

Doubling theorem and boundary states of five-dimensional Weyl semimetalJing-Yuan Chen¹,[✉] Biao Lian,² and Shou-Cheng Zhang^{1,3,*}¹*Stanford Institute for Theoretical Physics, Stanford University, Stanford, California 94305, USA*²*Princeton Center for Theoretical Science, Princeton University, Princeton, New Jersey 08544, USA*³*Stanford Center for Topological Quantum Physics, Stanford University, Stanford, California 94305, USA*

(Received 27 March 2019; published 6 August 2019)

We study the generic band structures of the five-dimensional (5D) Weyl semimetal, in which the band degeneracies are 2D Weyl surfaces in the momentum space, and may have nontrivial linkings with each other if they carry nonzero second Chern numbers. We prove a number of theorems constraining the topological linking configurations of the Weyl surfaces, which can be viewed as a 5D generalization of the celebrated doubling theorem for 3D Weyl semimetal. As a direct physical consequence of these constraints, the 5D Weyl semimetal hosts a rich structure of topological boundary states. We show that on the 4D boundary of the 5D Weyl semimetal, there are 3D chiral Fermi hypersurfaces protected by bulk Weyl surfaces. On top of that, for bulk Weyl surfaces that are linked and carry nonzero second Chern numbers, the associated boundary 3D Fermi hypersurfaces will shrink to singularities at certain energies, which trace out a protected 1D Weyl nodal arc, in analogy to the Fermi arc on the 3D Weyl semimetal surface.

DOI: [10.1103/PhysRevB.100.075112](https://doi.org/10.1103/PhysRevB.100.075112)**I. INTRODUCTION**

Topological states of matter are known as quantum states protected by topological invariants. Depending on whether the bulk states are gapped or not, they can be divided into gapped topological states and gapless topological states. While gapped topological phases have been extensively studied [1–4], the study of gapless topological states in various dimensions is still less understood and ongoing. In general, the spatial dimension and symmetries largely determine the classification of topological states [5,6]. In three dimensions, there are two well-known classes of gapless topological states: one class is the three-dimensional (3D) Weyl semimetal [7–17], which contains Weyl nodes in the band structure and needs no symmetry protection other than the lattice translational symmetry. Each Weyl node is a twofold degenerate point protected by the first Chern number, and gives rise to topologically protected Fermi arcs on the boundaries [8,18]. The other class is the nodal line semimetal or superconductor, which contain 1D nodal lines in the Brillouin zone (BZ) protected by crystal symmetry or chiral symmetry. The nodal lines may also have linking invariants [19–29], which give rise to linking related topological responses [30].

Extending the scope into general spatial dimensions has been proven a valuable approach in understanding gapped topological states [2,5,6,31]. This approach has also been employed in the study of gapless topological states. In spatial dimensions higher than 3, an intriguing gapless topological state is the 5D Weyl semimetal, which is simultaneously characterized by both the second Chern numbers and the 5D linking invariants [32,33]. Unlike the Weyl nodes in 3D Weyl semimetal which are zero dimensional, the band degeneracy

submanifolds in 5D Weyl semimetals become 2D closed surfaces in the momentum space, which are called the *Weyl surfaces* (WSs). Since the WSs are extended objects, they admit nontrivial global configurations, in particular linking numbers among each other. Remarkably, it is shown that the sum of linking numbers of a WS is equal to the second Chern number defined via the band Berry curvature [32]. Like the bulk-boundary correspondence of usual topological states of matter, the WS linking numbers in the bulk of 5D Weyl semimetal protect topological 1D *Weyl arcs* on the boundary of the system, as can be seen in explicit lattice models [32]. Moreover, the WS linking and the second Chern number are closely related to Yang monopole [31,34] in five dimensions when the system restores a TP symmetry [33]. Transition between gapped topological phases in five dimensions can be also understood by having the gapless 5D Weyl semimetal as intermediate stage [33], in analogy to the topological phase transitions in three dimensions [9].

One of the most well-known theorems on 3D Weyl nodes is the doubling theorem [35,36], which restricts the Weyl nodes in the BZ to appear in pairs of opposite chirality. This theorem has played an important role in the historical development of lattice quantum chromodynamics and Weyl semimetal. It is natural to ask whether a similar “doubling” constraint exists for the WSs in a 5D Weyl semimetal. While simple models of 5D Weyl semimetals have been constructed [32,33], the generic constraints of WSs in five dimensions have not been carefully studied yet. The aim of this paper is to answer this question in a generic way. We prove that there are nontrivial topological constraints governing the global configuration of the WSs and their linking. As we shall show, some of the constraints in five dimensions appear similar to the doubling theorem in three dimensions, however, their origins involve some substantial differences. Based on these constraints, we develop the correspondence between the bulk

*Deceased.

band topology and the gapless boundary states in a generic manner (as opposed to going into specific models), much like how the doubling theorem underlies the bulk-boundary correspondence in 3D Weyl semimetal.

We should note that extended band degeneracy submanifolds can also appear in other systems, most notably nodal line semimetals and superconductors in three dimensions, and the nodal lines may also have linking and other topological consequences. However, in these systems the extendedness of the degeneracy is protected by discrete space-time symmetries (space group symmetry, etc.), in contrast to the 5D Weyl semimetal in which the extendedness of the Ws is robust without symmetry (as long as lattice momentum is well defined). Therefore, in general, the nodal lines in 3D systems are not subjected to stringent topological constraints as in the doubling theorem or those that we are going to prove. Notably, however, when a 3D nodal line system carries nontrivial \mathbb{Z}_2 monopole charge [37,38], the nodal line configurations in three dimensions are subjected to similar constraints as our Weyl surfaces in five dimensions; we will discuss the connection as we proceed.

This paper is organized as follows. In Sec. II we review the well-known doubling theorem of Weyl nodes in three dimensions. In Sec. III we review the notion of the Weyl surfaces in 5D Weyl semimetal, and the mathematical description of their linking. From this discussion we raise a few questions, which lead to the topological constraints we present and prove in Sec. IV. Moreover, we will discuss how these constraints are related to the non-Abelian Yang monopoles. In Sec. V we consider a 5D Weyl semimetal with a 4D physical surface, which hosts rich surface states protected by the topological constraints in the 5D bulk. Finally, we conclude by a few further remarks.

II. 3D DOUBLING THEOREM

We first recall the physics of 3D Weyl semimetals and the doubling theorem for the Weyl nodes. We consider a band theory single-electron Hamiltonian $H(k)$ with N bands, labeled by momentum k in the BZ of $D = 3$ dimensions. We assume there is no symmetry other than lattice translational symmetry and time translational symmetry. The $N \times N$ Hamiltonian can be generically diagonalized as

$$H_{\beta}^{\alpha}(k) = \sum_{n=1}^N u_n^{\alpha}(k) E_n(k) u_{n\beta}^{*}(k), \quad (1)$$

which we may write as matrix factorization $H = U E U^{\dagger}$ for short. Without loss of generality, one can sort the energy levels so that [35]

$$E_n(k) \leq E_{n+1}(k), \quad (2)$$

which holds for any momentum k . Generically, without additional symmetry, the equality $E_n(k) = E_{n+1}(k)$ takes place only on a $(D - 3)$ -dimensional submanifold in the BZ. This is because when considering the degeneracy between two adjacent bands, one may project the Hamiltonian onto those two bands and use the Pauli matrices basis

$$H_{\text{proj}}(k) = h_0(k) + h_1(k)\sigma^1 + h_2(k)\sigma^2 + h_3(k)\sigma^3. \quad (3)$$

The two bands are degenerate if and only if $h_1 = h_2 = h_3 = 0$. This yields three conditions to be satisfied by the D components of k . When $D = 3$, the degeneracy takes places at points known as *Weyl nodes*. We shall denote the positions of the Weyl nodes between the n th and $(n + 1)$ th band by $k_{n+1/2}^i$ (the superscript labels each Weyl node and the subscript indicates between which two bands it lies). Each Weyl node is associated with a chirality $c_{n+1/2}^i = \pm 1$ (right or left handed), as determined by the Hamiltonian in the following way. Define the Abelian Berry connection and Berry curvature of the n th band as (we use the notation of differential forms and matrix multiplication)

$$A_n(k) \equiv -i u_n^{\dagger}(k) du_n(k),$$

$$F_n(k) \equiv dA_n(k) = -i du_n^{\dagger}(k) du_n(k). \quad (4)$$

A Weyl node can then be viewed as a ‘‘monopole’’ of Berry curvature:

$$dF_n(k) = -\star \sum_i \frac{c_{n+1/2}^i}{2} 4\pi\delta^3(k - k_{n+1/2}^i) + \star \sum_j \frac{c_{n-1/2}^j}{2} 4\pi\delta^3(k - k_{n-1/2}^j), \quad (5)$$

where now i runs over the Weyl nodes between the n th and $(n + 1)$ th band, and j runs over those between the $(n - 1)$ th and n th band; \star denotes Hodge dual. To better understand this expression, we can draw a small sphere S^i around a Weyl node $k_{n+1/2}^i$, then the *first Chern number* is given by the chirality,

$$(C_1)_n^{S^i} \equiv \oint_{S^i} \frac{F_n}{2\pi} = -c_{n+1/2}^i, \quad (6)$$

and moreover, $(C_1)_{n+1}^{S^i} = -(C_1)_n^{S^i} = c_{n+1/2}^i$. The monopole property (5) can be shown by linearizing H_{proj} in $k - k_{n+1/2}^i$ near $k_{n+1/2}^i$, and performing the explicit calculation.

The famous doubling theorem [35,36] asserts that the total chirality of the Weyl nodes between the n th and the $(n + 1)$ th band satisfies

$$\sum_i c_{n+1/2}^i = 0. \quad (7)$$

Namely, Weyl nodes have to appear in pairs of opposite chiralities. The proof is rather straightforward. Consider small spheres S^i enclosing each $k_{n+1/2}^i$ and also S^j enclosing each $k_{n-1/2}^j$. By Stoke’s theorem and Eq. (5), we have

$$\sum_{i,j} \oint_{S^{i,j}} \frac{F_n}{2\pi} = -\sum_i c_{n+1/2}^i + \sum_j c_{n-1/2}^j. \quad (8)$$

On the other hand, the BZ is a closed manifold, so we can equally well view the ‘‘outside’’ of all the spheres as the ‘‘inside’’ (up to a minus sign from reversing orientation). Then the integral of F_n is identically 0, since now the inside, which used to be the outside, contains no Weyl nodes. This indicates $\sum_i c_{n+1/2}^i = \sum_j c_{n-1/2}^j$. Besides, the lowest band $n = 1$ cannot have any Weyl nodes connected from below, as there are no lower bands, thus we have $\sum_j c_{1/2}^j = 0$. By iteration, we then arrive at the doubling theorem in Eq. (7).

The doubling theorem leads to important physical consequences in 3D Weyl semimetal, most remarkably the *Fermi arc* on the surface of the semimetal [8]. More precisely, the 2D spatial boundary of the 3D Weyl semimetal has an associated 2D momentum space. The Weyl nodes in the 3D momentum space can be projected on this 2D momentum space. It has been shown that the projected Weyl nodes must pair up with opposite chiralities, such that between each pair there connects a 1D Fermi arc, perpendicular to which a chiral boundary mode flows (since the mode is chiral, the 1D Fermi surface is an arc instead of a closed loop). This chiral surface mode is protected by the C_1 of the Weyl nodes at the ends of the Fermi arc. Therefore, the doubling theorem in the bulk give rise to well-defined topologically protected surface states on the surface of the system.

III. WEYL SURFACES AND LINKING IN FIVE DIMENSIONS

In this section we review the notion of Weyl surfaces (WSs) in a $D = 5$ -dimensional Weyl semimetal, and the mathematical relation between their linking number and the second Chern number [32]. Then we motivate the question of topological constraints for WS linking, which we shall prove in the next section.

The Hamiltonian and its band energies again take the form of Eqs. (1) and (2), except that the momentum k lives in $D = 5$ dimensions. As we mentioned below Eq. (2), the two-band degeneracy $E_n(k) = E_{n+1}(k)$ will generically take place on a $(D - 3)$ -dimensional submanifold in the D dimensional BZ. By a similar argument, a three-band degeneracy $E_{n-1}(k) = E_n(k) = E_{n+1}(k)$ will generically take place on a $(D - 8)$ -dimensional sub-manifold. Therefore, in $D = 5$ under consideration, we only have 2D submanifolds of two-band degeneracies, which is called the WSs. To be specific, we denote the WS between the n th band and $(n + 1)$ th band as $W_{n+1/2}$, which is generically a 2D closed manifold. It may consist of multiple disjoint connected components, and we can denote each connected component by $W_{n+1/2}^i$, with $\cup_i W_{n+1/2}^i = W_{n+1/2}$. Similar to the Weyl node in three dimensions, the WS in five dimensions can also be viewed as a ‘‘monopole’’ of Berry connection in five dimensions [32], satisfying an equation analogous to Eq. (5):

$$dF_n(k) = 2\pi \left(\int_{k' \in W_{n-1/2}^i} - \int_{k' \in W_{n+1/2}^j} \right) \star \delta^5(k - k'). \quad (9)$$

The left-hand side is a three-form, so is the right-hand side, because $\star \delta^5(k - k')$ is a five-form and $\int_{k' \in W_{n\pm 1/2}^i}$ are 2D integrals. This fixes the orientation on each connected WS $W_{n+1/2}^i$, similar to that Eq. (5) fixes the chirality $c_{n+1/2}^i$ for each 3D Weyl node.

Now we consider the linking between a WS $W_{n+1/2}^i$ and another WS $W_{n-1/2}^j$. In general, two submanifolds can link with each other if their sum of dimensions is equal to the total spatial dimension minus 1 (in this case $2 + 2 = 5 - 1$), so the WSs can indeed be linked together. An easy way to characterize their linking is the following: let $W_{n+1/2}^i = \partial \Sigma_{n+1/2}^i$ (the existence of such a 3D manifold $\Sigma_{n+1/2}^i$ will be

proven in the next section), then $W_{n+1/2}^i$ and $W_{n-1/2}^j$ will have linking number 1 if there is an intersection point of $\Sigma_{n+1/2}^i$ and $W_{n-1/2}^j$. This is a straightforward generalization of the picture of linked loops in three dimensions.

It has been shown that the linking number between $W_{n+1/2}^i$ and $W_{n-1/2}^j$ is related to the *second Chern number* of the Berry curvature in the n th band [32]. Here we show this using the above geometric picture of linking. Let's focus on the two WS connected components $W_{n+1/2}^i$ and $W_{n-1/2}^j$ which link with each other. Let ∂V be a 4D ‘‘narrow tube’’ (which is topologically $W_{n+1/2}^i \times S^2$) enclosing $W_{n+1/2}^i$ which is the boundary of a 5D region V . We can choose V narrow enough so that V does not touch any other WSs in addition to $W_{n+1/2}^i$. The second Chern number on ∂V , viewed in the Berry curvature of the n th band, is defined as

$$(C_2)_n^{\partial V} \equiv \oint_{\partial V} \frac{1}{2} \left(\frac{F_n}{2\pi} \right)^2. \quad (10)$$

Using Stoke's theorem and (9), we have

$$(C_2)_n^{\partial V} = \int_V \frac{dF_n}{2\pi} \frac{F_n}{2\pi} = - \oint_{W_{n+1/2}^i} \frac{F_n}{2\pi}, \quad (11)$$

which becomes a 2D integral on WS $W_{n+1/2}^i$. Now we have a problem, since F_n is singular on WS $W_{n+1/2}^i$. This is resolved as follows: since monopole is essentially a solution to the Poisson equation, one can separate the Berry curvature as $F_n = F_n^{(1)} + F_n^{(2)}$, such that $F_n^{(2)}$ is nonsingular on $W_{n+1/2}^i$, while $F_n^{(1)}$ is singular on $W_{n+1/2}^i$ but nonsingular on $W_{n-1/2}^j$. One may view $F_n^{(1)}$ and $F_n^{(2)}$ as produced by monopoles $W_{n-1/2}^j$ and $W_{n+1/2}^i$, respectively. Then we should understand the last F_n in Eq. (11) as the nonsingular part $F_n^{(2)}$. (In fact, the integral of the singular $F_n^{(1)}$ part gives self-linking number of $W_{n+1/2}^i$, which is identically zero in five dimensions. We will come back to this in the below [32]). Thus, using Stoke's theorem and Eq. (9) again we find

$$\begin{aligned} (C_2)_n^{\partial V} &= - \int_{\Sigma_{n+1/2}^i} \frac{dF_n^{(2)}}{2\pi} \\ &= - \int_{k \in \Sigma_{n+1/2}^i} \int_{k' \in W_{n-1/2}^j} \star \delta^5(k - k') \\ &\equiv -L_{W_{n+1/2}^i, W_{n-1/2}^j}, \end{aligned} \quad (12)$$

where $L_{W_{n+1/2}^i, W_{n-1/2}^j}$ counts the (signed) number of intersection points of $\Sigma_{n+1/2}^i$ and $W_{n-1/2}^j$, and is thus the linking number between $W_{n+1/2}^i$ and $W_{n-1/2}^j$. This process of computation is illustrated in Fig. 1, where the linking number takes values ± 1 depending on the orientation of the link.

The derivation in Eq. (12) motivates us to define two first Chern numbers associated with $W_{n+1/2}^i$ in the n th band, the singular first Chern number $(C_1^s)_n$ and the regular first Chern number $(C_1^r)_n$:

$$(C_1^s)_n = \oint_{S^2} \frac{F_n^{(1)}}{2\pi}, \quad (C_1^r)_n = \oint_{W_{n+1/2}^i} \frac{F_n^{(2)}}{2\pi}, \quad (13)$$

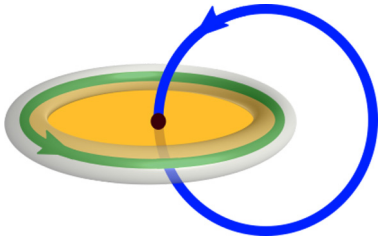


FIG. 1. Illustration of the computation relating the second Chern number to the linking number. The green loop represents the 2D WS component $W_{n+1/2}^i$, bounding the yellow region which represents the 3D $\Sigma_{n+1/2}^i$. The blue loop represents the 2D WS component $W_{n-1/2}^j$. The grey torus represents the 4D “narrow tube” ∂V , whose “inside” V contains the green loop. The linking number is witnessed by the black dot, the intersection point of $\Sigma_{n+1/2}^i$ and $W_{n-1/2}^j$.

where S^2 in the definition of $(C_1^s)_n$ is an infinitesimal 2D sphere in the three codimensions of the WS $W_{n+1/2}^i$ which links with $W_{n+1/2}^i$. In general, $(C_1^s)_n$ is always ± 1 , which also depends on the orientation of S^2 . The second Chern number is then given by $(C_2)_n^{\partial V} = -(C_1^s)_n (C_1^s)_n$ (recall that topologically $\partial V \cong W_{n+1/2}^i \times S^2$). Similarly, one can define two such Chern numbers for $W_{n+1/2}^i$ in the $(n+1)$ th band, for which we choose the orientation of S^2 so that $(C_1^s)_{n+1} = (C_1^s)_n$, and we also have $(C_2)_{n+1}^{\partial V} = -(C_1^s)_{n+1} (C_1^s)_{n+1}$. Note that $(C_1^s)_n = \pm 1$ simply implies the fact that a WS can be viewed as a Weyl node in the three codimensions orthogonal to the WS. As we shall prove in the next section, we always have $(C_2)_{n+1}^{\partial V} = -(C_2)_n^{\partial V}$, therefore under the above convention we have $(C_1^s)_{n+1} = -(C_1^s)_n$. The concepts of both of these first Chern numbers will appear in the surface-state physics in Sec. V.

The computation above shows a nice relation between the algebraic characterization C_2 and the geometric picture of linking. But it also raises several questions:

(1) We assumed every WS $W_{n+1/2}^i$ is a boundary $\partial \Sigma_{n+1/2}^i$ of some 3D manifold $\Sigma_{n+1/2}^i$. We need to prove such $\Sigma_{n+1/2}^i$ does always exist, so that the notion of “linking” and the computation above make sense.

(2) The linking we considered involves one WS component between the $(n-1)$ th and n th band, and the other WS component between the n th and $(n+1)$ th band. In principle, cannot one also consider both of them between the n th and $(n+1)$ th band? (Geometrically one can also think of linking between, say, some $W_{n+1/2}^i$ and some $W_{n-3/2}^j$. But such linking is not detected by any Berry curvature and can adiabatically pass through each other to unlink. So by “linking” we will not refer to such trivial possibilities.)

(3) There is an important distinction between 2D objects linked in five dimensions versus our familiar 1D objects linked in three dimensions. If one goes through the “counting intersection points” definition of L carefully, one finds that, for 1D loops linked in three dimensions, $L_{C,C'} = L_{C',C}$, but for 2D surfaces linked in five dimensions, $L_{W,W'} = -L_{W',W}$ (such an alternating sign repeats in higher odd dimensions mod four dimensions) [32]. This distinction has some important consequences in our problem:

(a) In three dimensions (mod four dimensions), when a loop is properly regularized, it has a notion of *self-linking*, which plays an important role in, e.g., relativistic flux attachment [39]. But in five dimensions (mod four dimensions), due to the negative sign above, the self-linking must vanish for self-consistency [40]. This justifies that in (11) we can regard the last F_n as $F_n^{(2)}$ and drop $F_n^{(1)}$, instead of managing to regularize the $F_n^{(1)}$ contribution.

(b) Suppose one generalizes the exercise (8) from three dimensions to five dimensions—that is, enclose each $W_{n+1/2}^i$ and each $W_{n-1/2}^j$ with a narrow tube, compute the $(C_2)_n$ over all these tubes, and then regard the outside of the tubes as the inside, in order to derive any topological constraint about the links in the BZ. This exercise leads to that the total $(C_2)_n$ must be 0, similar to (7). While (7) implies Weyl nodes must appear in pairs of opposite chiralities, the vanishing of the total $(C_2)_n$ in the present case provides no topological constraint on the links—the total $(C_2)_n$ being 0 is always *trivially* satisfied simply because $L_{W,W'} = -L_{W',W}$. Does this mean there is no topological constraint on the linking configurations in 5D Weyl semimetal?

These are the questions we will address next. Moreover, we will see that the answers to these questions are directly related to the structure of the gapless surface states of a 5D Weyl semimetal, similar to that the Fermi arcs are related to the doubling theorem in three dimensions.

IV. TOPOLOGICAL CONSTRAINTS ON WEYL SURFACES

A. Topological constraints

We shall show the following topological constraints on the WS and their linking:

(1) For each band index n , the WS $W_{n+1/2}$ must be a boundary,

$$W_{n+1/2} = \partial \Sigma_{n+1/2}. \quad (14)$$

Remark. $\Sigma_{n+1/2}$ may involve several connected components $\Sigma_{n+1/2}^i$, and we will denote $W_{n+1/2}^i = \partial \Sigma_{n+1/2}^i$. Note that we have modified our notation slightly—in the previous section $W_{n+1/2}^i$ refers to a connected component of $W_{n+1/2}$, but now it might involve several connected components such that they together form the boundary of a connected component $\Sigma_{n+1/2}^i$. [For instance, in case (b) of Fig. 2, the two connected components together form a boundary, but not individually.]

Notice that Eq. (14) resembles the 3D doubling theorem (a pair of Weyl nodes of opposite chiralities can be viewed as the ends of an arc), with Weyl nodes replaced by WS.

(2) For each $W_{n+1/2}^i$,

$$L_{W_{n+1/2}^i, W_{n+1/2}} = 0, \quad (15)$$

i.e., its linking number with the other WS components between the same pair of bands is zero.

(3) For each $W_{n+1/2}^i$,

$$L_{W_{n+1/2}^i, W_{n-1/2}} = -L_{W_{n+1/2}^i, W_{n+3/2}}, \quad (16)$$

i.e., its linking numbers with the WS one band lower and the WS one band higher must be opposite.

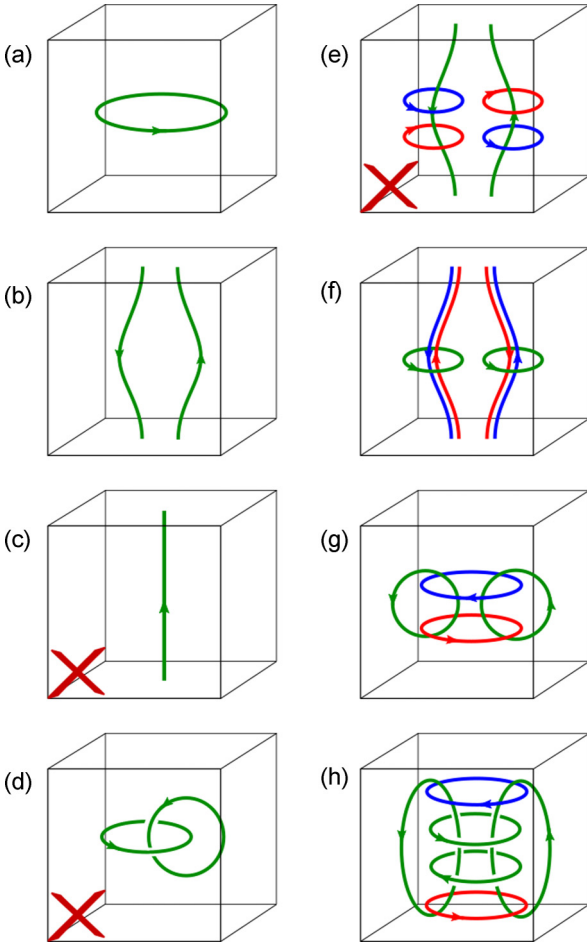


FIG. 2. Some examples of WS configuration satisfying/violating the topological constraints. We picture the 5D BZ by a 3D (periodic) cube, and 2D WS by 1D (oriented) loops. For simplicity we consider a four-band system, with $W_{3/2}$ in blue, $W_{5/2}$ in green, and $W_{7/2}$ in red. Configurations (a), (b), (f), (g), and (h) are allowed, while (c), (d), and (e) are not allowed for violating (14), (15), and (16) respectively. The explicit lattice model constructed in Ref. [32] has a WS configuration (f). [A Hopf link seen by the n th band, like that in Fig. 1, is only allowed if the number of bands $N > 4$, with $2 < n < N - 1$ due to (16). Moreover, the total number of Hopf links seen by the n th band must be even, due to (17).]

Remark. An important corollary is

$$L_{W_{n+1/2}, W_{n-1/2}} = 0, \quad (17)$$

i.e., for each n , WS links must appear in pairs with opposite linking numbers. This is because (16) taking the union over i leads to $L_{W_{n+1/2}, W_{n-1/2}} = -L_{W_{n+1/2}, W_{n+3/2}}$, then one can start with $n = 1$ and iterate. Notice that (17) also resembles the 3D doubling theorem, but [in contrast to (14)] with Weyl nodes replaced by WS links.

Some examples of allowed and not allowed WS configurations are illustrated in Fig. 2. The physical consequences of these constraints will be discussed in Sec. V. In the rest of this section, we shall prove these constraints. (The connection between these constraints and those in a special class of 3D nodal line systems [37] is discussed at the end of this section.)

B. Proof of the topological constraints

The first constraint (14) can be proven using some formality. Clearly the WS must be a closed manifold. In the spirit of induction, suppose we have already shown that $W_{n-1/2}$ is a boundary (the $n = 1$ case is trivially true). Now suppose $W_{n+1/2}$ is not a boundary. Then there must exist some closed, nonexact differential two-form G such that $\oint_{W_{n+1/2}} G \neq 0$, for the following reason. The closed $W_{n+1/2}$ not being a boundary means it is a nontrivial element in the homology group $H_2(BZ)$. On the other hand, closed differential m forms over the BZ are elements of the cohomology group $H^m(BZ, \mathbb{R})$. By the universal coefficient theorem,

$$H^m(BZ, \mathbb{R}) = \text{Hom}(H_m(BZ), \mathbb{R}), \quad (18)$$

where Hom is all homomorphisms. The right-hand side for $m = 2$ contains elements that map $W_{n+1/2}$ to nonzero real numbers, and thus guarantees the existence of G with the said properties. With such G , we can use the familiar method (8) from three dimensions. Consider narrow tube(s) ∂V enclosing $W_{n+1/2}$, then by Stoke's theorem, the closedness of G and (9), and we have

$$\oint_{\partial V} \frac{F_n}{2\pi} \frac{G}{2\pi} = \int_V \frac{dF_n}{2\pi} \frac{G}{2\pi} = - \oint_{W_{n+1/2}} \frac{G}{2\pi} \neq 0. \quad (19)$$

Now we regard the outside of ∂V as the inside and apply Stoke's theorem again: we have

$$- \oint_{\partial V} \frac{F_n}{2\pi} \frac{G}{2\pi} = \int_{BZ \setminus V} \frac{dF_n}{2\pi} \frac{G}{2\pi} = \oint_{W_{n-1/2}} \frac{G}{2\pi} = 0, \quad (20)$$

where the last equality is because G is closed and, by induction assumption, $W_{n-1/2}$ is a boundary. Thus a contradiction arises. This shows that $W_{n+1/2}$ must be a boundary.

The second constraint (15) is easily seen for a two-band system. Let u_1 be the first eigenvector. Then $F_1^2 = -(du_1^\dagger du_1)^2$. However, for a two-band system, u_1 and u_1^\dagger together only have three real parameters (the phase choice can be fixed without affecting F_1), and therefore F_1^2 , being a wedge product with four d 's acting on three real parameters, vanishes. Thus, for a two-band system, the second Chern number around each $W_{3/2}^i$ —whose integrand is proportional to F_1^2 —must vanish, leading to (15). For systems with more bands, the statement (15) only involves the n th and $(n + 1)$ th bands, so roughly speaking we can project the Hamiltonian as an effectively two-band system. However, in considering global topological effects such as linking, what we mean by “effectively two band” needs to be carefully addressed. We leave this technical detail to Appendix A.

In the third constraint (16), the left-hand side is detected by F_n and the right-hand side by F_{n+1} . We need to find a relation relating the two. Consider a narrow tube ∂V enclosing a WS component $W_{n+1/2}^i$ in its inside; we choose it to be so narrow that V is disjoint from any other WS. We use the fact that the sum of Chern numbers over all bands vanishes:

$$\sum_{n'=1}^N (C_2)_{n'}^{\partial V} = 0. \quad (21)$$

[This can be deduced by iterating (26), starting with $n = N$.] For $n' = n$, we have $(C_2)_n^{\partial V} = -L_{W_{n+1/2}^i, W_{n-1/2}}$. For $n' =$

$n + 1$, we have $(C_2)_{n+1}^{\partial V} = -L_{W_{n+1/2}, W_{n+3/2}}^i$. Other $(C_2)_{n'}^{\partial V}$ vanish because $W_{n+1/2}$ is not seen by other $F_{n'}$. This leads to (16).

This derivation of (16) and hence (17) appear very different from the familiar derivation of the 3D doubling theorem, despite that (17) is a ‘‘doubling theorem’’ for WS links. In particular, the derivation above does not involve ‘‘enclosing all the links and then viewing the outside as the inside.’’ Is it possible to rederive (17) from such a perspective? Another point is that there is an interesting relation between the WS link and the Yang monopole [33], but it seems not to manifest in the steps above. Below, we show that the answers to both these points are related. We first discuss the relation between the WS link and Yang monopole through *non-Abelian Berry curvature*, and using this notion we provide an alternative direct derivation of (17) that resembles the derivation of the 3D doubling theorem.

C. Weyl surface link, Yang monopole and non-Abelian Berry curvature

The doubling theorem (17) for WS links is an interesting result. In 3D nodal line materials in which linking also plays an interesting role [24–30], there is no topological constraint enforcing the nodal line links to appear in pairs (unless a more stringent constraint is applied [37,38] which we will return to in the end). Why in five dimensions must the WS links appear in pairs? This is because the nodal line in three dimensions is protected by discrete symmetry, while the WS in five dimensions is protected by topology. Below we provide an alternative view towards the constraint (17). One can motivate (17) by the relation between WS links and Yang monopoles [33]: a WS link can adiabatically arise from deforming a Yang monopole, and Yang monopoles themselves must appear in pairs, like Weyl nodes, so WS links must also appear in pairs. Now we proceed to the details.

While our previous proof of (17) only employed Abelian Berry curvature, in our alternative view to be introduced now we relate non-Abelian Berry curvature to Abelian ones. Let us introduce the notion of non-Abelian Berry curvature. Recall that the Hamiltonian is diagonalized as $H = UEU^\dagger$, and the eigenvalues in E are ordered from low to high energies. The matrix U has N columns, being the eigenvectors u_1, \dots, u_N . Let $U_{\leq n}$ be the $N \times n$ rectangular matrix consisting of the first n columns of U , i.e., the columns of this rectangular matrix is u_1, \dots, u_n . In other words, $U_{\leq n}$ is U projected to the first n bands. Note that $U_{\leq n}U_{\leq n}^\dagger$ is the projection matrix onto the first n bands, and $U_{\leq n}^\dagger U_{\leq n}$ is the identity $\mathbf{1}_{n \times n}$ acting on the first n bands. The non-Abelian Berry connection

$$A_{\leq n} \equiv -iU_{\leq n}^\dagger dU_{\leq n} \quad (22)$$

is therefore an $n \times n$ matrix-valued connection one-form. The associated non-Abelian Berry curvature

$$\begin{aligned} F_{\leq n} &\equiv dA_{\leq n} + iA_{\leq n}^2 \\ &= -i dU_{\leq n}^\dagger (1 - U_{\leq n}U_{\leq n}^\dagger) dU_{\leq n} \end{aligned} \quad (23)$$

is an $n \times n$ matrix-valued curvature two-form. (One may wonder, in the absence of band degeneracy, the system does not have the $U(n)$ symmetry rotating among the first n bands,

why would such $U(n)$ Berry curvature still be useful to define? The idea is that this Berry curvature captures the separation of the original N -dimensional vector bundle over the BZ into two sub-bundles consisting of the first n bands and the $N - n$ bands respectively. In doing so, we are viewing the first n bands as a whole and not worrying about the further separation within them.) Note that $F_{\leq n}$ is only singular on the WS $W_{n+1/2}$ between the n th and the $(n + 1)$ th band, but does not see, say, $W_{n-1/2}$, in contrast to F_n , because $F_{\leq n}$ treats the first n bands as a whole.

Having introduced the notion of non-Abelian Berry curvature, now let’s consider it in a four-band system with TP symmetry, such as the model constructed in Ref. [33]. The first and second bands are completely degenerate, so are the third and fourth bands. Yang monopoles exist between the two pairs of degenerate bands in the 5D BZ. A Yang monopole is characterized by the non-Abelian second Chern number

$$(C_2)_{\leq 2}^{S^4} \equiv \text{tr} \int_{S^4} \frac{1}{2} \left(\frac{F_{\leq 2}}{2\pi} \right)^2 \quad (24)$$

around an S^4 enclosing the Yang monopole taking the value ± 1 . The ± 1 Yang monopoles appear in pairs in analogy to the 3D doubling theorem, due to an analogous proof. A TP-breaking perturbation of this model does two things [33]: (1) it lifts the complete degeneracy between the first and second bands, leaving a WS degeneracy $W_{3/2}$ (and likewise for the third and fourth bands), and (2) it ‘‘stretches’’ each pointlike Yang monopole into a small spherical WS component $W_{5/2}^i$ linked to $W_{3/2}$ and $W_{7/2}$. The resulting WS configuration is like Fig. 2(f). If we turn on this TP-breaking perturbation adiabatically, $(C_2)_{\leq 2}^{S^4}$ should not change as long as the small $W_{5/2}^i$ sphere is still enclosed by the S^4 .

The example above motivates us to propose the following claim: $(C_2)_{\leq n}^{S^4}$ equals ± 1 if the S^4 encloses a small WS component $W_{n+1/2}^i$ that links with $W_{n-1/2}$. Note that $W_{n-1/2}$ may intersect with the S^4 , but this is no problem because $F_{\leq n}$, treating the first n bands as a whole, does not see (i.e., is nonsingular on) $W_{n-1/2}$. But now that $F_{\leq n}$ only sees $W_{n+1/2}$, we can deform the S^4 to a narrow tube ∂V enclosing $W_{n+1/2}^i$ (such that V is disjoint from any other WS), without changing the value of C_2 , i.e., $(C_2)_{\leq n}^{\partial V} = (C_2)_{\leq n}^{S^4}$. Thus, our claim becomes

$$(C_2)_{\leq n}^{\partial V} \equiv \text{tr} \int_{\partial V} \frac{1}{2} \left(\frac{F_{\leq n}}{2\pi} \right)^2 = \pm 1. \quad (25)$$

This idea is illustrated in Fig. 3. Moreover, with S^4 replaced by ∂V , we no longer need to require $W_{n+1/2}^i$ to be small so that it can be enclosed by an S^4 ; it can run across the BZ and still be enclosed by ∂V .

Let’s prove the claim (25). More particularly, we want to show that the non-Abelian second Chern number $(C_2)_{\leq n}^{\partial V}$ is in fact equal to the Abelian one $(C_2)_n^{\partial V}$ that counts the linking number between $W_{n+1/2}^i$ and $W_{n-1/2}$. To show this equality, in Appendix B we show that, on a 4D closed manifold \mathcal{M} that does not intersect $W_{n\pm 1/2}$, we have

$$\text{tr} \int_{\mathcal{M}} F_{\leq n}^2 = \int_{\mathcal{M}} F_n^2 + \text{tr} \int_{\mathcal{M}} F_{< n}^2. \quad (26)$$

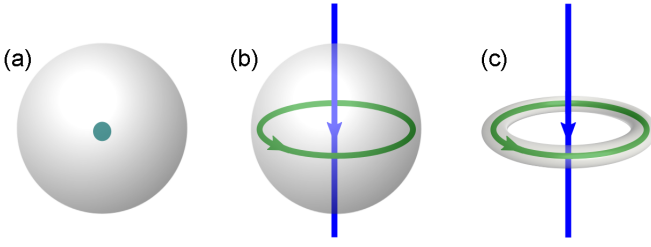


FIG. 3. The non-Abelian second Chern numbers $(C_2)_{\leq n}^{S^4}$ in (a), $(C_2)_{\leq n}^{S^4}$ in (b), and $(C_2)_{\leq n}^{\partial V}$ in (c) are all equal to the Abelian second Chern number $(C_2)_n^{\partial V} = 1$ in (c). From (a) to (b) the Yang monopole is “stretched out” into a WS link due to TP breaking. From (b) to (c) the sphere S^4 enclosing $W_{n+1/2}^i$ is deformed to the narrow tube ∂V .

(This is a special case of the Whitney sum formula for characteristic classes.) Taken $\mathcal{M} = \partial V$ where V contains $W_{n+1/2}^i$ but is disjoint from $W_{n-1/2}$, and we get

$$(C_2)_{\leq n}^{\partial V} = (C_2)_n^{\partial V} \quad (27)$$

because the $F_{<n}$ does not see $W_{n+1/2}$. This is the general case of the result $(C_2)_{\leq 2}^{\partial V} = (C_2)_2^{\partial V}$ explicitly computed in the four-band model in Ref. [33].

We can have a more intuitive understanding of the doubling Theorem (17) for WS links that resembles that of the 3D doubling theorem. Now let ∂V be narrow tubes that enclose all components of $W_{n+1/2}$. By (27), we have $(C_2)_{\leq n}^{\partial V} = -L_{W_{n+1/2}, W_{n-1/2}}$. Next we view the outside of ∂V as the inside and evaluate $(C_2)_{\leq n}^{\partial V}$ again; but there is no $W_{n+1/2}$ to be picked up, so the result must vanish (recall that by construction $F_{\leq n}$ only sees $W_{n+1/2}$, in contrast to F_n which also sees $W_{n-1/2}$). This proves (17).

The relation to Yang monopole also gives an intuitive understanding of (16). Upon TP breaking, a Yang monopole is “stretched out” into a component $W_{n+1/2}^i$ linked with $W_{n-1/2}$; but we can equally well replace $W_{n-1/2}$ with $W_{n+3/2}$, which resonates with (16). This argument, however, cannot be taken as an alternative proof of (16) because *a priori* we cannot assume each WS linking arises adiabatically from a Yang monopole.

Before we close, we would like to comment on an interesting connection between our constraints in 5D Weyl semimetal to those in certain 3D nodal line systems. It is shown [37,38] that in 3D nodal line systems protected by TP symmetry with $(\text{TP})^2 = +1$, one can define a \mathbb{Z}_2 monopole charge characterized by the second Stiefel-Whitney class of the band structure, and if the charge is nontrivial, the nodal lines must develop links, in a manner parallel to the constraints we proved in this section. Notably, one process to visualize [37] the relation between the \mathbb{Z}_2 monopole charge and the linking number is parallel to our visualization Fig. 3. Although the method employed there appears different from ours, an intuitive explanation to the similarity between the constraints is that the $(\text{TP})^2 = +1$ condition demands the Hamiltonian to be real valued, hence Refs. [37,38] and ours are considering the same kind of constraint problems for real versus complex valued Hamiltonians, in three and five dimensions respectively.

V. TOPOLOGICAL SURFACE STATES

The above topological constraints on WSs and their linkings lead to rich surface states on the 4D surface of a 5D Weyl semimetal, as we will derive in this section. In particular, we give a more intuitive and clearer understanding of the Weyl arc on the surface [32,33], which is known to be protected by the second Chern number of WSs.

We first recall that the 3D Weyl semimetal hosts topologically protected Fermi arcs on its 2D surface, each of which connects the surface projection of two Weyl nodes with monopole charges (first Chern numbers) $+1$ and -1 , respectively [8]. In other words, the constant electron energy contour of the surface states on a 2D surface of 3D Weyl semimetal is not a closed loop (the usual Fermi surface in two dimensions), but an open Fermi arc connecting two Weyl nodes with opposite monopole charges. Mathematically, one can show that the first Chern number of a Weyl node requires its projection on the 2D surface to be connected with a Fermi arc [33].

The 4D surface states of 5D Weyl semimetal exhibits a richer structure. There are two types of topological surface-state features: a 3D *Fermi hypersurface* protected by the first Chern number [Eq. (13)] of the WS, and, on top of it, a 1D *Weyl arc* protected by the second Chern number [Eq. (11)] of the WS. Now we explain them in detail.

First, consider a WS $W_{n+1/2}^i$ with its second Chern numbers $(C_2)_n^{\partial V} = -(C_2)_{n+1}^{\partial V} = 0$. According to the constraints we revealed in the last section, such a WS can exist alone, such as in Figs. 2(a) and 2(b). It is, however, not completely topologically trivial, since it still carries a singular first Chern number $(C_1)_n = +1$ [recall (13)]. As we have explained in Sec. III, the first Chern number $(C_1)_n$ simply implies that each point of the WS behaves as a Weyl node in the three codimensions to the 2D WS. Therefore, while in 3D Weyl semimetal each 0D Weyl node when projected to the 2D physical surface serves as an end point of a 1D Fermi arc [8] due to the first Chern number’s protection, for exactly the same mathematical reason, in 5D Weyl semimetal each 2D WS when projected to the 4D physical surface serves as the boundary of a 3D *Fermi hypersurface*, which occurs at a given Fermi energy near the energy of the WS. Such a 3D Fermi hypersurface with chiral 2D boundary (the projected WS) is the topological surface state protected by $(C_1)_n$, and can only exist on the 4D surface of a 5D system, because otherwise any Fermi surface in an intrinsically 4D system is necessarily a 3D closed manifold. Figure 4(a) gives an illustration of the 3D Fermi hypersurface in the surface momentum space, where we have omitted one dimension of the 4D momentum space, so that the solid circle stands for the projected 2D WS, while the surface connected to it represents the 3D Fermi hypersurface. It is important to note that, in the present case, because $(C_2)_n^{\partial V} = -(C_2)_{n+1}^{\partial V} = 0$, one has $(C_1)_n = (C_1)_{n+1} = 0$ [recall (13)], which is the Berry curvature integrated on the 2D WS. Accordingly, any 2D cross section on the 3D Fermi hypersurface which can continuously deform into (i.e., homotopic to) the 2D WS, e.g., the dashed circle in Fig. 4(a), will have a first Chern number $C_1^r = 0$, which is defined as the integration of Berry curvature of the 3D Fermi hypersurface states on the 2D cross section. Therefore, such a 2D cross section can shrink to zero, which

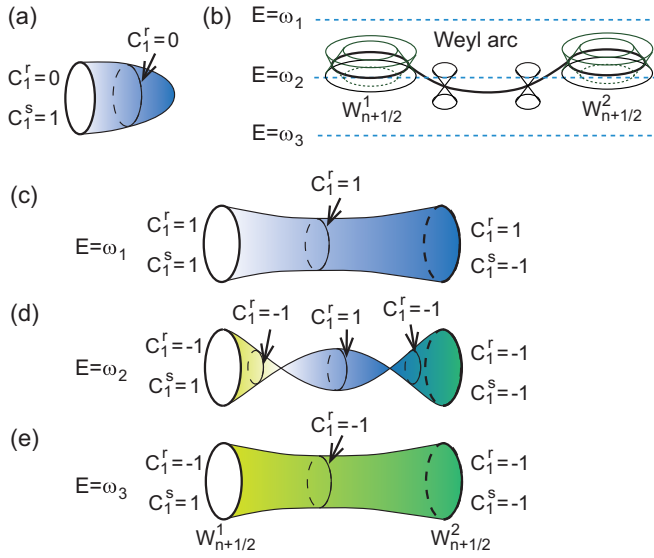


FIG. 4. Illustration of the topological surface states on the 4D surface of a 5D Weyl semimetal. (a) At a given fermi energy (near the WS), a WS $W_{n+1/2}^1$ with second Chern number $(C_2)_{n+1/2}^{\partial V} = 0$ (which implies $C_1^s = 1$ and $C_1^r = 0$) protects a 3D Fermi hypersurface in the 4D surface momentum space, whose 2D boundary is the surface projection of the WS (the solid circle). (b) Illustration of the energy dispersion of the surface states topologically protected by two WSs $W_{n+1/2}^1$ and $W_{n+1/2}^2$ with second Chern numbers $(C_2)_n^{(1)} = -1$ and $(C_2)_n^{(2)} = 1$, respectively. A Weyl arc arises in the surface states connecting the projection of the two WSs. (c)–(e) The topological surface states (3D Fermi hypersurfaces) at Fermi energies $E = \omega_1$, ω_2 , and ω_3 defined in panel (b), whose boundaries (2D) are the surface projection of the two WSs (the left and right solid circles). The regular and singular first Chern numbers of the WSs are labeled by the side of each WS. At energy $E = \omega_2$, the constant energy plane intersects with the Weyl arc in panel (b), leading to two Weyl points on the Fermi hypersurface in panel (d).

means the 3D Fermi hypersurface can close by itself away from the WS, forming a half 3D sphere topologically as shown in Fig. 4(a).

Then, we turn to WS $W_{n+1/2}^i$ with nonzero second Chern number $(C_2)_n^{\partial V} = -(C_2)_{n+1}^{\partial V} = 1$ [according to the topological constraints we have proved, one always has $(C_2)_n^{\partial V} = -(C_2)_{n+1}^{\partial V}$], and show that nonzero second Chern numbers imply the existence of 1D *Weyl arcs* [32] on top of the 3D Fermi hypersurfaces.

Consider two WSs $W_{n+1/2}^1$ and $W_{n+1/2}^2$ as shown in Fig. 4(b), which have second Chern numbers $(C_2)_n^{(1)} = -(C_2)_{n+1}^{(1)} = -1$ and $(C_2)_n^{(2)} = -(C_2)_{n+1}^{(2)} = 1$, respectively. For simplicity, we assume they are the only WSs between the n th band and the $(n+1)$ th band (as is allowed by the 5D doubling constraints). If we set the Fermi energy $E = \omega_1$ to be above the two WSs in the $(n+1)$ th band [Fig. 4(b)], we will expect a 3D Fermi hypersurface on the 4D surface of the 5D semimetal as shown in Fig. 4(c) (the left and right solid circles represent 2D WSs $W_{n+1/2}^1$ and $W_{n+1/2}^2$ projected on the 4D surface), which we shall explain below. Since the WS $W_{n+1/2}^1$ has a singular first Chern number $(C_1^s)_{n+1/2}^{(1)} = 1$, one would expect the projected $W_{n+1/2}^1$ to be a boundary

of a 3D Fermi hypersurface. However, when viewed in the $(n+1)$ th band, $(C_2)_{n+1}^{(1)} = 1$ implies the WS also has a regular first Chern number $(C_1^r)_{n+1}^{(1)} = 1$. Therefore, the surface states on a 2D cross section on the 3D Fermi hypersurface (the dashed circle in the middle) which is homotopic to $W_{n+1/2}^1$ will have the same regular first Chern number $C_1^r = 1$. Since C_1^r is nonzero, the 2D cross section cannot contract to zero by itself when moved continuously on the 3D Fermi hypersurface; instead it can only be continuously moved to another boundary of the Fermi hypersurface with regular first Chern number $C_1^r = 1$, which in this case has to be the surface projection of the other WS $W_{n+1/2}^2$ [which under proper orientation choice has $(C_1^s)_{n+1}^{(2)} = -1$ and $(C_1^r)_{n+1}^{(2)} = 1$]. Therefore, the 3D Fermi hypersurface has to connect $W_{n+1/2}^1$ and $W_{n+1/2}^2$ as shown in Fig. 4(c) due to nonzero second Chern number, which is clearly different from the case in Fig. 4(a).

Now assume we lower the Fermi energy to $E = \omega_2$, which is slightly below the WSs and enters the n th band. Upon entering into the n th band, the second Chern number of a WS changes sign [relative to that viewed in the $(n+1)$ th band], so the regular first Chern number of $W_{n+1/2}^1$ flips sign to $(C_1^r)_n^{(1)} = -1$ [while $(C_1^s)_n^{(1)} = 1$ remains unchanged] as shown in Fig. 4(d), and similarly for $W_{n+1/2}^2$. Accordingly, C_1^r of any 2D cross section on the 3D Fermi hypersurface that is homotopic to $W_{n+1/2}^1$ has to change sign, too. It may happen that certain 2D cross sections in the middle of the 3D Fermi hypersurface [the dashed circle in the middle in Fig. 4(d)] still has $C_1^r = +1$ (unflipped). Then the cross sections with opposite $C_1^r = +1$ on the 3D Fermi hypersurface has to be connected via a Weyl point on the 3D Fermi hypersurface, as shown in Fig. 4(d). Namely, two Weyl points will emerge from the two WSs and move towards each other on the 3D Fermi hypersurface as the Fermi energy is lowered. If one further lowers the Fermi energy to $E = \omega_3$, the two Weyl points will annihilate with each other in the middle of the 3D Fermi hypersurface, and the C_1^r of all 2D cross sections will be flipped, as shown in Fig. 4(e). Therefore, because of the fact that the nonzero second Chern number $(C_2)_{n+1}^{(i)} = -(C_2)_n^{(i)}$ flips sign from the n th band to the $(n+1)$ th band, the 3D Fermi hypersurface necessarily experiences the arising (at two WSs) and annihilation of two Weyl points when the Fermi energy is changed from the $(n+1)$ th band to the n th band.

If one plots the energy dispersion of the topological surface states on the 4D surface, one would expect to see the two Weyl points as a function of energy forming a 1D Weyl arc connecting the surface projection of the two WSs $W_{n+1/2}^1$ and $W_{n+1/2}^2$, as shown in Fig. 4(b). This is exactly the Weyl arc protected by the nonzero second Chern number of WSs, which is shown in earlier studies [32,33]. Besides, in the above we further show the 3D Fermi hypersurface is also nontrivial when the WSs have nonzero second Chern numbers, which have to connect two WSs as shown in Figs. 4(c) and 4(e).

We note that when the system has TP symmetry, the system will contain Yang monopoles instead of WSs, which are protected by non-Abelian second Chern numbers as we discussed in last section. In this case, one will have Weyl

arcs connecting Yang monopoles with opposite non-Abelian second Chern numbers [33].

VI. CONCLUSION

In this paper, we studied the 2D WS degeneracies in generic 5D Weyl semimetals. In particular, we showed that their topological configuration—most notably their linking configuration—must satisfy nontrivial constraints Eqs. (14)–(16), which are 5D analogs to the famous doubling theorem in 3D Weyl semimetal (but also with nontrivial distinctions). Furthermore, the relation between WS linking and Yang monopole in five dimensions is established by showing a general relation Eq. (26) between the Abelian and non-Abelian Chern numbers in topological band theory. More interestingly, when the 5D Weyl semimetal has a 4D surface, very rich topological surface states arise, including the 3D *Fermi hypersurface* (parallel to the 1D Fermi arc in 3D Weyl semimetal) protected by the WSs in the bulk, and, on top of that, the 1D *Weyl arc* protected by the linking of WSs in the bulk. The topological protection of the surface physics is closely related to the topological constraints in the bulk that we established. Brillouin zones with synthetic dimensions have been realized in cold atom systems [41], so the rich surface-state physics we derived maybe observed in such experiments.

We would like to make some final comments on the general mathematical framework behind the topology of band degeneracies. One early and deep result in this area is obtained for the stability of generalized Fermi surfaces (which include Weyl degeneracies) using K theory [42] (other studies of the stability of gapless fermionic ground states based on Green's functions include, e.g., [43,44]). More recently, the machinery of K theory and homotopy theory has been extensively applied to understand the topology of band degeneracies. Most particularly, as we have mentioned in Sec. IV, a homotopy study of $(\text{TP})^2 = +1$ 3D nodal line systems [37,38] has led to constraints analogous to ours in 5D Weyl semimetal. The connection between our simple Berry curvature method and the full machinery of homotopy theory might not be surprising—while the celebrated doubling theorem is usually presented using a Berry curvature computation (as we did), originally the theorem was established using homotopy [35]. Aside from homotopy theory, we also note that recently a standard construction in homology theory (in which computations are much simpler compared to homotopy theory), the Mayer-Vietoris sequence, has been applied to understand Weyl degeneracies [45,46]. At this point, this framework seems not to encompass some of the interesting features we discussed. In particular, it is unclear whether our constraints on WS linking can be detected with this method. Also, this method seems to be insensitive to the single component WS [Fig. 2(a)] which leads to a topologically protected 3D Fermi hypersurface; see Fig. 4(a). (Such a single WS can be realized by a simple model: take a 3D Weyl semimetal model with momentum k_μ , $\mu = 1, 2, 3$, and then make the Weyl node separation depend on k_4, k_5 so that they annihilate for large k_4, k_5 .) It would be interesting to understand what refinement of the Mayer-Vietoris approach is needed to capture the full topological information of the bulk degeneracies and surface states.

ACKNOWLEDGMENTS

J.-Y.C. thanks T. Bzdušek and J. Ahn for explaining their closely related results in 3D nodal line systems [37,38]. J.-Y.C. was supported by the Gordon and Betty Moore Foundation's EPiQS Initiative through Grant No. GBMF4302. B.L. was supported by the Princeton Center for Theoretical Science at Princeton University. S.-C.Z. was supported by the US Department of Energy, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515.

APPENDIX A

In this Appendix we explain in detail what is meant by “effectively two-band” in the proof of the constraint (15). Let's consider the n th and $(n+1)$ th band which comprise a two-dimensional subspace of the full Hilbert space. Recall in the diagonalization of the Hamiltonian $H = UEU^\dagger$ that the $N \times N$ unitary matrix U has its m th column being the eigenvector u_m . Now let $U_{\{n,n+1\}}$ be the $N \times 2$ matrix whose two columns are u_n and u_{n+1} . Let R be an arbitrary $U(2)$ matrix rotating in this subspace, $\tilde{U}_{\{n,n+1\}} = U_{\{n,n+1\}}R^\dagger$, $U_{\{n,n+1\}} = \tilde{U}_{\{n,n+1\}}R$. We let the $U(2)$ matrix R depend on the momentum k such that its two columns have the same singularity as $W_{n+1/2}^i$ in the original $U_{\{n,n+1\}}$, and thus $\tilde{U}_{\{n,n+1\}}$ does not have singularity on $W_{n+1/2}^i$ anymore. This way, $\tilde{U}_{\{n,n+1\}}$ carries the topological information about the separation of this two-dimensional subspace from the remaining of the Hilbert space, while R , the “effectively two-band” part, carries the topological information about the separation within the two-dimensional subspace.

Let r_n and r_{n+1} be the two columns of R . Then $u_n = \tilde{U}_{\{n,n+1\}} r_n$, $u_{n+1} = \tilde{U}_{\{n,n+1\}} r_{n+1}$. The Abelian Berry connection for u_n can be written as

$$A_n = -iu_n^\dagger du_n = r_n^\dagger (-id + \tilde{A}_{\{n,n+1\}}) r_n, \quad (\text{A1})$$

where $\tilde{A}_{\{n,n+1\}} = -iU_{\{n,n+1\}}^\dagger dU_{\{n,n+1\}}$ is the $U(2)$ Berry connection on the two-band sub-bundle. Let's denote the above as $A_n = A_n^1 + A_n^2$. The first term $A_n^1 = -ir_n^\dagger dr_n$ is an “effectively two-band” $U(1)$ Berry connection. We want to show that the second term $A_n^2 = r_n^\dagger \tilde{A}_{\{n,n+1\}} r_n$ only has a contribution to $(C_2)_n$ when $W_{n+1/2}^i$ is linked with $W_{n-1/2}$, but has no contribution regarding whether $W_{n+1/2}^i$ is linked with any component of $W_{n+1/2}$.

Consider the following geometric setup. Let $W_{n+1/2}^i = \partial \Sigma_{n+1/2}^i$ and let \mathcal{V} be the 5D vicinity of $\Sigma_{n+1/2}$. We know A_n is singular on 2D $W_{n\pm 1/2}$. Let $\mathcal{V}^* \equiv \mathcal{V} \setminus W_{n\pm 1/2}$ on which A_n is defined. A_n is not continuous over \mathcal{V}^* . Suppose \mathcal{V}^* is covered by a set of charts which are 5D. Across the boundaries between the charts (the boundaries are 4D), A_n in our expression (A1) is subjected to transition function, i.e., gauge transformation, of two kinds: the $U(1)$ gauge transformation

$$r_n \rightarrow r_n e^{i\lambda}, \quad (\text{A2})$$

characterizing how the n th band separates from the $(n+1)$ th band, and the $U(2)$ gauge transformation

$$U_{\{n,n+1\}} \rightarrow U_{\{n,n+1\}} \Lambda, \quad (\text{A3})$$

characterizing how these two bands as a whole separate from the other bands. The discontinuity from the $U(1)$ gauge

transformation extends from the singularity $W_{n+1/2}$ and is seen by A_n^1 , while the discontinuity from the $U(2)$ gauge transformation extends from the singularity $W_{n-1/2}$ and is seen by A_n^2 (note that $\tilde{A}_{(n,n+1)}$ sees the singularity $W_{n-1/2} \cup W_{n+3/2}$, but when projected by r_n , A_n^2 only sees the singularity $W_{n-1/2}$).

Thus, when considering the linking between $W_{n+1/2}^i$ with other components of $W_{n+1/2}$, we can limit ourselves to the “effectively two-band” part $A_n^1 = -ir_n^\dagger dr_n$ and ignore A_n^2 . Then, as mentioned in the main text, $(dA_n^1)^2 = (dr_n^\dagger dr_n)^2 = 0$ because r_n and r_n^\dagger together only have three real parameters (fixing an unimportant overall phase).

APPENDIX B

In this Appendix we show (26). Although this is a special case of the more general Whitney sum formula, the proof to the formula is nontrivial, therefore we shall show this special case via direct computation. It suffices to show the following.

$$\begin{aligned} \text{tr}(F_A^2) &= \text{tr}[dU_A dU_A^\dagger (\mathbf{1} - U_A U_A^\dagger) dU_A dU_A^\dagger (\mathbf{1} - U_A U_A^\dagger)] \\ &= \text{tr}[(dU_B dU_B^\dagger + dU_C dU_C^\dagger)(\mathbf{1} - U_B U_B^\dagger - U_C U_C^\dagger)(dU_B dU_B^\dagger + dU_C dU_C^\dagger)(\mathbf{1} - U_B U_B^\dagger - U_C U_C^\dagger)] \\ &= \text{tr}(F_A^2) + \text{tr}(F_C^2) - 2 \text{tr}(dU_B^\dagger dU_C dU_C^\dagger dU_B) - 2 \text{tr}(dU_C^\dagger dU_B U_B^\dagger dU_B U_B^\dagger dU_C) - 2 \text{tr}(dU_B^\dagger dU_C U_C^\dagger dU_C U_C^\dagger dU_B) \\ &\quad + 2 \text{tr}(U_C^\dagger dU_B dU_B^\dagger dU_B U_B^\dagger dU_C) + 2 \text{tr}(U_B^\dagger dU_C dU_C^\dagger dU_C U_C^\dagger dU_B) - 2 \text{tr}(U_C^\dagger dU_B U_B^\dagger dU_B dU_B^\dagger dU_C) \\ &\quad - 2 \text{tr}(U_B^\dagger dU_C U_C^\dagger dU_C dU_C^\dagger dU_B) \\ &= \text{tr}(F_A^2) + \text{tr}(F_C^2) - 2 dK_{B,C}, \end{aligned} \quad (\text{B3})$$

$$K_{B,C} = \text{tr}(U_B^\dagger dU_C dU_C^\dagger dU_B) + \text{tr}(U_C^\dagger dU_B U_B^\dagger dU_B U_B^\dagger dU_C) + \text{tr}(U_B^\dagger dU_C U_C^\dagger dU_C U_C^\dagger dU_B), \quad (\text{B4})$$

where from the second line to the third line we expanded the terms and many of them canceled because of $dU_B^\dagger U_B = -U_B^\dagger dU_B$, $dU_C^\dagger U_C = -U_C^\dagger dU_C$ (and likewise for $B \leftrightarrow C$).

It remains to show that if over a submanifold $\mathcal{M} \subset BZ$ the separation of A into B and C is well defined, then $K_{B,C}$ is smooth over \mathcal{M} . Let \mathcal{M} be covered by a set of charts. On each chart U_B , U_C , and hence $K_{B,C}$ are smooth. But discontinuity might arise at the boundary between two charts. Generally, across each chart boundary, U_B and U_C are subjected to a transition function, i.e., a gauge transformation in the gauge

Let $A \subset \{1, \dots, N\}$ be some subset of bands, which further separate into two disjoint subsets $B \cap C = \emptyset$, $B \cup C = A$. Let U_A be the $N \times |A|$ matrix whose columns are u_m , $m \in A$, and its associated non-Abelian Berry curvature is

$$F_A = -i dU_A^\dagger (\mathbf{1} - U_A U_A^\dagger) dU_A, \quad (\text{B1})$$

and likewise for B and C . We want to show

$$\text{tr}(F_A^2) = \text{tr}(F_B^2) + \text{tr}(F_C^2) - 2 dK_{B,C}, \quad (\text{B2})$$

where $K_{B,C}$ is a smooth three-form except on values of momenta k where the separation of A into B and C becomes not well defined (i.e., Weyl degeneracies between B and C). Choosing $A = \{1, \dots, n\}$, $B = \{n\}$, $C = \{1, \dots, n-1\}$ leads to (26).

In summing over the m indices in $(U_A)^\alpha_m$ for $m \in A$, we separate the summation into summations for $m \in B$ and $m \in C$. Thus,

group $U(|B|) \times U(|C|)$:

$$U_B \rightarrow U_B \Lambda_B, \quad U_C \rightarrow U_C \Lambda_C. \quad (\text{B5})$$

But it is straightforward to verify that $K_{B,C}$ is invariant under such gauge transformation (using $U_B^\dagger U_C = 0$). Hence $K_{B,C}$ is a smooth three-form over \mathcal{M} . This completes the proof to the claim (B2).

Such separation of topological characterization when a vector bundle can be unambiguously separated into two sub-bundles is the theme behind the general Whitney sum formula.

- [1] L. Fu, C. L. Kane, and E. J. Mele, *Phys. Rev. Lett.* **98**, 106803 (2007).
- [2] X.-L. Qi, T. L. Hughes, and S.-C. Zhang, *Phys. Rev. B* **78**, 195424 (2008).
- [3] X.-L. Qi and S.-C. Zhang, *Rev. Mod. Phys.* **83**, 1057 (2011).
- [4] M. Z. Hasan and C. L. Kane, *Rev. Mod. Phys.* **82**, 3045 (2010).
- [5] A. Kitaev, in *Advances in Theoretical Physics: Landau Memorial Conference*, edited by V. Lebedev and M. Feigel'man, AIP Conf. Proc. No. 1134 (AIP, Melville, NY, 2009), p. 22.
- [6] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig, *New J. Phys.* **12**, 065010.
- [7] L. Balents, *Physics* **4**, 36 (2011).
- [8] X. Wan, A. M. Turner, A. Vishwanath, and S. Y. Savrasov, *Phys. Rev. B* **83**, 205101 (2011).

- [9] S. Murakami, *New J. Phys.* **9**, 356 (2007).
- [10] S.-M. Huang, S.-Y. Xu, I. Belopolski, C.-C. Lee, G. Chang, B. Wang, N. Alidoust, G. Bian, M. Neupane, C. Zhang, S. Jia, H. Bansil, A. Lin, and M. Z. Hasan, *Nat. Commun.* **6**, 7373 (2015).
- [11] H. Weng, C. Fang, Z. Fang, B. A. Bernevig, and X. Dai, *Phys. Rev. X* **5**, 011029 (2015).
- [12] J. Liu and D. Vanderbilt, *Phys. Rev. B* **90**, 155316 (2014).
- [13] A. A. Zyuzin and A. A. Burkov, *Phys. Rev. B* **86**, 115133 (2012).
- [14] C.-X. Liu, P. Ye, and X.-L. Qi, *Phys. Rev. B* **87**, 235306 (2013).
- [15] D. T. Son and B. Z. Spivak, *Phys. Rev. B* **88**, 104412 (2013).
- [16] S.-Y. Xu, C. Liu, S. K. Kushwaha, R. Sankar, J. W. Krizan, I. Belopolski, M. Neupane, G. Bian, N. Alidoust, T.-R. Chang,

- H.-T. Jeng, C.-Y. Huang, W.-F. Tsai, H. Lin, P. P. Shibayev, F.-C. Chou, R. J. Cava, and M. Z. Hasan, *Science* **347**, 294 (2015).
- [17] S.-Y. Xu, I. Belopolski, N. Alidoust, M. Neupane, G. Bian, C. Zhang, R. Sankar, G. Chang, Z. Yuan, C.-C. Lee, S.-M. Huang, H. Zheng, J. Ma, D. S. Sanchez, B. Wang, A. Bansil, F. Chou, P. P. Shibayev, H. Lin, S. Jia, and M. Z. Hasan, *Science* **349**, 613 (2015).
- [18] M. V. Berry, *Proc. R. Soc. London, Ser. A* **392**, 45 (1984).
- [19] Y. X. Zhao and Z. D. Wang, *Phys. Rev. Lett.* **110**, 240404 (2013).
- [20] A. P. Schnyder and P. M. R. Brydon, *J. Phys.: Condens. Matter* **27**, 243201 (2015).
- [21] S. Matsuura, P.-Y. Chang, A. P. Schnyder, and S. Ryu, *New J. Phys.* **15**, 065001 (2013).
- [22] S. Yip, *Annu. Rev. Condens. Matter Phys.* **5**, 15 (2014).
- [23] C.-K. Chiu and A. P. Schnyder, *Phys. Rev. B* **90**, 205136 (2014).
- [24] X.-Q. Sun, B. Lian, and S.-C. Zhang, *Phys. Rev. Lett.* **119**, 147001 (2017).
- [25] W. Chen, H.-Z. Lu, and J.-M. Hou, *Phys. Rev. B* **96**, 041102(R) (2017).
- [26] M. Ezawa, *Phys. Rev. B* **96**, 041202(R) (2017).
- [27] Z. Yan, R. Bi, H. Shen, L. Lu, S.-C. Zhang, and Z. Wang, *Phys. Rev. B* **96**, 041103(R) (2017).
- [28] R. Bi, Z. Yan, L. Lu, and Z. Wang, *Phys. Rev. B* **96**, 201305(R) (2017).
- [29] G. Chang, S.-Y. Xu, X. Zhou, S.-M. Huang, B. Singh, B. Wang, I. Belopolski, J. Yin, S. Zhang, A. Bansil *et al.*, *Phys. Rev. Lett.* **119**, 156401 (2017).
- [30] B. Lian, C. Vafa, F. Vafa, and S.-C. Zhang, *Phys. Rev. B* **95**, 094512 (2017).
- [31] S.-C. Zhang and J. Hu, *Science* **294**, 823 (2001).
- [32] B. Lian and S.-C. Zhang, *Phys. Rev. B* **94**, 041105(R) (2016).
- [33] B. Lian and S.-C. Zhang, *Phys. Rev. B* **95**, 235106 (2017).
- [34] C. N. Yang, *J. Math. Phys.* **19** (1978).
- [35] H. B. Nielsen and M. Ninomiya, *Nucl. Phys. B* **185**, 20 (1981); **185**, 533 (1980).
- [36] H. B. Nielsen and M. Ninomiya, *Nucl. Phys. B* **193**, 173 (1981).
- [37] J. Ahn, D. Kim, Y. Kim, and B.-J. Yang, *Phys. Rev. Lett.* **121**, 106403 (2018).
- [38] Q. Wu, A. A. Soluyanov, and T. Bzdušek, *arXiv:1808.07469*.
- [39] A. M. Polyakov, *Mod. Phys. Lett. A* **3**, 325 (1988).
- [40] H. C. Tze, *Int. J. Mod. Phys. A* **3**, 1959 (1988).
- [41] A. Celi, P. Massignan, J. Ruseckas, N. Goldman, I. B. Spielman, G. Juzeliunas, and M. Lewenstein, *Phys. Rev. Lett.* **112**, 043001 (2014).
- [42] P. Horava, *Phys. Rev. Lett.* **95**, 016405 (2005).
- [43] G. E. Volovik, *Lect. Notes Phys.* **870**, 343 (2013).
- [44] M. A. Zubkov and G. E. Volovik, *Nucl. Phys. B* **860**, 295 (2012).
- [45] V. Mathai and G. C. Thiang, *J. Phys. A* **50**, 11LT01 (2017).
- [46] V. Mathai and G. C. Thiang, *Commun. Math. Phys.* **355**, 561 (2017).