Probing center vortices and deconfinement in SU(2) lattice gauge theory with persistent homology

Nicholas Sale[*] and Biagio Lucini[†]

Department of Mathematics, Swansea University, Bay Campus,
SA1 8EN Swansea, Wales, United Kingdom

Jeffrey Giansiracusa

Department of Mathematical Sciences, Durham University,
Upper Mountjoy Campus, Durham, DH1 3LE, United Kingdom

We investigate the use of persistent homology, a tool from topological data analysis, as a means to detect and quantitatively describe center vortices in SU(2) lattice gauge theory in a gauge-invariant manner. We provide evidence for the sensitivity of our method to vortices by detecting a vortex explicitly inserted using twisted boundary conditions in the deconfined phase. This inspires the definition of a new phase indicator for the deconfinement phase transition. We also construct a phase indicator without reference to twisted boundary conditions using a simple k-nearest-neighbors classifier. Finite-size scaling analyses of both persistence-based indicators yield accurate estimates of the critical β and critical exponent of correlation length ν of the deconfinement phase transition.

I. INTRODUCTION

Quantum chromodynamics (QCD) poses several outstanding problems in particle physics, including the mechanism of confinement and the deconfinement phase transition, mass-gap generation, and chiral symmetry breaking [1]. These phenomena are nonperturbative and are therefore typically investigated through the framework of lattice QCD. Supplementing this traditional approach, there is an emerging body of work exploring the use of machine learning and data analysis tools in extending Monte Carlo analysis of lattice QCD toward generating the insights required to tackle these open problems. References include [2–4] among others. For example, one application of machine learning is to classify phases of the theory based on sampled configurations, learning observables that function as order parameters for the phase transitions undergone by QCD, following the quantitative program outlined for spin models in [5].

Focusing on confinement, a compelling potential mechanism relies on the presence of topological defects called center vortices in confining gauge configurations [6–8]. Vortexlike configurations have been shown to exist in pure gauge theories (e.g., [9–35]) and have recently been observed in lattice simulations of QCD [36]. While order parameters have been constructed for the confinement-deconfinement phase transitions in Yang-Mills theories that are based on the topological symmetry related to the conservation of the number of vortices (see, e.g., [15,37,38]), identification of vortices proves to be a more challenging undertaking. In fact, existing methodologies for exposing vortices in gauge theories rely on performing gauge fixing and projection [39]. This procedure suffers from the problem of Gribov ambiguities [40,41], which, even with careful choices of the gauge fixing condition (for instance, following the prescription of [42]), can be mitigated only in part. Motivated by the physical appeal of a fully gauge-independent description, we investigate the possibility of instead analyzing vortices and performing phase classification in a gauge-invariant manner by making use of persistent homology [43] alongside other tools from topological data analysis, a novel approach to data (in our case, consisting of gauge configurations) that places emphasis on rigorous classifications of shapes of datasets based on their topological properties. Rather than considering full QCD, in this paper we instead develop a methodology for the pure gauge SU(2) lattice gauge theory, which also exhibits a deconfinement transition potentially

[*]nicholas.j.sale@gmail.com; https://nicksale.github.io/
[†]Also at Swansea Academy of Advanced Computing, Swansea University, Bay Campus, SA1 8EN, Swansea, Wales, United Kingdom.

driven by center vortices. In the context of vortex identification, this system represents a toy model of QCD where the quarks have been removed and the gauge group is simplified from SU(3).

Among other previous work on the use of persistent homology and computational topology to investigate systems from physics [44–54] (a useful survey is [55]), we remark on two that look at QCD and lattice gauge theories in particular. Kashiwa, Hirakida, and Kouno noted the sign problem in simulating dense QCD and instead considered an effective model in the form of a modified 3D Potts model where each $\mathbb{Z}_3$ Potts spin corresponds to a Polyakov loop [56], avoiding the need to construct a filtered complex from gauge fields directly. Computing the persistence of point clouds of lattice sites sharing the same Potts spin, they were able to probe the phase structure of the model using the average and maximum birth-death ratio of points in the resulting persistence diagrams. More recently, Sehayek and Melko investigated the 2D and 3D $\mathbb{Z}_2$ lattice gauge theories [57]. Given a configuration of the model, they place a point at the center of each spin-down lattice link and compute the Vietoris-Rips filtered complex of the resulting point cloud. The β_1 Betti curve then provides the number of closed strings of down spins and their sizes. These quantities and the filtered complex are gauge variant, but by averaging the loop count over many configurations they indirectly measure the density of vison defects which produces a clear indicator of the phase transition in the 3D model when plotted as a function of temperature.

Our main contributions are as follows.

(i) We introduce a gauge-invariant construction that produces a filtered complex (for input to persistent homology) from a field configuration for SU(2) lattice gauge theory, and we argue that this pipeline should see center vortices.

(ii) We demonstrate that the resulting persistent homology is able to detect an explicitly inserted vortex in the deconfined phase by showing that it distinguishes configurations generated using twisted boundary conditions. In particular, we show that a phase indicator can be recovered by comparing the PH_2 persistence diagrams of configurations generated using twisted boundary conditions and configurations sampled using the usual Wilson action, similarly to the vortex free energy order parameter. This phase indicator allows us to estimate the critical β and critical exponent of correlation length via finite-size scaling.

(iii) Using a k-nearest-neighbors classification of persistence images, we identify the deconfinement phase transition directly from configurations generated using the usual Wilson action. Moreover we accurately estimate the critical β and critical exponent of correlation length via finite-size scaling.

The rest of this paper is organized as follows. In Sec. II we introduce the SU(2) lattice gauge theory, the deconfinement phase transition and the center vortex picture for confinement. In Sec. III we review the techniques we use including persistent homology, k-nearest-neighbors classification and finite-size scaling analysis. In Sec. IV we introduce two different persistent homology-based phase indicators for deconfinement, meaning statistics that are zero in one phase and nonzero in the other, which we use to quantitatively analyze the phase transition. Finally in Sec. V we discuss our findings and identify potential directions for future work. The appendixes contain more detailed reviews of some of the tools we use.

II. SU(2) LATTICE GAUGE THEORY

A. The model

A configuration of the 4D SU(2) lattice gauge theory is specified by SU(2)-valued variables $U_\mu(x)$ located on each link (x, μ) of an $N_t \times N_s^3$ lattice Λ with periodic boundary conditions, where $\mu \in \{0, 1, 2, 3\}$ describes the direction in which the link emanates from the lattice site $x \in \Lambda$. In practice, $U_\mu(x)$ lies in the fundamental representation of SU(2), taking the form of a 2×2 complex matrix. To simulate at nonzero temperature we ensure $N_t \ll N_s$. The gauge symmetry is generated by gauge transformations $\Omega(x) \in$ SU(2)$^\Lambda$ sending each $U_\mu(x) \mapsto \Omega^\dagger(x) U_\mu(x) \Omega(x + \hat{\mu})$, where $x + \hat{\mu}$ denotes the lattice site one step in the μ direction from x. The observables that are invariant under these transformations are traces of products of the link variables along closed paths C, also known as Wilson loops $W(C)$. The simplest nontrivial example is the Wilson loop around a 1×1 plaquette (x, μ, ν) of the lattice:

$$W_{\mu,\nu}(x) = \frac{1}{2} tr[U_\mu(x) U_\nu(x + \hat{\mu}) U_\mu^\dagger(x + \hat{\nu}) U_\nu^\dagger(x)].$$

We use this to define the Wilson action given a configuration $\mathbf{U} = \{U_\mu(x)\}_{(x,\mu)}$ as

$$S(\mathbf{U}) = -\frac{\beta}{4} \sum_{x, \mu < \nu} W_{\mu,\nu}(x), \qquad (1)$$

where $\beta = 4/g^2$ and g is the gauge coupling parameter. This in turn allows us to define the vacuum expectation value of any given observable $A(\mathbf{U})$ as

$$\langle A \rangle = \frac{\int d\mathbf{U} A(\mathbf{U}) e^{-S(\mathbf{U})}}{\int d\mathbf{U} e^{-S(\mathbf{U})}}, \qquad (2)$$

where $d\mathbf{U} = \prod_{x,\mu} dU_\mu(x)$ is a product of Haar measures over SU(2) for each link variable. In practice we estimate expectations using Monte Carlo methods, where Eq. (2) becomes a simple mean of the observed values.

B. Deconfinement and center vortices

The model introduced above exhibits two phases—a confined phase at low β and a deconfined phase at high β—and the phase transition between these is known as the

deconfinement transition. This transition occurs only in the spatial continuum limit $N_s \to \infty$ and while we care also about the temporal limit $N_t \to \infty$, it is worth noting that a critical β is well defined for each finite N_t and it is these we will be estimating. Confinement in the SU(2) lattice gauge theory can be characterized in a number of ways [1], including the following.

(i) *Area law for Wilson loops.*—Let $W(C)$ denote the value of a Wilson loop around a closed curve $C = R \times T$, consisting of rectangle of edges R in a spacelike direction and T along the temporal direction. We consider the limit of large area $\mathcal{A}(C) = RT$. In the confined phase at low β we have that $\langle W(C) \rangle \propto \exp(-\sigma \mathcal{A}(C))$, where σ is known as the string tension. In the deconfined phase at high β we have that $\langle W(C) \rangle$ decays exponentially instead with $\mathcal{P}(C)$, the perimeter of C.

(ii) *Vanishing Polyakov loop.*—Define the Polyakov loop at a point x in the lattice as

$$P(x) = \frac{1}{2} tr[U_0(x)U_0(x + \hat{0})\ldots U_0(x - \hat{0})]$$

(note this only depends on the spatial coordinates of x). This represents a Wilson loop that has a nonzero winding number around the time direction of the periodic lattice. In the confined phase we have $\langle P(x) \rangle = 0$ whereas in the deconfined phase we have $\langle P(x) \rangle \neq 0$.

Several possible pictures of what drives the deconfinement transition, in both this model and QCD, have been proposed. Here we focus on the center vortex picture [6,7]. Fix a time slice at time t. Given two closed oriented contours C and C' in that three-dimensional slice with linking number m, a loop operator $B(C', t)$ can be defined that has the following commutation algebra with the Wilson loop $W(C, t)$:

$$W(C, t)B(C', t) - (-1)^m B(C', t)W(C, t) = 0. \quad (3)$$

This equation defines the 't Hoof algebra [6]. For simplicity, we consider planar nonintersecting curves C and C', for which $m = 0, 1$. The operator $B(C', t)$ is called the 't Hooft loop. When acting on a gauge configuration, $B(C', t)$ creates a magnetic flux with the resulting observable effect of multiplication by -1 of all Wilson loops having support on curves C with linking number 1 with C'. For this reason, the 't Hooft loop is said to be a vortex creation operator. Since the center of the group, which in our case is $Z(\text{SU}(2)) = \{I, -I\} \cong \mathbb{Z}_2$, plays a role in the 't Hooft algebra [as exposed by the factor $(-1)^m$], the vortices created by the 't Hooft loop operator are called center vortices. Fixing the curve C' for all time slices t, we see that a vortex traces out a surface in 4-space, closed by the periodic boundary conditions.

In the limit of weak fields, where the theory is deconfined, all Wilson loops are close to unity. Confining configurations are expected to have Wilson loops that largely deviate from unity. In particular, Wilson loops close to -1 can be obtained from a weak field configuration through the injection of center vortices generated with appropriate insertions of 't Hooft loop operators. Moving from this observation, operationally we can define a center vortex to be a collection of plaquettes in the dual lattice (in the sense of the dual graph) that form a closed surface (with the closedness being a consequence of the Bianchi identities) and that carry a nontrivial charge in $\mathbb{Z}_2$, corresponding to the -1 element. To carry a nontrivial center charge means that any Wilson loop in the lattice that topologically links with this surface is multiplied by that charge.

In the confined phase center vortices are found to form large surfaces, often wrapping around the periodic boundaries, that percolate throughout the lattice [58]. Therefore, given a particular Wilson loop $W(C)$, the number of vortices that link with C is proportional to the enclosed area $\mathcal{A}(C)$, leading to the area law for the suppression of $\langle W(C) \rangle$. In the deconfined phase, the center vortices become smaller and more sparse, ensuring that for sufficiently large Wilson loops, only those vortices close to the curve C have a chance of linking with it, leading to the perimeter law. Similarly we see that only in the confined phase, where vortices may wrap around the periodic boundary conditions of the lattice, is there a chance they may link with a Polyakov loop, suppressing its expectation. For an overview of the evidence supporting the center vortex picture see Ref. [59]. In practice, center vortices generated by the system have some finite thickness, so that only larger Wilson loops may fully link with them and obtain the full center charge. Loops that partially link may still obtain a partial charge, some factor lying between I and $-I$ in SU(2).

While the concepts of vorticity and creation of a vortex through the insertion of a 't Hooft loop are well understood in terms of symmetry and boundary conditions in a finite volume (see for instance [6,15,37,38]), quantum fluctuations make vortex identification a much more involved process, with currently used prescriptions not fully validated from first principles. A widely used method to detect and analyze these thick vortices is to transform configurations to the maximal center gauge, where each matrix $U_\mu(x)$ is as close to either I or $-I$ as possible, and then project the matrices onto whichever of I or $-I$ is closer. After projection, the Wilson loops of plaquettes take values either 1 or -1 and the latter are identified as projected vortices, or P-vortices. It has been shown that the locations of these correlate with the unprojected thick vortices [60]. However, the projection means that we lose gauge invariance, as well as geometric information such as the thickness of the vortices. In this work we introduce a method designed to look for thin vortices, but we argue that,

TABLE I. Estimates for the critical value of β for the deconfinement phase transition in the SU(2) lattice gauge theory for the values of N_t we consider in this paper. Reproduced from [61].

N_t	β_c
4	2.2986(6)
5	2.37136(54)
6	2.4271(17)

through the use of persistent homology, the results may tell us something about these thick vortices too.

There is no analytical formula for the critical value of β for each value of N_t, but detailed various numerical studies have shown good agreement with each other. We will compare our results to those found in [61] which we reproduce in Table I. We also estimate the critical exponent of correlation length ν which, as a consequence of the Svetitsky-Yaffe conjecture [62], is the same as for the 3D Ising model. We therefore compare our measurement of ν to the high-precision estimate $\nu = 0.629971(4)$ from [63].

C. Twisted boundary conditions

In order to test the sensitivity of our method to thin vortices we will make use of the trick of imposing twisted boundary conditions [64]. The idea is that we choose some coclosed collection of plaquettes in the lattice, i.e., plaquettes that link with some closed surface in the dual lattice, and negate their contribution to the action. See Fig. 1 for an illustration of coclosed collections of plaquettes. In our case we choose the plaquettes

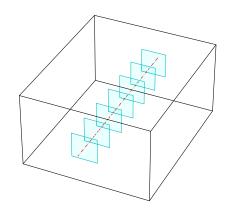$$T = \{(x, \mu, \nu) = ((0, 0, y, z), 0, 1) | 0 \le y, z < N_s\}$$

FIG. 1. A lower-dimensional illustration of a coclosed collection of plaquettes that wraps around the periodic boundary conditions of the lattice. Plaquettes in a 3D lattice link with edges in the dual lattice, so the condition of being coclosed means that the collection of those linking edges forms a closed loop. In this case the loop is closed by the periodic boundary conditions. Going to four dimensions, we imagine repeating the coclosed line of plaquettes along the new dimension, forming a coclosed surface of plaquettes.

corresponding to a surface wrapping around the latter two spatial dimensions of the lattice. The action with twisted boundary conditions becomes

$$S_T(\mathbf{U}) = -\frac{\beta}{4} \left[\sum_{\substack{x, \mu < \nu \\ (x, \mu, \nu) \notin T}} W_{x, \mu, \nu} - \sum_{\substack{x, \mu < \nu \\ (x, \mu, \nu) \notin T}} W_{x, \mu, \nu} \right], \quad (4)$$

which we refer to as the twisted action.

This modification of the action allows the lattice to support an odd number of center vortices wrapping in the yz plane, which is prohibited by the usual periodic boundary conditions of the Wilson action. It is important to note that we are talking about the boundary conditions of the gauge field on the lattice and not the lattice itself. We are not twisting the lattice and forming any kind of Möbius band; rather, it is the gauge field which obtains a factor of $-I$ as we loop around the lattice. We can alternatively think of this twisted action as explicitly inserting a thin vortex into the system on the surface defined by T, so that the system is forced to generate a (thick) vortex to cancel it out. We shall denote expectations calculated with respect to this twisted action by $\langle A \rangle_{\text{twist}}$, where A is a generic observable.

Twisted boundary conditions give us an alternative way to characterize confinement and the deconfinement transition. Magnetic and electric flux free energies can be defined in terms of the ratio of partition functions for the twisted and Wilson actions and the behavior of these can be shown to imply the area law decay for the Wilson loop [65] and therefore confinement.

III. METHODS

A. Background on persistent homology

Persistent homology is a computational topology tool introduced in its modern form in [43] and popularized in [66]. It is one of the main tools of topological data analysis. We shall give a brief overview here following that in [51], but for a more complete review of persistent homology useful references are [67–70].

Given a topological space, such as a manifold or a simplicial or cubical complex, homology is an algebraic way of describing the "holes" in the space. In particular, the spaces we consider will be cubical complexes. A very brief technical introduction to cubical complexes and their homology can be found in Appendix A. In general terms, the kth homology $H_k(C)$ of a cubical complex C, computed with coefficients in a field (here $\mathbb{Z}_2$), is a vector space that has a basis in 1-1 correspondence with the k-dimensional holes in C. As an example, consider removing a unit cube from $\mathbb{R}^3$. This leaves behind a two-dimensional hole since it is enclosed by a two-dimensional surface and as such would be recorded by H_2. Given a map of cubical complexes $f : C \to C'$ (e.g., an inclusion map), we obtain induced linear maps $f_k : H_k(C) \to H_k(C')$. The rank of f_k tells us how many of the k-dimensional holes survived after

being mapped into C', i.e., how many *persisted*. Given some data D, the idea of (cubical) persistence then is to construct a collection of cubical complexes indexed by the reals

$$F_D \colon \mathbb{R} \to \text{CubicalComplex}$$

using the data, so that we have inclusions

$$F_D(s) \subseteq F_D(t)$$

for all $s \leq t \in \mathbb{R}$. We call such an $\mathbb{R}$-indexed collection of complexes and inclusion maps a filtered complex. Since the $F_D(s)$ are each subcomplexes of the final complex $F_D(\infty)$, we can specify the filtered complex by assigning to each cube in the final complex the index at which it first appears, and then $F_D(s)$ is the subcomplex consisting of all cubes that have appeared at or before s.

Applying homology now yields an $\mathbb{R}$-indexed collection of vector spaces

$$H_k{\circ}F_D \colon \mathbb{R} \to \text{VectorSpace}$$

and linear maps

$$H_k(F_D(s)) \to H_k(F_D(t)),$$

and we may use the ranks of these maps to identify when new holes are born, how long they persist through the filtered complex, and when they die. We summarize this information as a multiset called a persistence diagram $PH_k(F_D) \subset \{(a,b) \in \mathbb{R} \times (\mathbb{R} \cup \{\infty\}) | a \leq b\}$ that contains a pair (b,d) every time a k-dimensional hole is born in $F_D(b)$ and dies in $F_D(d)$. In the case that a hole persists even in the final complex $F_D(\infty)$, we write $d = \infty$. We shall make use of such points later. It is said that a feature is born at b and dies at d and that its persistence is $d - b$. This can also be represented as a bar code (a multiset of intervals $[b,d)$). There are a few ways to define distances between persistence diagrams, but those which are most commonly used are the bottleneck and Wasserstein distances (for details see, e.g., [71]). For many typical choices of filtered complex a small change in the input data D leads to only a small change in the persistence diagram $PH_K(F_D)$ as measured by these distances. This property of persistent homology is known as stability and makes persistence a useful tool for dealing with real-world, noisy data.

In this work D is a single configuration of the SU(2) lattice gauge theory. We therefore obtain a persistence diagram for each sampled configuration and we can consider statistics computed from these diagrams.

In their form as multisets, persistence diagrams do not lend themselves to use directly as inputs for many standard machine learning models and can contain many points, taking up a large amount of computer memory. However

there are numerous methods to represent persistence diagrams as fixed-size vectors. In this work we make use of a particular vectorization called a persistence image [72]. Let $g_{\alpha,\beta} \colon \mathbb{R}^2 \to \mathbb{R}$ denote a 2D Gaussian of standard deviation σ centered at (α, β):

$$g_{\alpha,\beta}(x,y) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{(x-\alpha)^2 + (y-\beta)^2}{2\sigma^2}\right].$$

Given a persistence diagram $PH_k = \{(b_i, d_i)\}_{i \in I}$, its persistence surface is the function $\rho_k \colon \mathbb{R}^2 \to \mathbb{R}$ obtained by translating each point $(b,d) \in PH_k$ with $d \neq \infty$ into birth-persistence coordinates $(b, d - b)$ and then placing Gaussians with variance σ^2 on them, weighted by the persistence of the point:

$$\rho_k(x,y) = \sum_{(b,d) \in PH_k} (d - b)g_{b,d-b}(x,y).$$

The $(n_I)^2$-dimensional persistence image PI_k is obtained by discretizing a rectangular region of the domain of ρ_k into a collection of $n_I \times n_I$ pixels p_i and integrating ρ_k within each:

$$PI_k^i = \iint_{p_i} \rho_k(x,y)dxdy.$$

So long as we choose the same σ and discretization for each diagram, we can compute averages and variances componentwise. Besides emphasizing high-persistence points, the linear weighting by the persistence ensures the stability of the persistence image. Finally we note that, as discussed in [72], machine learning models trained on persistence images are generally insensitive to the resolution and variance parameters n_I and σ. We also found this to be the case in our previous work [51]. Therefore in this work, we shall fix the parameters with a resolution of 25×25 and σ equal to 5% of a pixel.

B. Filtered complex

As described in the previous section, to apply persistent homology we must choose how to define a filtered complex for a given configuration $U = \{U_\sigma(x)\}$. We present a filtered complex F_U which is constructed based on Wilson loops and which will therefore give gauge-invariant persistence diagrams.

The idea is to explicitly construct a cubical model of vortex surfaces, under the assumption that vortices are thin. Note that since we have periodic boundary conditions, spacetime in this model forms a 4-torus $S^1 \times S^1 \times S^1 \times S^1$. The lattice Λ defines a decomposition of the spacetime manifold into a cubical complex X in which there is a vertex for each lattice site, an edge for each link in the lattice, a 2-cube for each plaquette, etc. The dual lattice Λ^* determines a dual cubical complex decomposition Y of spacetime.

The vertices of Y are the centers of the 4-cubes of X. More generally, the d-cubes of Y are in bijection with the $(4 - d)$-cubes of X, where the bijection pairs a cube in X with the unique cube in Y for which the intersection is a single point.

Wilson loops live on the cubical complex X, while vortex sheets live on the dual complex Y. We will construct a filtration of Y by letting each 2-cube enter at a filtration index given by the value of the Wilson loop around the boundary of the dual 2-cube in X.

In more detail, denote by $c_A(y) = \prod_{\mu \in A}[y, y + \vec{\mu}]$ a cube in Y, where $[y, y + \vec{\mu}]$ is the line segment between lattice site $y \in \Lambda^*$ and $y + \vec{\mu}$. The cube $c_\varnothing(y)$ is just the point y itself. The dimension d of $c_A(y)$ is $|A|$ and we will refer to it as a d-cube. The boundary $\partial c_A(y)$ of a d-cube is the set of its $d - 1$-cube faces. For example $\partial c_{\{\mu\}}(y) = \{c_\varnothing(y), c_\varnothing(y + \vec{\mu})\}$. With this notation, our observation of the bijection between plaquettes in X and Y becomes that the 2-cube $c_{\{\mu,\nu\}}(y)$ is matched with the 2-cube $c_{\{\sigma,\tau\}}(x)$ in X (defined similarly) that is used to define the Wilson loop $W_{\sigma,\tau}(y + \vec{\mu} + \vec{\nu})$ where $\{\sigma,\tau\} \cap \{\mu,\nu\} = \varnothing$.

To define the filtered complex we will give a filtration index $f(c_A(y)) \in \mathbb{R}$ for each cube $c_A(y)$ in Y specifying when it appears. Then $F_U(s)$ is the subcomplex of Y consisting of all cubes c for which $f(c) \le s$. That is,

$$F_U(s) = f^{-1}(-\infty, s].$$

Since we are attempting to model vortex surfaces, we will initially specify when the 2-cubes are to enter the filtered complex and then introduce the cubes of other dimensions based on these.

Our construction of the function f is the following.

(1) We introduce each 2-cube $c_{\{\mu,\nu\}}(y)$ in our filtered complex at index

$$f(c_{\{\mu,\nu\}}(y)) = W_{\sigma,\tau}(y + \vec{\mu} + \vec{\nu}),$$

where $\{\sigma,\tau\} \cap \{\mu,\nu\} = \varnothing$, that is, at an index equal to the value of the Wilson loop around the plaquette in X paired with it by the bijection.

(2) Since a 2-cube is not allowed to be included before its constituent 1-cubes and 0-cubes in a cubical complex, we introduce these at the smallest index of all the 2-cubes they are incident to. So

$$f(c_A(y)) = \min\{f(C)|c_A(y) \in \partial C\}$$

when $|A| \le 1$.

(3) For the 3-cubes and 4-cubes we follow a cliquelike rule where we introduce a cube as soon as all of its boundary cubes are introduced. So

$$f(c_A(y)) = \max\{f(C)|C \in \partial c_A(y)\}$$

when $|A| \ge 3$.

Thus, for $s < -1$, $F_U(s)$ is the empty complex and, for $s \ge 1$, $F_U(s)$ is the filled in tiling homeomorphic to a 4-torus. Going between these values, the first cubes to enter F_U are surfaces made up of plaquettes in bijection with Wilson loops that are close to -1. The idea therefore is that thin vortex surfaces will enter the filtered complex early. Moreover, since small Wilson loops like those considered here still pick up a partial charge from thick vortices, surfaces representing those thick vortices ought to enter the filtered complex earlier than they otherwise would have. We expect to detect these closed surfaces in persistent H_2 (since we compute homology with $\mathbb{Z}_2$ coefficients, the orientability of the surfaces does not impact this). We may also see other topological features such as the presence of handles or holes in H_1, as well as the transient low-persistence points in persistent H_0 and H_1 that arise as the surface forms near the start of the filtration. An illustration of the connection between Wilson loops and the inclusion of vortices into the filtration is shown in Fig. 2. An illustration of how to imagine what the filtered complex is aiming to do is shown in Fig. 3.

It is worth noting the difference in approach from our previous work using persistent homology to identify vortices in 2D XY models [51]. Vortices there were point defects (located at vertices of the dual lattice), which we aimed to detect by constructing one-dimensional loops in the original lattice that encircled them. Here, we are modeling the vortex surfaces in the dual lattice directly.

It is straightforward to see that this filtered complex F_U is stable with respect to perturbations of the SU(2) link variables since the Wilson loop $W_{\mu,\nu}(y)$ is a linear map due to the linearity of the trace and is therefore Lipschitz
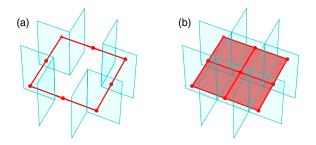

FIG. 2. A lower-dimension illustration of the idea behind the filtered complex. In three dimensions, center vortices form closed one-dimensional curves that link with two-dimensional plaquettes. In this setting we would include edges [dark (red)] according to the Wilson loop around the plaquette [light (cyan)] they intersect. (a) Early on in the filtered complex we include edges that link with plaquettes with negative Wilson loop values. In this way we build explicit cubical models of one-dimensional center vortices which are then detected in PH_1. (b) Later on we eventually fill in the rest of the edges and the 2-cubes between them, destroying the PH_1 features corresponding to the vortices. Moving from three dimensions to four, we are inserting plaquettes instead of edges and instead of a closed curve we obtain a closed surface which we detect with PH_2.
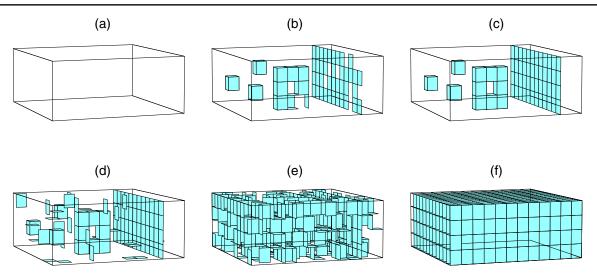
FIG. 3. A cartoon of a 3D slice of the filtered complex. (a) At the beginning of the filtration the complex is empty. (b) At low Wilson loop values we begin to form vortex surfaces with transient features detected in H_0 and H_1. (c) The vortex surfaces close, becoming detectable in H_2. (d) At higher Wilson loop values, vortex surfaces persist while other plaquettes begin to be included in the complex. (e) The vortex surfaces may become filled in if all plaquettes inside the surface are included, killing the corresponding H_2 class. (f) Eventually the whole 4-torus cubical complex is filled in.

continuous. The stability property of persistent homology [73] therefore ensures that a small perturbation of the link variables only results in a small perturbation of the resulting persistence diagram with respect to the bottleneck distance.

C. *k*-nearest-neighbor classification

In Sec. IV B we will make use of *k*-nearest-neighbors (*k*NN) classification to map the persistence images obtained from configurations onto phases as in [51]. For vector-valued data, a *k*NN classifier is a nonparametric model that models a categorical dependent variable $y(\mathbf{x}) \in \mathbb{N}$, where $\mathbf{x} \in \mathbb{R}^N$. The behavior of the model is determined by the training data $\{(\mathbf{x}_i, y_i)\}$ and a choice of the hyperparameter $k \in \mathbb{N}$. Given a new input $\mathbf{x}$, it finds the k indices $i_\mathbf{x}^1, \ldots, i_\mathbf{x}^k$ that minimize the Euclidean distance $\|\mathbf{x} - \mathbf{x}_i\|_2$. It then outputs the most common label among the $y_{i_\mathbf{x}^1}, \ldots, y_{i_\mathbf{x}^k}$.

Here $\mathbf{x}$ will be a persistence image, $y(\mathbf{x}) = 0$ will indicate the low β, confined phase, and $y(\mathbf{x}) = 1$ will indicate the high β, deconfined phase. We will train the model using data sampled from both phases close to the critical region. In the intermediate range of couplings where there are no training data, the *k*NN model will output an estimated classification $O_{k\mathrm{NN}} \in \{0, 1\}$. We may then treat $\langle O_{k\mathrm{NN}} \rangle$ as a phase indicator.

As in [51], we note that training the classifier directly on raw configurations is not computationally feasible. Doing so would require a vast number of samples to sufficiently fill out the configuration space and moreover the computational cost of the classification would grow too large. The mapping from configurations to persistence images concentrates the distribution near a low-dimensional subspace,

and hence *k*NN becomes effective with far fewer samples. We also note that vectorizing the persistence diagrams as persistence images is not necessary for using a *k*NN classifier since this only requires a notion of distance between samples, of which we have several for persistence diagrams. However these distances—such as the bottleneck distance or Wasserstein distance—are computationally expensive, especially for the large persistence diagrams we obtain in this application. Using the Euclidean distance between persistence images vastly speeds up the time taken to evaluate a *k*NN model and, as we will see later, maintains sufficient information to capture the phase transition.

D. Finite-size scaling analysis

The deconfinement phase transition in the SU(2) lattice gauge theory is known to be a second-order transition (again thanks to the Svetitsky-Yaffe conjecture [62]). As such, quantities such as the Polyakov loop susceptibility

$$\chi(\beta) = \langle P^2 \rangle_\beta - \langle P \rangle_\beta^2$$

diverge at the phase transition in a predictable way as we approach the continuum limit. On a finite lattice this susceptibility will remain analytic, displaying a pronounced peak at a pseudocritical β somewhere above or below the true critical inverse coupling β_c. Holding N_t (and therefore β_c) fixed, letting $N_s \to \infty$ causes the peak to narrow and move closer toward β_c. The way in which the susceptibility scales close to β_c as a function of N_s asymptotically approaches the form

$$\chi(N_s, b) = N_s^{\gamma/\nu} \hat{\chi}(N_s^{1/\nu} b), \tag{5}$$

where $\hat{\chi}$ is a dimensionless function, $b = \frac{\beta - \beta_c}{\beta_c}$ is the reduced inverse coupling, and γ and ν are the critical exponents for the susceptibility and correlation length, respectively. By simulating close to the phase transition on different lattice sizes N_s we can extract the heights and locations of the different peaks and then fit these to Eq. (5) to estimate β_c, γ and ν.

Similarly, persistent homology-derived observables may exhibit large variations at criticality. In Sec. IV A we look at the difference in expectations of the birth time of a specific point in the persistence diagram under the normal action and the twisted action. In Sec. IV B we look at the fluctuations in the output O_{kNN} of a trained kNN model, measuring the variance

$$\chi_{kNN}(\beta) = \langle O_{kNN}^2 \rangle_\beta - \langle O_{kNN} \rangle_\beta^2$$
$$= \langle O_{kNN} \rangle_\beta (1 - \langle O_{kNN} \rangle_\beta). \quad (6)$$

Note that the second equation follows since O_{kNN} takes values in $\{0, 1\}$. We find evidence that both these quantities display finite-size scaling behavior similar to Eq. (5) which we will use to estimate the critical inverse coupling β_c and the critical exponent of correlation length ν via a curve collapse approach, plotting the observable against $x = N_s^{1/\nu} b$ for multiple lattice sizes simultaneously and finding values of ν and β_c that minimize the distance between the curves using the Nelder-Mead method, as in the procedure described in [74].

If ν is known, we estimate β_c by fitting the peak temperatures $\beta_c(N_s)$ of χ_{kNN} obtained from multiple lattice sizes to the ansatz

$$\beta_c(N_s) - \beta_c(\infty) \propto \frac{1}{N_s^{1/\nu}}. \quad (7)$$

IV. ANALYSIS

A. Detecting twisted boundary conditions

We first investigate the ability of the persistent homology of our filtered complex to identify an inserted thin vortex, obtained using twisted boundary conditions, as a function of β. For $N_s \in \{12, 16, 20\}$, fixing $N_t = 4$, we generate 200 configurations using the Wilson action (1) and 200 configurations using the twisted action (4) for each $\beta \in \{1.5, 1.6, \ldots 2.9\}$. Configurations are generated using the HiRep software [75] with one heat bath step and four overrelaxation steps for each Monte Carlo step and a sample taken every 100 Monte Carlo steps.

Since the inserted vortex forms a closed surface, we expect to observe it in the PH_2 diagram. Moreover, the surface wraps round the periodic boundary of the lattice in the latter two spatial directions. If it is the first such surface to wrap around those dimensions to enter the filtered complex, then we will observe it as a point in PH_2 with

infinite death index since it encloses a two-dimensional hole which remains even in the final complex of the filtered complex (homeomorphic to a 4-torus). Otherwise, it would appear as a point with finite death index. In Fig. 4 we compare the persistence diagrams of individual sampled configurations in each phase using the Wilson action and the twisted action. In the confined phase there is no immediate distinction to be made between the persistence diagrams generated using the different actions. We claim that this is because vortices in this phase percolate throughout the system so there are likely to be many vortices that wrap around the periodic boundary conditions of the lattice. Our inserted vortex may then appear as a single point of finite persistence in PH_2, but the persistence of the system is not affected largely. However in the deconfined phase there is a clear difference. There are unlikely to be any system-generated vortex surfaces that wrap around the lattice so the inserted vortex becomes the first such surface to enter the filtered complex. We therefore observe that one of the PH_2 points with infinite death time has a much lower birth time than the others, allowing us to identify that this point represents our inserted vortex surface. We provide additional evidence for this claim in Appendix D by showing that the 2-cycle responsible for this point spans the same plane as the surface along which the boundary conditions are twisted. Moreover, we observe a significant change in PH_0 and PH_1 with many low persistence points appearing early on in the filtered
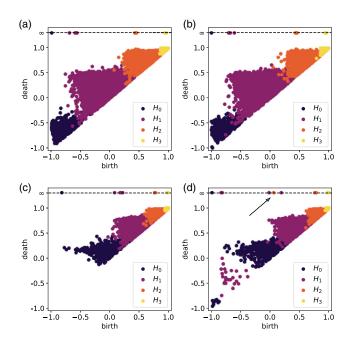

FIG. 4. Sample persistence diagrams of individual configurations obtained using the following actions and values of β: (a) Wilson, $\beta = 1.5$, (b) twisted, $\beta = 1.5$, (c) Wilson, $\beta = 2.9$, and (d) twisted, $\beta = 2.9$. The arrow in (d) indicates the point $(b, \infty) \in PH_2$ with the smallest birth index b. Note the distance between it and the others.
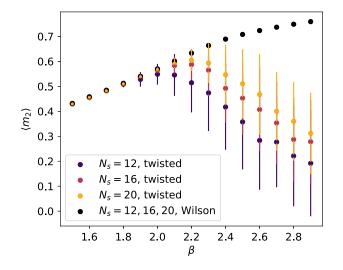
FIG. 5. The expectation value of the observable m_2 as a function of β plotted for different values of N_s and with the Wilson and twisted actions. The difference between the values for different lattice sizes using the Wilson action is not distinguishable at this scale, so they are plotted as the same points. Their error bars are also small enough so as to not be visible.

complex. These arise as different plaquettes of the inserted vortex enter the filtered complex at different indices, forming transient connected components and holes. This is supported by the fact that these points all die by the time we reach the birth index of the point in PH_2.

Following the discussion above, we define an observable based on the persistence diagram of a configuration:

$$m_2 = \min\{b | (b, \infty) \in PH_2\}.$$

The expectation value of m_2 for different lattice sizes with the Wilson action and twisted action are shown in Fig. 5. Note that there is no difference between the expectations estimated using the different actions well into the confined phase (low β), but in the deconfined phase (high β) the curves split apart. As the lattice size increases, the point at which the curves diverge approaches the critical β of the phase transition from below. These observations motivate measuring the difference between the expectation values using different actions

$$O_{m_2} = \langle m_2 \rangle - \langle m_2 \rangle_{\text{twist}}$$

as a phase indicator which will be zero in the confined phase and nonzero in the deconfined phase, similar to the definition of an order parameter but without the requirement to detect any symmetry breaking. A finite-size scaling analysis of this quantity yields the curve collapse in Fig. 6 with estimates of β_c and ν:

$$\beta_c = 2.291 \pm 0.019,$$
$$\nu = 0.614 \pm 0.079,$$

in agreement with the existing estimate $\beta_c = 2.2986(6)$ in Table I. Error estimates are obtained by performing 2000 bootstraps. While the error obtained is reasonably large, it should be stressed that these estimates were obtained using only 200 configurations. In Appendix D we look more closely at the distribution of m_2 as measured using the twisted action in order to better understand the behavior of O_{m_2}.

Note that we fixed the exponent of -2 for the scaling of O_{m_2} with N_s. Attempting to fit this exponent along with β_c and ν often led to the optimizer returning unrealistic large positive values for the exponent, spoiling the error estimation. The value of -2 was found by hand to give a good fit and we offer an heuristic argument for why. First note that N_s^2 is how the number of plaquettes in the inserted vortex surface scales with N_s, since the surface wraps around the periodic boundary conditions. In the case where it describes the formation of the inserted vortex surface, the value of m_2 is determined by the filtration index of the last plaquette to enter the surface. Now the larger the surface, the more likely it is that there will be at least one plaquette in the surface affected by noise or pierced by another vortex, causing it to enter the filtered complex late and dragging the value of m_2 closer to its average in the Wilson action. Assuming that this likelihood is independent for each plaquette or at least approximately linear in the number of plaquettes, we therefore obtain the quadratic scaling in N_s.

B. Investigating deconfinement without twisted boundary conditions

We now investigate whether the persistent homology is able to detect the deconfinement phase transition purely from the Wilson action by making use of a simple machine learning framework inspired by that in [51]. In particular, for $N_t \in \{4, 5, 6\}$ we attempt to estimate the critical inverse coupling β_c and the critical exponent of the correlation length ν via a finite-size scaling analysis of the output of a k-nearest-neighbors classifier trained on the persistence. In each case we repeat the following procedure for each $N_s \in \{12, 16, 20, 24\}$.

(1) Configurations are sampled from a range of values of β (specific values are given in each case below) using the HiRep software [75] with one heat bath step and four overrelaxation steps for each Monte Carlo step and a sample taken every 100 Monte Carlo steps.

(2) For each sample we compute their PH_0, PH_1, PH_2 and PH_3 persistent homology using our filtered complex and compute the corresponding persistence images with a resolution of 25×25 and σ equal to 5% of a pixel. We concatenate the four separate images and flatten them into a $4 \times 25 \times 25 = 2500$-dimensional vector.

(3) We train a kNN model ($k = 30$) to predict the phase of a configuration based on its concatenated persistence image vector by using vectors from well into the confined and deconfined phases.
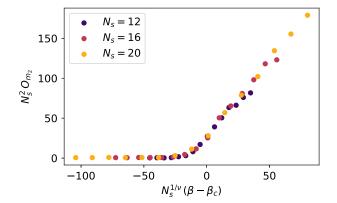
FIG. 6. The curve collapse of our phase indicator O_{m_2} using $\beta_c = 2.291$ and $\nu = 0.614$. Error bars are not shown for clarity but are comparable to those in Fig. 5.

(4) Using the trained classification model, we define an observable $O_{k\mathrm{NN}}$ which is the predicted phase of a given configuration.

(5) Applying multiple histogram reweighting to the variance

$$\chi_{k\mathrm{NN}} = \langle O_{k\mathrm{NN}}^2 \rangle - \langle O_{k\mathrm{NN}} \rangle^2$$
$$= \langle O_{k\mathrm{NN}} \rangle (1 - \langle O_{k\mathrm{NN}} \rangle)$$

(where the second equality follows since $O_{k\mathrm{NN}} \in \{0, 1\}$), we obtain an interpolated curve and a more precise estimate of the location of its peak.

By performing a finite-size scaling analysis of the locations of the peaks obtained for each value of N_s and a curve collapse of the different reweighted variance curves, we obtain estimates of β_c and ν for the deconfinement phase transition at the given value of N_t.

1. $N_t = 4$

For lattices of size $4 \times N_s^3$ with $N_s \in \{12, 16, 20, 24\}$, we train a k-nearest-neighbors classifier ($k = 30$) on the concatenated PH_0, PH_1, PH_2 and PH_3 persistence images of 200 configurations sampled at each β in the confined and deconfined regions given in Table II. The classifier is then used to produce a predicted classification $O_{k\mathrm{NN}}$ for 200

TABLE II. Values of β sampled at for the $N_t = 4$ phase transition.

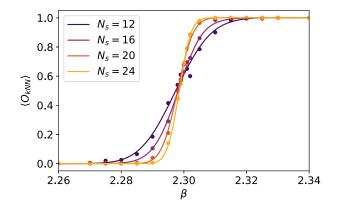| Region	β
Confined	2.2, 2.21, 2.22, 2.23, 2.24
Deconfined	2.36, 2.37, 2.38, 2.39, 2.4
Critical	2.25, 2.26, 2.27, 2.275, 2.28, 2.285, 2.29, 2.295, 2.298, 2.299, 2.3, 2.301, 2.302, 2.305, 2.31, 2.315, 2.32, 2.325, 2.33, 2.34, 2.35

FIG. 7. Plot showing our phase indicator $\langle O_{k\mathrm{NN}} \rangle$ as a function of β for $N_t = 4$. The points show the measured expectations and the curve is the output of histogram reweighting these measurements.

configurations sampled for each value of β from the critical region.

The resulting estimates of the expectation $\langle O_{k\mathrm{NN}} \rangle(\beta)$ are shown in Fig. 7 along with interpolating curves obtained via histogram reweighting. The variance curves $\chi_{k\mathrm{NN}}$ are shown in Fig. 8.

From here we proceed with two separate analyses. In the first we estimate β_c from a linear regression assuming the value of the critical exponent ν to be known. In the second we estimate β_c and ν concurrently via a curve collapse procedure. In both cases we perform two separate bootstraps to obtain 500 bootstrap samples from each. One bootstrap is carried out by resampling the configurations for each β used to train the k-nearest-neighbors classifier. The other is carried out by resampling the configurations for each β used to estimate $\langle O_{k\mathrm{NN}} \rangle$. Applying the finite-size scaling analysis to both collections of bootstraps yields two separate distributions for β_c and two for ν. The error in these quantities is therefore estimated by combining the standard deviation of the distributions coming from the different bootstrap procedures under the assumption that they are independent.
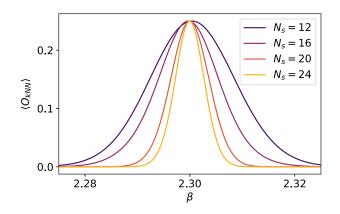

FIG. 8. The variance curves $\chi_{k\mathrm{NN}}$ of $O_{k\mathrm{NN}}$ for $N_t = 4$ to which we will apply our curve collapse procedure.
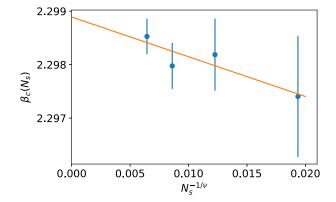
FIG. 9. Estimating β_c for $N_t = 4$. The pseudocritical values of β, obtained from locating the peaks of the variance curves in Fig. 8, are fitted to the ansatz (7). Error bars are estimated by bootstrapping.

By defining the pseudocritical inverse coupling $\beta_c(N_s)$ to be the point at which χ_{kNN} peaks, we can plot $\beta_c(N_s)$ against $N_s^{-1/\nu}$ using a previously estimated [63] value of $\nu = 0.629971$. The result is shown in Fig. 9. We observe that the points plotted with error bars of 1σ support a straight line fit. The intercept yields $\beta_c = 2.2989 \pm 0.0009$, supporting the previously obtained estimate of $\beta_c = 2.2986(6)$ in Table I.

To estimate β_c and ν concurrently we employ a numerical curve collapse procedure, plotting χ_{kNN} against $N_s^{1/\nu}(\beta - \beta_c)$ and tuning β_c and ν to minimize the distance between the curves using the Nelder-Mead method.

The resulting curve collapse is shown in Fig. 10, and the obtained estimates of β_c and ν

$$\beta_c = 2.2988 \pm 0.0007,$$
$$\nu = 0.634 \pm 0.014$$

are consistent with previous estimates.

To confidently claim that this methodology identifies the phase transition, we also tried using alternative values of β
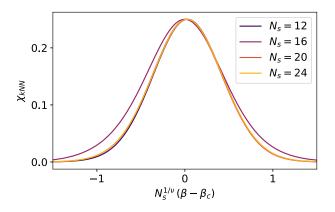

FIG. 10. The curve collapse of χ_{kNN} for $N_t = 4$ using $\beta_c = 2.2988$ and $\nu = 0.634$.

TABLE III. Alternative values of β sampled at for the $N_t = 4$ phase transition to test the sensitivity of the method to the choice of training data.

Region	β
Confined	2.21, 2.22, 2.23
Deconfined	2.38, 2.39, 2.4
Critical	2.24, 2.25, 2.26, 2.27, 2.275, 2.28, 2.285, 2.29, 2.295, 2.298, 2.299, 2.3, 2.301, 2.302, 2.305, 2.31, 2.315, 2.32, 2.325, 2.33, 2.34, 2.35, 2.36, 2.37

to train the kNN classifier, chosen further away from the transition point and so that the transition point is further from the center point between the highest β in the confined phase and the lowest β in the deconfined phase. The alternative training values are shown in Table III.

Using these training values we obtain estimates from the linear fit of

$$\beta_c = 2.2996 \pm 0.0008$$

and the curve collapse of

$$\beta_c = 2.2998 \pm 0.0007,$$
$$\nu = 0.638 \pm 0.013,$$

close to our previous ones and still compatible with our reference estimates.

2. $N_t = 5$

For lattices of size $5 \times N_s^3$ with $N_s \in \{12, 16, 20, 24\}$, we train a k-nearest-neighbors classifier ($k = 30$) on the concatenated PH_0, PH_1, PH_2 and PH_3 persistence images of 200 configurations sampled at each β in the confined and deconfined regions given in Table IV. The classifier is then used to produce a predicted classification O_{kNN} for 200 configurations sampled for each value of β from the critical region.

The resulting estimates of the expectation $\langle O_{kNN}\rangle(\beta)$ are shown in Fig. 11 along with interpolating curves obtained via histogram reweighting.

TABLE IV. Values of β sampled at for the $N_t = 5$ phase transition.

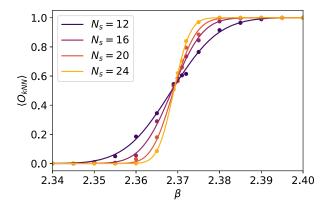| Region	β
Confined	2.29, 2.3, 2.31, 2.32, 2.33
Deconfined	2.41, 2.42, 2.43, 2.44, 2.45
Critical	2.34, 2.345, 2.35, 2.355, 2.36, 2.365, 2.369, 2.37, 2.371, 2.372, 2.375, 2.38, 2.385, 2.39, 2.395, 2.4

FIG. 11. Plot showing our phase indicator $\langle O_{k\mathrm{NN}} \rangle$ as a function of β for $N_t = 5$. The points show the measured expectations and the curve is the output of histogram reweighting these measurements.
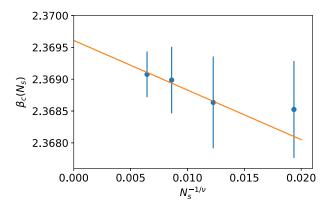

FIG. 12. Estimating β_c for $N_t = 5$. The pseudocritical values of β of the largest three lattice sizes, obtained from locating the peaks of the variance curves, are fitted to the ansatz (7). Error bars are estimated by bootstrapping.

The plot of the pseudocritical $\beta_c(N_s)$ against $N_s^{-1/\nu}$ using $\nu = 0.629971$ is shown in Fig. 12. Here we fit a straight line to the largest three lattice sizes since this gives a better fit than including $N_s = 12$. The intercept yields
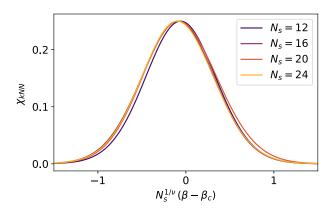

FIG. 13. The curve collapse of $\chi_{k\mathrm{NN}}$ for $N_t = 5$ using $\beta_c = 2.3697$ and $\nu = 0.634$.

TABLE V. Values of β sampled at for the $N_t = 6$ phase transition.

Region	β
Confined	2.33, 2.34, 2.35, 2.36, 2.37
Deconfined	2.49, 2.5, 2.51, 2.52, 2.53
Critical	2.38, 2.39, 2.4, 2.405, 2.41, 2.415, 2.42, 2.425, 2.426, 2.427, 2.428, 2.43, 2.435, 2.44, 2.445, 2.45, 2.455, 2.46, 2.47, 2.48

$\beta_c = 2.3696 \pm 0.0012$ which is just about compatible with the previously obtained estimate of $\beta_c = 2.37136(54)$ in Table I.

We also perform the curve collapse on only the highest three lattice sizes and the result is shown in Fig. 13. The obtained estimates of β_c and ν

$$\beta_c = 2.3697 \pm 0.0011,$$

$$\nu = 0.634 \pm 0.028$$

are consistent with previous estimates.

3. $N_t = 6$

For lattices of size $6 \times N_s^3$ with $N_s \in \{12, 16, 20, 24\}$, we train a k-nearest-neighbors classifier ($k = 30$) on the concatenated PH_0, PH_1, PH_2 and PH_3 persistence images of 200 configurations sampled at each β in the confined and deconfined regions given in Table V. The classifier is then used to produce a predicted classification $O_{k\mathrm{NN}}$ for 200 configurations sampled for each value of β from the critical region.

The resulting estimates of the expectation $\langle O_{k\mathrm{NN}} \rangle(\beta)$ are shown in Fig. 14 along with interpolating curves obtained via histogram reweighting.
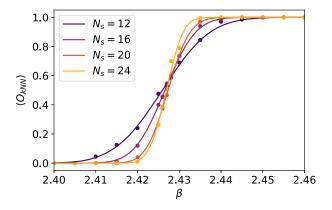

FIG. 14. Plot showing our phase indicator $\langle O_{k\mathrm{NN}} \rangle$ as a function of β for $N_t = 6$. The points show the measured expectations and the curve is the output of histogram reweighting these measurements.
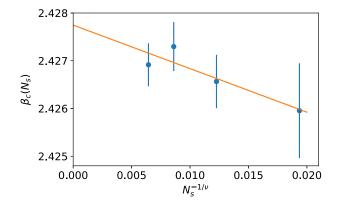
FIG. 15. Estimating β_c for $N_t = 6$. The pseudocritical values of β, obtained from locating the peaks of the variance curves, are fitted to the ansatz (7). Error bars are estimated by bootstrapping.

The plot of the pseudocritical $\beta_c(N_s)$ against $N_s^{-1/\nu}$ using $\nu = 0.629971$ is shown in Fig. 15. The intercept of the straight line fit yields $\beta_c = 2.4277 \pm 0.0008$, supporting the previously obtained estimate of $\beta_c = 2.4271(17)$ in Table I.

The result of the curve collapse is shown in Fig. 16. The obtained estimates of β_c and ν are

$$\beta_c = 2.4276 \pm 0.0008,$$

$$\nu = 0.666 \pm 0.016.$$

The estimate of β_c agrees with the previous estimate; however, the previous estimate of $\nu = 0.629971$ we refer to lies just over 2 standard deviations outside of our estimate. We ascribe this discrepancy to the use of smaller aspect ratios N_s/N_t as we increase N_t, remarking that the purpose of the investigation we have performed at various N_t is to show reasonable scaling with the lattice spacing rather than providing a precision study of critical properties of the system.
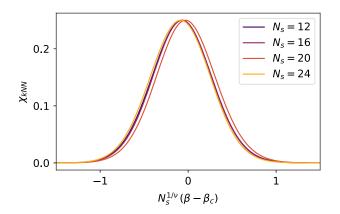

FIG. 16. The curve collapse of $\chi_{k\text{NN}}$ for $N_t = 6$ using $\beta_c = 2.4276$ and $\nu = 0.666$.

V. CONCLUSIONS AND DISCUSSION

In this paper we have developed a gauge-invariant method using persistent homology, a tool from topological data analysis, that is designed to detect center vortices in configurations of SU(2) lattice gauge theory. We defined two different persistence-based phase indicators for the deconfinement phase transition—one making use of twisted boundary conditions, and the other using simple machine learning—and successfully used them to estimate the critical β and exponent ν of the transition.

This method was designed to detect and capture the behavior of vortices and we provided preliminary evidence for this by showing it is possible to detect the use of twisted boundary conditions in the deconfined phase. A closer look at the relationship between the individual points in the resulting persistence diagrams and center vortices is left to a future analysis. In particular, it would be interesting to compare representative cycles for points in the H_2 persistence diagram with the vortex surfaces obtained by gauge fixing and projection. Moreover we showed that it is possible to detect the deconfinement phase transition from the persistent homology of samples using the original action alone. We argue that this has the potential to lend support to the center vortex picture of confinement once the connection between the persistent homology of our filtration and center vortices is fully established, at least by not ruling it out. For a stronger argument we would need to investigate the relationship between our method and other competing pictures. For example, looking at the monopole picture of confinement (see Refs. [76] or [1] for a summary), one could investigate the sensitivity of our method to monopoles. If the method were sensitive to vortices but not to monopoles, the fact that it captures the phase transition would be evidence for the vortex picture over the monopole picture. However, this is work yet to be done, which will require nontrivial adaptations of some of the steps used here. For instance, if the current method is indeed sensitive to monopoles, then one might instead attempt to devise a filtration that exposes monopolelike singularities but not vortices.

Besides a greater degree of interpretability, another advantage of this method over machine learning approaches based on deep learning is that we were able to obtain our results using only a small number of sampled configurations. This is particularly important in view of extending our methodology to full QCD, for which numerical computations for generating gauge configurations near the physical point are very demanding and hence the number of configurations one can use is generally limited.

ACKNOWLEDGMENTS

APPENDIX A: CUBICAL COMPLEXES AND HOMOLOGY

This is a very compressed version of the exposition found in [81]. An *elementary interval* is an interval of the form $[i, i+1] \subset \mathbb{R}$ (*nondegenerate*) or $[i, i] = \{n\}$ (*degenerate*) for some choice of $i \in \mathbb{Z}$. An *elementary cube* is a finite product of elementary intervals $Q = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$, where n is some fixed *embedding dimension*. Its *dimension* $\dim Q$ is the number of nondegenerate intervals in the product. A *cubical complex* C is a subset of $\mathbb{R}^n$ that is a union of elementary cubes. Specifying a field $\mathbf{F}$, we define $\mathbf{F}$-vector spaces $C_k = \{\sum \alpha_i Q_i | Q_i \subseteq C, \dim Q_i = k, \alpha_i \in \mathbf{F}\}$ for each $k \in \mathbb{N}$, consisting of finite formal sums of elementary cubes. The *boundary* of a nondegenerate elementary interval is given by the formal sum $\partial[i, i+1] = [i+1, i+1] - [i, i]$. For a degenerate elementary interval the boundary is zero. The boundary of an elementary cube $Q = (I_1 \times \ldots \times I_n)$ is a formal sum

$$\partial Q = \sum_{j=1}^{n} (-1)^{\sum_{i=1}^{j-1} \dim Q_i} (I_1 \times \cdots \times \partial I_j \times \cdots \times I_n), \quad \text{(A1)}$$

where we consider $\times$ as distributing over the formal summation. We can see that for $\dim Q \geq 1$ we have $\dim \partial Q = \dim Q - 1$. Therefore we can extend ∂ to linear maps $\partial_k : C_k \to C_{k-1}$ via the mapping $\sum \alpha_i Q_i \mapsto \sum \alpha_i (\partial Q_i)$. Since $\partial \partial I = 0$ for any elementary interval I, we also see that $\partial_k \circ \partial_{k+1} = 0$ for all $k \in \mathbb{N}$, so that $\mathrm{im} \partial_{k+1} \subseteq \ker \partial_k$. A sequence of linear maps

$$\cdots \to C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0 \xrightarrow{\partial_0} 0$$
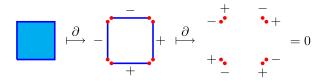

FIG. 17. Example of how the boundary operator ∂ acts on a simple cubical complex consisting of a single two-dimensional cube. Note how the sum in Eq. (A1) being alternating ensures that $\partial \partial = 0$.

with this property is called a *chain complex*. The kth *cubical homology* of C over $\mathbf{F}$ is defined to be the quotient vector space

$$H_k(C; \mathbf{F}) = \frac{\ker \partial_k}{\mathrm{im} \partial_{k+1}}.$$

This construction is functorial; i.e., given a suitable definition of a *cubical map* $f : C \to D$ between cubical complexes, there is an induced map $f_k : H_k(C; \mathbf{F}) \to H_k(D; \mathbf{F})$ for each $k \in \mathbb{N}$. We will not introduce these, in general, but will note that, given $C \subseteq D$, the inclusion map $C \hookrightarrow D$ is cubical and hence induces maps on homology.

APPENDIX B: HISTOGRAM REWEIGHTING

We make use of the correspondence between Euclidean field theory and statistical mechanics in order to apply histogram reweighting. This allows us to express the ensemble average of an observable O at inverse coupling β' in terms of averages at any other β according to the equation

$$\langle O \rangle_{\beta'} = \frac{\langle O e^{-(\beta'-\beta)S} \rangle_\beta}{\langle e^{-(\beta'-\beta)S} \rangle_\beta}, \quad \text{(B1)}$$

where S is the action of the configuration (slightly redefined to pull the factor of β outside) [82]. However, in practice we can only reweight so far, so that the sampled action distributions for β and β' have a sizable overlap. To reliably extrapolate to a wider region we can make use of multiple histogram reweighing [83] where we sample at multiple inverse couplings $\beta_1, \ldots, \beta_R$. Suppose we sample N_i configurations at β_i, and then we can iterate the equation

$$e^{-f_\beta} = \sum_{i=1}^{R} \sum_{a=1}^{N_i} \frac{g_i^{-1} e^{-\beta S_i^a}}{\sum_{j=1}^{R} N_j g_j^{-1} e^{-\beta_j S_i^a + f_j}}$$

to estimate the "free energies" $f_i = f_{\beta_i}$ at β_i up to an additive constant, where each g_i is a quantity related to the

integrated autocorrelation of the samples in run i. Given the f_i we can estimate

$$\langle O \rangle_{\beta'} = \sum_{i=1}^{R} \sum_{a=1}^{N_i} \frac{O_i^a g_i^{-1} e^{-\beta_k S_i^a + f_{\beta'}}}{\sum_{j=1}^{R} N_j g_j^{-1} e^{-\beta_j S_i^a + f_j}}.$$

APPENDIX C: BOOTSTRAP ERROR ESTIMATION

In order to make any reasonable conclusions from the results of our analysis we need to be able to estimate the error in any numerical values obtained. While the error in ensemble averages can be directly estimated from the sample, we also calculate various fits to the data. The way in which error propagates here is not necessarily easy to calculate directly. Recall that the idea of bootstrap analysis is to sidestep these concerns by estimating the sampling distribution of a statistic directly. Suppose we obtain N sampled configurations $S = \{\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N\}$ and calculate some numerical statistic $f(S)$ from the data. Given some preset integer N_B, bootstrap analysis proceeds by

(1) resampling S with replacement N_B times to obtain samples $S_1, \ldots, S_{N_B}$ each of size N and then
(2) computing $f(S_i)$ for each $i \in \{1, \ldots, N_B\}$.

For large enough N_B, the distribution of the $f(S_i)$ approximates the sampling distribution of f and we can estimate the standard error

$$\sigma_f \approx \sqrt{\frac{1}{N_B - 1} \sum_i (f(S_i) - \overline{f(S_j)})^2}.$$

APPENDIX D: DISTRIBUTION OF m_2

Motivated to understand the error bars for the estimates of $\langle m_2 \rangle_{\text{twist}}$ in Fig. 5, we look at the distribution of m_2 as measured using the twisted action on a 4×20^3 lattice in Fig. 18. We recenter the data using $\langle m_2 \rangle$ in order to compare the effect on O_{m_2} at different values of β.

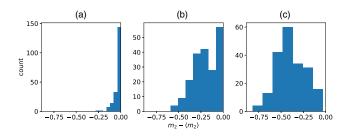We see that at lower values of β in the confined phase, the value of m_2 remains close to the average value

measured with the untwisted Wilson action. As β increases we observe a bimodal distribution, with some configurations maintaining an m_2 value close to the untwisted average and some joining a lower mode. A likely explanation for this behavior is that for those configurations in the zero mode, the H_2 generator of the 4-torus responsible for m_2 does not correspond to the inserted vortex in the yz plane, either because there is another vortex spanning a plane which is more easily observed or because the formation of a complete vortex surface along the twist in the filtered complex is being impeded, perhaps by intersection with other vortices. By the time we are firmly in the deconfined phase, the majority of configurations lie in the lower mode, signalling that the inserted vortex along the yz plane is responsible for the value of m_2.

To verify this picture, we can check if the H_2 generator responsible for m_2 is indeed represented by a yz plane. We do this by recomputing the persistent homology but with the cubical complex only being periodic in the y and z directions and open in the t and x directions. This makes the final complex homeomorphic to a 2-torus which has a single H_2 generator: the yz plane. Then we can simply check whether or not the birth time of this generator is the same as m_2. Denote by I_{yz} the indicator function for a configuration that is 1 if they match and 0 if they do not. That is, I_{yz} tells us if m_2 is determined by a H_2 generator which spans the yz plane. The mean and variance of this indicator as a function of β are shown in Fig. 19.

We see that at low β the inserted yz vortex sheet is never responsible for the value of m_2. There are likely to be many vortices and those which span planes including the t direction are smaller and likely to be picked up sooner in the persistence. Approaching and passing the transition point, the proportion of configurations for which m_2 describes the birth time of a yz plane increases until close to 1. There are fewer dynamically generated vortices and instead m_2 is determined by the one we inserted via the twisted boundary conditions. We note that peak in the variance of I_{yz} gives us an alternative marker for the phase transition.

FIG. 18. The distribution of $m_2 - \langle m_2 \rangle$ with the twisted action at (a) $\beta = 2.2$, (b) $\beta = 2.5$, and (c) $\beta = 2.8$. Note that $\langle m_2 \rangle$ is the expectation as measured with the Wilson action.
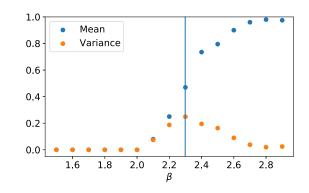

FIG. 19. The mean and variance of I_{yz} as a function of β. The vertical line marks the location of β_c.

[1] J. Greensite, *An Introduction to the Confinement Problem*, Lecture Notes in Physics (Springer, Berlin, 2011).

[2] S. J. Wetzel and M. Scherzer, Machine learning of explicit order parameters: From the Ising model to SU(2) lattice gauge theory, Phys. Rev. B **96,** 184410 (2017).

[3] D. L. Boyda, M. N. Chernodub, N. V. Gerasimeniuk, V. A. Goy, S. D. Liubimov, and A. V. Molochkov, Finding the deconfinement temperature in lattice Yang-Mills theories from outside the scaling window with machine learning, Phys. Rev. D **103,** 014509 (2021).

[4] A. Palermo, L. Anderlini, M. P. Lombardo, A. Kotov, and A. Trunin, Machine learning approaches to the QCD transition, Proc. Sci. LATTICE2021 (**2022**) 030.

[5] C. Giannetti, B. Lucini, and D. Vadacchino, Machine learning as a universal tool for quantitative investigations of phase transitions, Nucl. Phys. **B944,** 114639 (2019).

[6] G. 't Hooft, On the phase transition towards permanent quark confinement, Nucl. Phys. **B138,** 1 (1978).

[7] J. M. Cornwall, Quark confinement and vortices in massive gauge-invariant QCD, Nucl. Phys. **B157,** 392 (1979).

[8] G. 't Hooft, A property of electric and magnetic flux in nonabelian gauge theories, Nucl. Phys. **B153,** 141 (1979).

[9] L. Del Debbio, M. Faber, J. Greensite, and S. Olejnik, Center dominance and Z(2) vortices in SU(2) lattice gauge theory, Phys. Rev. D **55,** 2298 (1997).

[10] M. Faber, J. Greensite, and S. Olejnik, Casimir scaling from center vortices: Towards an understanding of the adjoint string tension, Phys. Rev. D **57,** 2603 (1998).

[11] L. Del Debbio, M. Faber, J. Giedt, J. Greensite, and S. Olejnik, Detection of center vortices in the lattice Yang-Mills vacuum, Phys. Rev. D **58,** 094501 (1998).

[12] R. Bertle, M. Faber, J. Greensite, and S. Olejnik, The structure of projected center vortices in lattice gauge theory, J. High Energy Phys. 03 (1999) 019.

[13] M. Faber, J. Greensite, S. Olejnik, and D. Yamada, The vortex finding property of maximal center (and other) gauges, J. High Energy Phys. 12 (1999) 012.

[14] M. Engelhardt, K. Langfeld, H. Reinhardt, and O. Tennert, Deconfinement in SU(2) Yang-Mills theory as a center vortex percolation transition, Phys. Rev. D **61,** 054504 (2000).

[15] P. de Forcrand and M. D'Elia, On the Relevance of Center Vortices to QCD, Phys. Rev. Lett. **82,** 4582 (1999).

[16] M. Engelhardt and H. Reinhardt, Center projection vortices in continuum Yang-Mills theory, Nucl. Phys. **B567,** 249 (2000).

[17] M. Engelhardt, Center vortex model for the infrared sector of Yang-Mills theory: Topological susceptibility, Nucl. Phys. **B585,** 614 (2000).

[18] R. Bertle, M. Faber, J. Greensite, and S. Olejnik, *P* vortices, gauge copies, and lattice size, J. High Energy Phys. 10 (2000) 007.

[19] K. Langfeld, H. Reinhardt, and J. Gattnar, Gluon propagators and quark confinement, Nucl. Phys. **B621,** 131 (2002).

[20] J. Greensite, The confinement problem in lattice gauge theory, Prog. Part. Nucl. Phys. **51,** 1 (2003).

[21] F. Bruckmann and M. Engelhardt, Writhe of center vortices and topological charge: An explicit example, Phys. Rev. D **68,** 105011 (2003).

[22] M. Engelhardt, M. Quandt, and H. Reinhardt, Center vortex model for the infrared sector of SU(3) Yang-Mills theory: Confinement and deconfinement, Nucl. Phys. **B685,** 227 (2004).

[23] P. Y. Boyko, V. G. Bornyakov, E. M. Ilgenfritz, A. V. Kovalenko, B. V. Martemyanov, M. Muller-Preussker, M. I. Polikarpov, and A. I. Veselov, Once more on the interrelation between Abelian monopoles and *P*-vortices in SU(2) LGT, Nucl. Phys. **B756,** 71 (2006).

[24] E.-M. Ilgenfritz, K. Koller, Y. Koma, G. Schierholz, T. Streuer, V. Weinberg, and M. Quandt, Localization of overlap modes and topological charge, vortices and monopoles in SU(3) LGT, Proc. Sci. LATTICE2007 (**2007**) 311 [arXiv:0710.2607].

[25] V. G. Bornyakov, E. M. Ilgenfritz, M. Muller-Preussker, B. V. Martemyanov, S. M. Morozov, and A. I. Veselov, Interrelation between monopoles, vortices, topological charge and chiral symmetry breaking: Analysis using overlap fermions for SU(2), Phys. Rev. D **77,** 074507 (2008).

[26] A. O'Cais, W. Kamleh, K. Langfeld, B. Lasscock, D. Leinweber, P. Moran, A. Sternbeck, and L. von Smekal, Preconditioning maximal center gauge with stout link smearing in SU(3), Phys. Rev. D **82,** 114512 (2010).

[27] M. Engelhardt, Center vortex model for the infrared sector of SU(3) Yang-Mills theory: Topological susceptibility, Phys. Rev. D **83,** 025015 (2011).

[28] P. O. Bowman, K. Langfeld, D. B. Leinweber, A. Sternbeck, L. von Smekal, and A. G. Williams, Role of center vortices in chiral symmetry breaking in SU(3) gauge theory, Phys. Rev. D **84,** 034501 (2011).

[29] E.-A. O'Malley, W. Kamleh, D. Leinweber, and P. Moran, SU(3) centre vortices underpin confinement and dynamical chiral symmetry breaking, Phys. Rev. D **86,** 054503 (2012).

[30] A. Trewartha, W. Kamleh, and D. Leinweber, Connection between center vortices and instantons through gauge-field smoothing, Phys. Rev. D **92,** 074507 (2015).

[31] A. Trewartha, W. Kamleh, and D. Leinweber, Evidence that centre vortices underpin dynamical chiral symmetry breaking in SU(3) gauge theory, Phys. Lett. B **747,** 373 (2015).

[32] J. Greensite, Confinement from center vortices: A review of old and new results, EPJ Web Conf. **137,** 01009 (2017).

[33] A. Trewartha, W. Kamleh, and D. B. Leinweber, Centre vortex removal restores chiral symmetry, J. Phys. G **44,** 125002 (2017).

[34] J. C. Biddle, W. Kamleh, and D. B. Leinweber, Gluon propagator on a center-vortex background, Phys. Rev. D **98,** 094504 (2018).

[35] F. Spengler, M. Quandt, and H. Reinhardt, Branching of center vortices in SU(3) lattice gauge theory, Phys. Rev. D **98,** 094508 (2018).

[36] J. Biddle, W. Kamleh, and D. Leinweber, Static quark potential from centre vortices in the presence of dynamical fermions, Phys. Rev. D **106,** 054505 (2022).

[37] L. Del Debbio, A. Di Giacomo, and B. Lucini, Vortices, monopoles and confinement, Nucl. Phys. **B594,** 287 (2001).

[38] L. Del Debbio, A. Di Giacomo, and B. Lucini, Monopoles, vortices and confinement in SU(3) gauge theory, Phys. Lett. B **500,** 326 (2001).

[39] R. Golubich and M. Faber, A possible resolution to troubles of SU(2) center vortex detection in smooth lattice configurations, Universe **7**, 122 (2021).

[40] V. Gribov, Quantization of non-abelian gauge theories, Nucl. Phys. **B139**, 1 (1978).

[41] J. D. Stack and W. W. Tucker, The Gribov ambiguity for maximal Abelian and center gauges in SU(2) lattice gauge theory, Nucl. Phys. B **94**, 529 (2001), Proceedings of the XVIIIth International Symposium on Lattice Field Theory.

[42] P. de Forcrand and M. Pepe, Center vortices and monopoles without lattice Gribov copies, Nucl. Phys. **B598**, 557 (2001).

[43] H. Edelsbrunner, D. Letscher, and A. Zomorodian, Topological persistence and simplification, Discrete Comput. Geom. **28**, 511 (2002).

[44] L. Speidel, H. A. Harrington, S. J. Chapman, and M. A. Porter, Topological data analysis of continuum percolation with disks, Phys. Rev. E **98**, 012318 (2018).

[45] T. Hirakida, K. Kashiwa, J. Sugano, J. Takahashi, H. Kouno, and M. Yahiro, Persistent homology analysis of deconfinement transition in effective Polyakov-line model (2018).

[46] Q. H. Tran, M. Chen, and Y. Hasegawa, Topological persistence machine of phase transitions, Phys. Rev. E **103**, 052127 (2021).

[47] B. Olsthoorn, J. Hellsvik, and A. V. Balatsky, Finding hidden order in spin models with persistent homology, Phys. Rev. Res. **2**, 043308 (2020).

[48] A. Cole, G. J. Loges, and G. Shiu, Quantitative and interpretable order parameters for phase transitions from persistent homology, Phys. Rev. B **104**, 104426 (2021).

[49] I. Donato, M. Gori, M. Pettini, G. Petri, S. De Nigris, R. Franzosi, and F. Vaccarino, Persistent homology analysis of phase transitions, Phys. Rev. E **93**, 052138 (2016).

[50] D. Spitz, J. Berges, M. K. Oberthaler, and A. Wienhard, Finding universal structures in quantum many-body dynamics via persistent homology, SciPost Phys. **11**, 060 (2021).

[51] N. Sale, J. Giansiracusa, and B. Lucini, Quantitative analysis of phase transitions in two-dimensional xy models using persistent homology, Phys. Rev. E **105**, 024121 (2022).

[52] G. Arias-Tamargo, Y.-H. He, E. Heyes, E. Hirst, and D. Rodriguez-Gomez, Brain webs for brane webs, Phys. Lett. B **833**, 137376 (2022).

[53] D. S. Berman, Y.-H. He, and E. Hirst, Machine learning Calabi-Yau hypersurfaces, Phys. Rev. D **105**, 066002 (2022).

[54] Y. He, S. Xia, D. G. Angelakis, D. Song, Z. Chen, and D. Leykam, Persistent homology analysis of a generalized Aubry-André-Harper model, Phys. Rev. B **106**, 054210 (2022).

[55] D. Leykam and D. G. Angelakis, Topological data analysis and machine learning (2022).

[56] K. Kashiwa, T. Hirakida, and H. Kouno, Persistent homology analysis for dense QCD effective model with heavy quarks, Symmetry **14**, 1783 (2022).

[57] D. Sehayek and R. G. Melko, Persistent homology of F_2 gauge theories, Phys. Rev. B **106**, 085111 (2022).

[58] M. Engelhardt, K. Langfeld, H. Reinhardt, and O. Tennert, Deconfinement in SU(2) Yang-Mills theory as a center

[59] vortex percolation transition, Phys. Rev. D **61**, 054504 (2000).

[59] Greensite Jeff, Confinement from center vortices: A review of old and new results, EPJ Web Conf. **137**, 01009 (2017).

[60] M. Faber, J. Greensite, and S. Olejník, Direct Laplacian center gauge, J. High Energy Phys. 11 (2001) 053.

[61] B. Lucini, M. Teper, and U. Wenger, The high temperature phase transition in SU(N) gauge theories, J. High Energy Phys. 01 (2004) 061.

[62] B. Svetitsky and L. G. Yaffe, Critical behavior at finite-temperature confinement transitions, Nucl. Phys. **B210**, 423 (1982).

[63] F. Kos, D. Poland, D. Simmons-Duffin, and A. Vichi, Precision islands in the Ising and $o(n)$ models, J. High Energy Phys. 08 (2016) 036.

[64] G. 't Hooft, A property of electric and magnetic flux in non-Abelian gauge theories, Nucl. Phys. **B153**, 141 (1979).

[65] E. T. Tomboulis and L. G. Yaffe, Finite temperature SU(2) lattice gauge theory, Commun. Math. Phys. **100**, 313 (1985).

[66] G. Carlsson, Topology and data, Bull. Am. Math. Soc. **46**, 255 (2009).

[67] G. Carlsson, Persistent homology and applied homotopy theory, arXiv:2004.00738.

[68] H. Edelsbrunner and J. Harer, Persistent homology—A survey, Discrete Comput. Geom. **453** (2008).

[69] N. Otter, M. Porter, U. Tillmann, P. Grindrod, and H. Harrington, A roadmap for the computation of persistent homology, EPJ Data Sci. **6** (2015).

[70] R. Ghrist, Barcodes: The persistent topology of data, Bull. Am. Math. Soc. **45**, 61 (2008).

[71] H. Edelsbrunner and J. Harer, *Computational Topology—An Introduction* (American Mathematical Society, Providence, 2010), pp. I–XII, 1–241.

[72] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier, Persistence images: A stable vector representation of persistent homology, J. Mach. Learn. Res. **18**, 218 (2017).

[73] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, *Stability of Persistence Diagrams* (2005), pp. 263–271.

[74] S. M. Bhattacharjee and F. Seno, A measure of data collapse for scaling, J. Phys. A **34**, 6375 (2001).

[75] L. Del Debbio, A. Patella, and C. Pica, Higher representations on the lattice: Numerical simulations, SU(2) with adjoint fermions, Phys. Rev. D **81**, 094503 (2010).

[76] G. t. Hooft and F. Bruckmann, Monopoles, instantons and confinement (2000).

[77] N. Sale, B. Lucini, and J. Giansiracusa, Probing center vortices and deconfinement in SU(2) lattice gauge theory with persistent homology—data release, Zenodo, 10.5281/zenodo.7060072 (2022).

[78] N. Sale, B. Lucini, and J. Giansiracusa, Probing center vortices and deconfinement in SU(2) lattice gauge theory with persistent homology—analysis workflow (2022), 10.5281/zenodo.7060223.

[79] G. Tauzin, U. Lupo, L. Tunstall, J. B. Pérez, M. Caorsi, A. Medina-Mardones, A. Dassatti, and K. Hess, giotto-tda: A topological data analysis toolkit for machine learning and data exploration, arXiv:2004.02551.

[80] M. R. Shirts and J. D. Chodera, Statistically optimal analysis of samples from multiple equilibrium states., J. Chem. Phys. **129**, 124105 (2008).

[81] T. Kaczynski, K. Mischaikow, and M. Mrozek, *Computational Homology*, Applied Mathematical Sciences (Springer, New York, 2004).

[82] A. M. Ferrenberg and R. H. Swendsen, New Monte Carlo Technique for Studying Phase Transitions, Phys. Rev. Lett. **61**, 2635 (1988).

[83] A. M. Ferrenberg and R. H. Swendsen, Optimized Monte Carlo Data Analysis, Phys. Rev. Lett. **63**, 1195 (1989).