ARTICLE   OPEN

Statistical learning of engineered topological phases in the kagome superlattice of AV_3Sb_5

Thomas Mertz[1]✉, Paul Wunderlich [1]✉, Shinibali Bhattacharyya [1], Francesco Ferrari [1] and Roser Valentí [1]✉

Recent experimental findings have reported the presence of unconventional charge orders in the enlarged (2×2) unit-cell of kagome metals AV_3Sb_5 (A = K, Rb, Cs) and hinted towards specific topological signatures. Motivated by these discoveries, we investigate the types of topological phases that can be realized in such kagome superlattices. In this context, we employ a recently introduced statistical method capable of constructing topological models for any generic lattice. By analyzing large data sets generated from symmetry-guided distributions of randomized tight-binding parameters, and labeled with the corresponding topological index, we extract physically meaningful information. We illustrate the possible real-space manifestations of charge and bond modulations and associated flux patterns for different topological classes, and discuss their relation to present theoretical predictions and experimental signatures for the AV_3Sb_5 family. Simultaneously, we predict higher-order topological phases that may be realized by appropriately manipulating the currently known systems.

npj Computational Materials (2022)8:66 ; https://doi.org/10.1038/s41524-022-00745-3

INTRODUCTION

The recent surge of interest in kagome materials, often discussed in the context of frustrated magnetism and spin liquid phases[1–7], has been boosted by the discovery of the kagome metals AV_3Sb_5 (A = K, Rb, Cs) undergoing successive charge density wave (CDW) and superconducting transitions upon lowering temperature[8–11]. The presence of flat bands, Dirac points, and van-Hove singularities in the electronic band structure of the ideal kagome lattice provides a playground for exotic topological properties and a variety of phases, ranging from superconductivity to charge, orbital momentum, and spin density waves[12–18]. Density functional theory calculations for AV_3Sb_5 have categorized the normal-state of this family as a $\mathbb{Z}_2$ topological metal with multiple protected Dirac crossings[10] and renormalization group analyses have proposed the occurrence of various complex CDW and charge bond order (CBO) phases[16,17,19,20]. Interestingly, reports of giant extrinsic anomalous Hall effect suggest nontrivial band topology in the absence of long range magnetic order[21], possibly driven by a CDW order with orbital currents, and high-resolution STM (scanning tunneling microscopy) measurements point to an unconventional intrinsic chiral charge[22–25] consistent with a doubling of the unit-cell $(2 \times 2$ superlattice)[26]. These observations imply the relevance of the ubiquitous chiral charge order present in the Haldane model[27], and the possibility of higher-order topological insulators, an avenue that demands further exploration. Although a thorough understanding of the various electronic orders warrants detailed microscopic investigations[16,17,19,28], we use the available plethora of experimental observations as our motivation to learn about the possible topological phases that can manifest within the electronic parameter space of the 2×2 kagome superlattice.

Another rapidly developing field of research is the application of machine learning to tackle physical problems[29], from variational representation of wave functions[30], to the detection of phase transitions[31,32]. Due to the absence of a local order parameter, topological phase transitions are generally more difficult to capture than symmetry-breaking phase transitions, although some progress has been achieved[33–35]. Additionally, an immediate physical interpretation of the results can turn out to be a complicated task in unbiased machine learning approaches. Yet, in a recent study[36] we proposed a statistical learning of topological models on a honeycomb lattice and showed that machine-assisted unbiased learning can differentiate between the electronic parameters that are most significant for the manifestation of the well-known topological Haldane phase in the underlying lattice structure. Making use of a generalization of this method, in this work we extract topological information for the generic 2×2 kagome superlattice, as "learned" from the statistics of data-sets of randomized tight-binding parameters, constrained only by specific crystal symmetries. The variations of the tight-binding parameters can be interpreted as modified hoppings arising from changes in lattice parameters, atomic mass, effects of strain, pressure, spin-orbit coupling, etc. We find topologically trivial Star-of-David (SoD)-like CBO phases and non-trivial chiral flux phases. Our results are compatible with present theoretical predictions and experimental observations for the AV_3Sb_5 family. Additionally, we predict higher-order topological phases which might be realized in future experimental work through appropriate manipulation of the kagome lattice.

RESULTS

Model

We consider a generic tight-binding Hamiltonian on the kagome lattice which can be taken, in a first approximation, as a minimal model to describe the low-energy electronic properties of the vanadium 3d bands in AV_3Sb_5[8,37]

$$H = \sum_i \epsilon_i c_i^\dagger c_i + \sum_{\langle i,j \rangle} t_{i,j} c_i^\dagger c_j \tag{1}$$

Here $\langle i, j \rangle$ runs over nearest-neighbor sites and c_i ($c_i^\dagger$) annihilates (creates) an electron at site i. ϵ_i denote onsite potentials, while $t_{i,j}$

[1]Institut für Theoretische Physik, Goethe-Universität, Frankfurt am Main, Germany. ✉email: mertz@itp.uni-frankfurt.de; wunderlich@itp.uni-frankfurt.de; valenti@itp.uni-frankfurt.de
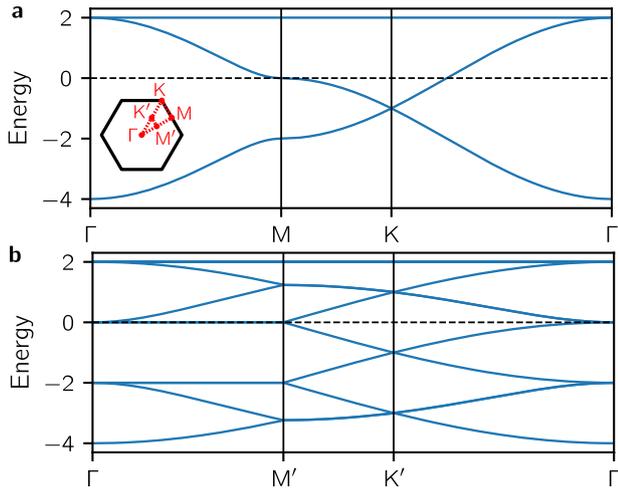
a

b

Fig. 1 Band structure. Metallic band structure along the high-symmetry path **a**. Γ-M-K-Γ for the kagome lattice, and **b**. Γ-M'-K'-Γ for the 2×2 kagome superlattice, for the tight-binding model with uniform nearest-neighbor hopping ($t_{i,j} = -1$) and $\epsilon_i = 0$. The dashed horizontal line indicates the Fermi energy at 5/12 filling. The inset of **a** shows the hexagonal Brillouin zone of the kagome lattice (in black) together with the location of the high-symmetry points and the chosen path (in red).
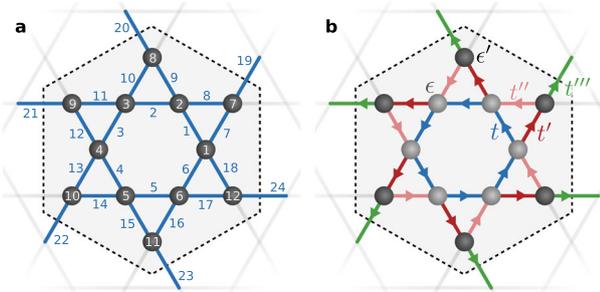
a   **b**

Fig. 2 2×2 kagome supercell. a Symmetric 2×2 kagome supercell. The unit cell, delimited by black dashed lines, contains 12 distinct sites. The 24 independent nearest-neighbor links are colored in blue. Sites and links are labeled counter-clockwise. **b** Schematic representation of the hoppings and onsite potentials in the 2×2 unit cell with C_6 symmetry. Spheres of different colors represent different onsite terms. Symmetries dictate equality of hoppings, as indicated by the colors of the bonds. The direction of the arrows signifies hopping from site j to site i (i.e., $c_i^\dagger c_j$), and has been chosen arbitrarily, however, adhering to the constraints imposed by the C_6 symmetry.

are hopping integrals between sites i and j. In the simple case of uniform hopping, i.e., $t_{i,j} = -1$, and zero onsite potentials, $\epsilon_i = 0$, the band structure of the system, shown in Fig. 1a, is characterized by a flat band at high energy and two lower-lying dispersive bands that touch each other in a Dirac point at the corners of the BZ (K points), and exhibit van Hove singularities at the M points.

Several works on AV3Sb5[16,17,19] have suggested that CDW instabilities at the van Hove fillings may cause a translational symmetry breaking of the perfect kagome lattice, leading to a lower periodicity described by a 2×2 supercell, analogous to that observed in STM experiments[22–25]. For this reason, we focus at a filling 5/12 with the Fermi energy lying at the higher van Hove singularity (see Fig. 1a). We assume our Hamiltonian to be periodic over the 2×2 enlarged unit cell illustrated in Fig. 2a which contains 12 sites (corresponding to the onsite terms $\epsilon_{i=1,\dots,12}$) arranged in a SoD pattern, and retains all the symmetries of the

point group of the kagome lattice (D_6). As a consequence of the superlattice periodicity, we are left with 24 independent nearest-neighbor hoppings, which we label as $t_{s=1,\dots,24}$ as depicted by the blue-colored links in Fig. 2a. The hopping parameters can be categorized in three distinct classes: the hoppings in the inner hexagon ($t_{s\leq6}$), the hoppings on the spikes of the SoD ($t_{7\leq s\leq18}$), and the hoppings connecting sites belonging to different unit cells ($t_{19\leq s\leq24}$). The various hoppings of each class can be mapped into each other by point group symmetry operations. In the uniform case ($t_s = -1, \forall s$) with zero onsite potential, one obtains 12 bands as shown in Fig. 1b for the BZ corresponding to the 2×2 superlattice. These bands can be unfolded to the 3 bands corresponding to the elementary 1×1 unit cell. By tuning the different hopping parameters, it is possible to open a topologically non-trivial gap at 5/12 filling We remark that the tight-binding model of Eq. (1) is not strictly bound to the description of the vanadium $3d$ bands of AV3Sb5 compounds, but retains a certain level of generality and could be applicable to other kagome systems at the van Hove filling. To investigate the possible topological phases of the 2×2 kagome superlattice, we employ the statistical approach described in the Methods section.

Statistical analysis

A completely unbiased analysis of the full nearest-neighbor 2×2 kagome superlattice involves sampling of 11 onsite and 24 hopping parameters (one onsite term is kept fixed to set the global energy scale). To improve tractability, we scale down the number of independent features by enforcing specific symmetry operations on the feature space, for instance, the point group C_6, which is a subgroup of the kagome point group D_6, lacking reflection symmetries. This specific choice is necessary in order to construct non-trivial tight-binding models, since the Chern number—which we set as our topological index—is odd under the effect of reflections, and thus a reflection-invariant Hamiltonian could only have $C = 0$.

Under C_6 symmetry, the onsite terms and hoppings are reduced to a set of six unique features as illustrated in Fig. 2b. These are real $\epsilon_{i\leq6} = \epsilon$ and $\epsilon_{i\geq7} = \epsilon'$, and complex $t_{s\leq6} = t$, $t_s = t'$ for $s \in \{7, 9, 11, 13, 15, 17\}$, $t_s = t''$ for $s \in \{8, 10, 12, 14, 16, 18\}$, and $t_{s\geq19} = p'''$. We choose $\mathbf{x}^{\text{ref}} = (\epsilon, \epsilon', t, t', t'', t''') = (0, 1, -1, -1, -1, -1)$ as reference point. Our choice for the sampling radius ensures that $\epsilon = 0$ for all samples. We generate a data set of $n_s = 2 \times 10^6$ samples, 67% (33%) of which are insulators (metals) (The classification of the samples in metals and insulators is performed numerically by computing the energy bands on a grid of 82×82 **k**-points, and checking whether the indirect gap at 5/12 filling is smaller or larger than an energy threshold (chosen to be $0.01|t_{\text{ref}}|$ here).). As shown by the pie chart in Fig. 3, 69.6% of the insulating samples are topologically non-trivial. Among the topological insulators, the largest fraction (60.4%) has $C = \pm1$, while the second largest fraction (9%) has $C = \pm2$. In the following, we discuss the properties of these topological insulators in detail.

In the analysis of the hopping parameters we will focus on the marginal probability distribution functions (PDFs) (see Eq. (2) of Methods) of the onsite energy $p_C(\text{Re}[\epsilon'])$ (Fig. 3a), imaginary parts of the hoppings $p_C(\text{Im}[t_s])$ [Fig. 3b–d], which determine the hopping direction, and PDFs of their moduli $p_C(|t_s|)$ [Fig. 3e–g], which describe the overall hopping strength. These features are those which provide most of the information about the topological character of the samples. Due to the inherent symmetry of the kagome lattice, the PDFs for t' and t'' show the same behavior. Hence, the distribution for only one of these hoppings, i.e., t', is shown. All PDFs are provided in Supplementary Figure 1. First, we analyze the PDF of the onsite term $\epsilon' = \epsilon_{i\geq7}$ (Fig. 3a) that distinguishes between trivial (grey line) and topological phases (colored lines). For the trivial phase, ϵ' tends to be larger than zero, i.e., the outer ring of the spikes tend to be
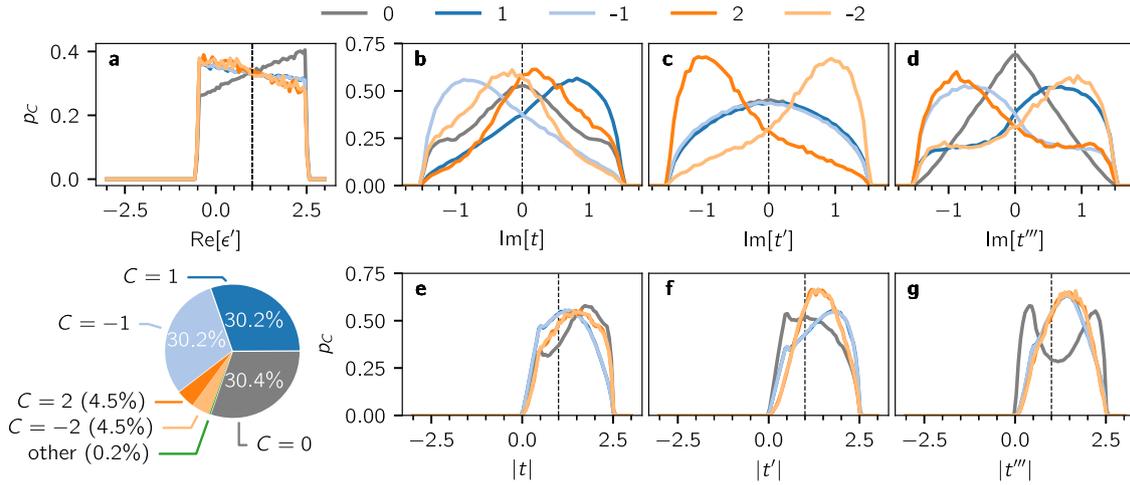
Fig. 3 Chern numbers and marginal probability distributions. Distribution of Chern numbers (pie chart) and **a**–**g** marginal PDFs for the most descriptive features of insulators within the 2×2 kagome superlattice with C_6 symmetry. The pie chart shows the percentages of trivial ($C = 0$) and topological ($C \neq 0$) samples obtained out of all insulating samples. For the PDF plots corresponding to different feature components, grey lines denote trivial phases ($C = 0$) and colored lines denote non-trivial ($C \neq 0$) phases as indicated by the legend on top. Dashed vertical lines indicate the reference point.

"heavier" compared to the inner hexagon. By contrast, in the topological phase ϵ' tends to have smaller values. This behavior is well known from the Haldane model[38] and reflects the fact that large $|\epsilon'|$ eventually turns the system into a trivial insulator.

Trivial phases (C = 0)

In the trivial phase ($C = 0$) we observe that the PDFs $p_0(\mathrm{Im}[t])$, $p_0(\mathrm{Im}[t'])$ and $p_0(\mathrm{Im}[t'''])$ [Fig. 3b–d] shows a maximum at zero and perfectly symmetric behavior around it. Hence, no particular hopping direction is preferred. The moduli $|t|$ and $|t'|$ [Fig. 3e–f] tend to be slightly larger than 1 (the reference value), and their PDFs do not show any significant structure. On the other hand, $p_0(|t'''|)$ (Fig. 3g) possesses two local maxima of similar magnitude. By restricting the data set to the samples with $|t'''| < 1.25$ and $|t'''| > 1.25$ [corresponding to the approximate midpoint between the two local maxima of $p_0(|t'''|)$, see Supplementary Figure 2], we identify two distinct dominant configurations with $C = 0$, which are illustrated schematically in Fig. 4 (top row). One of them shows strong $|t|$ and $|t'''|$, and weaker $|t'|$, consistent with an inverse Star of David (iSoD)-like CBO pattern (Fig. 4, top row, left panel). The fraction of samples with this configuration amounts to 48% of the trivial cases. The remaining 52% of the trivial samples show an opposite pattern similar to the SoD-like CBO pattern, with larger $|t'|$, and smaller $|t|$ and $|t'''|$ (Fig. 4, top row, right panel). Such CBO patterns have also been predicted by phenomenological analyses of possible electronic instabilities at the van Hove filling[13,16,19,39], and STM experiments have hinted towards the presence of chiral charge order patterns[22] in KV_3Sb_5 with an iSoD-like CBO as observed for the trivial phase. The $C = 0$ phase of our analysis, however, is not chiral, since the *real* hoppings of the tight-binding model fulfill all mirror symmetries of the kagome superlattice. On the other hand, the topological phases discussed in the remainder of the paper possess a chiral character due to complex hoppings, which induce non-trivial fluxes with a specific handedness (see Fig. 4).

Topological phases (C = ±1)

In contrast to the trivial phase, the hoppings in the topological phases display a preference for certain winding directions, as can be seen from the distributions $p_{C \neq 0}(\mathrm{Im}[t])$, $p_{C \neq 0}(\mathrm{Im}[t'])$ and $p_{C \neq 0}(\mathrm{Im}[t'''])$ in Fig. 3b–d. Phases with positive Chern number can be distinguished from the corresponding phases with

negative Chern number by the sign of the imaginary parts of the hoppings, since their respective PDFs are mirror images of each other. The PDFs of the moduli [Fig. 3e–g], instead, are equal for phases with positive and negative Chern number, and hence, do not distinguish between them. By restricting the data set to specific feature values, we analyze the most likely configurations for the respective Chern numbers.

We evaluate the *importance score* $D_B(p_C(x_i), p_0(x_i))$ (see Eq. (3) of the Methods Section) to identify the most descriptive features x_i that distinguish the non-trivial phases from the trivial one. The results are shown in Fig. 5. The importance score of ϵ is trivially zero since it is always kept at a constant value of $\epsilon = 0$. Due to the large overlap of $p_0(\epsilon')$ and $p_{C \neq 0}(\epsilon')$, the importance of ϵ' is rather low. However, as described earlier, the PDFs show clear peaks revealing ϵ' as distinguishing parameter between topological and trivial phases. Next, we infer from Fig. 5 that t' and t'' have the same importance, since their PDFs show the same behavior. t and t''' have higher importance than t' (and t'') for differentiating $C = \pm 1$ phases from the $C = 0$ phase, while t' (and t'') and t''' are more important than t for differentiating $C = \pm 2$ phases from the $C = 0$ phase. This importance with respect to the differentiation between the Chern classes is reflected further in the PDFs of these features in a qualitative manner, as discussed in the following.

For $C = \pm 1$ phases, we find that the moduli $|t|$, $|t''|$ and $|t'''|$ behave similarly [Fig. 3e–g], and we infer that the relative bond strengths may not be a strong distinguishing feature for the topological phases. This is depicted in Fig. 4 (middle row) by equal thickness of the blue, red and green bonds for $C = \pm 1$ phases. On the other hand, we gain crucial insight from the imaginary parts of hoppings in this phase [Fig. 3(b-d)]. For $C = 1$, both $\mathrm{Im}[t]$ and $\mathrm{Im}[t''']$ tend to be larger than zero, which corresponds to a counter-clockwise winding of the hoppings of the inner hexagon and a clockwise winding of the hoppings forming the outer triangles (connecting different 2×2 cells), as schematically illustrated by arrows in Fig. 4 (middle row). The sign of $\mathrm{Im}[t']$ and $\mathrm{Im}[t'']$, instead, does not discriminate between $C = 1$ and $C = -1$ due to missing contrast between the corresponding probability distributions. Hence, the orientations of t' and t'' bonds are not shown in Fig. 4 for $C = \pm 1$. We note that a large fraction of $C = 1$ topological insulators (49%) shows this configuration, while the remaining samples are distributed incoherently.
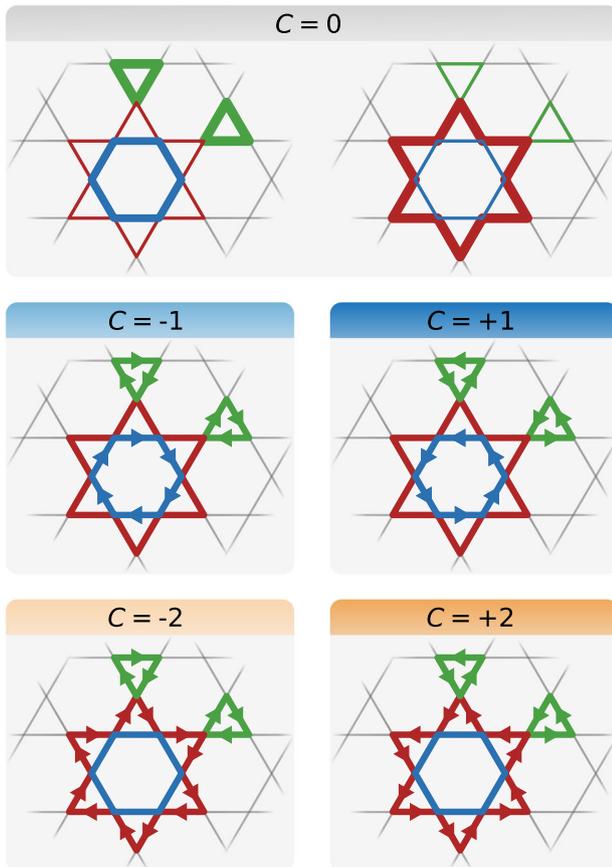
Fig. 4 Insulating phases of the kagome superlattice. Overview of the characteristics of possible insulating phases of the 2×2 kagome superlattice with C_6 symmetry that is extracted from an unbiased data set. The hopping parameter t is colored in blue, t' and t' in red and t'' in green. t''' connects different unit cells and are arranged in triangular manner. Arrows illustrate the hopping flux direction and the line thickness indicates the relative magnitude of $|t|$, $|t'| = |t''|$ and $|t''|$. For $C = 1$ and $C = 2$ the hopping flux direction is mirrored with respect to $C = -1$ and $C = -2$, respectively. The handedness of the complex hopping patterns testifies the chiral nature of the topological phases.

Our characterization of the $C = \pm 1$ phase shares similarities with the "chiral flux phase" (CFP) proposed in ref. [40] as a minimal model for the time-reversal symmetry breaking which is observed in muon spin relaxation experiments in KV_3Sb_5[41] and CsV_3Sb_5[42], and for the giant anomalous Hall effect measurements[21] in KV_3Sb_5. The CFP phase, which represents a possible electronic instability of the kagome metal at the van Hove filling[16], is described by a C_6-symmetric tight-binding model, which breaks time-reversal, but is invariant under the simultaneous action of time reversal and lattice reflections[40], analogous to the Haldane model on the honeycomb lattice[27]. As opposed to the CFP phase of ref. [40], our results for the $C = \pm 1$ phase suggest that the imaginary parts of t' and t'' hoppings may not play a relevant role in the topological character of this phase.

Topological phases ($C = \pm 2$)

In the $C = \pm 2$ phases, the moduli of all features behave similarly to the $C = \pm 1$ phases. On the other hand, the sign of $\mathrm{Im}[t]$ does not discriminate between $C = 2$ and $C = -2$ due to low contrast between the PDFs $p_2(\mathrm{Im}[t])$ and $p_{-2}(\mathrm{Im}[t])$ [Fig. 3 (b)]. Phases with positive and negative Chern number are differentiated by the signs of $\mathrm{Im}[t']$, $\mathrm{Im}[t'']$ and $\mathrm{Im}[t''']$, which leads to their relatively higher importance score. For $C = 2$, the hoppings along the outer spikes of the SoD (t' and t'') show clockwise winding, while the
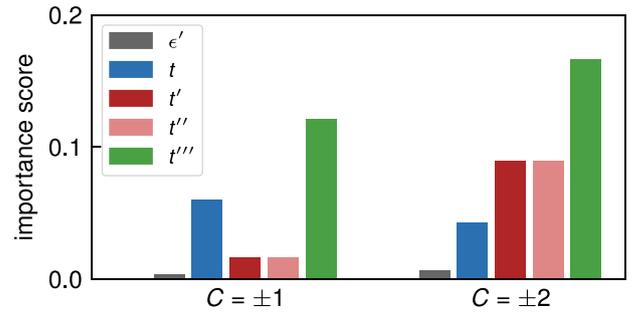

Fig. 5 Importance score. Importance score defined by the Bhattacharyya distance $D_B(p_C(x_i), p_0(x_i))$ (Eq. (3)) for phases with $C = \pm 1$ (left) and $C = \pm 2$ (right) of the C_6 symmetric model for the 2×2 kagome superlattice. The distinct features x_i are ϵ, ϵ', t, t', t'' and t''' as defined in the main text.

hoppings in the outer triangles (t''') show counter-clockwise winding, as illustrated in Fig. 4 (bottom row). The largest coherent group of samples of topological insulators with Chern number $C = 2$ (56%) shows this particular configuration.

Further analysis

Motivated by recent experimental results detecting signatures of rotational symmetry breaking in the electronic properties of some AV_3Sb_5 materials[18,43], we investigate the fate of the topological phases of Fig. 4 when the symmetry of the tight-binding model is reduced from C_6 to C_2. We repeat our statistical analysis by forcing the Hamiltonian on the 2×2 superlattice to be invariant only under rotations of 180°, thus increasing the feature space of the model to 18 distinct parameters, i.e., 6 onsite potentials and 12 hoppings. We start from a reference point (as explained in the Methods Section) with uniform hoppings ($t = -1$) and zero onsite potentials. Among our samples, only a fraction of 13% represent topological insulators (vs. 46% in the analysis with C_6 symmetry), most of them possessing $C = \pm 1$ Chern number (98.6%). While a smaller fraction of topological insulators can be expected as a consequence of the enlargement of the feature space, the strong reduction of $C = \pm 2$ samples (1.4% of the total number of topological insulators) may suggest that this phase is rather fragile to rotational symmetry breaking. In contrast, the $C = \pm 1$ insulating phase is considerably less affected, and thus seemingly more stable.

It is worth emphasizing that, although our analysis has been performed on a model of spinless electrons which explicitly breaks time-reversal symmetry (due to the presence of complex hoppings), our results can provide direct information about what one shall expect for a time-reversal invariant Hamiltonian of spinful electrons. Indeed, in analogy to the generalization of the Haldane model to the Kane-Mele model[44], we took two copies of our topological tight-binding Hamiltonians of Fig. 4 to construct a time-reversal invariant model for spinful electrons. The samples with odd Chern number in the spinless case, i.e., those belonging to the $C = \pm 1$ phase, yield a non-trivial $\mathbb{Z}_2$ invariant in the case of spinful electrons, which is characteristic of quantum spin Hall phases[44].

DISCUSSION

By employing machine-assisted unbiased statistical learning constrained only by specific crystal symmetries, we extract meaningful topological information concerning the 2×2 kagome superlattice. The highlights of our procedure are three-fold: first, one is able to tune through a large parameter space to find non-trivial topology in the kagome superlattice, second, specific crystal symmetries can constrain these parameters resulting in certain flux patterns concomitant with CBO/CDW orders, and third, one retains high levels of physical interpretability of the results. For the kagome

superlattice with C_6 symmetry, we infer possible SoD/iSoD-like CBO patterns and topologically non-trivial flux patterns from the large data sets of randomized hopping parameters. Our findings for the trivial and topological phases share similarities with recent experimental observations and theoretical predictions for the intensely discussed AV_3Sb_5 kagome materials. Moreover, we infer that additional topological phases with higher Chern index ($C = \pm2$) might exist. Furthermore, by reducing the crystal symmetry to C_2, whose signatures were found in AV_3Sb_5 in recent experiments, we examined the stability of topological phases. While $C = \pm1$ appears to be stable, the discovered $C = \pm2$ phases seem to be rather fragile. We also extended our analysis to spinful Hamiltonians, which show quantum spin Hall states. Our results provide a repository of knowledge that can guide future engineering endeavors to build kagome materials (or modify existing ones) with a desirable topological phase. In this regard, a foreseeable extension of the present work consists of pursuing a material-specific analysis, searching for topological phases in the feature space of a tight-binding model obtained by ab initio calculations for a specific target material. In the case of AV_3Sb_5 compounds, this may involve a multi-orbital description, featuring additional vanadium d-orbitals and antimony p-orbitals[37], and the inclusion of spin-orbit coupling effects. Furthermore, the investigation of a layered superlattice geometry, such as the $2 \times 2 \times 4$ structure recently observed in Raman spectroscopy[45] and x-ray diffraction[46] experiments in CsV_3Sb_5, represents a viable future direction. In both cases, the addition of physically motivated ingredients in the tight-binding Hamiltonian could lead to an improved understanding of the actual physical origin of the chiral topological phases.

METHODS

Definitions

Our statistical approach[36] uses random number generators to yield a data set of n_s different tight-binding Hamiltonians (samples) for a given lattice. Each sample is characterized by a vector of features, $\mathbf{x} = (x_1, \ldots, x_{n_f})$, grouping the n_f distinct parameters of the model, and is classified by the label, which is a function of the features, i.e., $l = f(\mathbf{x})$. For the current system, the onsite terms ϵ_i and the (complex) hopping parameters t_i act as features. Samples are categorized into metals and insulators based on the presence of a finite band gap at the filling to be considered. After omitting metallic samples, the first Chern number C[47–49] is then chosen as the label for insulating samples. Hence, a feature vector for a sample is given by $\mathbf{x} = (\epsilon_1, \epsilon_2, \ldots, t_1, t_2, \ldots)$ with label $l = C[H(\mathbf{x})] \in \mathbb{Z}$.

Data generation

Each sample in the data set is generated by randomly picking a value for each feature from a uniform probability distribution function (PDF). Specifically, for a given complex feature x_i, where i indexes different features, we sample the uniform PDF restricted to a sphere in the complex plane centered at a given reference point x_i^{ref}, with radius $a|x_i^{\text{ref}}|$, where $a \in \mathbb{R}$. Throughout this work, we choose $a = 1.5$. This choice for the sampling space ensures physically reasonable configurations, since extreme hopping values are excluded. For gaining maximum insight into the data, we can decompose the complex features x_i into real features, namely the real part $\text{Re}[x_i]$, the imaginary part $\text{Im}[x_i]$, the magnitude $|x_i|$, and the phase $\varphi[x_i] = \arg[x_i]$.

Statistical analysis

To understand which features play a major role in determining the topological properties of the model, we calculate the PDFs $p_l(x_i)$ of each feature x_i for each label l. This is achieved by integrating out all other features $x_j \neq x_i$ from the bare distributions of the topological class $\rho_l(\mathbf{x})$

$$p_l(x_i) = \int \ldots \int \rho_l(\mathbf{x}) \prod_{j \neq i} \mathrm{d} x_j. \tag{2}$$

For a given feature x_i, the comparison of the PDFs for different labels l, i.e., for different Chern numbers, provides information on the importance of x_i

for the topological properties of the tight-binding model. To quantify the difference between two PDFs, we make use of the Bhattacharyya distance[50], defined for a complex feature as

$$D_B(p, q) = -\log\left[\int \sqrt{p(x_i)q(x_i)} \, \mathrm{d} x_i\right]. \tag{3}$$

Here, $p(x_i)$ and $q(x_i)$ are generic PDFs and $D_B(p, q)$ is always larger than zero unless $p = q$.

The measure represented by D_B acts as an indicator of the descriptiveness of features x_i through $D_B(p_l(x_i), p_0(x_i))$, which quantifies the difference of the respective PDFs for Chern labels $l \neq 0$ w.r.t. the trivial case. Larger values contribute most to the topological character. Based on this, one can simplify the investigation of the feature space by focusing only on the most descriptive features with high importance score, which amounts to a dimensionality reduction. A complementary strategy makes use of symmetries that are either based on observed behavior of the PDFs or physical motivation.

The combined approach can generally take several iterations of re-sampling and analyzing the obtained data sets. The interplay of different features can be assessed by computing statistical correlations among them. A straight-forward estimator of linear correlations is provided by the Pearson correlation coefficient[51]. A complementary way to investigate correlations is to restrict the data set to samples where certain features have specific values, e.g., $\text{Im}[x_i] < 0$, and afterwards investigating the PDFs of the restricted data set, as done here.

Discussions

Summarizing, this approach tackles an n_f-dimensional phase space by sampling hopping parameters and computing the Chern number of the resulting Hamiltonians. From the average properties of the distributions of the different Chern numbers, we are able to reconstruct a posteriori an effective description of the topological phases and their properties. This method not only yields information on the symmetry of the topological phases, but also provides crucial insights on which hoppings play a relevant role in determining the topological character. For example, the statistical analysis of the $C = \pm1$ phases identified in our work indicates that the imaginary parts of t' and t'' hoppings are not important to determine the topological character of the state, as discussed in the main text.

REFERENCES

1. Balents, L. Spin liquids in frustrated magnets. *Nature* **464**, 199–208 (2010).
2. Savary, L. & Balents, L. Quantum spin liquids: a review. *Rep. Prog. Phys.* **80**, 016502 (2016).
3. Mendels, P. et al. Quantum magnetism in the paratacamite family: towards an ideal kagomé lattice. *Phys. Rev. Lett.* **98**, 077204 (2007).
4. Han, T.-H. et al. Fractionalized excitations in the spin-liquid state of a kagome-lattice antiferromagnet. *Nature* **492**, 406–410 (2012).
5. Norman, M. R. Colloquium: Herbertsmithite and the search for the quantum spin liquid. *Rev. Mod. Phys.* **88**, 041002 (2016).
6. Broholm, C. et al. Quantum spin liquids. *Science* **367**, eaay0668 (2020).
7. Jeschke, H. O., Salvat-Pujol, F. & Valentí, R. First-principles determination of Heisenberg Hamiltonian parameters for the spin-1 2 kagome antiferromagnet $ZnCu_3(OH)_6Cl_2$. *Phys. Rev. B* **88**, 075106 (2013).
8. Ortiz, B. R et al. New kagome prototype materials: discovery of KV_3Sb_5, RbV_3Sb_5, and CsV_3Sb_5. *Phys. Rev. Mater.* **3**, 094407 (2019).
9. Ortiz, B. R. et al. Superconductivity in the Z_2 kagome metal KV_3Sb_5. *Phys. Rev. Mater.* **5**, 034801 (2021).

10. Ortiz, B. R. et al. CsV$_3$Sb$_5$: A Z$_2$ Topological Kagome metal with a superconducting ground state. *Phys. Rev. Lett.* **125**, 247002 (2020).
11. Yin, Q. et al. Superconductivity and normal-state properties of kagome metal RbV$_3$Sb$_5$ single crystals. *Chin. Phys. Lett.* **38**, 037403 (2021).
12. Syôzi, I. Statistics of Kagomé Lattice. *Prog. Theor. Phys.* **6**, 306–308 (1951).
13. Kiesel, M. L., Platt, C. & Thomale, R. Unconventional Fermi surface instabilities in the Kagome Hubbard model. *Phys. Rev. Lett.* **110**, 126405 (2013).
14. Mazin, I. et al. Theoretical prediction of a strongly correlated Dirac metal. *Nat. Commun.* **5**, 1–7 (2014).
15. Guterding, D., Jeschke, H. O. & Valentí, R. Prospect of quantum anomalous Hall and quantum spin Hall effect in doped kagome lattice Mott insulators. *Sci. Rep.* **6**, 1–8 (2016).
16. Park, T., Ye, M. & Balents, L. Electronic instabilities of kagome metals: saddle points and Landau theory. *Phys. Rev. B* **104**, 035142 (2021).
17. Denner, M. M., Thomale, R. & Neupert, T. Analysis of charge order in the kagome metal AV$_3$Sb$_5$ (A = K, Rb, Cs). *Phys. Rev. Lett.* **127**, 217601 (2021).
18. Jiang, K. et al. Kagome superconductors AV$_3$Sb$_5$ (A = K, Rb, Cs). Preprint at https://arxiv.org/abs/2109.10809 (2021).
19. Lin, Y.-P. & Nandkishore, R. M. Complex charge density waves at Van Hove singularity on hexagonal lattices: Haldane-model phase diagram and potential realization in the kagome metals AV$_3$Sb$_5$ (A = K, Rb, Cs). *Phys. Rev. B* **104**, 045122 (2021).
20. Jiang, B. et al. Experimental observation of non-Abelian topological acoustic semimetals and their phase transitions. *Nat. Phys.* **17**, 1239–1246 (2021).
21. Yang, S.-Y. et al. Giant, unconventional anomalous Hall effect in the metallic frustrated magnet candidate KV$_3$Sb$_5$. *Sci. Adv.* **6**, eabb6003 (2020).
22. Jiang, Y.-X. et al. Unconventional chiral charge order in kagome superconductor KV$_3$Sb$_5$. *Nat. Mater.* **20**, 1353–1357 (2021).
23. Shumiya, N. et al. Intrinsic nature of chiral charge order in the kagome superconductor RbV$_3$Sb$_5$. *Phys. Rev. B* **104**, 035131 (2021).
24. Wang, Z. et al. Electronic nature of chiral charge order in the kagome superconductor CsV$_3$Sb$_5$. *Phys. Rev. B* **104**, 075148 (2021).
25. Zhao, H. et al. Cascade of correlated electron states in the kagome superconductor CsV$_3$Sb$_5$. *Nature* **599**, 216–221 (2021).
26. Tsirlin, A. A. et al. Role of Sb in the superconducting kagome metal CsV$_3$Sb$_5$ revealed by its anisotropic compression. *SciPost Phys.* **12**, 49 (2022).
27. Haldane, F. D. M. Model for a Quantum Hall effect without Landau levels: condensed-matter realization of the "Parity Anomaly". *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
28. Christensen, M. H., Birol, T., Andersen, B. M. & Fernandes, R. M. Theory of the charge density wave in av$_3$sb$_5$ kagome metals. *Phys. Rev. B* **104**, 214513 (2021).
29. Carrasquilla, J. Machine learning for quantum matter. *Adv. Phys.: X* **5**, 1797528 (2020).
30. Carleo, G. & Troyer, M. Solving the quantum many-body problem with artificial neural networks. *Science* **355**, 602–606 (2017).
31. Carrasquilla, J. & Melko, R. G. Machine learning phases of matter. *Nat. Phys.* **13**, 431–434 (2017).
32. Zhang, Y. et al. Machine learning in electronic-quantum-matter imaging experiments. *Nature* **570**, 484–490 (2019).
33. Beach, M. J. S., Golubeva, A. & Melko, R. G. Machine learning vortices at the Kosterlitz-Thouless transition. *Phys. Rev. B* **97**, 045207 (2018).
34. Rodriguez-Nieva, J. F. & Scheurer, M. S. Identifying topological order through unsupervised machine learning. *Nat. Phys.* **15**, 790–795 (2019).
35. Greplova, E. et al. Unsupervised identification of topological phase transitions using predictive models. *New J. Phys.* **22**, 045003 (2020).
36. Mertz, T. & Valentí, R. Engineering topological phases guided by statistical and machine learning methods. *Phys. Rev. Res.* **3**, 013132 (2021).
37. Gu, Y., Zhang, Y., Feng, X., Jiang, K. & Hu, J. Gapless excitations inside the fully gapped kagome superconductors AV$_3$Sb$_5$. *Phys. Rev. B* **105**, L100502 (2022).
38. Haldane, F. D. M. Model for a Quantum Hall Effect without Landau levels: condensed-matter realization of the "Parity Anomaly". *Phys. Rev. Lett.* **61**, 2015–2018 (1988).
39. Feng, X., Zhang, Y., Jiang, K. & Hu, J. Low-energy effective theory and symmetry classification of flux phases on the kagome lattice. *Phys. Rev. B* **104**, 165136 (2021).
40. Feng, X., Jiang, K., Wang, Z. & Hu, J. Chiral flux phase in the Kagome superconductor AV$_3$Sb$_5$. *Sci. Bull.* **66**, 1384–1388 (2021).
41. Mielke, C. et al. Time-reversal symmetry-breaking charge order in a kagome superconductor. *Nature* **602**, 245–250 (2022).
42. Yu, L. et al. Evidence of a hidden flux phase in the topological kagome metal CsV$_3$Sb$_5$. Preprint at https://arxiv.org/abs/2107.10714 (2021).
43. Li, H. et al. Rotation symmetry breaking in the normal state of a kagome superconductor KV$_3$Sb$_5$. *Nat. Phys.*1-6 (2022).
44. Kane, C. L. & Mele, E. J. Z$_2$ topological order and the Quantum Spin Hall effect. *Phys. Rev. Lett.* **95**, 146802 (2005).
45. Wu, S. et al. Charge density wave order in kagome metal AV$_3$Sb$_5$ (A = Cs, Rb, K). Preprint at https://arxiv.org/abs/2201.05188 (2022).
46. Xiao, Q. et al. Coexistence of Multiple Stacking Charge Density Waves in Kagome Superconductor CsV$_3$Sb$_5$. Preprint at https://arxiv.org/abs/2201.05211 (2022).
47. Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. A* **392**, 45–57 (1984).
48. Wilczek, F. & Zee, A. Appearance of gauge structure in simple dynamical systems. *Phys. Rev. Lett.* **52**, 2111–2114 (1984).
49. Fukui, T., Hatsugai, Y. & Suzuki, H. Chern numbers in discretized brillouin zone: efficient method of computing (Spin) hall conductances. *J. Phys. Soc. Jpn* **74**, 1674–1677 (2005).
50. Bhattacharyya, A. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* **35**, 99–109 (1943).
51. Pearson, K. Note on regression and inheritance in the case of two parents. *Proc. R. Soc. London Series I* **58**, 240–242 (1895).

AUTHOR CONTRIBUTIONS

T.M. and P.W. performed the calculations and contributed equally to the work. S.B. and F.F. contributed to the analysis of the statistical data and the implementation of symmetries. R.V. supervised the project. All authors made contributions to the development of the approach and wrote the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-022-00745-3.

Correspondence and requests for materials should be addressed to Thomas Mertz, Paul Wunderlich or Roser Valentí.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.