Persistent Homology of Convection Cycles in Network Flows

Minh Quang Le and Dane Taylor

\textsuperscript{1}Department of Mathematics, University at Buffalo, State University of New York, Buffalo, NY 14260, USA

(Dated: March 15, 2022)

Convection is a well-studied topic in fluid dynamics, yet it is less understood in the context of networks flows. Here, we incorporate techniques from topological data analysis (namely, persistent homology) to automate the detection and characterization of convective/cyclic/chiral flows over networks, particularly those that arise for irreversible Markov chains (MCs). As two applications, we study convection cycles arising under the PageRank algorithm, and we investigate chiral edges flows for a stochastic model of a bi-monomer’s configuration dynamics. Our experiments highlight how system parameters—e.g., the teleportation rate for PageRank and the transition rates of external and internal state changes for a monomer—can act as homology regularizers of convection, which we summarize with persistence barcodes and homological bifurcation diagrams. Our approach establishes a new connection between the study of convection cycles and homology, the branch of mathematics that formally studies cycles, which has diverse potential applications throughout the sciences and engineering.

PACS numbers: 89.75.Hc, 02.50.Ga, 87.15.hj, 84.35.+i

I. INTRODUCTION

One of the main goals of topological data analysis (TDA) is to characterize the structure of an object—usually a point cloud—through its topological features. In particular, persistent homology \[ \text{[2][22]} \] is a family of techniques that detect and summarize multiscale topological features and has been applied to a wide variety of applications including time-series data \[ \text{[10][25]} \], image processing \[ \text{[32]} \], machine learning \[ \text{[20]} \], and artificial intelligence \[ \text{[3][18]} \]. Complementing the study of point-cloud data, another line of research involves utilizing the TDA toolset to study complex systems, for which applications include the analysis of spreading processes over social networks \[ \text{[29]} \], network neuroscience \[ \text{[5][11][26]} \], mechanical-force networks \[ \text{[13]} \], jamming in granular material \[ \text{[14]} \], molecular structure \[ \text{[17]} \], and DNA folding \[ \text{[7]} \]. In this paper, we employ techniques from TDA to study Markov chains (MCs), which provide a foundation to numerous areas of science and engineering including queuing theory \[ \text{[9]} \], population dynamics \[ \text{[12]} \], as well as statistical (and machine learning) models that rely on Markov chain Monte Carlo \[ \text{[4]} \], hidden Markov models \[ \text{[31]} \], and Markov decision process \[ \text{[24]} \].

We utilize the mathematical framework of persistent homology to automate the detection (and summarize the multiscale properties) of convection cycles that arise for the stationary flows of irreversible MCs. Notably, while convection cycles have been extensively studied in fluid dynamics, they are less understood in the context of flows over networks. For example, it was recently observed that the coupling together of reversible MCs can give rise to an irreversible MC with convection cycles that are an emergent property \[ \text{[29]} \]. Emergent convection cycles have also been recently found to describe the phenomenon of “chiral edge flows” \[ \text{[27]} \], providing new insights into the quantum Hall effect, configurational dynamics of monomers, and biological (e.g., circadian) rhythms.

Given the inherent prevalence of convection cycles in MCs and other network flows, it is important that we place their study on a stronger mathematical, computational, and theoretical footing.

We study convection using a branch of mathematics called \textit{homology} and the related field \textit{computational homology} \[ \text{[8]} \]. Both are concerned with studying the absence/presence of \textit{k}-dimensional “holes” (and their connectivity) within a topological space such as a simplicial complex. Importantly, cycles on a graph are 1-dimensional (1D) holes, and so persistent homology is a natural fit to analyze convection cycles. We construct filtrations of graphs by including edges according to the stationary flows across them (which is done in descending order so that the last edges to be included are those with the smallest stationary flows), and we summarize the persistent homology of the filtered graphs’ associated clique complexes. See Fig. 1 for a graph and its associated clique complex. Computationally, we implement these techniques by building on a popular TDA framework called Gudhi \[ \text{[30]} \], which we adapt to implement \textit{edge-value clique (EVC) filtrations} that are defined over the edges of a graph.

We apply this technique to two applications. First, we study convection cycles arising under the PageRank algorithm \[ \text{[23]} \], examining the role of the \textit{teleportation parameter}. Second, we study chiral edge flows that emerge for a 4-state model that describes the configurational changes of a bi-monomer \[ \text{[27]} \], examining the roles of the external and internal transition rates. These parameters significantly affect convection cycles arising for these respective applications, and we show that they act as “homology regularizers” of convection. We introduce “homological bifurcation diagrams” to summarize these effects. Our methods provide mathematically principled (and automated) tools to gain a deeper understanding of the structural patterns of convection on networks, and they are expected to be useful to myriad applications across the physical, social, biological and computational sciences.

The remainder of this paper is organized as follows: We present background information in Sec. \text{[II]} our methodology in Sec. \text{[III]} applications in Sec. \text{[IV]} and a discussion in Sec. \text{[V]}.
II. BACKGROUND INFORMATION

Here, we present introductory material about simplicial complexes and homology (Sec. II A), persistent homology of graphs (Sec. II B), and discrete-time MCs (Sec. II C).

A. Simplicial complexes (SCs) and their homology

We first define an undirected graph \( G = (V, E) \), where \( V = \{1, \ldots, N_0\} \) is a set of \( N_0 \) vertices and \( E \subset V \times V \) is a set of edges. Note that each vertex is specified by a single index \( i \in V \), and each edge is specified by an unordered pair \( (i, j) \in V \times V \). More generally, we define a \((k+1)\)-tuple of vertices \( \sigma = (i_0, i_1, \ldots, i_k) \) as a \( k \)-dimensional simplex, or \( k \)-simplex \([21]\). Vertices and edges are equivalent to 0-simplices and 1-simplices, respectively. An abstract SC is a set of simplices of possibly different dimensions, and it is a generalization of an undirected graph. It is also a type of hypergraph with a constraint on which simplices can exist. That is, for any \( k \)-simplex \( (i_0, i_1, \ldots, i_k) \) in an SC, its faces are the \((k-1)\) simplices in which one of the vertices is omitted (e.g., \( i_1 \) is omitted to yield \((i_0, i_2, \ldots, i_k)\)). The cofaces of a \((k-1)\)-simplex are the \( k \)-simplices for which it is a face. Note that the faces of an edge \((i, j)\) are the vertices \( i \) and \( j \), and likewise, \((i, j)\) is a coface of each of these vertices.

With these definitions, we state the two restrictions that are required for an SC: (i) for any face, its faces must be included in the SC; and (ii) the intersection of any two faces is either a face of both, or it is an empty set. The dimension of an SC is the maximum dimension of its simplices, and an undirected graph is a 1-dimensional SC—it contains 0-simplices and 1-simplices, and for any edge \((i, j)\), the vertices \( i \) and \( j \) must exist. We will focus on a particular type of SC that can be generated from a graph and is called a clique complex. A clique complex \( K(G) \) of a graph \( G \) is the SC in which there is a 1-to-1 correspondence between the \((k+1)\)-cliques in the graph and the \( k \)-simplices in the SC. (Recall that an \( n \)-clique is a complete subgraph on \( n \) vertices of a graph.) Given this 1-to-1 correspondence, the map from \( G \) to \( K(G) \) is invertible, and \( G \) can be recovered as the 1-skeleton of \( K(G) \). (A \( k \)-skeleton of an SC is the SC that is obtained after removing all simplices having dimensions that are greater than \( k \).)

We next discuss simplicial homology, which will lead to a formal definition of “homological” cycles. To this end, we consider vector spaces defined over the \( k \)-simplices in a SC. A \( k \)-chain, \( \sum_{n=1}^{N_k} \sigma_n \alpha_n \), is a linear combination of \( k \)-simplices \( \{\sigma_n\} \) with weights \( \{\alpha_n\} \). (Note that \( N_0 \) and \( N_1 \) are the numbers of vertices and edges, respectively.) If a SC contains \( N_k \) different \( k \)-simplices, then the vector space of \( k \)-chains is \( \mathbb{R}^{N_k} \)-dimensional, and it is isomorphic to \( \mathbb{R}^{N_k} \) if one assumes \( \alpha_k \in \mathbb{R} \). We now consider a simplicial map \( f : X_k \rightarrow X_{k-1} \) between \( X_k \). If \( X_k \) is a SC of dimension \( k \), and \( X_{k-1} \), which is a SC of dimension \( k-1 \) that contains the faces of simplices in \( X_k \). Considering the vector space \( C_k \) of \( k \)-chains defined over \( k \)-simplices in \( X_k \) and vector space \( C_{k-1} \) of \((k-1)\)-chains defined over their cofaces in \( X_{k-1} \), we define the linear boundary map \( \partial_k : C_k \rightarrow C_{k-1} \), where the action of \( \partial_k \) on any \( k \)-simplex is given by

\[
\partial_k(i_0, \ldots, i_k) = \sum_{j=0}^{k} (-1)^j (i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_k).
\]

The boundary map allows one to relate vectors in \( C_k \) to those in \( C_{k-1} \). For example, the boundary of a 2-simplex (i.e., triangle) \((i, j, k)\) is the signed combination if the associated edges, \( \partial_2(i, j, k) = (j, k) - (i, k) + (i, j) \). Notably, the boundary of any closed path is zero, which yields an algebraic definition of a \( k \)-cycle: any \( k \)-chain that lies within the subspace \( Z_k \), where \( Z_k = \ker(\partial_k) \subseteq C_k \) is the vector space of \( k \)-cycles.

Notably, \( k \)-cycles can arise for different reasons, and we distinguish two types. The boundary map satisfies the property \( \partial_k \circ \partial_{k+1} = 0 \), which essentially states that the boundary of a boundary is zero. [For the triangle, \( \partial_1 \circ \partial_2(i, j, k) = \partial_1(j, k) - \partial_1(i, k) + \partial_1(i, j) = (k-j) - (k-i) + (j-i) = 0 \] Thus we define \( B_k = \text{image}(\partial_{k+1}) \) as the subspace of \((k+1)\)-boundaries, and it follows that \( B_k \subseteq Z_k \). In other words, some cycles arise simply because they are boundaries of \((k+1)\)-simplices. For example, observe in Fig. II B that there are two “triangular” cycles that exist around the two 2-simplices, but that there are other cycles that also exist. The \( k \)-th simplicial homology is defined as the quotient space \( \text{H}_k = Z_k / B_k \), and it represents the subspace of \( k \)-dimensional cycles (i.e., \( k \)-cycles) that do not arise simply as the boundary of a \((k+1)\)-simplex.

The \( k \)-th simplicial homology can be represented by the span of homology generators, which are a linearly independent set of \( k \)-chains that span \( H_n \) and represent the associated \( k \)-cycles. The number of linearly independent homology generators is called a Betti number

\[
\beta_k = \dim H_k = \dim(Z_k) - \dim(B_k).
\]

Informally, \( \beta_0 \) is the number of connected components; \( \beta_1 \) is...
the number of 1-dimensional cycles or “loops” (that is, not including the triangular boundaries of 2-simplices); and $\beta_2$ is the number of 2-dimensional holes or “voids” (e.g., the interior of a triangulated sphere). For the SC shown in Fig. 1(B), $\beta_0 = 1$ since there’s one connected component, and $\beta_1 = 2$ since there are two cycles that are not simply the boundaries of 2-simplices.

By formulating $k$-cycles algebraically, one can consider the linear dependence and independence of $k$-cycles. As such, one can not only identify cycles, but also investigate the relations/connectivity between cycles, which we find to be instrumental for understanding pattern formation for cycle. We also highlight that a given homological $k$-cycle can potentially have more than one homological generator. Such generators are said to be homologically equivalent, and they can be obtained by considering linear combinations of $k$-cycles (including both homological $k$-cycles and boundaries). We will later show that this complicates the investigation of convection cycles through the lens of homological $k$-cycles.

### B. Persistent homology of scalar functions defined over edges

One of the greatest tools of topological data analysis is the study of persistent homology [22]. Here, we examine how the homology of a topological object changes as it undergoes a filtration to yield a monotonically increasing sequence $X_0 \subseteq X_1 \subseteq X_1 \subseteq \ldots$ (e.g., of simplicial complexes $\{X_i\}$). We consider filtrations in which one has a scalar function $f : E \to \mathbb{R}$ over the edges, and each edge $(i, j) \in E$ is retained/removed according to $f(i, j)$. The values $f(i, j)$ could be edge weights for a weighted graph, but in general they can encode any other property. We visualize such a graph and the values $f(i, j)$ in Fig. 2(A).

We call the process an edge-value clique (EVC) filtration, and we construct it as follows. Given a graph $G(V, E)$ and a filtration function $f$, we define the subsets $E_\epsilon = \{(i, j) | f(i, j) > \epsilon\}$, and in practice we assume $\epsilon_A > \max_{(i, j) \in E} f(i, j)$ and $\epsilon_B < \min_{(i, j) \in E} f(i, j)$. It then follows that $E_\epsilon = \emptyset$ is an empty set of edges when $\epsilon \geq \epsilon_A$, and $E_\epsilon = E$ (i.e., all edges are retained) when $\epsilon \leq \epsilon_B$. See [16] for our codebase that implements EVC filtrations by adapting the TDA framework called Gudhi [30], and which reproduces the results of this paper.

In Fig. 2(B), we visualize a sequence of filtered clique complexes $\{K_\epsilon\}$ that are associated with the filtered graphs $\{G_\epsilon\}$ that are defined with the edge sets $\{E_\epsilon\}$. In Fig. 2(C), we summarize the persistent homology of $\{K_\epsilon\}$ in a persistence barcode, which reveals how homology changes with $\epsilon$. Observe that when $\epsilon$ is sufficiently large, $K_\epsilon$ contains vertices but no edges. On the other hand, when $\epsilon$ decreases to be sufficiently small, then $K_\epsilon$ recovers the original clique complex [recall Fig. 1(B)]. The values of $\epsilon$ that were used to create Fig. 2(B) are indicated by the vertical dotted lines in Fig. 2(C).

Each horizontal bar in the persistence barcode shown in Fig. 2(C) indicates the lifetime of a homological 1-cycle—that is, the values of $\epsilon$ for which it exists. The red and blue bars reflect 0-homology and 1-homology respectively. The dimensions of the homology spaces (i.e., Betti numbers) can be found by counting the number of homological 1-cycles at any particular $\epsilon$. For example, one can observe that $\beta_1 = 0$ when $\epsilon = 3.5$, $\beta_1 = 1$ when $\epsilon = 2.5$, and $\beta_1 = 2$ when $\epsilon = 1.5$. Clearly, the homological 1-cycles are undergoing bifurcations as $\epsilon$ varies. A persistence barcode is convenient to identify for each generator: the value $\epsilon_b$ of $\epsilon$ when it is “born” (i.e., the homological $k$-cycle does not exist when $\epsilon > \epsilon_b$); the value $\epsilon_d$ of $\epsilon$ when it “dies” (i.e., the homological $k$-cycle...
does not exist when $\epsilon < \epsilon_d$; its lifetime $(\epsilon_d, \epsilon_b]$; and lifespan $(\epsilon_d - \epsilon_b]$. A cycle's lifespan quantifies its persistence under the filtration, and it is often interpreted as a proxy for the cycle's significance (although short-lifetime cycles can also be important in certain contexts).

C. Convection cycles for irreversible Markov chains (MCs)

We will apply persistence homology to study convection cycles in irreversible MCs [19], which we now briefly summarize. A discrete-time MC is a "memoryless" random process in which for time steps $t = 0, 1, 2, \ldots$, the system state $S_t \in \mathcal{V}$ satisfies the Markov property $P[S_{t+1} = i|S_0 = i_0, ..., S_t = i_t] = P[S_{t+1} = i|S_t = i_t]$, which implies that the probability of a state occurring at the next time step only depends on the current state and not earlier states. In our case, we consider MCs that correspond to a random walk on a (possibly) weighted and directed graph having an adjacency matrix $A$ in which $A_{ij} \in \mathbb{R}$ is nonzero if $(i, j)$ is an edge, $(i, j) \in \mathcal{E}$, and $A_{ij} = 0$ otherwise. We similarly define a transition matrix, $P = D^{-1}A$, where $D$ is a diagonal matrix with entries that encode the (possibly) weighted vertex degrees $D_{ii} = \sum_j A_{ij}$. For directed graphs, each $(i, j)$ is considered to be an ordered pair, and each $D_{ii}$ encodes the out-degree of vertex $i$. Each matrix element $P_{ij}$ gives the probability for a random walk to transition from vertex $i$ to $j$. Letting $x_i(t)$ denote the probability that the system is in state $i$ (or equivalently, the probability that a random walker is at vertex $i$) at time $t$, one can utilize the Markov property to obtain the linear discrete-time system $x_i(t+1) = \sum_j x_j(t)P_{ij}$. By defining $x(t) = [x_1(t), \ldots, x_{N_0}]^T$, one equivalently has

$$x(t+1)^T = x(t)^T P.$$  

(3)

Since $x(t)$ is a vector of probabilities, we assume that it is normalized in 1-norm, $\sum_i x_i = 1$.

Herein, we focus on network flows after a system reaches a stationary state, in which case $x(t)$ converges to a limiting vector $\pi = \lim_{t \to \infty} x(t)$ that satisfies the eigenvalue equation $\pi^T = P\pi$. By construction, $\pi$ is a vector of probabilities and contains nonnegative entries. Furthermore, as a row-stochastic matrix, $P$ has an eigenvalue equal to one (i.e., the largest eigenvalue) and its right dominant eigenvector is the vector containing 1's as entries. Our assumption of convergence requires that matrix $P$ is an irreducible and aperiodic [1] or that the initial condition $x(0)$ lies in a converging subspace. In the stationary state, the stationary flow across each edge $(i, j)$ per time step is given by

$$F_{ij} = \pi_i P_{ij}.$$  

(4)

We study convection cycles using an approach that was developed in [28]. Specifically, for each edge we define the stationary flow imbalance

$$\Delta_{ij} = F_{ij} - F_{ji}.$$  

(5)

By construction, $\Delta_{ij} = -\Delta_{ji}$, and we say that the imbalance direction is from $i$ to $j$ when $\Delta_{ij} > 0$. Importantly, the defining feature of a reversible MC is that $\Delta_{ij} = 0$ for all $i$ and $j$. That is, the directional flows match $\pi_i P_{ij} = \pi_j P_{ji}$ for any edge $(i, j)$. This is the case for any undirected graph, since in this case $A = A^T$, and it follows that $\pi_i = D_{ii}/\sum_j D_{ij}$. In contrast, an irreversible MC yields asymmetric stationary flows and $\Delta_{ij}$ is nonzero for some edges. To formally define convection cycles, we consider a new graph $G_{\Delta}(\mathcal{V}, \mathcal{E}_\Delta)$ such that each positive value $\Delta_{ij}$ gives rise to a directed edge $(i, j, \Delta_{ij}) \in \mathcal{E}_\Delta$ having weight $\Delta_{ij}$. We then define a convection cycle to be any non-intersecting closed path in $G_{\Delta}(\mathcal{V}, \mathcal{E}_\Delta)$.

In Fig.3, we illustrate for an example MC how flow imbalances manifest as a pattern of convection cycles. In Figs.3(A), 3(B), and Fig. 3(C), we use edge colors to indicate the stationary distribution $\pi$, stationary edge flows $F_ij$, and flow imbalances $\Delta_{ij}$, respectively. Observe that some of the arrows in Figs.3(A)–(B) are bidirectional, since some of the graph's edges are bidirectional. In contrast, the arrows in Fig.3(C) are exclusively directed since they now indicate the directions of flow imbalances. There exists an edge $i \rightarrow j$ only if $\Delta_{ij} > 0$, which also implies $j \rightarrow i$ is not an edge since $\Delta_{ji} = -\Delta_{ij}$. Observe in Fig.3(C) that this yields five convection cycles. In
FIG. 4. Persistent homology of convection cycles. (A) Visualization of an EVC filtration applied to flow imbalances arising for the irreversible MC shown in Fig. 3, and we use the magnitude $|\Delta_{ij}|$ of flow imbalance as the filtration function $f : E \rightarrow \mathbb{R}$. We indicate flow imbalances’ directions with arrows, noting that the clique complexes that are constructed by the filtration are undirected, since the filtration does not incorporate information about edge directions. (B) Persistence barcodes for homological 1-cycles. Observe that the 1-cycle that first appears dies before the other 1-cycles are born.

Sec. III B, we will further discuss these convection cycles and their relation to homological 1-cycles.

Before continuing, we highlight that convection cycles revealed through flow imbalances [28] do not take into account the probability of transitioning to or away from a convection cycle, and so they are not necessarily “cyclic traps.” That is, the presence of a convection cycle does not imply that it is unlikely for a random walker to leave (or move in an opposite direction as) the cycle. For example, observe in Fig. 3 that the counter-clockwise flow around convection cycle $A \rightarrow B \rightarrow C \rightarrow D$ is approximately 0.025, yet there is a flow of approximately 0.02 that leaves the cycle at node D, and a flow of approximately 0.15 moves in the opposite direction from node C to B. Future research will likely uncover complementary notions of convection with different advantages/disadvantages, and our proposed techniques using persistent homology can likely be similarly extended.

III. HOMOLOGICAL ANALYSES OF CONVECTION

We now employ persistent homology to automate the detection, characterization, and summarization of the homological patterns of convection cycles. In Sec. III A we study the MC that was presented in Fig. 3. In Sec. III B, we discuss the relation between convection cycles and homological 1-cycles.

A. Persistent homology of convection cycles

Recall from Sec. III B that EVC filtrations were defined for an undirected graph with a scalar function defined on the edges. Therefore, given an MC corresponding to a (potentially) directed and weighted graph, we first consider the associated undirected graph. Then we study homology under an EVC filtration in which the filtration function $f : E \rightarrow \mathbb{R}$ is given by the magnitudes of the flow imbalances

$$f(i, j) = |\Delta_{ij}|.$$

In this way, the persistent homology that is revealed corresponds to the convection cycles that arise under flow imbalances.

In Fig. 4, we visualize persistence barcodes for an EVC filtration associated with the convection cycles shown in Fig. 3. Note that this figure is analogous to Fig. 2 C), where we had previously chosen the filtration function $f$ to be the edge weights. Since we now use a different function $f$, the cycles now have different births, deaths, lifetimes and lifespans. Interestingly, the 1-cycle involving vertices $\{A, B, C, D\}$ is now born and dies before the other two 1-cycles are born. While there is an obvious connection between the EVC homology of a graph induced by edge weights and that which is induced by convective flows, this relation remains unclear and should be explored in future work.

We note that one could also construct EVC filtrations by increasing $\epsilon$ and retaining edges $(i, j)$ for which $|\Delta_{ij}|$ is smaller than $\epsilon$. In Appendix A, we provide an example illustrating why EVC filtrations with decreasing $\epsilon$ are superior to those with increasing $\epsilon$ for the goal of studying convection cycles. In particular, EVC filtrations that decrease $\epsilon$ focus on 1-cycles that are associated with large-flow convection cycles (i.e., large values of $|\Delta_{ij}|$), which we consider to be the ones that are more significant. In contrast, EVC filtrations that increase $\epsilon$ focus on 1-cycles that are associated with small-flow convection cycles (i.e., small values of $|\Delta_{ij}|$), which we consider to be less significant.

B. Comparing convection cycles and homological 1-cycles

We propose to study pattern formation for convection cycles using persistent homology techniques for homological
1-cycles; however, one should keep in mind that these are
different notions for cycles. Homological 1-cycles are
1-dimensional holes for a topological space, and k-cycles
generalize to higher dimensional by representing higher-
dimensional holes (Sec. II A). In contrast, we define convec-
tion cycles to be closed non-backtracking paths in a directed
graph that encodes flow imbalances (Sec. II C). In this section,
we will clarify the relationship between homological k-cycles
and convection cycles, thereby revealing the capabilities and
disadvantages of existing persistent homology techniques for
studying convection cycles. Continuing with the previous ex-
ample [see Figs. 3-4], we find that flow imbalances give rise to
five convection cycles, which we enumerate I–V and visual-
ize in Fig. 5(A). In contrast, we identify three homological 1-
cycles using persistent homology with EVC filtrations, which
we enumerate (i)–(iii) and visualize in Fig. 5(B).

Observe that there is a one-to-one correspondence between
convection cycles I and II and homological 1-cycles (i) and
(ii). Also observe that homological 1-cycle (iii) has three ho-
mologically equivalent generators, and any of them can be
used to represent the 1-cycle (which again, is defined as a
1-dimensional hole). Each subsequent generator can be ob-
tained via a topological retraction in which a 2-simplex is col-
lapsed down onto one of its edges. Interestingly, the first two
homological generators for 1-cycle (iii) correspond to convec-
tion cycles III and IV. In contrast, the third generator corre-
sponds to a loop that is not a convection cycle, since the flow-
imbalances’ directions do not point in a consistent direction
along the cycle (i.e., clockwise or counter-clockwise). Finally,
observe that convection cycle V is a boundary of a 2-simplex,
and it therefore does not contribute to the 1-dimensional sim-
plcial homology.

Thus, it is important to not misinterpret one notion of cy-

FIG. 5. Relation between convection cycles and homological 1-cycles. (A) The flow imbalances shown in Fig. 3(C) give rise to five
convection cycles, which we label I–V. (B) Persistent homology using EVC filtrations applied to a network of flow imbalances reveals three
homological 1-cycles, which we label (i)–(iii). Each homological 1-cycle represents a “1-dimensional hole” and can be represented by one
or more homological generator (recall Sec. II A). Observe that there is a one-to-one correspondence between convection cycles I and II and
homological 1-cycles (i) and (ii). In contrast, there are three homologically equivalent generators for 1-cycle (iii) as shown. Two of the
generators correspond to convection cycles III and IV. The third generator does not correspond to a convection cycle, because the edge
directions are not consistently in the same orientation (i.e., always clockwise or counter-clockwise).

IV. APPLICATIONS

In this section, we apply our approach to two applications. In Sec. IV A we study MCs arising for the Google PageRank
algorithm, exploring how convection cycles are effected by the teleportation parameter α. In Sec. IV B we study a type
of emergent convection cycle called a chiral edge flow.

A. Teleportation is a homology regularizer for PageRank

We now study the persistent homology of convection cycles
arising for the PageRank algorithm [15, 23], which is a popu-
lar technique to rank the importance of vertices in graphs. It
has been applied to numerous applications (see survey [6],

1-cycles for the other. At the same time, our findings in Fig. 5
also highlight that there is a need for new persistent homology
techniques that cater specifically to convection cycles and
directed graphs. For example, if one were to omit the 2-simplex
that involves vertices A, B and D from the clique complexes
that arise under an EVC filtration, then convection cycle V
would coincide with a homological 1-cycle. However, the
aim of this paper is not to develop new methods for persist-
ent homology. Instead, we proposed to begin this pursuit by
studying convection cycles using existing methods for persist-
ent homology. Even though there is not an exact one-to-one
match between convection cycles and homological 1-cycles,
because they are closely related, we find that persistent homol-
ogy can effectively detect and summarize convection cycles’
patterns.
but most notably, for many years it was utilized by Google to rank websites and facilitate web search. The PageRank of a vertex \( i \) is given by the stationary density \( \pi_i(\alpha) \) of the Markov chain (MC) having a transition matrix of the form

\[
P(\alpha) \equiv \alpha P + (1 - \alpha) N^{-1} 11^T,
\]

where \( P \) is the transition matrix described in Sec. II C and \( \alpha \in (0, 1) \) is the teleportation parameter. As \( \alpha \to 1 \), the second term vanishes and \( P(\alpha) \to P \). Usually, \( \alpha \) is chosen to be near 1 (often 0.85) so that the second term can be considered as a small perturbation that improves the mathematical characteristics of \( P \)—or more formally, it is a “regularization” of matrix \( P \). In particular, when \( \alpha \in (0, 1) \) the matrix \( P(\alpha) \) is guaranteed to be irreducible, aperiodic and positive, and the Perron-Frobenius theorem ensures that its dominant left eigenvector \( \pi \) is unique and has positive entries (i.e., \( \pi_i(\alpha) > 0 \) for all \( i \)). In other words, the PageRanks are well-defined for all vertices.

We now show that the introduction of teleportation also regularizes the homology of convection cycles. In this experiment, we construct EVC filtrations with the filtration function \( f(i, j) = |\Delta_{ij}(\alpha)| \), which now depends on \( \alpha \). In Fig. 6(A), we illustrate for several choices of \( \alpha \) the flow imbalances that arise under PageRank, which we apply to the graph from Fig. 3(A). In Fig. 6(B), we visualize their associated persistence barcodes, which we create using EVC filtrations. Note that the choice \( \alpha = 1 \) recovers the transition matrix, stationary distribution, flow imbalances, and persistence barcodes that were previously studied in Figs. 3 and 4.

Observe that the homological patterns of convection cycles significantly change with \( \alpha \). For example, when \( \alpha \) is sufficiently small, the homological 1-cycle \( \{A, B, C, D\} \) vanishes—it is “washed out” by the introduction of teleportation. In other word, \( \alpha \) is a homology regularizer. This is further illustrated in Fig. 7(A), where we plot the birth and death times of homological 1-cycles versus \( \alpha \). For comparison, we also plot the PageRanks \( \pi_i(\alpha) \) in Fig. 7(B). The vertical line highlights that one of the 1-cycles vanishes when \( \alpha \) decreases (approximately) below \( \alpha^* \approx 0.54 \).

In Appendix A, we present additional experiments that explore convection cycles arising under PageRank with \( \alpha = 0.8 \). We show that homological 1-cycles arising for EVC filtrations with decreasing filtration parameter \( \epsilon \) reveal patterns for large-
The system is contains two monomers of sizes \( s_1 \) and \( s_2 \), respectively, giving the external state \((s_1, s_2)\). Moreover, there are four internal states: a, b, c, and d. Transitions involving changes to external and internal states occur at rates \( \gamma_{\text{ex}} \) and \( \gamma_{\text{in}} \), respectively.

flow convection cycles. In contrast, when EVC filtrations are constructed with increasing \( \epsilon \), we find that the resulting homological 1-cycles relate to small-flow convection cycles, and in particular, those involving low probability teleportation transitions.

**B. Persistent homology of chiral edge flows**

Our second application investigates homological patterns of convection cycles that arise for an MC that models the stochastic configuration dynamics of two monomers. We adopt the same notation as in [27], which motivated our experiment. The monomer configuration (i.e., “external state”) is given by the number of monomers of each type, \((s_1, s_2)\), whereas the “internal state” is one of four possibilities: a, b, c, or d. Transitions that involve a change of internal state occur at rate \( \gamma_{\text{in}} \), whereas transitions between involving external states (i.e., the addition or removal of a monomer) occur at rate \( \gamma_{\text{ex}} \). The resulting MC can be visualized as a 2-dimensional lattice, which we visualize in Fig. 8.

In Fig. 8(A), we visualize flow imbalances for transitions between the external states. We fix \( \gamma_{\text{in}} = 0.01 \) and consider several \( \gamma_{\text{ex}} \). Observe that as \( \gamma_{\text{ex}} \) increases, a large counterclockwise convection cycle emerges on the boundary (i.e., “edge”) of the lattice. This type of convection cycle is called a chiral edge flow, and such convection cycles have important implications for the quantum Hall effect, biological rhythms, and the dynamics of monomers [27]. In the limit \( \gamma_{\text{ex}} \gg \gamma_{\text{in}} \), there is an emergence of a chiral edge flow, i.e., a convection cycle around the lattice’s outer boundary. (B) The corresponding persistence barcodes capture the emergence of this prominent convection cycle and other convection cycles within the lattice.

**FIG. 9. Persistent homology chiral edge flow.** (A) Flow imbalances \( \Delta_{ij} \) between external states for the bi-monomer shown in Fig. 8 with \( \gamma_{\text{in}} = 0.01 \) and different \( \gamma_{\text{ex}} \). In the limit \( \gamma_{\text{ex}} \gg \gamma_{\text{in}} \), there is an emergence of a chiral edge flow. (B) The corresponding persistence barcodes capture the emergence of this prominent convection cycle.
tent homology methods to study convection cycles as well as other functions and signals defined on directed graphs. In Sec. II B, we discuss the relation between convection cycles and homological 1-cycles, and we showed that these are two closely related, but notably different, notions of cycles. Sometimes there is a one-to-one correspondence between these cycles, and sometimes the relation is more complicated, due in part to the fact that a given homological k-cycle can be equivalently represented by possibly more than one homological generator. Such generators may or may not correspond to convection cycles. Moreover, convection cycles can also correspond to the boundaries of 2-simplices, and as such, they will not be identified via the traditional tools of persistent homology. Developing persistent homology techniques that cater to convection cycles, and which specifically account for edge directions, remains an important open challenge for the applied mathematics and physics communities.

Our work opens up several other new lines of research that are also worth noting. First, convection cycles were recently found to be an emergent property of multiplex Markov chains [28] in which a set of (intralayer) Markov chains are coupled together by another set of (interlayer) Markov chains. It would be interesting to employ persistent homology to gain a deeper understanding of this phenomenon. Second, chiral edge flows are known to be important to other applications including the quantum Hall effect and biological rhythms [27], and future work could utilize our methods to investigate these exciting applications. Notably, our methods can reveal convection cycles that exist in addition to a chiral edge flow, which may specifically account for edge directions, remains an important open challenge for the applied mathematics and physics communities.

See [16] for a codebase that reproduces our results and can be used to study the persistent homology for convection cycles arising for other applications.

Appendix A: Convection cycles are better revealed by filtrations that decrease the filtration parameter $\epsilon$ versus increase $\epsilon$

In Sec. II B, we defined EVC filtrations in which one decreases a filtration parameter $\epsilon$, retaining edges for which $f(i,j) < \epsilon$. Our numerical experiments that study convection cycles using persistent homology use this approach and let the filtration be given by the flow imbalances $f(i,j) = \left| \Delta_{ij} \right|$. By decreasing $\epsilon$, the cycles that are first revealed correspond to large-flow convection cycles, which we consider to be the ones that are more significant. One could also construct EVC filtrations by increasing $\epsilon$ and retaining edges for which $f(i,j) > \epsilon$. Here, we show that this latter filtration reveals 1-cycles that relate to small-flow convection cycles, which we consider to be insignificant.

In Fig. 10, we study EVC filtrations applied to flow imbalances arising under the PageRank algorithm with $\alpha = 0.8$ for the same MC that we investigated in Sec. [IV A] In Fig. 10(A) and Fig. 10(B), we illustrate EVC filtrations with decreasing and increasing $\epsilon$, respectively. Observe in Fig. 10(A) that the 1-cycles revealed by decreasing $\epsilon$ correspond to large-flow convection cycles. In contrast, observe in Fig. 10(B) that the 1-cycles revealed by increasing $\epsilon$ are small-flow cycles that relate to low-probability transitions that occur due to teleportation.

ACKNOWLEDGMENTS

MQL and DT were supported in part by the National Science Foundation (DMS-2052720 and EDT-1551069) and the Simons Foundation (grant #578333).

[1] Ravindra B Bapat. A max version of the perron-frobenius theorem. Linear Algebra and its Applications, 275:3–18, 1998.
[2] Herbert Edelsbrunner and John Harer. Computational topology: an introduction. American Mathematical Soc., 2010.
[3] Maxime Gabella. Topology of learning in artificial neural networks. arXiv preprint arXiv:1902.08160, 2019.
[4] Walter R Gilks, Sylvia Richardson, and David J Spiegelhalter. Introducing markov chain monte. Markov chain Monte Carlo in practice, page 1, 1995.
[5] Chad Giusti, Eva Pastalkova, Carina Curto, and Vladimir Itskov. Clique topology reveals intrinsic geometric structure in neural correlations. Proceedings of the National Academy of Sciences, 112(44):13455–13460, 2015.
[6] David F Gleich. Pagerank beyond the web. siam REVIEW, 57(3):321–363, 2015.
[7] Takashi Ichinomiya, Ippei Obayashi, and Yasuaki Hiraoka. Protein-folding analysis using features obtained by persistent homology. Biophysical Journal, 118(12):2926–2937, 2020.
[8] Tomasz Kaczynski, Konstantin Mischaikow, and Marian Mrozek. Computational homology, volume 157. Springer Science & Business Media, 2006.
[9] David G Kendall. Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain. The Annals of Mathematical Statistics, pages 338–354, 1953.
[10] Firas A Khasawneh and Elizabeth Munch. Chatter detection in turning using persistent homology. Mechanical Systems and Signal Processing, 70:527–541, 2016.
[11] Bengier Ulgen Kilic and Dane Taylor. Simplicial cascades are orchestrated by the multidimensional geometry of neuronal complexes. arXiv preprint arXiv:2201.02071, 2022.
[12] John FC Kingman. Markov population processes. Journal of Applied Probability, pages 1–18, 1969.
[13] L Kondic, A Goullet, CS O’Hern, M Kramar, Konstantin Mischaikow, and RP Behringer. Topology of force networks in compressed granular media. EPL (Europhysics Letters), 97(5):54001, 2013.
[14] M Kramar, Arnaud Goullet, Lou Kondic, and Konstantin Mischaikow. Persistence of force networks in compressed granular media. Physical Review E, 87(4):042207, 2013.
[15] Amy N Langville and Carl D Meyer. Updating markov chains with an eye on google’s pagerank. SIAM journal on matrix
FIG. 10. **Comparing EVC filtrations with decreasing and increasing filtration parameter** $\epsilon$. Extending our study in Sec. IV A that uses persistent homology to study convection cycles arising for a MC under the PageRank algorithm with $\alpha = 0.8$, we now study homological 1-cycles obtained via two different EVC filtrations. (A) Similar to our results in Fig. 6, we construct EVC filtrations by including edges for which $|\Delta_{ij}| > \epsilon$ while decreasing $\epsilon$. Observe that the 1-cycles reveal large-flow convection cycles that are associated with large values of $|\Delta_{ij}|$. (B) For comparison, we construct EVC filtrations by including weighted edges $|\Delta_{ij}| < \epsilon$ while increasing $\epsilon$. Observe that these 1-cycles now correspond to small-flow convection cycles that are associated with small values of $|\Delta_{ij}|$. They primarily describe low-probability transitions that occur due to teleportation. In this work we focus on EVC filtrations with decreasing $\epsilon$, since we consider high-flow convection cycles to be the ones that are most important.

[16] Minh Quang Le. Codebase for persistent homology of convection cycles in network flows https://github.com/minhquan89/Persistent-Homology-of-Convection-Cycles.
[17] Jie Liang, Herbert Edelsbrunner, Ping Fu, Pambdihantam V Sudhakar, and Shankar Subramaniam. Analytical shape computation of macromolecules: I. molecular area and volume through alpha shape. *Proteins: Structure, Function, and Bioinformatics*, 33(1):1–17, 1998.
[18] Shusen Liu, Di Wang, Dan Maljovec, Rushil Anirudh, Jayaraman J Thiagarajan, Sam Ade Jacobs, Brian C Van Essen, David Hysom, Jae-Seung Yeon, Jim Gaffney, et al. Scalable topological data analysis and visualization for evaluating data-driven models in scientific applications. *IEEE transactions on visualization and computer graphics*, 26(1):291–300, 2019.
[19] László Lovász et al. Random walks on graphs: A survey. *Combinatorics, Paul Erdős is eighty*, 2(1):1–46, 1993.
[20] Francis Motta, Christopher Tralie, Rossella Bedini, Fabiano Bini, Gilberto Bini, Hamed Eramian, Marcio Gameiro, Steve Haase, Hugh Haddox, John Harer, et al. Hyperparameter optimization of topological features for machine learning applications. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pages 1107–1114. IEEE, 2019.
[21] The notion of “dimension” is more interpretable for the case of a non-abstract simplicial complex, for which the vertices correspond to locations in a Euclidean metric space. In that case, every $k$-simplex is defined as the $k$-dimensional surface that is contained by its faces, which themselves are $(k−1)$-dimensional surfaces. For example, a 2-simplex is a triangle defined as the interior of 3 line segments (which are the 3 1-dimensional cofaces of the 2-simplex).
[22] Nina Otter, Mason A Porter, Ulrike Tillmann, Peter Grindrod, and Heather A Harrington. A roadmap for the computation of persistent homology. *EPJ Data Science*, 6(1):17, 2017.
[23] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.

[24] Ronald Parr and Stuart Russell. Reinforcement learning with hierarchies of machines. Advances in neural information processing systems, pages 1043–1049, 1998.

[25] Jose A Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. Foundations of Computational Mathematics, 15(3):799–838, 2015.

[26] Giovanni Petri, Paul Expert, Federico Turkheimer, Robin Carhart-Harris, David Nutt, Peter J Hellyer, and Francesco Vaccarino. Homological scaffolds of brain functional networks. Journal of The Royal Society Interface, 11(101):20140873, 2014.

[27] Evelyn Tang, Jaime Agudo-Canalejo, and Ramin Golestanian. Topology protects chiral edge currents in stochastic systems. Physical Review X, 11(3):031015, 2021.

[28] Dane Taylor. Multiplex markov chains: Convection cycles and optimality. Physical Review Research, 2(3):033164, 2020.

[29] Dane Taylor, Florian Klimm, Heather A Harrington, Miroslav Kramár, Konstantin Mischaikow, Mason A Porter, and Peter J Mucha. Topological data analysis of contagion maps for examining spreading processes on networks. Nature communications, 6:7723, 2015.

[30] The GUDHI Project. GUDHI User and Reference Manual. GUDHI Editorial Board, 3.4.1 edition, 2021.

[31] Luke Tierney. Markov chains for exploring posterior distributions. the Annals of Statistics, pages 1701–1728, 1994.

[32] Ying-Jie Xin and Yuan-Hua Zhou. Topology on image processing. In Proceedings of ICSIPNN’94. International Conference on Speech, Image Processing and Neural Networks, pages 764–767. IEEE, 1994.