Persistent homology of quantum entanglement

Bart Olsthoorn ⬤[*]

Nordita, KTH Royal Institute of Technology and Stockholm University, Hannes Alfvéns väg 12, SE-106 91 Stockholm, Sweden

Structure in quantum entanglement entropy is often leveraged to focus on a small corner of the exponentially large Hilbert space and efficiently parametrize the problem of finding ground states. A typical example is the use of matrix product states for local and gapped Hamiltonians. We study the structure of entanglement entropy using persistent homology, a relatively new method from the field of topological data analysis. The inverse quantum mutual information between pairs of sites is used as a distance metric to form a filtered simplicial complex. Both ground states and excited states of common spin models are analyzed as an example. Furthermore, the effect of homology with different coefficients and boundary conditions is also explored. Beyond these basic examples, we also discuss the promising future applications of this modern computational approach, including its connection to the question of how space-time could emerge from entanglement.

I. INTRODUCTION

Problems in quantum physics are often difficult to solve due to an exponentially large Hilbert space (e.g., $d = 2^N$ for N spins), which limits exact diagonalization to small systems only. However, due to the remarkable entanglement scaling properties found in many physical systems, it is often possible to focus only on a small corner of the Hilbert space. For example, the problem of finding a ground state of a one-dimensional local and gapped Hamiltonian can be parameterized efficiently with matrix product states [1]. Beyond this class of Hamiltonians, the entanglement properties have also proven useful in projected entangled-pair states [2] and the multiscale entanglement renormalizaton ansatz (MERA) [3].

Phases of matter can differ in their entanglement scaling properties, and the critical phase transition is typically a point of special interest. The area law states that the entanglement entropy only depends on the surface between two subregions of the system [4,5]. The transverse field Ising model (TFIM) exhibits a quantum phase transition (QPT) from the ferromagnet with area law entanglement to a paramagnet that also follows an area law. At the quantum critical point (QCP), the entanglement entropy diverges logarithmically with the subsystem size (volume law) [6–8], thus allowing the identification of the phase transition. Another example of a phase transition that is studied through entanglement properties is the many-body localization (MBL) transition [9,10]. In gen-

eral, a randomly selected state in Hilbert space typically follows a volume law.

Considering the daunting complexity of Hilbert space, various large-scale computational techniques are employed: Phase transitions and order parameters are now often analyzed and constructed with machine learning [11,12].

Here we outline the proposal to use persistent homology to identify entanglement structures and QPTs. Persistent homology analysis has been used to identify phase transitions in classical spin models [13–18] and some quantum systems [14,19,20]. These examples also illustrate the evolving views on the importance and unique role that entanglement plays as an indicator of qualitative changes in quantum systems. Our approach focuses on the objects that are derivative of quantum states (simplicial complexes) and function as a key descriptor of global properties of the system. The main technique used in this paper is persistent homology, a relatively new method from the field of topological data analysis (TDA) that computes the shapes present in data [21–24]. In short, TDA often involves turning a point cloud in data space into a filtered simplicial complex where the filtration is done with a chosen length scale parameter. At each filtration stage, the homology groups are computed and compared. Homology group elements that persist over many stages lead to persistent homology and can be visualized in a barcode or persistence diagram. For more details and a practical introduction to persistent homology, see Ref. [25].

In this paper, we study the geometric and topological features of the entanglement entropy with persistent homology. The logic of the approach is schematically shown in Fig. 1. A state $|\psi\rangle$ in Hilbert space is converted to a corresponding barcode that efficiently describes its entanglement structure. The barcode is the result of applying persistent homology to a distance matrix that quantifies the entanglement between all subsystems. Using this approach, we discuss the changes in barcode space comparing ground states, excited states, and phase transitions. As an example, we focus on two common
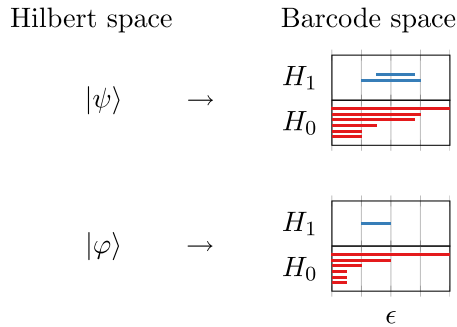
*[*]bartol@kth.se

Hilbert space

Barcode space

FIG. 1. A given state ψ in Hilbert space is mapped to a barcode that describes the topological structure of the entanglement of its subsystems. The proximity parameter ϵ is computed from the quantum mutual information and the homology groups indicate features in the emergent geometry.

quantum spin chains in a transverse field: The Ising model and XXZ model. The former is the simplest example of a QPT, and the latter is a model commonly used to study MBL [10,26].

The remainder of this paper is structured as follows. In Sec. II, we outline how persistent homology captures the entanglement structure of a quantum state. An introduction to simplicial homology, persistence, and our distance metric are presented in Sec. III–V. A demonstration is carried out for two different spin models in Secs. VI and VII, followed by a discussion in Sec. VIII. We conclude in Sec. IX.

II. ENTANGLEMENT AND PERSISTENT HOMOLOGY

The foundation of our paper is the description of the entanglement structure of a quantum state using persistent homology. Given a state $|\psi\rangle$, we divide the quantum system into N subsystems. For each pair of subsystems, the quantum mutual information (MI) M_{ij} is computed. A structure appears when entangled subsystems (with large MI) form clusters. A distance metric is defined as the inverse MI between two subsystems. This brings strongly entangled subsystems close, while nonentangled subsystems are far apart. Given a distance matrix between all subsystems, the computation of persistent homology groups makes it possible to study topological and geometrical features of the entanglement structure. This information can be plotted as a barcode, where long bars indicate topological features that persist over large length scales. The length scale refers to the entanglement through the defined distance metric. The barcode captures the global entanglement structure of the quantum state and gives a detailed view of the phase portrait of the system. How to compute this in practice is outlined in the next section.

The barcode is a summary of the topological and geometrical features of the entanglement structure. Each quantum state has a barcode and it can be used in a number of applications. In our paper, we demonstrate that changes in the barcode reveal phase transitions. At the same time, the barcode reveals the length scale of the entanglement in the state which is important when developing suitable wave function ansatze. Finally, the barcode describes the emergent geometry of the entanglement, which could provide a starting point to study the emergence of space-time from entanglement.

The barcodes corresponding to quantum states can be compared in a number of ways. Distance measures such as the bottleneck distance [27] or the Wasserstein distance [28] operate directly on the complete barcodes rather than derived quantities. However, depending on the application, it can be sufficient to simply count the number of bars at a specific length scale (i.e., the Betti number β_k). We focus on Betti numbers in our paper and demonstrate its sensitivity to QPTs. However, the computational algorithm outlined here is general and can be used in a variety of applications.

III. SIMPLICIAL HOMOLOGY

Simplices are the building blocks for topological spaces and are convenient due to their combinatorial nature. A k-simplex consists of $k+1$ vertices, i.e., $S = [v_0 v_1 \ldots v_k]$ (see Fig. 3). Each simplex can be oriented in two ways, $-S$ or $+S$, however, this feature is not always used in persistent homology, as we will see. Combining simplices and connecting them forms a simplicial complex that represents a topological space. This section introduces the concepts of simplicial homology, namely, chains, boundaries, cycles, and homology groups.

For each definition, we also study an example. The example simplicial complex K corresponds to topological space that is a disk glued to a circle,

$$K = \{ \underbrace{[012]}_{\text{2-simplex}}, \underbrace{[01], [12], [20], \ldots}_{\text{1-simplices}}, \underbrace{[0], [1], \ldots}_{\text{0-simplices}} \}$$

(1)

as becomes clear once we compute its homology.

A. Chains

A k-chain is a sum of k-simplices in K. All the possible k-chains form a group. For the example in Eq. (1), there are three chain groups: C_0, C_1, and C_2. The 1-chains are, for example,

$$C_1 = \{a_{01}[01] + a_{12}[12] + a_{20}[20] + a_{13}[13] + a_{23}[23]\}$$
$$= \langle [01], [12], [20], [13], [23] \rangle, \qquad (2)$$

where the 1-simplices form a basis. An element in the chain group is denoted by $c \in C_k$, and this element represents a specific k-chain.

B. Boundaries

The boundary operator ∂ for a k-simplex is defined as

$$\partial S = \sum_{j=0}^{k} (-1)^j [v_0 v_1 \ldots \hat{v}_j \ldots v_k], \qquad (3)$$

where $\hat{v}_j$ is removed from the sequence. In other words, it is a sum of the faces of the simplex, where the faces are $(k-1)$-simplices. The *Fundamental Lemma of Homology* [29] states
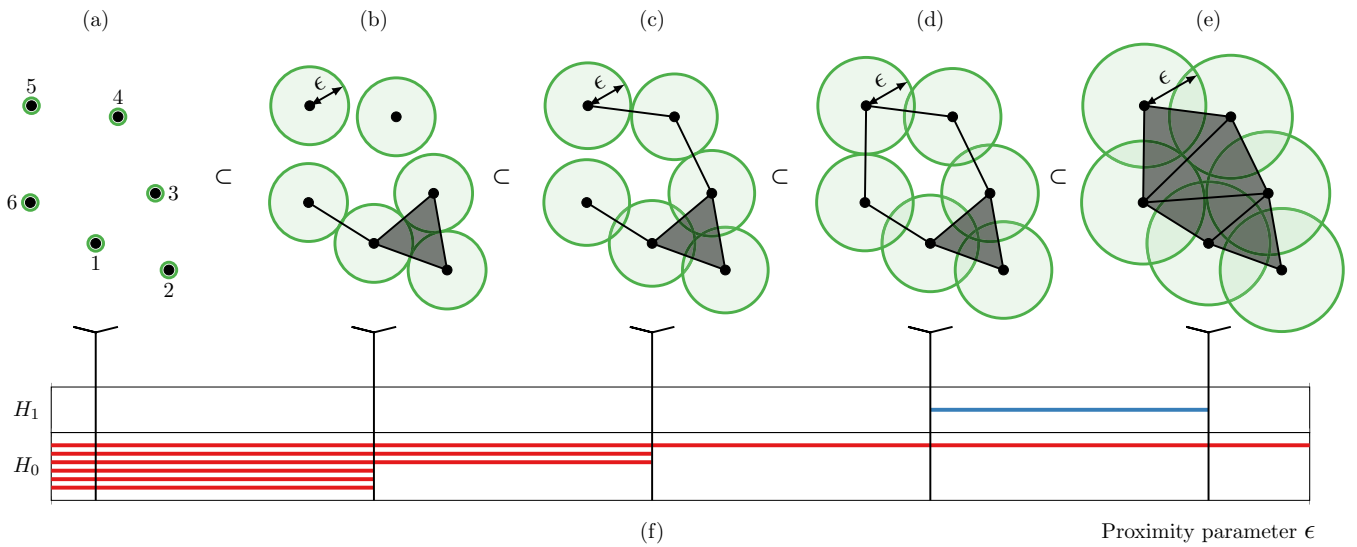
FIG. 2. Vietoris-Rips complex of a point cloud with six data points (a) at increasing proximity parameter ϵ (b)–(e). The barcode (f) shows the persistent homology.

that applying the boundary twice always leads to zero, i.e., $\partial\partial S = 0$. The boundary operator is typically applied to a chain group C_k, leading to a boundary group B_k. For Eq. (1), the boundary group:

$$
\begin{aligned}
B_0 &= \partial_1 C_1 \\
&= \{a_{01}([1] - [0]) + a_{12}([2] - [1]) \\
&\quad + a_{20}([0] - [2]) + a_{13}([3] - [1]) \\
&\quad + a_{23}([3] - [2])\} \\
&= \{(-a_{01} + a_{20})[0] + (a_{01} - a_{12} - a_{13})[1] \\
&\quad + (a_{12} - a_{20} - a_{23})[2] + (a_{13} + a_{23})[3]\}.
\end{aligned} \tag{4}
$$

C. Cycles

A k-cycle is a k-chain with an empty boundary, i.e., $\partial_k c = 0$. This group is denoted by Z_k, where Z stands for the German *zyklus* (cycle). By definition, the k-cycles are a subset of all the k-chains, i.e., $Z_i \subset C_i$. Mathematically, it corresponds to taking the kernel of the boundary operator, i.e., all the elements that are mapped to the identity element. In
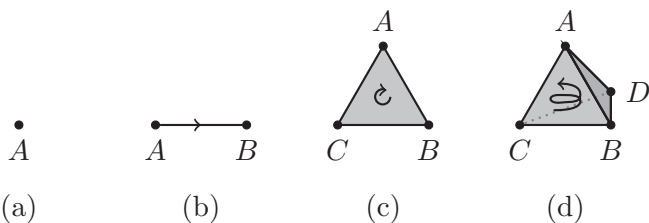

(a)   (b)   (c)   (d)

FIG. 3. Here we show (a) 0-simplex (point), (b) 1-simplex (line segment), (c) 2-simplex (filled triangle), (d) 3-simplex (filled tetrahedron). Higher-dimensional simplices exist but are not shown here for simplicity. Simplices are glued together to form a simplicial complex and its topological properties are described by simplicial homology. Entangled quantum subsystems A, B, C and D are represented by 0-simplices. A k-simplex consists of $k + 1$ subsystems.

other words, the k-cycle group is

$$
Z_k = \text{Ker}(\partial_k) = \{c \in C_k \mid \partial(c) = 0\} \tag{5}
$$

because 0 is the identity element. For Eq. (2), it means the coefficients $a \in \mathbb{Z}$ are constrained by $\partial(c) = 0$, narrowing it down to two generating elements for the 1-cycle group:

$$
Z_1 = \left\langle \; \boxed{[01] + [12] + [20]}, \; \boxed{[12] - [13] + [23]} \; \right\rangle
$$

$$\cong \tag{6}$$

Note that the larger cycle of $0 - 1 - 3 - 2$ can be constructed by a linear combination of this basis.

D. Homology groups

The homology groups H_k are defined as the quotient of cycles and boundaries,

$$
H_k = \frac{Z_k}{B_k} = \frac{\text{Ker } \partial_k}{\text{Im } \partial_{k+1}}, \tag{7}
$$

where Z_k are the k-cycles and B_k the k-boundaries. This is sometimes referred to as "cycles mod boundaries." The first homology group for our example contains one generating element,

$$
\begin{aligned}
H_1(K) &= \frac{\langle [01] + [12] + [20], [12] - [13] + [23] \rangle}{\langle [12] - [02] + [01] \rangle} \\
&\cong \langle [12] - [13] + [23] \rangle \cong \mathbb{Z},
\end{aligned} \tag{8}
$$

which captures the one-dimensional hole (of the empty triangle $1 - 3 - 2$). The 1-cycle between the points $0 - 1 - 2$ is modded out because it is the boundary of the 2-simplex (filled triangle). A similar computation leads to the other homology groups. In summary, the homology groups for the

example K are $H_0(K) \cong \mathbb{Z}$, $H_1(K) \cong \mathbb{Z}$ and $H_k(K) \cong 0$ for all $k > 1$. In other words, the topological space K contains one connected component (as described by H_0) and contains one one-dimensional hole (as described by H_1).

In this example, the homology computation is performed over the integers because the coefficients are chosen to be $a \in \mathbb{Z}$, but most practical codes use booleans, i.e., $a \in \mathbb{Z}_2$. For more examples of homology computations and the effect of different coefficients, see Appendix C.

IV. PERSISTENT HOMOLOGY

The persistent homology of a set of discrete data characterizes the shapes that are present in the data at different length scales. In practice, this means the simplicial homology computations described in the previous section with a method of *filtration*. The filtration refers to the choice of a distance metric and how this metric leads to a set of nested simplicial complexes. A simple example of six data points in the two-dimensional Euclidean plane is shown in Fig. 2. The chosen proximity parameter ϵ is, in this case, given by simply the Euclidean distance and visualized by the growing disks. As soon as disks overlap, a new simplex is formed. When three points are connected, a 2-simplex (filled triangle) is formed. In general, for k points, a (k-1)-simplex is formed. This particular filtration leads to a so-called Vietoris-Rips complex. The simplicial homology at every level of filtration is summarized in the barcode Fig. 2(f). More specifically, a generating element of a homology groups that spans multiple length scales is indicated by a bar with a start (birth) and end (death). The length of the bar is also referred to as the lifetime.

In this paper, the discrete data are the quantum subsystems and the distance metric is based on MI as a measure of entanglement. The distance matrix between all quantum subsystems i and j, i.e., D_{ij}, forms a sequence of simplicial complexes, $K_1 \subset K_2 \subset \cdots \subset K_n$. For each simplicial complex K_i, the homology groups H_k are computed, where k refers to k-dimensional simplices (see Fig. 3). The homology groups H_k are computed for each K_i and sequences of identical homology lead to so-called persistent homology. The sequence of homology groups are typically visualized in a persistence barcode or persistence diagram. Both show the same information and reveal shapes present in the quantum MI. A barcode is a collection of line segments that represent the generating elements of a homology group that span multiple length scales. When two elements become homologous, the older one survives, as described by the *elder rule* [29].

In practical applications, since the barcode itself is not a scalar quantity, it is common to either use scalar properties derived from the barcode or construct a scalar distance between two barcodes. For example, the Betti number is an integer,

$$\beta_k = \text{rank}(H_k), \tag{9}$$

that counts the rank of the kth homology group. Another example is the lifetime of a specific bar in the barcode. Bars with a long lifetime indicate that a feature (like a one-dimensional hole) persists over a large range of length scale. Features with a short lifetime are sometimes considered to be noise in the input point cloud. However, these features can also indicate

the curvature of the manifold that the point cloud was sampled from Ref. [30].

There are many different codes that implement the computation of persistent homology given a distance matrix. An example of the algorithm is shown in Appendix B. We use the GUDHI Python module for the construction of a Vietoris-Rips complex from the distance matrix and the computation of the barcode [31].

In our method, each state $|\psi\rangle$ has a corresponding barcode and a set of homology groups. The barcode is a fingerprint of the entanglement structure of the state. Abrupt changes can indicate phase transitions, and barcodes can also differentiate between phases.

V. DISTANCE METRIC BASED ON ENTANGLEMENT

The N particles (or spins) form a discrete point cloud, where the distance metric between spins is the (additive) inverse of the quantum MI. MI is commonly used to identify clusters of entanglement and as a probe for phase transitions (e.g., MBL and QPTs) [10,14,32,33]. The MI between two sites i and j is defined as

$$0 \leqslant M_{ij} = S_i + S_j - S_{ij} \leqslant 2\ln 2. \tag{10}$$

Here, the entanglement entropy S_i and S_{ij} refer to taking only sites i and i, j in subsystem A, respectively.

Theorem V.1. Let D_{ij} be the inverse of the MI between sites i and j:

$$2\ln 2 \geqslant D_{ij} = 2\ln 2 - M_{ij} \geqslant 0. \tag{11}$$

This is a distance metric that brings strongly entangled sites ($M \to 2\ln 2$) close together ($D \to 0$), while nonentangled sites ($M \to 0$) are far apart ($D \to 2\ln 2$).

Proof. The distance metric defined in Eq. (11) satisfies the axioms for a metric because it is symmetric and satisfies the triangle inequality. This triangle inequality relies on properties of quantum entropy, and we start by writing the inequality in terms of entropy:

$$D_{xy} \leqslant D_{xz} + D_{zy},$$
$$2S_z + S_{xy} - S_{xz} - S_{zy} \leqslant 2\ln 2. \tag{12}$$

To show that the inequality holds, we first note that S_{xz} and S_{zy} are non-negative and the term $X = S_{xz} + S_{zy}$ can be replaced by $S_x + S_y$, which is guaranteed to be equal or smaller by the strong subadditivity of quantum entanglement [Eq. (A5)]. Furthermore, the term S_{xy} is replaced by an expression that is guaranteed to be larger, $S_{xy} \leqslant S_x + S_y$ [Eq. (A4)]. This leads to

$$2S_z + S_x + S_y - S_x - S_y \leqslant 2\ln 2, \tag{13}$$

$$2S_z \leqslant 2\ln 2, \tag{14}$$

which is true because S_z has a maximum value of $\ln 2$. Finally, the distance of point i with itself, D_{ii}, is never evaluated when constructing homology, and we consider it to be zero. ∎

Given a set of discrete data (quantum subsystems) and the distance matrix (inverse MI), we can now study its homology (see Fig. 4).
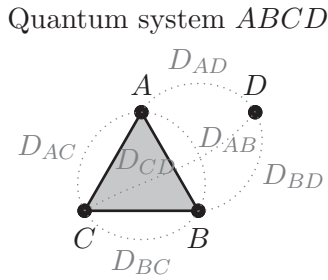
Quantum system $ABCD$

FIG. 4. Example of a simplicial complex created from a quantum system comprising four subsystems: A, B, C, and D. The entanglement entropy between all subsystems is used to compute the distance D_{ij} [Eq. (11)]. This complex corresponds length scale ϵ, where in this case subsystems A, B, and C form a 2-simplex. Subsystem D is not connected to the ABC component, indicating that the distances between D and A, B, and C is larger than ϵ. Here, the Betti numbers are $\beta_0 = 2$ (i.e., two connected components: ABC and D) and $\beta_1 = 0$ (no one-dimensional holes, since the 1-cycle ABC is also the boundary of the 2-simplex ABC).

VI. EXAMPLE 1: ISING CHAIN IN TRANSVERSE FIELD

To demonstrate our approach to persistent homology of quantum states, we study the simplest model of a QPT, the TFIM on a one-dimensional lattice. The Hamiltonian $\mathcal{H}$ of the TFIM is [8]

$$\mathcal{H} = -\sum_i \sigma_i^x \sigma_{i+1}^x - h \sum_i \sigma_i^z, \qquad (15)$$

where σ are the Pauli matrices. The exact diagonalization and entanglement entropy computation is implemented using the QuSpin package [34]. There is a QPT that can be measured by the magnetization $M = |1/N \sum_i \sigma_i^x|$. For $h = 0$, the ferromagnetic ground state is the basis state with all spins pointing along the x direction. For large h, the ground state is a quantum paramagnet, a superposition of all basis states.

For $h = 0$, there is a degenerate ground state that is a product state of all spin up $|\uparrow\uparrow \ldots \uparrow\rangle$ or all down $|\downarrow\downarrow \ldots \downarrow\rangle$. For a small but finite h, these levels split with order h^N and the ground state is a macroscopic superposition ($|\uparrow\uparrow \ldots \uparrow\rangle + |\downarrow\downarrow \ldots \downarrow\rangle$) with $\ln 2$ half-chain entanglement entropy [see Fig. 5(a)]. The system undergoes a QPT at $h = 1$, and the entanglement structure forms a 1-cycle, as will be discussed. At $h = \infty$, the ground state is a a quantum paramagnet, i.e., superposition of all basis states $\frac{1}{\sqrt{2^N}} \sum_i^{2^N} |i\rangle$ and zero half-chain entanglement entropy.

The quantum ground state is converted into a barcode by computing the distance matrix D between all sites. This constitutes constructing Vietoris-Rips complexes from this distance matrix and computing its homology groups. The barcodes are shown in Fig. 6. At zero length scale $\epsilon = 0$, the number of H_0 bars (i.e., β_0) indicates the number of quantum subsystems (spins in this case) and the dimensionality of the matrix D. At large length scales, at least one H_0 bar survives, indicating one large connected component, similar to the situation shown in the simple two-dimensional Euclidean example (Fig. 2).

Both the low ($h < 1$) and high ($h > 1$) field ground states have constant MI M_{ij} for all pairs (see Fig. 5). This causes all
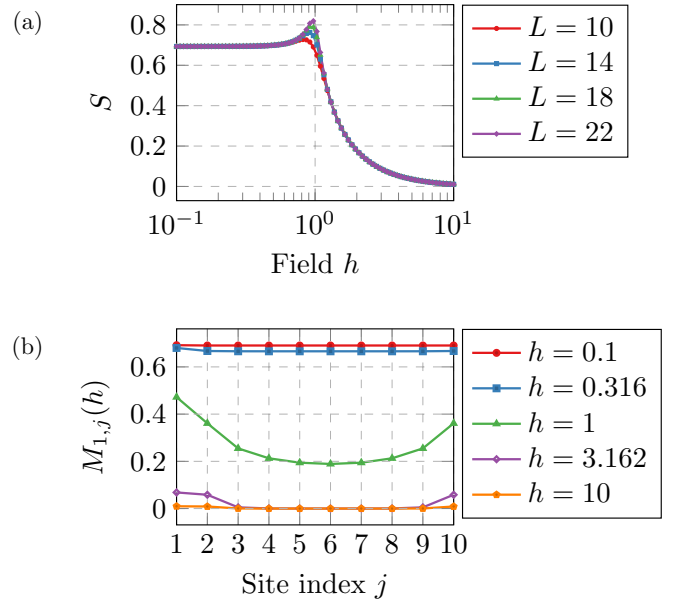

FIG. 5. (a) Half-chain entanglement entropy S as a function of field h in the 1D TFIM. For small fields h, the entanglement is constant, $\ln(2)$. At the critical point of $h = 1$, S grows as $\ln(N)$. (b) Quantum mutual information between first site 1 and site j as a function of field h (logarithmically spaced) in the 1D TFIM. The mutual information M_{1j} is largest between nearby sites, indicating the local interactions of the Hamiltonian. Periodic boundary conditions are used and lead to increased M_{1j} for large j.

sites to pair up at the same length scale in the barcode, as indicated by the end of the H_0 bars, also referred to as the death of the feature. For the ferromagnetic phase, all the sites have the same pairwise distance and are connected at the length scale of $\epsilon = \ln 2$ [see Fig. 6(a)]. For the quantum paramagnet, the spins are maximally separated, which is essentially a rescaled version of having points being infinitely separated in space, and each site survives as a single component up to the maximum length scale $\epsilon = 1 \cdot 2 \ln 2$ see Fig. 6(c).

Around the QCP ($h = 1$), neighboring sites are more strongly entangled than next-nearest neighbors. Due to the symmetry of the Hamiltonian and the periodic boundary conditions (PBCs), the distances are the same for each site, e.g., $D_{12} = D_{23}$. This means we have a one-dimensional hole to form inside the ring of spins that persists over a finite length scale [see Fig. 6(b)]. At slightly longer length scales, next-nearest neighbors are connected, after that next-next-nearest neighbors, and so on. This process continues until furthest-separated spins on the chain are connected and the hole closes. This is a unique feature of the critical point due to the entanglement structure of the critical state, and it can be characterized by the start (birth) and end (death) of the bar in H_1, as shown in Fig. 6(d).

Figure 7 shows the birth and death value of H_1 at $h = 1$ for different chain lengths. This model is exactly solvable (see Appendix D), and we can therefore compare the results to the infinite chain. As described earlier, the cycle is formed when the nearest neighbors connect, and this is set by the distance D_{01}. The hole closes when maximally separated spins connect, which in the case of an infinite chain is set by $D_{0\infty}$.
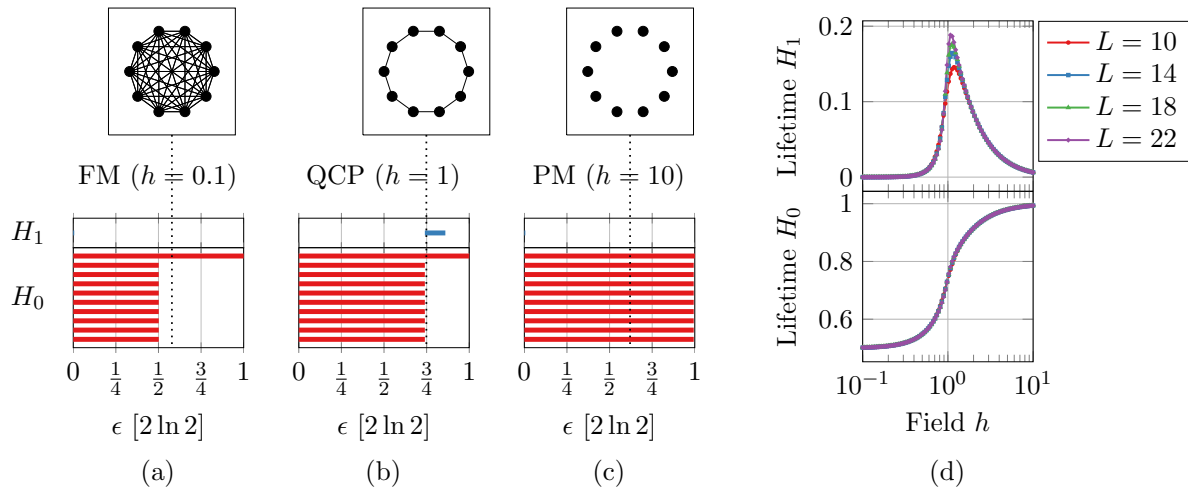
FIG. 6. Barcodes for transverse-field Ising chain with $L = 10$ sites at (a) $h = 0.1$, (b) 1, and (c) 10. The proximity parameter ϵ refers to the length scale set by the distance matrix $D(i, j)$. The dotted lines and panels above show the simplicial complex for a given ϵ. At the quantum critical point, the spin chain forms a cycle of entanglement with a one-dimensional hole. (d) Lifetime of H_1 and H_0 bars. The persistence of the 1-cycle at the QCP increases with system size L.

For the infinite chain, the distances are $D_{01}/2 \ln 2 \approx 0.744293$ and $D_{0\infty}/2 \ln 2 = 1$, which matches the birth and death curves shown in Fig. 7.

VII. EXAMPLE 2: XXZ SPIN CHAIN IN TRANSVERSE FIELD

To study multiscale entanglement, we now focus on the 1D XXZ spin chain in tranverse field with Hamiltonian

$$H = -\sum_i \left[(\sigma_i^+ \sigma_{i+1}^- + \sigma_i^- \sigma_{i+1}^+) + \frac{\Delta}{2} \sigma_i^z \sigma_{i+1}^z + h_i \sigma_i^z \right], \quad (16)$$

where $\sigma_i^\pm = \sigma_i^x \pm i\sigma_i^y$, Δ is the interaction strength and h_i is a random magnetization. We take $\Delta = 1$ and the field uniformly random $h_i = [-W, W]$. This model is often used to study the nature of the MBL phase transition [10,26]. The critical value of disorder reported in the literature varies, usually close to $W_c \approx 3.8$ as obtained through large-scale ($L = 26$) exact diagonalization [35]. However, more recent studies suggest a slightly larger value closer to $W_c \approx 5$ [36,37]. In the ergodic phase (small W), the excited states have volume law entan-
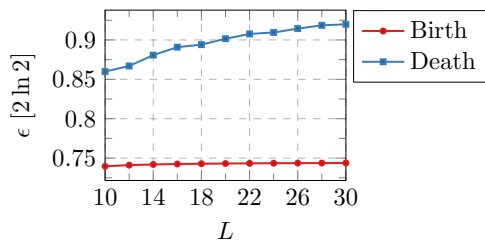
glement. In the MBL phase (large W), the spins are mostly weakly entangled, but the entanglement structure is the object of our interest. Clusters (of spins) are defined as a subsystem that has stronger entanglement internally than with the rest of the system.

Using exact diagonalization, we select $k = 50$ eigenstates from the middle of the spectrum for $N_d = 3000$ disorder realizations (i.e., changing the random magnetization h_i). For each eigenstate, we construct a Vietoris-Rips complex with the distance metric of Eq. (11). The resulting persistence barcodes are merged by discarding the information from which realization each barcode originated. Since this leads to barcodes with many bars, we use Betti numbers [Eq. (9)] to summarize and investigate the results. Figure 8 shows the Betti numbers for varying disorder strengths W. As the disorder strength W
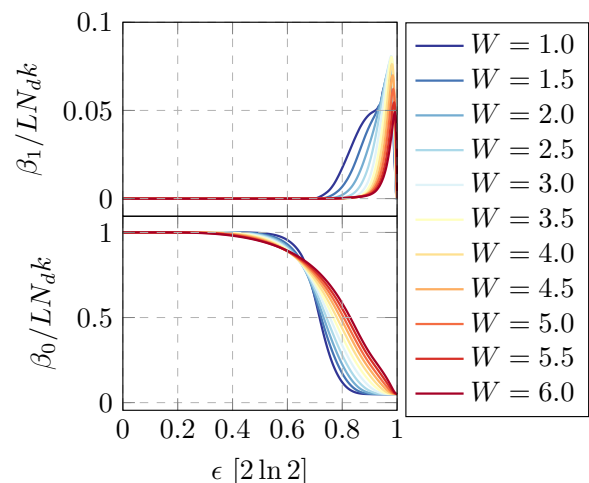

FIG. 7. Birth and death of the H_1 persistence bar at the quantum critical point $h = 1$ of the transverse-field Ising chain of varying size L. The birth value is set by the distance between neighboring sites and converges to $D_{01}/2 \ln 2 \approx 0.744293$. The death value is set by the distance between two maximally separated sites and converges to $D_{0\infty}/2 \ln 2 = 1$ (see Appendix D).

FIG. 8. Betti numbers β_0 and β_1 for varying disorder strength W. XXZ spin chain with $L = 16$ sites, $N_d = 3000$ realizations and $k = 50$ states.
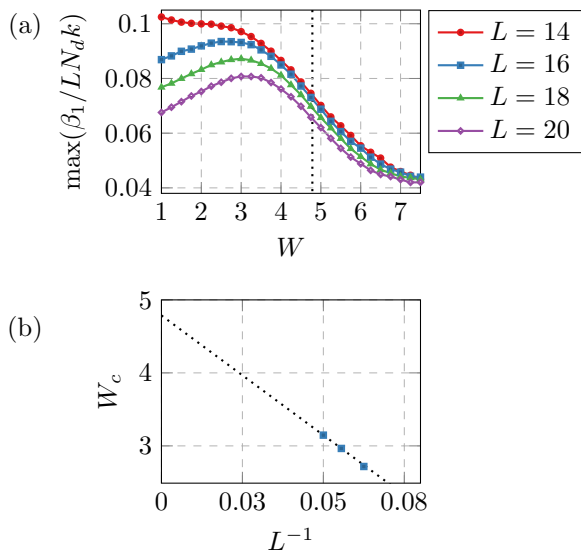
FIG. 9. (a) Maximum Betti number of the first homology group for varying disorder strengths W. 3000 disorder realizations. (b) The disorder strength with largest Betti number, i.e., peak in (a), that extrapolates to a critical strength of $W_c \approx 4.79$ (marked by dotted line).

decreases, there are fewer connected components as counted by the normalized β_0.

Figure 9 shows the maximum of β_1 (see Fig. 8) for varying disorder strengths W. This maxima occur at a length scale around $\epsilon \approx 0.95 \cdot 2 \ln 2$. The Betti number β_1 essentially counts the number of spin clusters that are more strongly entangled (strength set by ϵ) within the cluster than the environment. The minimum set of spins to form a 1-cycle is four spins (because three connected spins would form a 2-simplex in a Vietoris-Rips complex). For smaller system sizes $L \leqslant 14$, the maximum is at small disorder strength W. However, for larger systems, a maximum in the expected range of $3 - 5$ appears. Extrapolating the peak leads to a predicted critical strength of $W_c \approx 4.79$, which is in line with the aforementioned values reported in literature.

VIII. DISCUSSION

Here we lay out a few examples of related work and future directions to explore. Computing the entanglement entropy of all possible partitions is numerically impractical due to the factorial number of possibilities. For single pairs of N particles (or spins), the number is N choose 2 [scales as $O(N^2)$]. However, due to the nonadditivity of entropy, this does not necessarily capture all the structure present in the entanglement of the state.

Rather than starting from individual spins, it is also possible to perform a recursive bipartitioning of the spin chain (as used in Ref. [10]). This turns the wave function into a binary tree, where each node is a set of spins. Compared to our method, which is bottom up from spin pairs, this could be considered a top-down approach. Given the binary tree, the distance between leaf nodes (single spins) can be used to compute homology. This would be an alternative to our

bottom-up approach based on MI and an interesting direction to explore in the future.

Another interesting future direction would be to connect barcodes to matrix product states and tensor networks. In matrix product states, the bond dimension refers to the dimension of matrices used to represent the quantum state, and a small bond dimension comes with low computational complexity. For example, ground states of gapped Hamiltonians obey the area law and have constant bond dimension [1]. The bond dimension is dependent on the decay of correlations in the model, and these correlations are measured by the barcode, providing a potential bridge between the two. Regarding tensor networks, the MERA [3] is used for infinite 1D quantum critical states and represents an interesting case to study through the lens of homology. Another state to consider is the rainbow state, which is a ground state of the inhomogeneous Ising transverse field chain that also exhibits a volume law [38,39].

We compute entanglement entropy using singular value decomposition, but diagonalizing the partial trace of a density matrix (constructed as the outer product of the pure state) has similar computational complexity. The purity (or linear entropy) of a quantum state could be an alternative to the von Neumann entropy that is used here, and it has the benefit of not requiring diagonalization of the density matrix.

The first example of the Ising chain discussed in this paper can be solved efficiently using the Jordan-Wigner transformation. However, this is a special case and is not generally efficient for all interacting Hamiltonians, so we do not rely on this transformation in this work. Furthermore, the PBCs are essential for forming a 1-cycle in the transverse-field Ising chain and this appears to limit the usefulness of barcodes. In Appendix E, we show an example with open boundary conditions, where a phase transition is still captured.

The von Neumann entropy is typically not experimentally accessible, whereas the Rényi entropy of order $\alpha = 2$ is experimentally accessible in certain setups [40]. This would make it possible to compute the barcode of an experimental setup.

Regarding persistent homology, we note that our homology groups are computed with $\mathbb{Z}_2$ (boolean) coefficients, the most common choice in TDA. The choice of coefficients determines whether a simplex can occur multiple times in a chain or whether the direction of travel (represented by the sign of the coefficient) matters. The change in barcodes for entanglement structure when computing homology with different coefficients is also an interesting direction to explore. For examples of manifolds and the effect of homology with different coefficients, see Appendix C.

TDA is a powerful tool that can be applied to the questions of emergent structures in the topology of the Hilbert space describing the system. This connects directly to a deeper question: Does space-time emerge from entanglement? This is one of the questions that is part of the Simons Collaboration: *It from Qubit* [41]. The study of the emergent geometry from entanglement is still a topic of active research [42–45]. For example, Cao *et al.* have proposed the use of MI and classical multidimensional scaling to form emergent geometry [46].

IX. CONCLUSION

Persistent homology is a tool for dealing systematically with the entanglement structure in quantum states. The topological features that it computes can be used to answer a number of scientific questions. First, the barcode of a quantum state changes dramatically when the state undergoes a QPT. It is therefore a unique type of quantum order parameter. We have demonstrated its capability for two basic examples of the Ising chain and XXZ spin chain in a transverse field. Second, it can guide the development of suitable wave function ansatze with, for example, tensor networks. Third, it provides a numerical approach for studying the emergent geometry of entanglement. This is very relevant to the scientific question whether space-time emerges from entanglement.

The focus in this multidisciplinary work is on the application of persistent homology to equilibrium QPTs, the first of the aforementioned uses of the method. However, it is also possible to examine many quantum phenomena through the lens of homology and Betti numbers of the entanglement structure. The effect of adding time dependence to the presented scheme is also a promising future direction to explore, and recent work has already introduced some persistent homology observables for quantum many-body dynamics [19].

APPENDIX A: BIPARTITE ENTANGLEMENT ENTROPY

The distance metric used in our paper is based on quantum MI and the properties of entanglement entropy. We therefore recall the key properties of entanglement entropy. Bipartite entanglement entropy is computed by performing Schmidt decomposition on a quantum state. We start by breaking down the Hilbert space in two parts, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, and form a matrix C_{ij} such that

$$|\psi\rangle = \sum_{i=1} \sum_{j=1} C_{ij} |i\rangle_A \otimes |j\rangle_B, \tag{A1}$$

where $\{|i\rangle_A\}$ and $\{|j\rangle_B\}$ are the basis sets for $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively [47]. Schmidt decomposition is essentially singular value decomposition $C = U\Sigma V^\dagger$, where λ_α forms the diagonal of Σ. Assuming $N_B > N_A$, then the state can be expressed in the following form:

$$|\psi\rangle = \sum_{\alpha}^{2^{N_A}} \lambda_\alpha |\alpha\rangle_A \otimes |\alpha\rangle_B. \tag{A2}$$

If there is more than one nonzero singular value λ_α, then the state is entangled. The singular values can also be used to compute the entanglement entropy

$$S = -\sum_j |\lambda_j|^2 \ln(|\lambda_j|^2), \tag{A3}$$

which is zero if the state $|\psi\rangle$ is a product state (without entanglement). It is also upper bounded to $\ln(d)$, where d is the Hilbert space dimension. This is $N \ln 2$ in the case of N spins. Two other properties of quantum entropy that are used to form the distance metric are the triangle inequality,

$$|S_i - S_j| \leqslant S_{ij} \leqslant S_i + S_j, \tag{A4}$$

and strong subadditivity [48],

$$S_x + S_z \leqslant S_{xy} + S_{yz}. \tag{A5}$$

In the main text, we show how the bipartite von Neumann entanglement entropy S is used to compute MI between subsystems and how it sets the length scale for the corresponding persistence barcodes.

APPENDIX B: PERSISTENT HOMOLOGY ALGORITHM

The computation of persistent homology for a sequence simplicial complexes is, in practice, simply a row reduction of a matrix. More specifically, the boundary matrix encodes the boundaries present in the complex and each row and column represent a simplex. The rows and columns are also ordered by the appearance time of the simplices in the complex, such that the first column/row appears before the next in the sequence of complexes, and so on. In other words, this ordering originates from the input distance metric (e.g., Euclidean distance), as visualized in Fig. 2.

The *standard algorithm* (sometimes called the *column algorithm*) reduces the boundary matrix, see Algorithm 1 below [21,29].

The worst-case time complexity of this algorithm is cubic in the number of simplices, and a sparse matrix implementation is possible [21]. From the reduced boundary matrix, we can read off the barcode. For other coefficients than $\mathbb{Z}_2$, see Ref. [22].

Algorithm 1 Column algorithm

1: **for** $j = 1$ to m **do**
2:     **while** there exists $j_0 < j$ with $low(j_0) = low(j)$ **do**
3:         add column j_0 to column j
4:     **end while**
5: **end for**

We now demonstrate the algorithm on an example simplicial complex:

$$K = \qquad \qquad \text{(B1)}$$
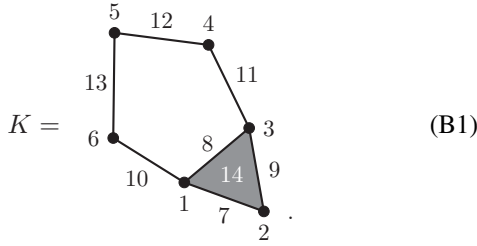
This corresponds to a space of a disk (2-simplex No. 14) glued to a one-dimensional hole (cycle of one-simplices, 8-11-12-13-10). First, deconstruct the simplicial complex into a boundary matrix,

$$B = \qquad \qquad \text{(B2)}$$

where the nonzero rows indicate the boundary of the corresponding column. Although we did not specify the ordering when defining the simplicial complex, the simplices in the columns and rows are also ordered by their appearance time in the complex.

Next, reduce the boundary matrix following Algorithm 1 and indicate the lowest 1s with a box:

$$B = \qquad \qquad \text{(B3)}$$

The persistence pairs can be harvested from the reduced boundary matrix. An all-zero column indicates the *birth* of a persistence pair. It *dies* by pairing it with a lowest 1. For example, column 9 is all-zero and initiates a pair. For the corresponding row 9, it pairs with column 14, leading to persistence pair $[9, 14)_1$, where the subscript indicates the birth simplex is one-dimensional. The all-zero columns that cannot

be paired live until infinity. In our example, all the persistence pairs are

$$[2, 7)_0, \qquad \qquad \text{(B4)}$$

$$[3, 8)_0, \qquad \qquad \text{(B5)}$$

$$[9, 14)_1, \qquad \qquad \text{(B6)}$$

$$[6, 10)_0, \qquad \qquad \text{(B7)}$$

$$[4, 11)_0, \qquad \qquad \text{(B8)}$$

$$[5, 12)_0, \qquad \qquad \text{(B9)}$$

$$[13, \infty)_1, \qquad \qquad \text{(B10)}$$

$$[1, \infty)_0. \qquad \qquad \text{(B11)}$$

As expected, only one connected component (0 dimensionality hole) survives to infinity: $[1, \infty)_0$.

The persistence pairs are typically visualized in a barcode (see Fig. 2), where the birth and death indicate the start and end of the bar. Note that the explicit length scales associated to the appearance of each simplex is not part of the persistent homology algorithm, only the ordering is. The birth and death simplices can be translated into birth and death length scales when plotting a barcode by referring to the input sequence of simplicial complexes that was generated using a distance metric.

APPENDIX C: NONORIENTABLE SURFACES AND TORSION

The entanglement structure of a quantum state is converted into a topological space using MI as a metric. This space is then studied using simplicial homology with $\mathbb{Z}_2$ coefficients, which means that simplices either are included or excluded when talking about loops in this space. This is computationally easier, however, the more general case of homology with integer coefficients (i.e., integral homology) is able to capture more information. In this section, we study this in detail and show that manifolds may have different homology groups, depending on the choice of coefficients.

Nonorientable surfaces are twisted and turn clockwise objects to counterclockwise, with common examples being the Möbius strip, Klein bottle, and the real projective plane. A surface is orientable if it has a consistent notion of clockwise rotation as we move around. It turns out that nonorientability is connected to the notion of *torsion*. Algebraically, torsion means that a group element has *finite order* (i.e., $g^n = e$ for a positive integer n) and a group where all elements fit this condition is called a torsion group. A torsion-free group has no elements of finite order other than the identity element, e.g., $(\mathbb{Z}, +)$. Both orientable (e.g., $\mathbb{RP}^3$) and nonorientable (e.g., $\mathbb{RP}^2$) manifolds may have torsion in its integral homology. However, there is a theorem that states that if M is a closed connected n-manifold that is nonorientable, the homology group H_{n-1} contains torsion [49], Corollary 3.28.
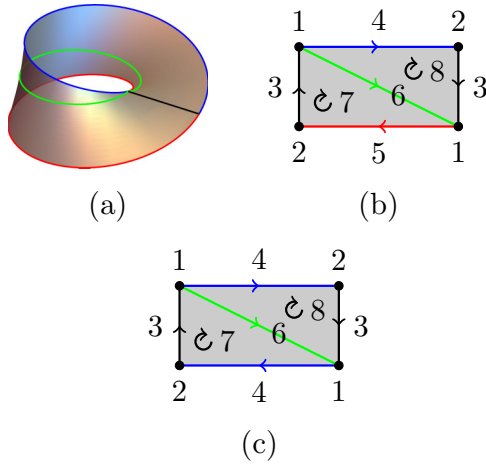
FIG. 10. (a) Möbius strip that is triangulated by gluing together 2-simplices. The faces are four unique 1-simplices and two 0-simplices. (b) Simplicial complex of the Möbius strip showing the orientation of the simplices. (c) Simplicial complex of the real projective plane that is constructed by gluing the sides of the Möbius strip together—this can be done in $\mathbb{R}^4$ without the surface intersecting itself.

In general, the k-homology group over the integers for a simplicial complex K has the following form:

$$H_k(K, \mathbb{Z}) = \overbrace{\mathbb{Z} \oplus \cdots \oplus \mathbb{Z}}^{\beta_k} \oplus \mathbb{Z}_{k_1} \oplus \cdots \oplus \mathbb{Z}_{k_n}, \quad (C1)$$

where β_k is the Betti number and k_i are the torsion coefficients. However, it is common to compute homology over $\mathbb{Z}_2$ rather than the integers, leading to a different kind of homology, namely, $H_k(K, \mathbb{Z}_2)$. To illustrate the absence or presence of torsion in nonorientable surfaces, we start with the Möbius strip and then proceed to the real projective plane.

Figures 10(a) and 10(b) show the Möbius strip M and its corresponding simplicial complex. We now compute its one-homology over the integers. The chain groups for the Möbius strip M are

$$C_0 = \langle [1], [2] \rangle, \quad (C2)$$

$$C_1 = \langle [3], [4], [5], [6] \rangle, \quad (C3)$$

$$C_2 = \langle [7], [8] \rangle, \quad (C4)$$

where the simplices form a basis and the coefficients are hidden, e.g., $C_0 = \{a_1[1] + a_2[2]\}$. The 1-cycles are a subgroup of C_1, namely,

$$Z_1 = \{c \in C_1 \mid \partial(c) = 0\} \quad (C5)$$

$$= \{a_3[3] + a_4[4] + a_5[5] + a_6[6]\}, \quad (C6)$$

where the coefficients $a_i \in \mathbb{Z}$ are constrained by $\partial(c) = 0$. To find the space of 1-cycles, we apply the boundary operator [Eq. (3)] to the 1-simplices,

$$\partial[3] = [1] - [2], \quad (C7)$$

$$\partial[4] = [2] - [1], \quad (C8)$$

$$\partial[5] = [2] - [1], \quad (C9)$$

$$\partial[6] = [1] - [1] = 0, \quad (C10)$$

and row reduce this system of equations. The space of 1-cycles is generated by

$$Z_1 = \langle [3] + [4], [3] + [5], [6] \rangle. \quad (C11)$$

Next, to find the boundary group $B_1 = \text{Im } \partial_2 = \partial_2(C_2)$, we apply the boundary operator to the 2-simplices:

$$\partial[7] = [3] + [6] + [5], \quad (C12)$$

$$\partial[8] = [3] - [6] + [4]. \quad (C13)$$

Therefore, the boundary group B_1 is generated by

$$B_1 = \langle [3] + [6] + [5], [3] - [6] + [4] \rangle. \quad (C14)$$

Finally, we find that the one-homology (cycles mod boundaries) over the integers for the Möbius strip M is

$$H_1(M, \mathbb{Z}) = \frac{\langle [3] + [4], [3] + [5], [6] \rangle}{\langle [3] + [6] + [5], [3] - [6] + [4] \rangle} \quad (C15)$$

$$\cong \frac{\langle [3] + [4] + [6], [3] + [5] - [6], [6] \rangle}{\langle [3] + [6] + [5], [3] - [6] + [4] \rangle} \quad (C16)$$

$$\cong \langle [6] \rangle \quad (C17)$$

$$\cong \mathbb{Z}. \quad (C18)$$

The number of one-dimensional holes is described by the Betti number $\beta_1 = 1$. Geometrically, we can see this loop in Fig. 10(a) in the middle of the Möbius strip and observe that group H_1 is torsion-free.

In the case of $\mathbb{Z}_2$ coefficients, the boundary operator is

$$\partial S = \sum_{j=0}^{k} [v_0 v_1 \dots \hat{v}_j \dots v_k], \quad (C19)$$

with $\hat{v}_j$ simplex being left out. Simplices appearing twice mod out to zero, and again $\partial^2 S = 0$ is true. The Betti number β_k is now defined as the number of copies of $\mathbb{Z}_2$. Performing the same steps as before, we arrive at the one-homology with $\mathbb{Z}_2$ coefficients for the Möbius strip,

$$H_1(M, \mathbb{Z}_2) = \frac{\langle [3] + [4], [3] + [5], [6] \rangle}{\langle [3] + [6] + [5], [3] + [6] + [4] \rangle} \quad (C20)$$

$$\cong \frac{\langle [3] + [4] + [6], [3] + [5] + [6], [6] \rangle}{\langle [3] + [6] + [5], [3] + [6] + [4] \rangle} \quad (C21)$$

$$\cong \langle [6] \rangle \quad (C22)$$

$$\cong \mathbb{Z}_2, \quad (C23)$$

where the Betti number β_k is now defined as the number of copies of $\mathbb{Z}_2$ (i.e., $\beta_1 = 1$ in this case).

The real projective plane $\mathbb{RP}^2$ can be constructed by embedding the Möbius strip in $\mathbb{R}^4$ and gluing the sides together with the right orientation, see Figs. 10(b) and 10(c). The chain groups for this manifold are the same except for the 1-chains and 1-cycles

$$C_1 = \langle [3], [4], [6] \rangle, \quad (C24)$$

$$Z_1 = \{c \in C_1 \mid \partial(c) = 0\} \quad (C25)$$

$$= \{a_3[3] + a_4[4] + a_6[6]\}. \quad (C26)$$

Again, the coefficients are constrained by the boundary operator $\partial(c) = 0$ and the space of the one-cycles is generated by

$$Z_1 = \langle [3] + [4], [6] \rangle. \tag{C27}$$

The boundary of the 2-simplices are now

$$\partial[7] = [3] + [6] + [4], \tag{C28}$$

$$\partial[8] = [3] - [6] + [4]. \tag{C29}$$

The boundary group B_1 is generated by

$$B_1 = \langle [3] + [6] + [4], [3] - [6] + [4] \rangle. \tag{C30}$$

The one-homology group for the real projective plane:

$$H_1(\mathbb{RP}^2, \mathbb{Z}) = \frac{\langle [3] + [4], [6] \rangle}{\langle [3] + [6] + [4], [3] - [6] + [4] \rangle} \tag{C31}$$

$$\cong \frac{\langle [3] + [4] + [6], [6] \rangle}{\langle [3] + [6] + [4], 2[6] \rangle} \tag{C32}$$

$$\cong \frac{\langle [6] \rangle}{\langle 2[6] \rangle} \cong \frac{\mathbb{Z}}{2\mathbb{Z}} \equiv \mathbb{Z}_2. \tag{C33}$$

The real projective plane has two-torsion and a Betti number $\beta_1 = 0$. Geometrically, it means that the single move from [1] to [1] is a loop, but it cannot be contracted to a point, however, twice the loop (i.e., walking around the rectangular perimeter of the simplicial complex in Fig. 10) is contractible to zero.

In the case of $\mathbb{Z}_2$ coefficients, the real projective plane has a one-homology of

$$H_1(\mathbb{RP}^2, \mathbb{Z}_2) = \frac{\langle [3] + [4], [6] \rangle}{\langle [3] + [6] + [4], [3] + [6] + [4] \rangle} \tag{C34}$$

$$\cong \frac{\langle [3] + [4] + [6], [6] \rangle}{\langle [3] + [6] + [4], 2[6] \rangle} \tag{C35}$$

$$\cong \langle [6] \rangle \cong \mathbb{Z}_2. \tag{C36}$$

This is the same result as for the Möbius strip [Eq. (C23)], with Betti number $\beta_1 = 1$, so it is clear that homology over $\mathbb{Z}_2$ coefficients cannot distinguish between these two topological spaces. In contrast, the integral homology is different between the two spaces due to the presence of torsion in the real projective plane. However, it is also important to note that orientable manifolds (e.g., $\mathbb{RP}^3$) can have torsion in its integral homology groups.

More generally, there is the *universal coefficient theorem* that relates the integral homology to homology with different coefficients A, and it turns out that the integral homology completely determines the homology groups for any Abelian group A [49], Sec. 3.A.

Our study of entanglement structures only considers homology with coefficients $\mathbb{Z}_2$ since it is the most computationally straightforward. However, as shown here, integral homology groups are sometimes more informative. There exist algorithms that efficiently compute persistent homology with different coefficients and deduce the torsion subgroups of the integral homology [50]. Whether torsion provides important information in the study of entanglement structures is unclear and the topic of future work.

APPENDIX D: TRANSVERSE-FIELD ISING CHAIN: ENTANGLEMENT ENTROPY AT CRITICALITY

The demonstration of persistent homology on the one-dimensional transverse-field Ising chain (with PBCs) relied on the general approach of exact diagonalization, and this scales to finite-size systems up to about $L = 30$ sites. However, this model is exactly solvable and this makes it possible to compute the entanglement entropy where a finite number of sites are in subsystem A while the rest of the infinitely long chain is in subsystem B. This also provides exact results for the entanglement entropy and the MI used in our paper.

Persistent homology captures the structure of the entanglement as described by the two-site MI M_{ij}. Translational invariance means that the MI only depends on the distance between two sites, i.e., $M_{ij} = M_{0r}$ with $r = |i - j|$. The distance metric used in persistent homology is simply the inverse of the MI, such that high MI means two sites are close, see Eq. (11).

At the critical point, the asymptotic behavior of the persistent homology for this model is governed by two special cases. The first case sets the birth length scale of the one-homology when neighboring spins are connected and form a cycle [Fig. 6(b)]. This is set by the amount of MI between neighboring spins, i.e., M_{01}. The second case is the point when the one-hole is closed and it sets the death length scale. This is set by the amount of MI between spins on opposite sides of the chain, which in a case of an infinite chain corresponds to $M_{0\infty}$. Therefore, we need to compute the entanglement entropies S_1, S_{01}, and $S_{0\infty}$ [see Eq. (10)]. The exact one-site and two-site reduced density matrix and the spin-spin correlation functions for the infinite transverse-field Ising chain (at criticality) are known [51–54].

Starting with the case of the single site in subsystem A, the reduced density matrix is

$$\rho_1 = \frac{I + \langle \sigma^z \rangle \sigma^z}{2}, \tag{D1}$$

where I and σ^z are the identity matrix and the Z Pauli matrix, respectively. For the ground state, the transverse magnetization is [53]

$$\langle \sigma^z \rangle = \frac{1}{\pi} \int_0^\pi d\phi \frac{1 + \lambda \cos \phi}{\sqrt{1 + \lambda^2 + 2\lambda \cos \phi}} = \frac{2}{\pi}. \tag{D2}$$

This gives the exact entanglement entropy,

$$\frac{S_1}{\ln 2} = \frac{-\text{Tr}(\rho_1 \ln \rho_1)}{\ln 2} \tag{D3}$$

$$= \frac{2\pi \ln \left(\frac{4\pi^2}{\pi^2 - 4} \right) - 8 \tanh^{-1} \left(\frac{2}{\pi} \right)}{\pi \ln(16)} \tag{D4}$$

$$\approx 0.68376, \tag{D5}$$

normalized by $\ln(2)$ such that it ranges from 0 to 1.

The two-site case takes two spins i and j in subsystem A,

$$\rho_{0r} = \frac{I_{0r} + \langle \sigma^z \rangle (\sigma_0^z + \sigma_r^z) + \sum_{k=1}^{3} \langle \sigma_0^k \sigma_r^k \rangle \sigma_0^k \sigma_r^k}{4}, \tag{D6}$$
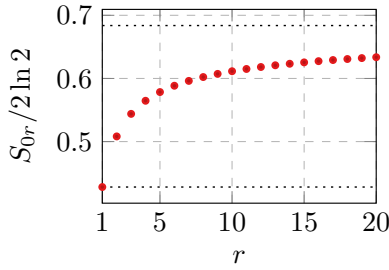
FIG. 11. Exact two-site entanglement entropy $S_{0r}/2\ln 2$ (where $r = |i - j|$) for an infinitely long transverse-field Ising chain with periodic boundary conditions. The dotted lines indicate the minimum S_{01} and the maximum $S_{0\infty}$.

which can be described with $r = |i - j|$ due to translational symmetry. The spin-spin correlation functions are known at criticality,

$$\langle \sigma_0^x \sigma_r^x \rangle = \left(\frac{2}{\pi} \right)^r 2^{2r(r-1)} \frac{H(r)^4}{H(2r)}, \tag{D7}$$

$$\langle \sigma_0^y \sigma_r^y \rangle = -\frac{\langle \sigma_0^x \sigma_r^x \rangle}{4r^2 - 1}, \tag{D8}$$

$$\langle \sigma_0^z \sigma_r^z \rangle = \langle \sigma_0^z \rangle \langle \sigma_r^z \rangle + \frac{4}{\pi^2} \frac{1}{4r^2 - 1}, \tag{D9}$$

with $H(r) = 1^{r-1} 2^{r-2} \ldots (r - 1)$. The two-site entanglement entropy for neighboring sites is

$$\frac{S_{01}}{2\ln 2} = \frac{-\text{Tr}(\rho_{01} \ln \rho_{01})}{2\ln 2} \tag{D10}$$

$$\approx 0.42805. \tag{D11}$$

For the two-site case with limit $r \to \infty$, the correlation functions are $\langle \sigma_0^x \sigma_r^x \rangle = 0$, $\langle \sigma_0^y \sigma_r^y \rangle = 0$ and $\langle \sigma_0^z \sigma_r^z \rangle = \langle \sigma_0^z \rangle \langle \sigma_r^z \rangle = 4/\pi^2$, and the entanglement entropy,

$$\frac{S_{0\infty}}{2\ln 2} = \frac{-\text{Tr}(\rho_{0\infty} \ln \rho_{0\infty})}{2\ln 2} \tag{D12}$$

$$= \frac{S_1}{\ln 2} \tag{D13}$$

$$\approx 0.68376. \tag{D14}$$

For completeness, Fig. 11 shows the entanglement entropy as a function of r. This shows that neighboring sites with $r = 1$ have the lowest entanglement entropy and this leads to the smallest MI distance. This matches what we see for a small chain with $h = 1$ in Fig. 5(b).

Finally, we find that the MI [Eq. (10)] for the case of two neighboring spins is

$$\frac{M_{01}}{2\ln 2} = \frac{S_0 + S_1 - S_{01}}{2\ln 2} \approx 0.255707. \tag{D15}$$

Therefore, the distance [Eq. (11)] $D_{01} = 1 - M_{01} \approx 0.744293$, which matches the birth length scale shown in Fig. 7. When the spins are far apart, the MI is zero,

$$\frac{M_{0\infty}}{2\ln 2} = \frac{S_0 + S_1 - S_{0\infty}}{2\ln 2} = 0, \tag{D16}$$

meaning that the distance is the maximum $D_{0\infty} = 2\ln 2$. This refers to the death length scale Fig. 7. In other words, the birth
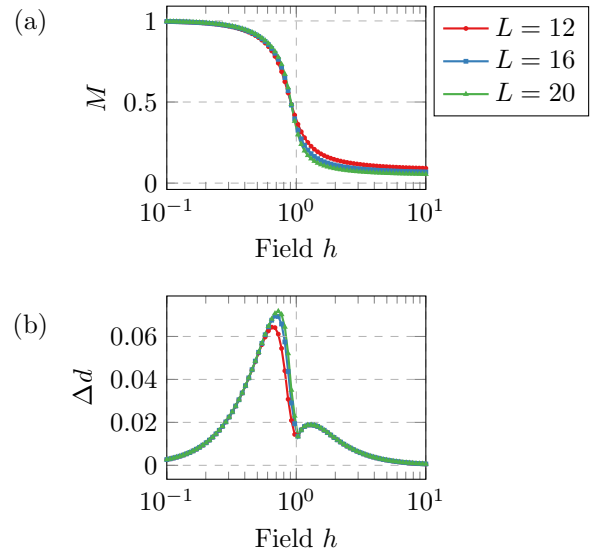

FIG. 12. The ground-state magnetization $M = |1/N \sum_i \sigma_i^x|$ of the transverse-field Ising chain with open boundary conditions (a). Range of the set of zero-homology deaths, where the maximum range coincides with the transition close to $h = 1$.

and death of the one-homology of this model at criticality converge to their expected values.

APPENDIX E: TRANSVERSE-FIELD ISING CHAIN: PERSISTENT HOMOLOGY WITH OPEN BOUNDARY CONDITIONS

The entanglement structure of a quantum state and its corresponding (persistence) barcode depend on the Hamiltonian and its boundary conditions. For example, the existence of the one-cycle for the tranverse-field Ising chain at criticality [Fig. 6(b)] relies on PBCs being present. This raises the question whether persistent homology is able to detect transitions without PBCs. Here we show that the model with open boundary conditions also shows a structure in its persistence barcode. However, it is the zero-homology instead of the one-homology that indicates a phase transition.

Without translational invariance due to PBCs, the MI between pairs of neighboring spins is not always exactly the same. For example, close to the critical transverse field of $h = 1$, we see that some pairs are closer as defined by the MI distance metric [Eq. (11)], which means that the spins connect and form one-simplices earlier. In terms of zero-homology (connected components), this coincides with the end (typically referred to as death) of a bar, e.g., Fig. 2. Note that zero-homology differs from one-homology in that the bars always start (typically referred to as birth) at zero, indicating a disconnected data point (spin in our case). Therefore, the only information is the set of death values. Since the zero-homology bars are all born at zero, this is equivalent to the set of lifetimes. The changes in zero-homology can be summarized in many ways, and here we plot the difference between the maximum and minimum of the set of death values Fig. 12(b), and this indicates the transition around $h = 1$ at its peak, similarly to the traditional magnetization [Fig. 12(a)]. This example shows that persistent homology is a general tool that does not necessarily rely on PBC being present.

[1] M. B. Hastings, J. Stat. Mech.: Theory Exp. (2007) P08024.

[2] F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, Phys. Rev. Lett. **96**, 220601 (2006).

[3] G. Vidal, Phys. Rev. Lett. **101**, 110501 (2008).

[4] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).

[5] M. Srednicki, Phys. Rev. Lett. **71**, 666 (1993).

[6] C. Holzhey, F. Larsen, and F. Wilczek, Nucl. Phys. B **424**, 443 (1994).

[7] P. Calabrese and J. Cardy, J. Stat. Mech.: Theory Exp. (2004) P06002.

[8] A. Dutta, G. Aeppli, B. K. Chakrabarti, U. Divakaran, T. F. Rosenbaum, and D. Sen, *Quantum Phase Transitions in Transverse Field Spin Models* (Cambridge University Press, Cambridge, England, 2015).

[9] B. Bauer and C. Nayak, J. Stat. Mech.: Theory Exp. (2013) P09005.

[10] L. Herviou, S. Bera, and J. H. Bardarson, Phys. Rev. B **99**, 134205 (2019).

[11] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).

[12] D. Zvyagintseva, H. Sigurdsson, V. K. Kozin, I. Iorsh, I. A. Shelykh, V. Ulyantsev, and O. Kyriienko, Commun. Phys. **5**, 8(2022).

[13] I. Donato, M. Gori, M. Pettini, G. Petri, S. De Nigris, R. Franzosi, and F. Vaccarino, Phys. Rev. E **93**, 052138 (2016).

[14] Q. H. Tran, M. Chen, and Y. Hasegawa, Phys. Rev. E **103**, 052127 (2021).

[15] B. Olsthoorn, J. Hellsvik, and A. V. Balatsky, Phys. Rev. Res. **2**, 043308 (2020).

[16] A. Cole, G. J. Loges, and G. Shiu, Phys. Rev. B **104**, 104426 (2021).

[17] N. Sale, J. Giansiracusa, and B. Lucini, Phys. Rev. E **105**, 024121 (2022).

[18] D. Sehayek and R. G. Melko, Phys. Rev. B **106**, 085111 (2022).

[19] D. Spitz, J. Berges, M. Oberthaler, and A. Wienhard, SciPost Phys. **11**, 060 (2021).

[20] Y. He, S. Xia, D. G. Angelakis, D. Song, Z. Chen, and D. Leykam, Phys. Rev. B **106**, 054210 (2022).

[21] H. Edelsbrunner, D. Letscher, and A. Zomorodian, in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science, FOCS '00* (IEEE Computer Society, Redondo Beach, CA, 2000), p. 454.

[22] A. Zomorodian and G. Carlsson, Discrete Comput. Geom. **33**, 249 (2005).

[23] V. Robins, Topol. Proc. **24**, 503 (1999).

[24] P. Frosini and C. Landi, Pattern Recognition and Image Analysis **9**, 596 (1999).

[25] N. Otter, M. A. Porter, U. Tillmann, P. Grindrod, and H. A. Harrington, EPJ Data Sci. **6**, 17(2017).

[26] M. Žnidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B **77**, 064426 (2008).

[27] P. Bubenik, V. de Silva, and J. Scott, Found. Comut. Math. **15**, 1501 (2015).

[28] C. Villani, *Optimal Transport* (Springer, Berlin, 2009).

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

[30] P. Bubenik, M. Hull, D. Patel, and B. Whittle, Inverse Probl. **36**, 025008 (2020).

[31] C. Maria, P. Dlotko, V. Rouvreau, and M. Glisse, in *GUDHI User and Reference Manual*, 3.4.1 ed. (GUDHI Editorial Board, 2021).

[32] G. De Tomasi, S. Bera, J. H. Bardarson, and F. Pollmann, Phys. Rev. Lett. **118**, 016804 (2017).

[33] E. Iyoda and T. Sagawa, Phys. Rev. A **97**, 042330 (2018).

[34] P. Weinberg and M. Bukov, SciPost Phys. **2**, 003 (2017).

[35] F. Pietracaprina, N. Macé, D. J. Luitz, and F. Alet, SciPost Phys. **5**, 045 (2018).

[36] T. Devakul and R. R. P. Singh, Phys. Rev. Lett. **115**, 187201 (2015).

[37] E. V. H. Doggen, F. Schindler, K. S. Tikhonov, A. D. Mirlin, T. Neupert, D. G. Polyakov, and I. V. Gornyi, Phys. Rev. B **98**, 174202 (2018).

[38] G. Ramírez, J. Rodríguez-Laguna, and G. Sierra, J. Stat. Mech. (2015) P06002.

[39] N. Samos Sáenz de Buruaga, S. N. Santalla, J. Rodríguez-Laguna, and G. Sierra, Phys. Rev. B **104**, 195147 (2021).

[40] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, Science **364**, 260 (2019).

[41] It from qubit: Simons collaboration on quantum fields, gravity and information, https://www.simonsfoundation.org/mathematics-physical-sciences/it-from-qubit/, accessed June 21, 2021.

[42] M. V. Raamsdonk, Gen. Relativ. Gravit. **42**, 2323 (2010).

[43] S. S. Roy, S. N. Santalla, J. Rodríguez-Laguna, and G. Sierra, in *Quantum Theory and Symmetries* (Springer International Publishing, Cham, Switzerland, 2020), pp. 347–357.

[44] V. Allori, *Quantum Mechanics and Fundamentality Naturalizing Quantum Theory between Scientific Realism and Ontological Indeterminacy*, Synthese Library (Springer International Publishing, 2022).

[45] A. Ney, in *Philosophy Beyond Spacetime*, edited by C. Wüthrich, B. L. Bihan, and N. Huggett (Oxford University Press, Oxford, 2020).

[46] C. J. Cao, S. M. Carroll, and S. Michalakis, Phys. Rev. D **95**, 024031 (2017).

[47] J. Preskill, *Lecture Notes for Physics 229:Quantum Information and Computation* (CreateSpace Independent Publishing Platform, 1998).

[48] E. H. Lieb and M. B. Ruskai, J. Math. Phys. **14**, 1938 (1973).

[49] A. Hatcher, *Algebraic Topology* (Cambridge University Press, Cambridge, England, 2001).

[50] J.-D. Boissonnat and C. Maria, J. Appl. Comput. Topology **3**, 59 (2019).

[51] P. Pfeuty, Ann. Phys. **57**, 79 (1970).

[52] E. Barouch and B. M. McCoy, Phys. Rev. A **3**, 786 (1971).

[53] T. J. Osborne and M. A. Nielsen, Phys. Rev. A **66**, 032110 (2002).

[54] Anyon, https://physics.stackexchange.com/users/49107/anyon, Entanglement entropy of the infinite transverse field ising chain at the critical point, Physics Stack Exchange, https://physics.stackexchange.com/q/720404, accessed July 31, 2022.