Persistent Homology of Convection Cycles in Network Flows

Minh Quang Le1, *∗* and Dane Taylor1, *†*

1*Department of Mathematics, University at Buffalo, State University of New York, Buffalo, NY 14260, USA*

(Dated: April 18, 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 PageR- ank 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 con- vection, 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

1. **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, per- sistent homology [2, 22] is a family of techniques that de- tect and summarize multiscale topological features and has been applied to a wide variety of applications including time- series data [10, 25], image processing [32], machine learn-

ing [20], and artificial intelligence [3, 18]. Complementing the study of point-cloud data, another line of research in- volves utilizing the TDA toolset to study complex systems, for which applications include the analysis of spreading processes over social networks [29], network neuroscience [5, 11, 26], mechanical-force networks [13], jamming in granular mate- rial [14], molecular structure [17], and DNA folding [7]. In this paper, we employ techniques from TDA to study Markov chains (MCs), which provide a foundation to numerous ar- eas of science and engineering including queuing theory [9], population dynamics [12], as well as statistical (and machine learning) models that rely on Markov chain Monte Carlo [4], hidden Markov models [31], and Markov decision process [24].

We utilize the mathematical framework of persistent ho- mology to automate the detection (and summarize the mul- tiscale properties) of convection cycles that arise for the sta- tionary 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 [29]. Emer- gent convection cycles have also been recently found to de- scribe the phenomenon of “chiral edge flows” [27], providing new insights into the quantum Hall effect, configurational dy- namics of monomers, and biological (e.g., circadian) rhythms.

*∗* [minhquan@buffalo.edu](mailto:minhquan@buffalo.edu)

*†* [danet@buffalo.edu](mailto:danet@buffalo.edu)

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 theo- retical footing.

We study convection using a branch of mathematics called *homology* and the related field *computational homology* [8]. Both are concerned with studying the absence/presence of *k*- dimensional “holes” (and their connectivity) within a topolog- ical 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. Com- putationally, we implement these techniques by building on a popular TDA framework called Gudhi [30], which we adapt to implement *edge-value clique (EVC) filtrations* of scalar- functions that are defined over the the edges of a graph.

We apply this technique to two applications. First, we study convection cycles arising under the PageRank algorithm [23], examining the role of the *teleportation parameter*. Second, we study chiral edge flows that emerge for a 4-state model that describes the configurational changes of a bi-monomer [27], examining the roles of the external and internal transi- tion 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 ex- pected 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. II, our methodology in Sec. III, applications in Sec. IV, and a discussion in Sec. V.

1. **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).

1. **Simplicial complexes (SCs) and their homology**

We first define an *undirected graph G* = *{V, E}*, where *V* = *{*1*, . . . , N*0*}* is a set of *N*0 vertices and *E ⊂ V × V* is a set of edges. Note that each vertex is specified by a single

(*i, j*) *∈ V × V*. More generally, we define a (*k* + 1)-tuple of vertices *σ* = (*i*0*, i*1*, . . . , ik*) as a *k*-dimensional simplex, index *i ∈ V* , and each edge is specified by an unordered pair

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 *hy-*

is, for any *k*-simplex (*i*0*, i*1*, . . . , ik*) in an SC, its *faces* are the (*k −* 1) simplices in which one of the of indices is omitted *pergraph* with a constraint on which simplices can exist. That (e.g., *i*1 is omitted to yield (*i*0*, i*2*, . . . , ik*)). 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 a 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 undi- rected 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

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 be generated from a graph and is called a *clique complex*. A

is a complete subgraph on *n* + 1 vertices of a graph.) Given this 1-to-1 correspondence, the map from *G* to *K*(*G*) is in- graph and the *k*-simplices in the SC. (Recall that an *n-clique* vertible, and *G* can be recovered as the 1-skeleton of *K*(*G*). (A *k-skeleton* of a SC is the SC that is obtained after removing

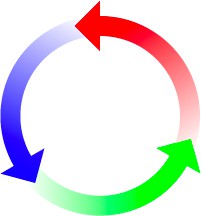
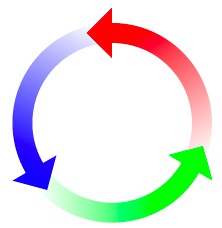
all simplices having dimensions that are greater than *k*.)

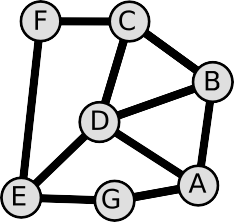
We next discuss *simplicial homology*, which will lead to a formal definition of “homological” cycles. To this end, we consider vector spaces defined over the *k*-simplices in a SC. A

*k-chain*, Σ*Nk αnσn*, is a linear combination of *k*-simplices

*n*=1

(A) (B)



FIG. 1. **A graph and an associated simplicial complex (SC).**

(A) Graph *G* with *N* = 7 vertices and *M* = 10 undirected, un- weighted edges. (B) The corresponding *clique complex S*, where each *k*-clique gives rise to a (*k −* 1)-simplex. Each triangle (i.e., 3-clique) gives rise to a 2-simplex (see shaded triangles). The SC is a topological space and has associated vector spaces. Consider the space R7 of real-valued functions defined over the vertices. Since there is just 1 connected component, the SC’s 0-dimensional (0D) homology is a 1D subspace of R7. Similarly, there are two “homo- logical” 1-cycles that are not a boundary of a 2-simplex, and so the 1D homology is a 2D subspace of R10 (i.e., the space of real-valued functions defined over the ten edges).

on any *k*-simplex is given by

*∂k*(*i*0*, ..., ik*) = Σ(*−*1)*j*(*i*0*, ..., ij−*1*, ij*+1*, ..., ik*)*.* (1)

*k*

*j*=0

The boundary map allows one to relate vectors in *Ck* to those in *Ck−*1. For example, the *boundary* of a 2-simplex (i.e., trian-

gle) (*i, j, k*) is the signed combination if the associated edges,

*∂*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 *Zk*, where

*Zk* = ker(*∂k*) *⊆ Ck* is the *vector space of k-cycles*.

Notably, *k*-cycles can arise for different reasons, and we

erty *∂k ◦ ∂k*+1 = 0, which essentially states that the boundary of a boundary is zero. [For the triangle, *∂*1 *◦ ∂*2(*i, j, k*) = distinguish two types. The boundary map satisfies the prop-

*∂*1(*j, k*) *− ∂*1(*i, k*) + *∂*1(*i, j*) = (*k − j*) *−* (*k − i*) + (*j − i*) =

0.] Thus we define *Bk* = image(*∂k*+1) as the *subspace of* (*k* + 1)*-boundaries*, and it follows that *Bk ⊆ Zk*. In other of (*k* + 1)-simplices For example, observe in Fig. 1(B) that words, some cycles arise simply because they are boundaries

there are two “triangular” cycles that exist around the two

2-simplices, but that there are two other cycles that also ex- ist. The *k-th simplicial homology* is defined as the quotient

space *H* = *Z /B* , and it represents the subspace of *k*-

*k*

*k*

*k*

*{σn}* with weights *{αk}*. (Note that *N*0 and *N*1 are the num-

bers of vertices and edges, respectively.) If a SC contains *Nk* different *k*-simplices, then the vector space of *k*-chains is *Nk*-dimensional, and it is isomorphic to R*Nk* if one assumes

*αk ∈* R. We now consider a simplicial map *f* : *Xk → Xk−*1

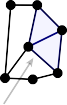
is a SC of dimension *k−* 1 that contains the faces of simplices between *Xk*, which is a SC of dimension *k*, and *Xk−*1, which over *k*-simplices in *Xk* and vector space *Ck−*1 of (*k −* 1)- in *Xk*. Considering the vector space *Ck* of *k*-chains defined ear *boundary map ∂k* : *Ck −→ Ck−*1, where the action of *∂k* chains defined over their cofaces in *Xk−*1, we define the lin-

the boundary of a (*k* + 1)-simplex. dimensional cycles (i.e., *k*-cycles) that do not arise simply as

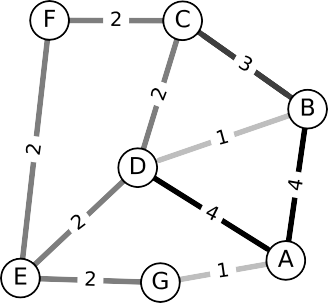
The *k*-th simplicical homology can be represented by the

span of *homology generators*, which are a linearly indepen- dent set of *k*-chains that span *Hn* and represent the associated *k*-cycles. The number of linearly independent homology gen- erators is called a *Betti number*

*βk* = dim *Hk* = dim(*Zk*) *−* dim(*Bk*)*.* (2) Informally, *β*0 is the number of connected components; *β*1 is

(A) (B)



(C)

cycle birth cycle death

1-cycle id



4

3

2

1

FIG. 2. **Persistent homology of a graph according to an *edge-value clique (EVC) filtration*.** (A) An undirected graph *G*(*V, E* ) with a scalar function *ƒ* : *E →* R defined over the edges *E* . We apply an EVC filtration to the graph by considering a monotonically decreasing *filtration parameter g* = (0*,* 5], and by considering a filtered sequence of graphs *G*(*V, Ec*), where *Ec* = *{*(*i, j*)*|ƒ* (*i, j*) *c g}* is the subset of edges for which *ƒ* (*i, j*) are larger than a threshold *g*. (B) Visualization of the graphs’ associated clique complexes *Kc ≡ K*(*G*(*V, Ec*)) for several

4

*g*. (C) A *persistence barcode* summarizes how the 0-dimensional (red) and 1-dimensional (blue) homological *k*-cycles of *Kc* change with decreasing *g*. The arrows highlight two events: at *g* = *gb*, a homological 1-cycle involving four edges is *born*; at *g* = *gd*, the 1-cycle *dies*, since it is “filled in” by a 1-simplex and two 2-simplices.

the number of 1-dimensional cycles or “loops” (that is, not including the triangular boundaries of 2-simplices); and *β*2 is the number of 2-dimensional holes or “voids” (e.g., the inte- rior of a triangulated sphere). For the SC shown in Fig. 1(B),

*β*0 = 1 since there’s one connected component, and *β*1 = 2

of 2-simplices. since there are two cycles that are not simply the boundaries

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 re- lations/connectivity between cycles, which we find to be in- strumental 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 ob- tained by considering linear combinations of *k*-cycles (includ- ing both homological *k*-cycles and boundaries). We will later show that this complicates the investigation of convection cy- cles through the lens of homological *k*-cycles.

1. **Persistent homology of scalar functions defined over edges**

One of the greatest tools of topological data analysis is the study of *persistent homology* [2, 22]. Here, we exam- ine how the homology of a topological object changes as it undergoes a *filtration* to yield a monotonically increasing se- quence *X*0 *⊆ X*1 *⊆ X*1 *⊆ ...* (e.g., of simplicial complexes

tion *ƒ* : *E −→* R over the edges, and each edge (*i, j*) *∈ E* is retained/removed according to *ƒ* (*i, j*). The values *ƒ* (*i, j*) *{Xt}*). We consider filtrations in which one has a scalar func-

could be edge weights for a weighted graph, but in general

graph and the values *ƒ* (*i, j*) in Fig. 2(A). they can encode any scalar property. We visualize such a

We call the process an *edge-value clique (EVC) filtra-*

*tion*, and we construct it as follows. Given a graph *G*(*V, E* )

and a *filtration function ƒ* , we define the subsets *Eє* =

*{*(*i, j*)*|ƒ* (*i, j*) *> ϵ}* in which one retains edges only for which *ƒ* (*i, j*) is sufficiently large. Note that the subsets *{Eє}* are We must specify a range over which to decrease *ϵ ∈* (*ϵB, ϵA*], and in practice we assume *ϵA >* max(*i,j*)*∈E ƒ* (*i, j*) and *ϵB <* non-decreasing as *ϵ* decreases (i.e., *Eє ⊆ Eє'* for any *ϵ′ < ϵ*). min(*i,j*)*∈E ƒ* (*i, j*). It then follows that *Eє* = *∅* is an empty set of edges when *ϵ ≥ ϵA*, and *Eє* = *E* (i.e., all edges are retained) when *ϵ ≤ ϵB*. 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 com- plexes *{Kє}* that are associated with the filtered graphs *{Gє}* that are defined with the edge sets *{Eє}*. In Fig. 2(C), we sum- marize the persistent homology of *{Kє}* in a *persistence bar- code*, which reveals how homology changes with *ϵ*. Observe that when *ϵ* is sufficiently large, *Kє* contains vertices but no edges. On the other hand, when *ϵ* decreases to be sufficiently small, then *Kє* recovers the original clique complex [recall Fig. 1(B)]. The values of *ϵ* that were used to create Fig. 2(B) are indicated by the vertical dotted lines in Fig. 2(C).

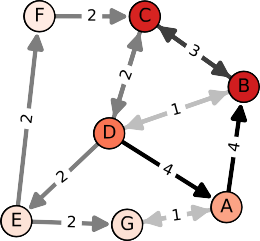
Each horizontal bar in the persistence barcode shown in Fig. 2(C) indicates the *lifetime* of a homological 1-cycle— that is, the values of *ϵ* for which it exists. The red and blue

bars reflect 0-homology and 1-homology respectively. The

dimensions of the homology spaces (i.e., Betti numbers) can

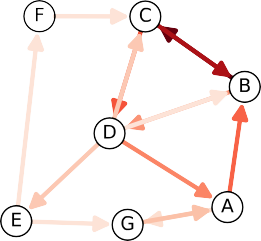
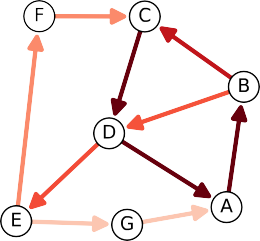
any particular *ϵ*. For example, one can observe that *β*1 = 0 when *ϵ* = 3*.*5, *β*1 = 1 when *ϵ* = 2*.*5, and *β*1 = 2 when be found by counting the number of homological 1-cycles at *ϵ* = 1*.*5. Clearly, the homological 1-cycles are undergoing bi- furcations as *ϵ* varies. A persistence barcode is convenient to

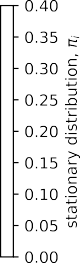
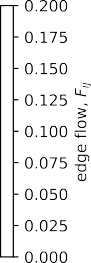
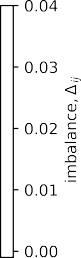
identify for each generator: the value *ϵb* of *ϵ* when it is “born” (i.e., the homological *k*-cycle does not exist when *ϵ > ϵb*); the value *ϵd* of *ϵ* when it “dies” (i.e., the homological *k*-cycle

(A)

(B)

(C)

FIG. 3. **Stationary distribution, edge flows, and convection cycles for an irreversible MC.** We study a discrete-time random walk over a directed, weighted graph that resembles the undirected graph in Fig. 2, except that the edges are now either directed or bidirectional. (Recall that undirected graphs give rise to reversible MCs that lack convection cycles.) (A) The color of each vertex indicates the *stationary distribution πi* of random walkers at each vertex *i*. (B) Edge colors indicate the *stationary flows Fij* = *πiPij* across edges, i.e., the stationary fraction of random walkers that traverse each directed edge. (C) *Flow imbalances* ∆*ij* = (*Fij − Fji*) manifest as a pattern of *convection cycles*. By construction, ∆*ij* = *−*∆*ji*, and we use arrows to indicate the directions of imbalance, e.g., *i → j* if ∆*ij c* 0.



does not exist when *ϵ < ϵd*); its *lifetime* (*ϵd, ϵb*]; and *lifespan*

*|ϵd − ϵb|*. A cycle’s lifespan quantifies its persistence under

the filtration, and it is often interpreted as a proxy for the cy- cle’s significance (although short-lifetime cycles can also be important in certain contexts).

1. **Convection cycles for irreversible Markov chains (MCs)**

We will apply persistence homology to study convection cycles in irreversible MCs [19], which we now briefly sum- marize. A discrete-time MC is a “memoryless” random pro-

cess in which for time steps *t* = 0*,* 1*,* 2*, . . .* , the system state

*St ∈ V* satisfies the Markov property P[*St*+1 = *i|S*0 =

bilities 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 con-

vergence requires that matrix **P** is an irreducible and aperiodic

[1] or that the initial condition **x**(0) lies in a converging sub- edge (*i, j*) per time step is given by space. In the stationary state, the *stationary flow* across each

*Fij* = *πiPij.* (4)

We study *convection cycles* using an approach that was de-

veloped in [28]. Specifically, for each edge we define the sta- tionary *flow imbalance*

*i*0*, ..., St* = *it*] = P[*St*+1 = *i|St* = *it*], which implies that

the probability of a state occurring at the next time step only

∆*ij*

= *Fij*

*— Fji*

*.* (5)

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 *Aij ∈* R is nonzero if (*i, j*) is an edge,

(*i, j*) *∈ E* , and *Aij* = 0 otherwise. We similarly define a

By construction, ∆*ji* = *−*∆*ij*, and we say that the *imbalance direction* is from *i* to *j* when ∆*ij >* 0. Importantly, the defin- ing feature of a *reversible MC* is that ∆*ij* = 0 for all *i* and

*j*. That is, the directional flows match *πiPij* = *πjPji* for any

edge (*i, j*). This is the case for any undirected graph, since

*transition matrix*, **P** = **D***−*1**A**, where **D** is a diagonal ma-

trix with entries that encode the (possibly) weighted vertex

in this case **A** = **A***T* , and it follows that *πi*

= *Dii*

*/* Σ*j*

*Djj*.

degrees **D***ii*

= Σ*j*

*Aij*

. For directed graphs, each (*i, j*) is

In contrast, an *irreversible MC* yields asymmetric stationary

flows and ∆*ij* is nonzero for some edges. To formally de-

considered to be an ordered pair, and each **D***ii* encodes the

out-degree of vertex *i*. Each matrix element *Pij* gives the probability for a random walk to transition from vertex *i* to *j*.

Letting *xi*(*t*) denote the probability that the system is in state

*i* (or equivalently, the probability that a random walker is at

tain the linear discrete-time system *xj*(*t* + 1) = *i xi*(*t*)*Pij*. By defining **x**(*t*) = [*x*1(*t*)*, . . . , xN* ]*T* , one equivalently has vertex *i*) at time *t*, one can utilize the Markov property to ob-

Σ

0

**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, *i xi*(*t*) = 1.

Σ

a *stationary state*, in which case **x**(*t*) converges to a limiting vector *π* = lim*t→∞* **x**(*t*) that satisfies the eigenvalue equa- Herein, we focus on network flows after a system reaches tion *πT* = *πT* **P**. By construction, *π* is a vector of proba-

fine convection cycles, we consider a new graph *G*∆(*V, E*∆) such that each positive value ∆*ij* gives rise to a directed edge (*i, j,* ∆*ij*) *∈ E*∆ having weight ∆*ij*. We then define *G*∆(*V, E*∆). a *convection cycle* to be any non-intersecting closed path in

In Fig. 3, we illustrate for an example MC how flow imbal-

ances manifest as a pattern of convection cycles. In Figs. 3(A), 3(B), and Fig. 3(C), we use edge colors to indicate the sta- tionary distribution *π*, stationary edge flows *Fij*, and flow im-

balances ∆*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 → j* only if ∆*ij >* 0, which also implies *j → i* is not an edge since ∆*ji* = *−*∆*ij*. Observe in Fig. 3(C) that this yields five convection cycles. In

(B)



cycle id

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 *|*∆*ij|* of flow imbalance as the filtration function *ƒ* : *E →* 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.

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

Before continuing, we highlight that convection cycles re-

vealed 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 cy- cle *A → B → C → 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 di- rection from node C to B. Future research will likely uncover

complementary notions of convection with different advan- tages/disadvantages, and our proposed techniques using per- sistent homology can likely be similarly extended.

1. **HOMOLOGICAL ANALYSES OF CONVECTION**

We now employ persistent homology to automate the detec- tion, 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 rela-

tion between convection cycles and homological 1-cycles.

* 1. **Persistent homology of convection cycles**

Recall from Sec. II 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 (poten- tially) directed and weighted graph, we first consider the as- sociated undirected graph. Then we study homology under an

EVC filtration in which the filtration function *ƒ* : *E −→* R is

given by the magnitudes of the flow imbalances

*ƒ* (*i, j*) = *|*∆*ij|.* (6)

In this way, the persistent homology that is revealed corre-

sponds to the convection cycles that arise under flow imbal- ances.

In Fig. 4, we visualize persistence barcodes for an EVC filtration associated with the convection cycles shown in Fig. 3(C). Note that this figure is analogous to Fig. 2(C),

where we had previously chosen the filtration function *ƒ* (*i, j*)

to be the edge weights. Since we now use a different func-

tion *ƒ* , 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 be- tween 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.

creasing *ϵ* and retaining edges (*i, j*) for which *|*∆*ij|* is smaller We note that one could also construct EVC filtrations by in- than *ϵ*. In Appendix A, we provide an example illustrating

why EVC filtrations with decreasing *ϵ* are superior to those with increasing *ϵ* for the goal of studying convection cycles. In particular, EVC filtrations that decrease *ϵ* focus on 1-cycles that are associated with large-flow convection cycles (i.e.,

large values of *|*∆*ij|*), which we consider to be the ones that

are more significant. In contrast, EVC filtrations that increase

vection cycles (i.e., small values of *|*∆*ij|*), which we consider *ϵ* focus on 1-cycles that are associated with small-flow con- to be less significant.

* 1. **Comparing convection cycles and homological 1-cycles**

We propose to study pattern formation for convection cy- cles using persistent homology techniques for homological



convection cycle I convection cycle II convection cycle III convection cycle IV convection cycle V



(B)



inconsistent



homological 1-cycle (i)

homological 1-cycle (ii) homologically equivalent generators for 1-cycle (iii)

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 correespond 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).

two different notions for cycles. Homological 1-cycles are 1-cycles; however, one should keep in mind that these 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- plicial homology.

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

cle 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 di- rected 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 persis- tent homology. Instead, we proposed to begin this pursuit by studying convection cycles using existing methods for persis- tent 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.

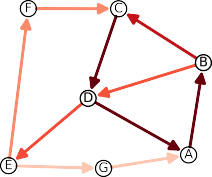
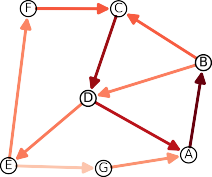
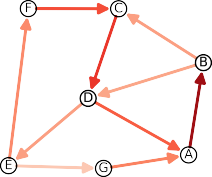
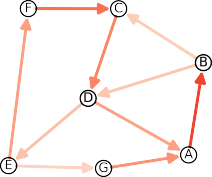
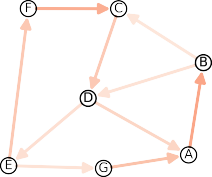
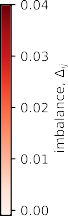
1. **APPLICATIONS**

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.

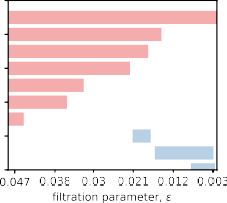
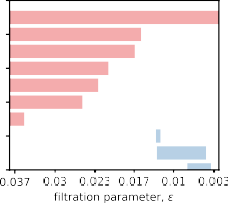
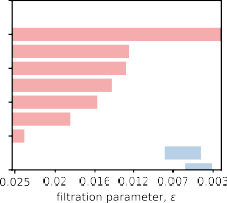
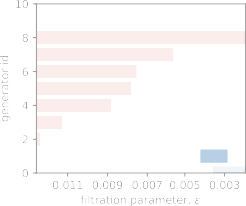
1. **Teleportation is a homology regularizer for PageRank**

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]),

(A)

(B)



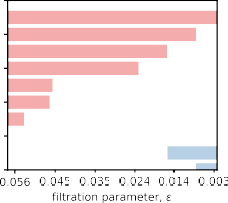
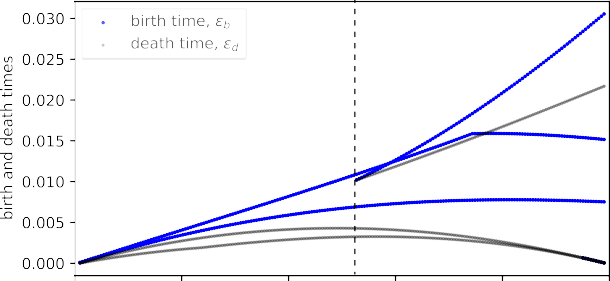


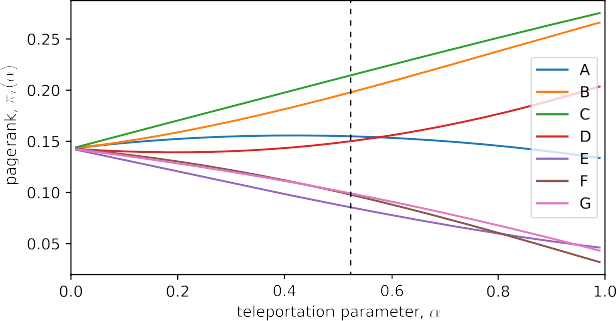
FIG. 6. **Persistent homology of convection cycles arising under PageRank.** We study the MC associated with PageRank for the directed graph from Fig. 3(A) under several choices of *α*. (A) Flow imbalances ∆*ij* (*α*) give rise to convection cycles. For clarity, we do not visualize flow imbalances for transitions due to teleportation. (B) Their homology changes with *α*, which is summarized by persistence barcodes.

but most notably, for many years it was utilized by Google to

vertex *i* is given by the stationary density *πi*(*α*) of MC having rank website and facilitate web search. The PageRank of a a transition matrix of the form

**P**(*α*) *≡ α***P** + (1 *− α*)*N−*1**11***T ,* (7)

(A)



(B)

FIG. 7. **Bifurcation diagram summarizes homological changes onset by** *α***.** (A) Birth and death times *gb,d* of 1-cycles arising under PageRank versus *α*. (B) For comparison, we depict the vertices’ PageRanks *πi*(*α*). Vertical dashed lines near *α∗* = 0*.*54 highlight that there are three cycles when *α c α∗*, but only two when *α < α∗*.

*α ∈* (0*,* 1) is the *teleportation parameter*. As *α →* 1, the sec- ond term vanishes and **P**(*α*) *→* **P**. Usually, *α* is chosen to where **P** is the transition matrix described in Sec. II C and

be near 1 (often 0.85) so that the second term can be consid-

ered as a small perturbation that improves the mathematical characteristics of **P**—or more formally, it is a “‘regulariza-

tion” of matrix **P**. In particular, when *α ∈* (0*,* 1) the matrix **P**(*α*) is guaranteed to be irreducible, aperiodic and positive, and the Perron-Frobenius theorem [1] ensures that its domi-

*πi*(*α*) *>* 0 for all *i*). In other words, the PageRanks are well- nant left eigenvector *π* is unique and has positive entries (i.e., defined for all vertices.

We now show that the introduction of teleportation also reg- ularizes the homology of convection cycles. In this experi- ment, we construct EVC filtrations with the filtration function

*ƒ* (*i, j*) = *|*∆*ij*(*α*)*|*, which now depends on *α*. In Fig. 6(A),

we illustrate for several choices of *α* 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 persis-

that the choice *α* = 1 recovers the transition matrix, stationary tence barcodes, which we create using EVC filtrations. Note distribution, flow imbalances, and persistence barcodes that

were were previously studied in Figs. 3 and 4.

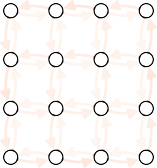
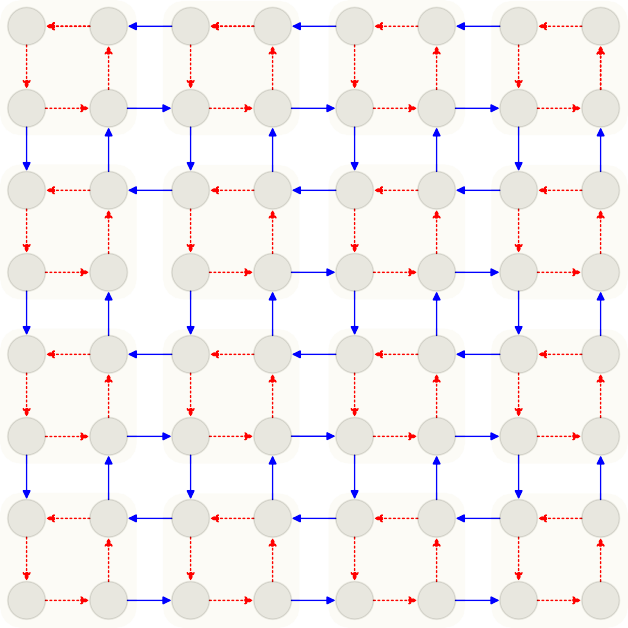
Observe that the homological patterns of convection cy- cles significantly change with *α*. For example, when *α* is sufficiently small, the homological 1-cycle *{A, B, C, D}* vanishes—it is “washed out” by the introduction of teleporta- tion. In other word, *α* is a *homology regularizer*. This is fur- ther illustrated in Fig. 7(A), where we plot the birth and death times of homological 1-cycles versus *α*. For comparison, we

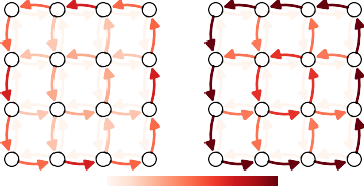
also plot the PageRanks *πi*(*α*) in Fig. 7(B). The vertical line

(approximately) below *α∗* = 0*.*54 highlights that one of the 1-cycles vanishes when *α* decreases

plore convection cycles arising under PageRank with *α* = 0*.*8. In Appendix A, we present additional experiments that ex- We show that homological 1-cycles arising for EVC filtrations

with decreasing filtration parameter *ϵ* reveal patterns for large-

(A)



|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| d  c | (0,3) | a  b | d  c | (1,3) | a  b | d  c | (2,3) | a  b | d  c | (3,3) | a  b |
| d |  | a | d |  | a | d |  | a | d |  | a |
|  | (0,2) |  |  | (1,2) |  |  | (2,2) |  |  | (3,2) |  |
| c |  | b | c |  | b | c |  | b | c |  | b |
| d | (0,1) | a | d | (1,1) | a | d | (2,1) | a | d | (3,1) | a |
| c |  | b | c |  | b | c |  | b | c |  | b |
| d |  | a | d |  | a | d |  | a | d |  | a |
|  | (0,0) |  |  | (1,0) |  |  | (2,0) |  |  | (3,0) |  |
| c |  | b | c |  | b | c |  | b | c |  | b |

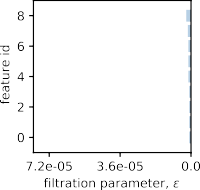
(B)

FIG. 9. **Persistent homology chiral edge flow.** (A) Flow imbalances

∆*ij* between external states for the bi-monomer shown in Fig. 8 with *γin* = 0*.*01 and different *γex*. In the limit *γex γin* [27], 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.

external transitions internal transitions

FIG. 8. **MC model for configuration dynamics of a monomer.** 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 *γex* and *γin*, respectively.

flow convection cycles. In contrast, when EVC filtrations are constructed with increasing *ϵ*, we find that the resulting homo- logical 1-cycles relate to small-flow convection cycles, and in particular, those involving low probability teleportation tran- sitions.

1. **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 ex-

is given by the number of monomers of each type, (*s*1*, s*2), periment. The monomer configuration (i.e., “external state”) whereas the “internal state” is one of four possibilities: a, b,

c, or d. Transitions that involve a change of internal state oc- cur at rate *γin*, whereas transitions between involving external states (i.e., the addition or removal of a monomer) occur at rate *γex*. The resulting MC can be visualized as a 2-dimensional lattice, which we visualize in Fig. 8.

between the external states. We fix *γin* = 0*.*01 and consider In Fig. 9(A), we visualize flow imbalances for transitions several *γex*. Observe that as *γex* increases, a large counter-

clockwise 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 Fig. 9(B), we visual- ize persistence barcodes for EVCs filtrations constructed us- ing the method that we described in Sec. III A. The chiral edge cycle corresponds to the 1-cycle with having the largest lifes- pan, and its homology becomes more persistent (i.e., promi- nent) in the limit *γex γin*.

1. **DISCUSSION**

In this paper, we examined the patterns of convection cycles that arise under irreversible Markov chains (MCs) from the perspective of persistent homology. Our approach required formalizing a type of filtration (EVC filtration) for scalar func- tions that are defined on the edges of a graph, and we studied convection cycles by choosing the filtration function to be an MC’s flow imbalances in the stationary state. Because Markov chains are crucial to so many diverse applications, we expect our methods to be broadly applicable across the sciences and engineering. Herein, we highlighted two such applications: the PageRank algorithm for centrality analysis and chiral edge flows that arise for the configuration dynamics of monomers. Our experiments revealed how system properties can act as homology regularizers of convection cycles, and we intro- duced homological bifurcation diagrams to summarize these changes. This approach automates the detection, summary, and examination of convection cycles over networks, places it on stronger mathematical and computational foundations, and paves the way for further investigation into convective flows on networks.

Additionally, our work highlights the need for new persis-

tent homology methods to study convection cycles as well as other functions and signals defined on directed graphs. In Sec. III 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. Some- times there is a one-to-one correspondence between these cy- cles, and sometimes the relation is more complicated, due in part to the fact that a given homological k-cycle can be equiv- alently represented by possibly more than one homological generator. Such generators may or may not correspond to con- vection cycles. Moreover, convection cycles can also corre- spond to the boundaries of 2-simplices, and as such, they will not be identified via the traditional tools of persistent homol- ogy. Developing persistent homology techniques that cater to convection cycles, and which specifically account for edge di- rections, 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 cy- cles that exists in addition to a chiral edge flow, which may lead to new insights for these applications and other appli- cations (e.g., reinforcement learning) that rely on irreversible Markov chains.

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** *g* **versus increase** *g*

In Sec. II B, we defined EVC filtrations in which one de- creases a filtration parameter *ϵ*, retaining edges for which

*ƒ* (*i, j*) *> ϵ*. Our numerical experiments that study convection

the filtration be given by the flow imbalances *ƒ* (*i, j*) = *|*∆*ij|*. cycles using persistent homology use this approach and let By decreasing *ϵ*, 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 *ϵ* and retaining edges for which

*ƒ* (*i, j*) *< ϵ*. Here, we show that this latter filtration reveals

1-cycles that relate to small-flow convection cycles, which we

consider to be insignificant.

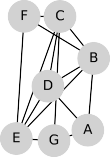
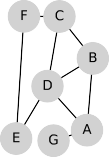
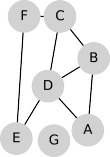
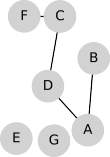
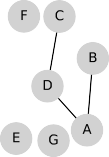
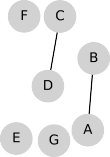
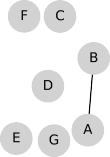
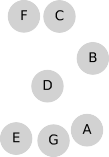
ances arising under the PageRank algorithm with *α* = 0*.*8 for In Fig. 10, we study EVC filtrations applied to flow imbal- the same MC that we investigated in Sec. IV A. In Fig. 10(A)

and Fig. 10(B), we illustrate EVC filtrations with decreasing and increasing *ϵ*, respectively. Observe in Fig. 10(A) that the 1-cycles revealed by decreasing *ϵ* correspond to large-flow convection cycles. In contract, observe in Fig. 10(B) that the 1-cycles revealed by increasing *ϵ* are small-flow cycles that relate to low-probability transitions that occur due to telepor- tation.

**ACKNOWLEDGMENTS**

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

1. Ravindra B Bapat. A max version of the perron-frobenius the- orem. *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 net- works. *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 It- skov. 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 Sci- ence & 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, 2012.
14. M Kramar, Arnaud Goullet, Lou Kondic, and Konstantin Mis- chaikow. 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*

(A)



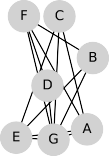
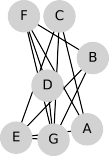
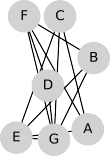
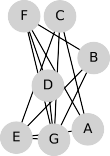
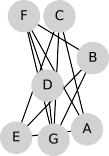
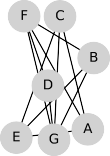
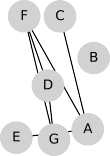
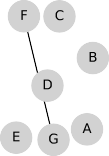
(B)



FIG. 10. **Comparing EVC filtrations with decreasing and increasing filtration parameter** *g***.** Extending our study in Sec. IV A that uses persistent homology to study convection cycles arising for a MC under the PageRank algorithm with *α* = 0*.*8, we now study homological 1- cycles obtained via two different EVC filrations. (A) Similar to our results in Fig. 6, we construct EVC filtrations by including edges for which

*|*∆*ij| c g* while decreasing *g*. Observe that the 1-cycles reveal large-flow convection cycles that are associated with large values of *|*∆*ij|*. (B) For comparison, we construct EVC filtrations by including weighted edges *|*∆*ij| < g* while increasing *g*. Observe that these 1-cycles now correspond to small-flow convection cycles that are associated with small values of *|*∆*ij|*. They primarily describe low-probability transitions that occur due to teleportation. In this work we focus on EVC filtrations with decreasing *g*, since we consider high-flow convection cycles to be the ones that are most important.

*analysis and applications*, 27(4):968–987, 2006.

1. Minh Quang Le. Codebase for persistent homology of convection cycles in network flows https://github.com/minhquan89/ Persistent-Homology-of-Convection-Cycles.
2. Jie Liang, Herbert Edelsbrunner, Ping Fu, Pamidighantam V

Sudhakar, and Shankar Subramaniam. Analytical shape com- putation of macromolecules: I. molecular area and volume through alpha shape. *Proteins: Structure, Function, and Bioin- formatics*, 33(1):1–17, 1998.

1. Shusen Liu, Di Wang, Dan Maljovec, Rushil Anirudh, Jayara- man J Thiagarajan, Sam Ade Jacobs, Brian C Van Essen, David Hysom, Jae-Seung Yeom, Jim Gaffney, et al. Scalable topolog- ical data analysis and visualization for evaluating data-driven models in scientific applications. *IEEE transactions on visual- ization and computer graphics*, 26(1):291–300, 2019.
2. La´szlo´ Lova´sz et al. Random walks on graphs: A survey. *Com- binatorics, Paul erdos is eighty*, 2(1):1–46, 1993.
3. Francis Motta, Christopher Tralie, Rossella Bedini, Fabiano Bini, Gilberto Bini, Hamed Eramian, Marcio Gameiro, Steve Haase, Hugh Haddox, John Harer, et al. Hyperparameter opti- mization of topological features for machine learning applica- tions. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*, pages 1107–1114. IEEE, 2019.
4. The notion of “dimension” is more interpretable for the case of a non-abstract simplicial complex, for which the vertices cor- respond 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) dimen- sional 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).
5. 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.
6. Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Wino- grad. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab, 1999.
7. Ronald Parr and Stuart Russell. Reinforcement learning with hierarchies of machines. *Advances in neural information pro- cessing systems*, pages 1043–1049, 1998.
8. Jose A Perea and John Harer. Sliding windows and persistence: An application of topological methods to signal analysis. *Foun- dations of Computational Mathematics*, 15(3):799–838, 2015.
9. Giovanni Petri, Paul Expert, Federico Turkheimer, Robin Carhart-Harris, David Nutt, Peter J Hellyer, and Francesco Vac- carino. Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, 11(101):20140873, 2014.
10. Evelyn Tang, Jaime Agudo-Canalejo, and Ramin Golestanian. Topology protects chiral edge currents in stochastic systems. *Physical Review X*, 11(3):031015, 2021.
11. Dane Taylor. Multiplex markov chains: Convection cycles and optimality. *Physical Review Research*, 2(3):033164, 2020.
12. Dane Taylor, Florian Klimm, Heather A Harrington, Miroslav Krama´r, Konstantin Mischaikow, Mason A Porter, and Peter J Mucha. Topological data analysis of contagion maps for ex- amining spreading processes on networks. *Nature communica- tions*, 6:7723, 2015.
13. The GUDHI Project. *GUDHI User and Reference Manual*. GUDHI Editorial Board, 3.4.1 edition, 2021.
14. Luke Tierney. Markov chains for exploring posterior distribu- tions. *the Annals of Statistics*, pages 1701–1728, 1994.
15. Ying-Jie Xin and Yuan-Hua Zhou. Topology on image process- ing. In *Proceedings of ICSIPNN’94. International Conference on Speech, Image Processing and Neural Networks*, pages 764– 767. IEEE, 1994.