even if there is only the $Z_2$ symmetry, here both the $\mathcal{T}$ and $Z_2$ symmetries are responsible for the anomaly.

G.10.3 Anomaly of $E_{bTT'} M_{b-}$

In this subsection we show the anomaly of $E_{bTT'} M_{b-}$, by showing that no SPT made of $M_{b-}$ will become $E_{bTT'} M_{b-}$ after gauging. As a reminder, in this case $Z_2$ does not act as a charge-conjugation, and it anticommutes with $\mathcal{T}$ on $M_{b-}$. As before, we will first condense double charge on the surface to get a surface superfluid, whose minimal trivial vortex is the $4\pi$ vortex. We will then proliferate these $4\pi$ vortices to restore the full symmetry and get a surface topological order. Again, there are two scenarios for the surface superfluid: it can either be a simple superfluid, or this superfluid has to coexist with another anomalous topological order. We will discuss these cases in turn.

**Simple superfluid**

We begin with the case of a simple superfluid. Using similar argument as before, we see the symmetric surface topological order obtained by condensing $4\pi$ vortices can be written as

$$\{1, X, N, X^{-1}, v, E_{bTT'}, v^{-1}\} \times \{1, M_{b-}\}$$  \hspace{1cm} (G.130)
The symbols stand for parallel meanings as before, while now the $2\pi$ vortex is $E_{bTT'}$, a Kramers doublet under both $T$ and $T'$.

Similar arguments as before show that the braiding and fusion are similar to the surface state obtained from the simple superfluid for the corresponding SPT of $(E_{bT}M_{bT'})_-$, and the symmetry assignment is shown in Table G.10. In particular, in order for $E_{bTT'} = v \times v$ to be a Kramers doublet under both $T$ and $T'$, both of the $T$ and $T'$ partners of $v$ should be $vN$. Notice $Xv$ will become $XvM^\dagger_{b-}$ under $T$, and it is invariant under $Z_2$. Below we show this is disallowed.

Table G.10: Symmetry assignments of the surface topological order from the simple superfluid surface of the corresponding SPT of $E_{bTT'}M_{b-}$. The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the charges of these excitations under $U(1)$. The third row lists the $T$ partners of these excitations. The fourth row lists the $Z_2$ partners of these excitations. The fifth row lists the $T'$ partners of these excitations.

<table>
<thead>
<tr>
<th>$U(1)$</th>
<th>$T$</th>
<th>$T'$</th>
<th>$Z_2$</th>
<th>$T_{b-}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>$X^{-1}$</td>
<td>$X^{-1}$</td>
<td>$N$</td>
<td>$vN$</td>
</tr>
<tr>
<td>$y$</td>
<td>$N$</td>
<td>$N$</td>
<td>$vN$</td>
<td>$E_{bTT'}$</td>
</tr>
<tr>
<td>$z$</td>
<td>$E_{bTT'}$</td>
<td>$E_{bTT'}$</td>
<td>$E_{bTT'}$</td>
<td>$E_{bTT'}$</td>
</tr>
</tbody>
</table>

For notational simplicity, denote $Xv$ and $XvM^\dagger_{b-}$ by $x$ and $y$, respectively. Suppose the $T$ action is

$$T : x_i \rightarrow T_{ij}y_j, \quad y_i \rightarrow \tilde{T}_{ij}x_j$$ (G.131)

and the $Z_2$ action is

$$Z_2 : x_i \rightarrow Z_{ij}x_j, \quad y_i \rightarrow \tilde{Z}_{ij}y_j$$ (G.132)

with $T$, $\tilde{T}$, $Z$ and $\tilde{Z}$ four invertible matrices.

Because, for any matrix $M$, $x_i^* M_{ij} x_j$ is local and neutral, $Z_2$ and $T$ should commute on this operator. This gives a condition

$$T \tilde{Z} = e^{i\phi} Z^* T$$ (G.133)
Because, for any matrix $M$, $x_i^* M_{ij} y_j$ is local and carries charge-1, $Z_2$ and $T$ should anti-commute on this operator, because the system is made of $M_{b-}$. This, together with the above condition, yields

$$\tilde{T} Z = e^{i\phi_2} \tilde{Z}^* \tilde{T} \quad (G.134)$$

with $e^{i(\phi_1 - \phi_2)} = -1$.

Combining these two equations, we get

$$Z = e^{i\phi_1} T^* \tilde{Z}^* T^{-1} = e^{i\phi_2} \tilde{T}^{-1} \tilde{Z}^* \tilde{T} \quad (G.135)$$

Using that $e^{i(\phi_1 - \phi_2)} = -1$, the above equation yields

$$\tilde{Z}^* = -\tilde{T} T^* \tilde{Z}^* T^{-1} \tilde{T}^{-1} \quad (G.136)$$

Now notice $y_i^* M_{ij} y_j$ should have Kramers parity 1 for any matrix $M$, this implies that $\tilde{T}^* T = e^{i\phi_3}$, or $\tilde{T} T^* = e^{-i\phi_3}$, where $\phi_3$ is another phase. Plugging this into the above equation yields

$$\tilde{Z}^* = -\tilde{Z}^* \quad (G.137)$$

so $\tilde{Z} = 0$. This is disallowed.

The above argument shows that in the scenario of a simple superfluid, there is no SPT made of $M_{b-}$ that can become $E_{bT'T} M_{b-}$ after gauging.

**Superfluid coexisting with another anomalous topological order**

Next we turn to the more complicated case where the superfluid has to coexist with another anomalous topological order. Again, this happens when the bulk remains as a nontrivial SPT when the bulk symmetry is also reduced. For this purpose, let us first clarify what the reduced symmetry is. The reduced symmetry group has a $Z_2^T$ subgroup, with an anti-unitary generator $t$ that satisfies $t^2 = 1$. It also has a $Z_2$ subgroup, with a unitary generator $g$ that satisfies $g^2 = 1$. 
However, $gt + tg = 0$. Denote $t' = gt$, then $t'^2 = gtgt = -1$, so $t'$ generates a $Z_4^{T'}$ symmetry. Notice $t'$ gets inverted when conjugated with both $t$ and $g$, that is, $gt'g = gtgt = -gt = -t'$ and $tt't = tgtt = -gt = -t'$. So we will denote this group by $D_4^T \equiv Z_4^{T'} \times Z_2 = Z_4^{T'} \times Z_2$.

Now it is also easy to get the full symmetry, which also has $U(1)$ charge conservation. In the gauge where $t^2$ and $g^2$ are fixed to be identity, because each charge-1 boson has $Z_2$ and $T$ anticommuting, the $\pi$ rotation of the $U(1)$ is locked with $t'^2$, and the full symmetry group can be written as $(U(1) \times D_4^T)/Z_2$.

### 3D bosonic SPTs with $D_4^T$ symmetry

Now let us discuss 3D bosonic SPTs with $D_4^T$ symmetry. By group cohomology, these SPTs are classified by $Z_2^3$. There should be another root state, $efmf$, which is beyond group cohomology. So the full classification of these SPTs is expected to be $Z_2^4$. Among the other root states, $eTmT$ should be one of them. The $U(1)$ symmetry can be added to $efmf$ and $eTmT$ with a trivial action, so reduced bulk SPTs with such anomalies can be easily cancelled, and we will ignore these two states from now on. We propose two other root states: $(eTT'_i mT'_i)_{ZT'}TT'$ and $(eT'_i mT'_i)_{ZT'}$.

The root state $(eTT'_i mT'_i)_{ZT'}TT'$ is actually the descendant of the corresponding SPT of $E_{fTT'} M_{b-}$ (viewed from the perspective of $M_{b-}$), when the $U(1)$ symmetry is broken to its $Z_2$ subgroup generated by its $\pi$ rotation. The surface state of this corresponding SPT of $E_{fTT'} M_{b-}$ can be a $Z_2$ topological order, with symmetries assigned as in Table G.11. These symmetry assignments can be derived from the corresponding SPT of $E_{fTT'} M_{b-}$ viewed from the perspective of $E_{fTT'}$, which is described in Appendix G.10.1. Notice the fusion rules are $e \times e = m \times m = M_{b-}$.

When the $U(1)$ symmetry is reduced to its $Z_2$ subgroup generated by its $\pi$ rotation, the symmetry becomes $D_4^T$. The resulting state is still anomalous, which can be seen by checking its $Z_4^{T'}$ domain wall. Consider breaking the $Z_4^{T'}$ symmetry in two different ways on the two sides of a 2D domain wall, while keeping a $Z_2$ subgroup generated by $t'^2$ intact across the entire system. Then this domain wall has a $Z_2 \times Z_2$ symmetry, and by relating it to the corresponding SPT of $E_{fTT'} M_{b-}$, we see this time reversal domain wall is a Levin-Gu state (protected by the $Z_2$ gen-
Table G.11: Symmetry assignments of the $\mathbb{Z}_2$ surface topological order of the bosonic SPT made of $M_b^-$ that will become $E_{fTT'}M_b^-$ after gauging. The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the charges of these excitations under $U(1)$. The third row lists the time reversal partners of these excitations. The fourth row lists the values of $T^2$ of these excitations, with empty entries representing that $T^2$ is not well-defined. The fifth row lists the $\mathbb{Z}_2$ partners of these excitations. The sixth row lists the values of $Z_2^2$ of these excitations. The seventh row lists the $T'$ partners of these excitations. And the last row lists the values of $T^{t^2}$ of these excitations, with $\pm i$ standing for that $T^{t^2} = -1$.

<table>
<thead>
<tr>
<th></th>
<th>$e$</th>
<th>$m$</th>
<th>$\epsilon = em^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U(1)$</td>
<td>$\frac{1}{2}$</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
</tr>
<tr>
<td>$T$</td>
<td>$e^\dagger$</td>
<td>$m^\dagger$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>$T^2$</td>
<td>$-1$</td>
<td>$1$</td>
<td>$-1$</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>$m$</td>
<td>$e$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>$Z_2^2$</td>
<td></td>
<td></td>
<td>$1$</td>
</tr>
<tr>
<td>$T'$</td>
<td>$m^\dagger$</td>
<td>$e^\dagger$</td>
<td>$\epsilon$</td>
</tr>
<tr>
<td>$T^t$</td>
<td>$\pm i$</td>
<td>$\pm i$</td>
<td>$-1$</td>
</tr>
</tbody>
</table>

The existence of this decorated domain wall shows the descendant state is still a nontrivial SPT with the $D_{4T}$ symmetry.

Next we turn to $(eT'_i mT'_i)_{ZT'}$, whose symmetry assignments are shown in Table G.12. Or more precisely, the symmetry assignments are

$$
T: e \rightarrow ie, \ m \rightarrow m
$$

$$
Z_2: e \rightarrow m, \ m \rightarrow e
$$

(G.138)

Notice the fusion rules are $e \times e = m \times m = M_{b^-}$.

To show that this state is anomalous, consider breaking the $D_{4T}$ symmetry to $Z_2 \times Z_2^T$, where the first $Z_2$ is generated by $t^2$, and the $Z_2^T$ is generated by $t$. Notice in this case the $Z_2$ generated by $g$ is broken. This can be done, for example, by considering that $T$ acts on the local boson as

$$
M_1 \rightarrow M_2, \ M_2 \rightarrow M_1
$$

(G.139)
Table G.12: Symmetry assignments of the $Z_2$ surface topological order $(eT'_i mT'_i Z)_{TT'}eTZ$. The first row lists all nontrivial excitations, from which the symmetry assignments on all their bound states can be inferred. The second row lists the $T$ partners of these excitations. The third row lists the values of $T^2$ of these excitations. The fourth row lists the $Z_2$ partners of these excitations. The fifth row lists the values of $Z_2^2$ of these excitations. The sixth row lists the $T'$ partners of these excitations. And the last row lists the values of $T'^2$ of these excitations, with $±i$ standing for that $T'^4 = -1$.

<table>
<thead>
<tr>
<th></th>
<th>$e$</th>
<th>$m$</th>
<th>$ε = em^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>$e$</td>
<td>$m$</td>
<td>$ε$</td>
</tr>
<tr>
<td>$T^2$</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$Z_2$</td>
<td>$m$</td>
<td>$e$</td>
<td>$ε$</td>
</tr>
<tr>
<td>$Z_2^2$</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>$T'$</td>
<td>$m$</td>
<td>$e$</td>
<td>$ε$</td>
</tr>
<tr>
<td>$T'^2$</td>
<td>$±i$</td>
<td>$±i$</td>
<td>1</td>
</tr>
</tbody>
</table>

and $Z_2$ acts on the local boson as

$$M_1 \rightarrow M_1, \quad M_2 \rightarrow -M_2 \quad \text{(G.140)}$$

Giving $M_1 M_2$ a nonzero expectation value will break the $D^T_4$ to $Z_2 \times Z^T_2$ in the above way.

With this reduced symmetry, this state becomes $eZmZ$, where the symbol $Z$ stands for half charge under the $Z_2$ (generated by $t'^2$). This reduced state is anomalous. To see this, define $\tilde{t} = t'^2 \cdot t$, which generates another anti-unitary symmetry, $\tilde{T}$. The state $eZmZ$ can then be relabelled as $e\tilde{T}m\tilde{T}$, so it is anomalous. Enhancing the symmetry back to $D^T_4$ will not add Kramers doublet boson under $\tilde{T}$ to the system, so it will not remove this anomaly. This also shows that $(eT'_i mT'_i)_{ZT'}$ is distinct from the previous $(eTT'_i mT'_i)_{ZT'}eTT'$, because when the symmetry is reduced to $Z_2 \times T$ in the above way, the latter becomes $eZTmZ$, which is a distinct anomalous state.

For later purpose, let us now consider the $U(1)$ charge of $e$ and $m$ when the $4\pi$ vortices are condensed and the full symmetry is restored. Again, because the $\pi$ rotation of the $U(1)$ is locked with $t'^2$, $e$ and $m$ should carry half charge in the topological order.

The above discussion implies that, in the scenario where the surface superfluid has to coexist with another anomalous topological order, in order to show there is no SPT made of $M_6$ that
can become $E_{TT'}M_{b-}$ after gauging, it is sufficient to show:

1. In the scenario of a simple superfluid surface, there is no SPT made of $M_{b-}$ that can become $E_fM_{b-}$.

2. In the scenario of $(eT'_i m T'_i)_{ZT'}$ reduced bulk SPT, there is no SPT made of $M_{b-}$ that can become $E_{TT'}M_{b-}$.

3. In the scenario of $(eT'_i m T'_i)_{ZT'}$ reduced bulk SPT, there is no SPT made of $M_{b-}$ that can become $E_fM_{b-}$.

Below we will show these three statements in turn.

\textit{Simple superfluid surface for $E_fM_{b-}$}

We start with the first statement. Using similar arguments as before, the surface topological order in this case can be written as

$$\{1, X, N, X^{-1}, v, E_f, v^{-1}\} \times \{1, M_{b-}\}$$  \hspace{1cm} (G.141)

The symbols stand for parallel meanings as in the previous cases, and now the $2\pi$ vortex, $E_f$, is a fermion and is a Kramers singlet under both $T$ and $T'$.

In this case, most of the topological data will be the same as in the case with simple superfluid surface for $E_{TT'}M_{b-}$. The only difference is in the fusion rule of $v$. The similar arguments as before implies that the fusion rule of two $v$’s in this case is

$$v \times v = E_fN$$  \hspace{1cm} (G.142)

otherwise the antiparticle of $v$ will have an opposite topological spin as itself. Because the right hand side is a fermion, the topological spin of $v$ will be $\theta_v = \exp \left(i \left(\frac{\pi}{4} + \frac{n\pi}{2}\right)\right)$, with $n$ an integer.

Now let us consider the $T$ partner of $v$. As $T$ should keep the vorticity and flip the charge, there are two possible $T$ partners of $v$: $v$ and $vN$. Because $\theta_v = \exp \left(i \left(\frac{\pi}{4} + \frac{n\pi}{2}\right)\right)$ and $\theta_{v,N} = -1$, neither of these options will conjugate $\theta_v$ under $T$. So this is inconsistent.
This establishes the first statement: in the scenario of a simple superfluid surface, there is no SPT made of $M_{b-}$ that can become $E_f M_{b-}$.

$$(eT'_i m T'_i)_{ZT'} \text{ reduced bulk SPT for } E_{bTT'} M_{b-}$$

Next we turn to the second statement. In this case, similar arguments as before show the surface topological order can be written as

$$\{1, X, N, X^{-1}, v, E_{bTT'}, v^{-1}, e, m\} \times \{1, M_{b-}\} \quad \text{(G.143)}$$

with the symbols standing for parallel meanings as before.

Recall that both $e$ and $m$ will carry half charge under $U(1)$. By charge conservation and vorticity conservation, the possible fusion products of two $v$’s are

$$E_{bTT'}, E_{bTT'} N, E_{bTT'} e, E_{bTT'} e N$$

$$E_{bTT'} e X^{-1}, E_{bTT'} e X^{-1} N, \quad \text{(G.144)}$$

$$E_{bTT'} m X^{-1}, E_{bTT'} m X^{-1} N$$

Similar arguments as before imply that there are two possible fusion rules for $v$:

$$v \times v = E_{bTT'} e \quad \text{(G.145)}$$

or

$$v \times v = E_{bTT'} e N \quad \text{(G.146)}$$

In both cases, the bound state of two $v$’s is fermionic, so the topological spin of $v$ should be $\theta_v = \exp \left( i \left( \frac{\pi}{4} + \frac{n\pi}{2} \right) \right)$, with $n$ an integer. In order for $T$ to conjugate $\theta_v$, the $T$ partner of $v$ should be one of

$$ve X^{-1}, ve X^{-1} N, vm X^{-1}, vm X^{-1} N \quad \text{(G.147)}$$
From the symmetry actions on \(e\) and \(m\) in the superfluid phase, by multiplying something condensed in the superfluid to be consistent with charge conservation, we get the symmetry actions on \(e\) and \(m\) in the topologically ordered phase:

\[
\mathcal{T} : \ e \rightarrow eM_{b_{-}}^{\dagger}N, \ m \rightarrow mM_{b_{-}}^{\dagger}N \\
Z_{2} : \ e \rightarrow m, \ m \rightarrow e
\]  \hspace{1cm} (G.148)

Combining this with that \(\theta_{v,N} = -1\), we see no matter which one of the four excitations is the \(\mathcal{T}\) partner of \(v\), \(\theta_{e,v}\) cannot be conjugated by \(\mathcal{T}\), which is disallowed.

This establishes the second statement: in the scenario of \((eT'_{i}mT'_{i})_{ZT'}\) reduced bulk SPT, there is no SPT made of \(M_{b_{-}}\) that can become \(E_{bTT'}M_{b_{-}}\).

Finally we come to the last statement. Similar arguments as before imply the topological order can be written as

\[
\{1, X, N, X^{-1}, v, E_{f}, v^{-1}, e, m\} \times \{1, M_{b_{-}}\}
\]  \hspace{1cm} (G.149)

with symbols standing for parallel meanings as before. Notice the \(2\pi\) vortex, \(E_{f}\), is a fermion, and it is a Kramers singlet under both \(\mathcal{T}\) and \(\mathcal{T}'\).

Most of the topological data can be easily determined using similar arguments as before:

\[
\theta_{X,v} = -i, \theta_{X,E_{f}TT'} = -1, \theta_{v,e} = \pm 1, \theta_{v,m} = \pm 1, \\
X \times X = NM_{b_{-}}, N \times N = 1, E_{bTT'} \times E_{bTT'} = 1
\]  \hspace{1cm} (G.150)

There are two possible fusion rules for \(v\):

\[
v \times v = E_{f}\epsilon
\]  \hspace{1cm} (G.151)

or

\[
v \times v = E_{f}\epsilon N
\]  \hspace{1cm} (G.152)
In both cases, the bound state of two \( v \)'s is a boson, so \( v \) is a boson, fermion, semion or anti-semion.

Knowing that \( \mathcal{T} \) should flip the charge and keep the vorticity, the \( \mathcal{T} \) partner of \( v \) can be one of

\[
v, vN, ve, veN,
\]

\[
veX^{-1}, veX^{-1}N, vmX^{-1}, vmX^{-1}N
\]

Because \( v \) is a boson, fermion, semion, or anti-semion, and \( \theta_{X,v} = -i \), \( \theta_{v,e} = \pm 1 \) and \( \theta_{v,m} = \pm 1 \), the last four options can be ruled out, because in those cases \( \mathcal{T} \) will not conjugate \( \theta_{v} \). Because the \( \mathcal{T} \) partner of \( e \) is \( eM_{b-}^{\dagger} \), in order to conjugate \( \theta_{e,v} \), the \( \mathcal{T} \) partner of \( v \) cannot be \( v \) or \( vN \). That is, the \( \mathcal{T} \) partner of \( v \) is either \( ve \) or \( veN \).

If \( v \times v = Ef_{\epsilon} \) and the \( \mathcal{T} \) partner of \( v \) is \( ve \), then \( v \times ve = Ef \). Because \( v \) and \( ve \) are \( \mathcal{T} \) partners and \( Ef \) is a fermion, \( Ef \) must be a Kramers doublet under \( \mathcal{T} \), which contradicts the original assumption. The same reasoning rules out the possibility that \( v \times v = Ef_{\epsilon}N \) and the \( \mathcal{T} \) partner of \( v \) is \( veN \).

If \( v \times v = Ef_{\epsilon} \) and the \( \mathcal{T} \) partner of \( v \) is \( veN \), then \( v \times veN = EfN \). This means \( N \) is a Kramers doublet under \( \mathcal{T} \). On the other hand, the \( \mathcal{T} \) partner of \( e \) is \( eM_{b-}^{\dagger} \), so \( N \) can be viewed as the bound state of \( e \) and its \( \mathcal{T} \) partner. This implies that \( N \) is a Kramers singlet under \( \mathcal{T} \), which leads to a contradiction. The same reasoning also rules out the possibility that \( v \times v = Ef_{\epsilon}N \) and the \( \mathcal{T} \) partner of \( v \) is \( ve \).

Putting all these analysis together, we have established the last statement: in the scenario of \( (eT_{m}^{\dagger}mT_{i}^{\dagger})_{ZT} \) reduced bulk SPT, there is no SPT made of \( M_{b-} \) that can become \( EfM_{b-} \).

Therefore, we have shown that no SPT made of \( M_{b-} \) can become \( E_{bTT}M_{b-} \) after gauging, which means \( E_{bTT}M_{b-} \) is anomalous with \( Z_{2} \times \mathcal{T} \) symmetry.

As discussed before, we have already shown all states in Table G.3 and Table 7.10 are anomalous.
G.11 $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids with $\theta = \pi$ and $Z_2$ not acting as a charge conjugation

In this appendix, we discuss $Z_2 \times T$ symmetric $U(1)$ quantum spin liquids with $\theta = \pi$ and $Z_2$ not acting as a charge conjugation.

As discussed in Sec. 7.7.1, in this case the quantum numbers of the $(\frac{1}{2}, 1)$ dyon determine the phase. Since there is no nontrivial projective representation of the $Z_2 \times T$ symmetry on the $(\frac{1}{2}, 1)$ dyon, we expect only one state: $(E_f T T' M_f)_\theta$. In this state, the electric charge must be a Kramers doublet under both $T$ and $T'$, and $M$ (the $(0, 2)$ dyon in this context) has $Z_2$ and $T$ commuting with each other. One may wonder whether it is possible to have $Z_2$ and $T$ anti-commuting with each other in this case. Below we show this is not possible.

Denote the $(\frac{1}{2}, 1)$ dyon by $D^{(+)}$, and its time reversal partner, the $(\frac{1}{2}, -1)$ dyon, by $D^{(-)}$. Notice $M$ is a bound state of $D^{(+)}$ and $D^{(-)\dagger}$. Suppose the $Z_2$ action on $D^{(+)}$ and $D^{(-)}$ is

$$Z_2 : D_i^{(+)} \to Z_{ij} D_j^{(+)} , \quad D_i^{(-)} \to \bar{Z}_{ij} D_j^{(-)}$$

and the $T$ action on $D^{(+)}$ and $D^{(-)}$ is

$$T : D_i^{(+)} \to T_{ij} D_j^{(-)} , \quad D_i^{(-)} \to T_{ij} D_j^{(+)}$$

For any matrix $M$, the operator $D_i^{(+)}\dagger M j D_j^{(+)}$ is local, so the actions of $Z_2$ and $T$ should commute on it. This gives the condition

$$(Z^* T)^\dagger M^* (Z^* T) = (T \bar{Z})^\dagger M^* (T \bar{Z})$$

In order for this equation to be satisfied by any matrix $M$, we need to have

$$Z^* T = e^{i\phi_1} T \bar{Z}$$
or, equivalently,

\[ Z = e^{-i\phi_1}T^* \tilde{Z}^*(T^*)^{-1} \]  \hspace{1cm} (G.158)

Consider the local operator \( D_i^{(-)\dagger} M_{ij} D_j^{(-)} \) in a similar way, we get the condition

\[ \tilde{Z}^* \tilde{T} = e^{i\phi_2} \tilde{T} Z \]  \hspace{1cm} (G.159)

or, equivalently,

\[ Z = e^{-i\phi_2} \tilde{T}^{-1} \tilde{Z}^* \tilde{T} \]  \hspace{1cm} (G.160)

On the other hand, acting time reversal twice on the local operator \( D_i^{(+)*} M_{ij} D_j^{(+)} \) should result in a trivial action, which implies that

\[ T^* \tilde{T} = e^{i\phi_3} \]  \hspace{1cm} (G.161)

or, equivalently,

\[ \tilde{T} = e^{i\phi_3} (T^*)^{-1} \]  \hspace{1cm} (G.162)

Combining this equation and (G.160) yields

\[ Z = e^{-i\phi_2} T^* \tilde{Z}^* (T^*)^{-1} \]  \hspace{1cm} (G.163)

Comparing this equation and (G.158) yields

\[ e^{i\phi_1} = e^{i\phi_2} \]  \hspace{1cm} (G.164)

Now consider the \((0,2)\) dyon, which is represented as \( D_i^{(+)\dagger} M_{ij} D_j^{(-)\dagger} \). Acting on this operator by \( Z_2 \) and \( T \) with different orders using (G.154) and (G.155), and using the constraints (G.158), (G.160) and (G.164), we see that the \( Z_2 \) and \( T \) commute on \( D_i^{(+)\dagger} M_{ij} D_j^{(-)\dagger} \), which proves the
The above argument can also be applied to $O(2) \times T$ symmetric $U(1)$ quantum spin liquids discussed in Sec. G.12.1. If such a spin liquid has $\theta = \pi$ and the improper $Z_2$ component not acting as a charge conjugation, then on $M$ the actions of $Z_2$ and $T$ should commute.

G.12 $U(1)$ QUANTUM SPIN LIQUIDS WITH SOME OTHER SYMMETRIES

In this appendix we briefly discuss $U(1)$ quantum spin liquids with some other symmetries. The classifications of these symmetry enriched $U(1)$ quantum spin liquids are quite complicated, which we leave for future work. In this appendix we only lay out the principle of enumerating the putative states and make some comments, but we will not attempt to finish the procedure of examining the anomalies.

G.12.1 $O(2) \times T$ SYMMETRY

First we consider the case where the $SO(3) \times T$ symmetry is broken down to $O(2) \times T \cong (U(1) \rtimes Z_2) \times T$. Physically, here $U(1)$ can represent spin rotations around one axis, while the $Z_2$ transformation is a $\pi$ spin rotation around another axis perpendicular to this one. This case is rather complicated, and we do not attempt to complete the anomaly detection and determine the final classification. Instead, we will just give a way to systematically list all putative (possibly anomalous) states and make some comments.

The structure of projective representations of $O(2) \times T$ is rich. On the electric charge, it is classified by $Z_3^2$, and the three root projective representations physically correspond to having half charge under the $U(1)$ subgroup, being a Kramers doublet under $T$ and being a Kramers doublet under $T'$ (the anti-unitary symmetry whose generator is the product of the generators of $Z_2$ and $T$). If $\theta = 0$, then on the magnetic monopole the projective representations are classified by $Z_2^2$, which physically correspond to having half charge under the $U(1)$ subgroup and having the discrete $Z_2$ anti-commuting with $T$. So at $\theta = 0$, if the discrete $Z_2$ symmetry does not act as a charge conjugation, there are $3 \times 2^4 \times 2^2 = 96$ putative states. We will not write down the long list of all these states, since it is straightforward and not particularly illuminating. Notice some
of these are descendant states of an $SO(3) \times T$ symmetric state. Because the anomaly argument for the $SO(3) \times T$ symmetric states should also apply to these $O(2) \times T$ symmetric states, their anomalies can be determined immediately.

At $\theta = \pi$, there are only two putative states, $(E_{fT}T(M_f)_{\theta})$ and $(E_{fT}T(M_f)_{\theta \frac{1}{2}})$, which are the descendents of $(E_{fT}M_f)_{\theta}$ and $(E_{fT}M_f)_{\theta \frac{1}{2}}$ with $SO(3) \times T$ symmetry, respectively. In the former state the $(\frac{1}{2},1)$ dyon carries integer $U(1)$ charge, while in the latter it carries half charge. As discussed in Sec. 7.3, the former state is anomaly-free while the latter is anomalous even with $O(2)$ symmetry. Here the actions of $Z_2$ and $T$ commute on $M$ (the $(0,2)$ dyon in this context), and it is shown in Appendix G.11 that, in this case, it is impossible to have the actions of $Z_2$ and $T$ anti-commuting on $M$.

In the case where the discrete $Z_2$ symmetry acts as a charge conjugation, the possible fractional quantum numbers on the electric charge have a structure of $Z_2^3$: having half charge under the $U(1)$ subgroup, being a Kramers doublet under $T$, and having charge conjugation squaring to $-1$. If $\theta = 0$, the possible fractional quantum numbers on the magnetic monopole also have a structure of $Z_2^2$: being a Kramers doublet under $T'$, and having charge conjugation squaring to $-1$. So there are $3 \times 2^3 \times 2^2 = 96$ such putative states. At $\theta = \pi$, there are two putative states: $(E_{fT}M_{fT'})_{\theta -}$ and $(E_{fT}M_{fT'})_{\theta -Z}$. In the former, the $(\frac{1}{2},1)$ dyon carries a linear representation of the symmetry, while this dyon has charge conjugation squaring to $-1$ for the latter. This former state is anomaly-free, and it can be obtained by equipping its $Z_2 \times T$ symmetric cousin with a further $U(1)$ symmetry. The latter is anomalous, because even if the symmetry is broken to $Z_2 \times T$ it is still anomalous.

We finish this subsection by briefly commenting on a few models that realize $O(2) \times T$ symmetric $U(1)$ quantum spin liquids. Ref. Motrunich2002, HermeleFisherBalents2004, Motrunich2005 studied a couple of different lattice models that realize three dimensional $U(1)$ quantum spin liquid phases with $O(2) \times T$ symmetry, and the particular phase realized in these works is $(E_{bM_{b\frac{1}{2}}})_-$. Ref. Levin2006 constructed two models of $O(2) \times T$ symmetric $U(1)$ quantum spin liquids, where one of them has a bosonic monopole and the other has a fermionic monopole. These two states are $(E_{bM_{b}})_-$ and $(E_{bM_{f}})_-$, respectively.
G.12.2 $Z_2 \times Z_2 \times \mathcal{T}$ symmetry

Now consider the case where the $SO(3) \times \mathcal{T}$ symmetry is broken down to $Z_2 \times Z_2 \times \mathcal{T}$, where these two $Z_2$’s can represent $\pi$ spin rotations around two perpendicular axes. It is known that the projective representations (on the electric charge) of $Z_2 \times Z_2 \times \mathcal{T}$ symmetry are classified by $\mathbb{Z}_4^2$, where two of the four $Z_2$’s are descendants of the projective representations of $SO(3) \times \mathcal{T}$, and the other two $Z_2$’s come from the nontrivial interplay between $Z_2 \times Z_2$ and time reversal. More precisely, each of the two $Z_2$’s together with time reversal can form a new anti-unitary symmetry, and the other two root nontrivial projective representations can be viewed as having Kramers doublets under such new anti-unitary symmetries. If $\theta = 0$, on the magnetic monopole the projective representations are classified by $\mathbb{Z}_2^2$, and the nontrivial root projective representations physically correspond to the two $Z_2$’s and $\mathcal{T}$ anti-commuting.

If $\theta = \pi$, the phase is determined by the $(\frac{1}{2}, \pm 1)$ dyons, which has one nontrivial projective representation that corresponds to that the two $Z_2$ symmetries anti-commute. The state with the two $Z_2$ symmetries commuting is a descendant of the $SO(3) \times \mathcal{T}$ symmetric $(E_{fT}M_f)_{\theta}$, so it is still anomaly-free. The state with the two $Z_2$ symmetries anti-commuting is a descendant of the $SO(3) \times \mathcal{T}$ symmetric $(E_{fT}M_f)_{\theta \frac{1}{2}}$, which we conjecture is still anomalous with $Z_2 \times Z_2 \times \mathcal{T}$ symmetry.

The descendants of the 15 non-anomalous spin liquid states with $SO(3) \times \mathcal{T}$ still remain non-anomalous and distinct, and we conjecture all the anomalous states remain anomalous even if the symmetry is broken to $Z_2 \times Z_2 \times \mathcal{T}$. This is of course just a partial classification, because, on the one hand, states labelled by the other projective representations of $Z_2 \times Z_2 \times \mathcal{T}$ should be taken into account, and on the other hand, states where $Z_2 \times Z_2$ can act as charge conjugation should also be considered. This is not attempted in this paper.

G.12.3 $Z_2 \times Z_2$ symmetry

Parallel considerations as in Sec. 7.8.2 can be applied to the case with $Z_2 \times Z_2$ symmetry. Here the projective representations are only classified by $Z_2$, and the nontrivial projective representation is the descendant of the projective representation of $SO(3)$. Therefore, the descendants of
$E_b M_b$ and $E_b^{1/2} M_b$ will remain distinct and non-anomalous when the symmetry is broken from $SO(3)$ to $Z_2 \times Z_2$. We conjecture that the descendant of $E_b^{1/2} M_b^{1/2}$ remains anomalous. Again, to have a complete classification, states where $Z_2 \times Z_2$ permutes fractional excitations should be taken into account. We do not attempt it here.
Appendices of chapter 8

H.1 Time reversal symmetric generalizations of the mirror symmetric TCI

In this appendix we introduce two time reversal symmetric generalizations of the MTCI discussed in Sec. 8.3.2. Because MTCI has nontrivial monopoles and is stable under interactions even without a time reversal symmetry, this will remain the case in the presence of a further time reversal. Our main purpose is then to check if the time reversal symmetry gives further nontrivial quantum numbers to the monopoles.

The low energy surface Hamiltonian is still given by (8.2), and the \(U(1)\) symmetry and \(\mathcal{M}\) symmetry are assigned as (8.10). We will further equip the TCIs with a time reversal symmetry \(\mathcal{T}\). Depending on whether the fermion is a Kramers singlet or a Kramers doublet under \(\mathcal{T}\), we
will have two different generalizations.

H.1.1 Kramers singlet fermions

First consider the $\mathcal{T}$ action

$$
\mathcal{T} : \psi \rightarrow \sigma_y \tau_y \psi \quad \text{(H.1)}
$$

where the fermions are Kramers singlets under $\mathcal{T}$. Notice, combining this and (8.10), we see $Z^2 \equiv (\mathcal{T}\mathcal{M})^2 = -1$ for the fermions. This generalization of MTCI will be labeled as TMTCI-1.

Now we check the quantum numbers of the monopoles. Under $\mathcal{T}$

$$
M_1 \sim f_1^\dagger|0\rangle \rightarrow \bar{i}f_1^\dagger|\bar{0}\rangle \sim -iM_1^\dagger
\quad \text{(H.2)}
$$

$$
M_2 \sim f_2^\dagger|0\rangle \rightarrow \bar{i}f_1|\bar{0}\rangle \sim -iM_2^\dagger
\quad \text{(H.2)}
$$

Combining this with (8.11) yields

$$
M_1 \rightarrow -iM_2, \quad M_2 \rightarrow iM_1 \quad \text{(H.3)}
$$

under $\mathcal{T}\mathcal{M}$, which means $(\mathcal{T}\mathcal{M})^2 = -1$ for the monopole.

After gauging, we identify the electric charge of the resulting $U(1)$ QSL as the fermions of the TCI, and the magnetic monopole as the monopole of the TCI. Then the charge is a Kramers singlet under $\mathcal{T}$, and it has $Z^2 = -1$. The monopole has $\mathcal{M}^2 = -1$ and $Z^2 = -1$. So this $U(1)$ QSL is denoted as $(E_f Z M b M Z)_-$. 

H.1.2 Kramers doublet fermions

Next consider the $\mathcal{T}$ action

$$
\mathcal{T} : \psi \rightarrow i\sigma_y \psi \quad \text{(H.4)}
$$
Notice, combining this and (8.10), we see $Z^2 \equiv (\mathcal{T}\mathcal{M})^2 = 1$ for the fermions. This generalization of MTCI will be labeled as TMTCI-2.

Under $\mathcal{T}$ the monopoles transform as in (8.8), and under $\mathcal{T}\mathcal{M}$

$$M_1 \rightarrow -M_1, \quad M_2 \rightarrow -M_2$$

so $(\mathcal{T}\mathcal{M})^2 = 1$ on the monopoles.

After gauging, we identify the electric charge of the resulting $U(1)$ QSL as the fermions of the TCI, and the magnetic monopole as the monopole of the TCI. Then the charge is a Kramers doublet under $\mathcal{T}$, and it has $Z^2 = 1$. The monopole has $\mathcal{M}^2 = -1$ and $Z^2 = 1$. So this $U(1)$ QSL is denoted as $(E_{fT}\mathcal{M}_{hM})$.

In this appendix we have derived the quantum numbers of the monopoles from a symmetric gapless surface state of these TCIs, and it is instructive to reproduce these results by studying a symmetric STO of them. This will be done in Appendix H.2.

### H.2 Surface topological orders of various TCIs

In this appendix we construct symmetric STOs of the DTI insulator described in Sec. 8.3.3, and those of TMTCI-1 and TMTCI-2 discussed in Appendix H.1. Our method is based on that in Ref. [74, 356, 378, 386].

#### H.2.1 STO of the doubled topological insulator

Let us begin with the STO of the DTI discussed in Sec. 8.3.3. Starting from the surface theory (8.2) together with the symmetry (8.12), we can introduce an extra term to the Hamiltonian

$$\delta H = m(x)\psi^\dagger \sigma_y \tau_y \psi$$

where $m(x) = m_0\text{sgn}(x)$ represents a mass domain wall. This term respects all symmetries, and it gaps out the surface except at $x = 0$, which now hosts a pair of helical Dirac fermions. In fact, these helical Dirac fermions are identical to those in the edge state of a 2D topological insulator.
Therefore, a DTI is also characterized by having a 2D topological insulator on its mirror plane. Recall that a WTI can be viewed as a stack of 2D topological insulators. This then implies that, starting from a WTI protected by time reversal and $T_z$, the translation symmetry along the $z$ direction, one can obtain a DTI by suitably breaking $T_z$ while preserving a mirror symmetry with respect to the $z = 0$ plane, such that the $z = 0$ plane hosts a 2D topological insulator.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{example_diagram.png}
\caption{Gapping out the helical fermions by a $Z_4$ STO.}
\end{figure}

In order to fully gap out the surface, we can introduce two mutually mirror symmetric $Z_4$ topological orders on the two sides of $x = 0$, such that (the topological part of) the theory near $x = 0$ is described by the following chiral Luttinger liquid theory \[ (H.7) \]

\[
\mathcal{L} = \frac{K_{IJ}}{4\pi} \partial_t \Phi_I \partial_x \Phi_J
\]

where $\Phi = (\phi_{L1}, \phi_{L2}, \phi_{R1}, \phi_{R2}, \phi_{m1}, \phi_{m2})^T$ satisfies the following Kac-Moody algebra

\[
[\partial_y \Phi_I(y), \partial_{y'} \Phi_J(y')] = 2\pi i (K^{-1})_{IJ} \partial_y \delta(y - y')
\]

\[ (H.8) \]

In our case, the $K$-matrix is

\[
K = \begin{pmatrix}
4\sigma_x & & \\
& 4\sigma_x & \\
& & \sigma_z
\end{pmatrix}
\]

\[ (H.9) \]
where the first (second) block represents the $Z_4$ topological order on the left (right) side of $x = 0$, and the last block represents the helical fermion modes at $x = 0$ (see Figure H.1). Notice in defining the chiral Luttinger liquid theory, normal directions of the left and right STOs need to be chosen. In writing down the above $K$-matrix, we choose the normal direction for the left STO to be opposite to that for the right STO.

We assign the symmetries as

$$\begin{pmatrix} \phi_{L1}, \phi_{L2}, \phi_{R1}, \phi_{R2}, \phi_{m1}, \phi_{m2} \end{pmatrix} \xrightarrow{U(1)} \begin{pmatrix} \phi_{L1} - \frac{\theta}{2}, \phi_{L2}, \phi_{R1} - \frac{\theta}{2}, \phi_{R2}, \phi_{m1} + \theta, \phi_{m2} + \theta \end{pmatrix},$$

$$\begin{pmatrix} \phi_{L1}, \phi_{L2}, \phi_{R1}, \phi_{R2}, \phi_{m1}, \phi_{m2} \end{pmatrix} \xrightarrow{T} \begin{pmatrix} -\phi_{L1} + t \cdot \frac{\pi}{4}, \phi_{L2}, -\phi_{R1} + t \cdot \frac{\pi}{4}, \phi_{R2}, -\phi_{m1} + \phi_{m2} + \pi \end{pmatrix},$$

$$\begin{pmatrix} \phi_{L1}, \phi_{L2}, \phi_{R1}, \phi_{R2}, \phi_{m1}, \phi_{m2} \end{pmatrix} \xrightarrow{M} \begin{pmatrix} \phi_{R1}, \phi_{R2}, \phi_{L1}, \phi_{L2}, \phi_{m1} + \theta, \phi_{m2} + y \pi \end{pmatrix},$$

with $t = 1$, $x = 1$ and $y = 0$. It is straightforward to check this is a consistent symmetry assignment. For example, one can check all local charge-1 objects are fermions, and they are Kramers doublets under $T$ and also have $(TM)^2 = -1$.

To gap out the helical fermions, consider introducing the following term to the Lagrangian (H.7)

$$\delta \mathcal{L} = -U \cos(4\phi_{L2} + 4\phi_{R2} - 2\phi_{m1} + 2\phi_{m2})$$

$$- V \cos(4\phi_{L1} + \phi_{m1} + \phi_{m2})$$

$$- \eta V \cos(4\phi_{R1} + \phi_{m1} + \phi_{m2})$$

with $\eta = 1$, and $U$ and $V$ positive. It is straightforward to check that $\delta \mathcal{L}$ respects all symmetries, the arguments in the cosines in $\delta \mathcal{L}$ mutually commute, and pinning the values of these cosines by making $U$ and $V$ large does not break any symmetry spontaneously. Furthermore, in this strong-coupling limit the two $Z_4$ topological orders collapse into a single $Z_4$ topological order due to the coherent propagation of anyons across $x = 0$. Therefore, this TCI can indeed have a symmetric
Now we take a closer look at the symmetry actions on the anyons of this $Z_4$ STO. We interpret $e^{i\phi L_1}$ and $e^{i\phi R_1}$ as the topological sector of the $Z_4$ charge $e$, and $e^{i\phi L_2}$ and $e^{i\phi R_2}$ as the topological sector of the $Z_4$ flux $m$. From the above symmetry assignment, the $e$ carries charge-1/2 under $U(1)$, while the $m$ is neutral. Based on the principle in Sec. 8.2.2, $m^2$, which correspond to $e^{2i\phi L_2}$ and $e^{2i\phi R_2}$, are the surface avatars of the bulk monopole. To determine the quantum numbers of the monopole, we just need to determine the quantum numbers of $m^2$.

To check their quantum number under $\mathcal{M}$ and $\mathcal{T}\mathcal{M}$, we use the method in Ref. [414]. First we consider a mirror symmetric string operator of this anyon

$$W = \exp \left( i \left( 2\phi_{L2} + 2\phi_{R2} - \phi_{m1} + \phi_{m2} \right) \right) \quad (H.12)$$

The reason to consider this string operator is because it is able to coherently move an anyon $m^2$ from a point at $x < 0$ to a point $x > 0$, given that in the strong-coupling limit $\delta\mathcal{L}$ pins the value of $\langle 4\phi_{L2} + 4\phi_{R2} - 2\phi_{m1} + 2\phi_{m2} \rangle = 2\pi N$ with $N$ an integer. Under $\mathcal{M}$, this string operator does not change. This means $\mathcal{M}^2 = 1$ for this anyon. Under $\mathcal{T}\mathcal{M}$,

$$W \rightarrow -\exp \left( -i \left( 2\phi_{L2} + 2\phi_{R2} - \phi_{m1} + \phi_{m2} \right) \right) \quad (H.13)$$

In the strong-coupling limit, this becomes

$$W \rightarrow -W \langle e^{-i(4\phi_{L2}+4\phi_{R2}-2\phi_{m1}+2\phi_{m2})} \rangle = -W \quad (H.14)$$

Therefore, $(\mathcal{T}\mathcal{M})^2 = -1$ for this anyon.

The above results are consistent with that obtained Sec. 8.3.3 based on a symmetric gapless surface: the monopole has $\mathcal{M}^2 = 1$ and $(\mathcal{T}\mathcal{M})^2 = -1$. We note that other interesting properties of this STO can also be read off from the above construction, which is beyond the purpose of this paper and we refer interested readers to Ref. [356, 386].
H.2.2 STO of TMTCI-1

Very similar as the above, a symmetric $Z_4$ STO can be constructed for TMTCI-1, and it is described by (H.7), (H.10) and (H.11), but now with \( t = 0, x = 0, y = 1 \) and \( \eta = -1 \). Again, the surface avatar of the bulk monopole corresponds to \( e^{2i\phi_{L2}} \) and \( e^{2i\phi_{R2}} \).

Now consider the string operator in (H.12). Under \( \mathcal{M} \)

\[
W \rightarrow -W
\]

which means \( \mathcal{M}^2 = -1 \) for the monopole. This is consistent with the result in Sec. 8.3.2. Under \( \mathcal{T}\mathcal{M} \)

\[
W \rightarrow -\exp \left( -i \left( 2\phi_{L2} + 2\phi_{R2} - \phi_{m1} + \phi_{m2} \right) \right)
\]

and in the strong-coupling limit we have \( W \rightarrow -W \). Therefore, \( (\mathcal{T}\mathcal{M})^2 = -1 \) for the monopoles. This is consistent with the results obtained based on the symmetric gapless surface in Appendix H.1.1.

H.2.3 STO of TMTCI-2

TMTCI-2 can also have a symmetric $Z_4$ STO described by (H.7), (H.10) and (H.11), but now with \( t = 1, x = 1, y = 1 \) and \( \eta = -1 \). Again, the surface avatar of the bulk monopole corresponds to \( e^{2i\phi_{L2}} \) and \( e^{2i\phi_{R2}} \).

Now consider the string operator in (H.12). Under \( \mathcal{T}\mathcal{M} \) it becomes

\[
W \rightarrow \exp \left( -i \left( 2\phi_{L2} + 2\phi_{R2} - \phi_{m1} + \phi_{m2} \right) \right)
\]

and in the strong-coupling limit we have \( W \rightarrow W \). Therefore, \( (\mathcal{T}\mathcal{M})^2 = 1 \) for the monopoles. This is consistent with the results obtained based on the symmetric gapless surface in Appendix H.1.2.
References


[59] The literature is too extensive to cite, but an early and important work is J.L. Tallon *et al.*, physica status solidi (b), **215**, 531 (1999).


N. Kikugawa, A. Rost, C. Hicks, A. Schofield, and A. Mackenzie, “Ca$_3$ru$_2$o$_7$: Density wave formation and quantum oscillations in the hall resistivity,” *Journal of the Physical Society of Japan*, vol. 79, 2 2010.


[238] X.-Y. Song and et. al. (in prepare)


O. I. Motrunich, “Orbital magnetic field effects in spin liquid with spinon fermi sea: Possible application to \( \kappa-(\text{ET})_2\text{cu}_2(\text{CN})_3 \),” *Phys. Rev. B*, vol. 73, p. 155115, Apr 2006.


433


