Reorderability of node-filtered order complexes

Ann Sizemore Blevins [1] and Danielle S. Bassett [1,2,3,4,5,6,*]

[1]*Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania,
Philadelphia, Pennsylvania 19104, USA*
[2]*Department of Physics and Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[3]*Department of Electrical and Systems Engineering, School of Engineering and Applied Science, University of Pennsylvania,
Philadelphia, Pennsylvania 19104, USA*
[4]*Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[5]*Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA*
[6]*Santa Fe Institute, Santa Fe, New Mexico 87501, USA*

Growing graphs describe a multitude of developing processes from maturing brains to expanding vocabularies to burgeoning public transit systems. Each of these growing processes likely adheres to proliferation rules that establish an effective order of node and connection emergence. When followed, such proliferation rules allow the system to properly develop along a predetermined trajectory. But rules are rarely followed. Here we ask what topological changes in the growing graph trajectories might occur after the specific but basic perturbation of permuting the node emergence order. Specifically, we harness applied topological methods to determine which of six growing graph models exhibit topology that is robust to randomizing node order, termed global reorderability, and robust to temporally local node swaps, termed local reorderability. We find that the six graph models fall upon a spectrum of both local and global reorderability, and furthermore we provide theoretical connections between robustness to node pair ordering and robustness to arbitrary node orderings. Finally, we discuss real-world applications of reorderability analyses and suggest possibilities for designing reorderable networks.

I. INTRODUCTION

Growing networks can embody myriad developing systems spanning from maturing organisms in ecology and biology to expanding connections in economics and sociology. For each of these systems, the representative growing network captures the normal growth process by noting when new actors, or nodes, join the mass and where relations between actors, or edges, form. Such a representation could, for example, allow us to understand spreading synchronization of coupled dynamical systems [1] and crystallization processes in materials science [2,3]. Outside of normal growth processes, we can also use growing networks to understand abnormality in maturation, enlargement, or spreading processes such as the propagation of a seizure through a brain network [4] or the dissipation of β-amyloid in Alzheimer's disease [5,6]. Many of these examples include mechanisms or models that describe the specific growing process and as such inherently suggest a proper order of nodes necessary to achieve the desired final network architecture.

But what if, instead, the growth process was perturbed? Can normal development recover if, for example, those objects that are usually last to emerge were to instead to appear first in the growth order? Or how might the growth process react to a larger perturbation such as an entirely random re-ordering of the nodal emergence order? Intuition suggests that a networked system's robustness to such perturbations may depend upon some intrinsic property of the system, and that different systems will exhibit differing levels of robustness. In some systems, we may expect robustness to these types of perturbations; for example, a monodisperse physical system undergoing crystallization will achieve its final ordered form regardless of growth process minutiae, due to underlying thermodynamics [7]. Thus, any small reordering of particles joining the crystal will not change the outcome. In other systems, we expect sensitivity to these types of perturbation; for example, the rate of disease spread across a network can largely depend on the network organization surrounding those nodes that were infected earliest [8], and therefore the global system outcome would indeed be susceptible to the order in which nodes were infected.

Here we aim to understand a piece of the above phenomenon by formalizing and investigating the stability of a growing graph's developing architecture in response to perturbation of node addition order. We will refer to this notion of stability as the *reorderability* of a growing graph. Thinking generally, a reordered growing graph with a developing architecture that is similar to that produced by the original node ordering may share many properties with the original growing graph. Still, perhaps the simplest property that the two would share is a similar evolving global topology. Quantitatively tracking and characterizing that evolving global topology is made tractable by emerging tools from

*Corresponding author: dsb@seas.upenn.edu

applied topology [9–11], specifically, persistent homology which records the pattern of topological cavity emergence throughout network growth. Such tools allow us to quantitatively compare the evolving topology of two or more growing graphs [12], facilitating inferences regarding their reorderability.

In this work, we examine the reorderability of six growing graph models chosen to span a range of topologies relevant for social systems, material systems, spatially embedded systems, and dynamical systems. We first ask if each of these growing graph models exhibits stable persistent homology after randomly reordering the node addition order, and we refer to the degree of such stability as the global reorderability of a growing graph model. We next ask which growing graph models show stable persistent homology in the face of temporally localized swaps in the node ordering, and we refer to the degree of stability as the local reorderability of a growing graph. Finally, the third direction of this study investigates how the local reorderability of a growing graph may (or may not) constrain its global reorderability. We observe that a spectrum of both global and local reorderability exists within the graph models tested. Furthermore, we find that our notions of local and global reorderability are surprisingly distinct, which provides an opportunity to design growing processes with specific reorderability characteristics.

II. METHODS

A. Building intuitions

Consider, for example, the birth of neurons in the developing brain, which is a strictly ordered process [13,14]. Recent studies revealed that small amounts of stochasticity actually are advantageous for early developing populations of neurons [15–17]. These results suggest that if only a few temporally close neurons swap birth times, healthy development will proceed as expected. Still, since overall neuronal development is extremely stereotyped [14,18] and since cell function so linked to birth time [19,20], we might reasonably speculate that drastic birth order swaps, such as spawning progenitor cells last, would impede proper maturation. Ultimately, the response of a growth process to a reordering of node birth times depends on both the network connections (the topology) and the original order of the nodes (our baseline to which we compare). Our goal in this work is to understand how the connections and node order influence the stability of growth processes.

To make our goal more concrete, we consider the growing graph in the top row of Fig. 1. Each node a, b, c, d, e, and f is added in turn, with new edges added only between the new node and nodes already present according to some graph rule or predefined structure. After adding the first three nodes a, b, and c, our graph has only exhibited a tree structure. As we move forward in time, node d creates a cycle (loop), node e forms a second cycle, and then finally node f tessellates the top cycle. Next we ask how the network grows after a perturbation to node order. Specifically, how would the maturing graph architecture change if we swapped nodes b

and d in the ordering so that we build the same final graph, but the nodes are now reordered (row 2 of Fig. 1)? In building the same final graph, we mean that if an edge exists between nodes v and u in the originally grown graph, this edge will exist in our reordered growing graph as soon as both u and v are added. We observe that this perturbation does not seem to have changed the growing graph much—after the first three nodes, we have only seen tree structures; then we see a four-node loop after the fourth node, a second loop after the fifth node, and a tessellated cycle after the sixth node. At each slice (node addition), the original and perturbed growing graphs are quite similar (in fact, isomorphic in this example), so we can say that our growing graph is stable with respect to *this particular* node swap. But perhaps we did not change the evolving graph much because nodes b and d were only two steps apart in the original ordering. Let us instead swap d and f (also two steps apart in the original ordering) and observe the effects. Again, the node swap does not affect the identity of the edges that we assign, only the order of their inclusion; we will always have the same edges at the end of the growing process. We see in the third row of Fig. 1 that this $d \rightleftarrows f$ switch drastically changes the growing graph. We see triangles much earlier than in the original ordering, differing numbers of edges, and by the time node d is added forming the loop with a, b, and c, that loop has already been tessellated.

The example just described suggests that the reorderability problem is at least more complicated than simply understanding how far nodes move between the reordered growing graph and the original. Furthermore, not all growing graphs exhibit change after node reordering. As discussed in Example 2 in the Appendix, growing trees, cross polytopes, and cliques all exhibit total reorderability (in at least one dimension; see Example 2 for more details). So indeed some growing graphs can always have their growth reordered without a change to their growing topology, while others (as in Fig. 1) cannot. We note also that if we have a reorderable graph such as an octahedron, we can often find node orderings that do not preserve the average degree or clustering coefficient at each step in the growth process. Thus, preserving the topology of a growing graph fundamentally differs from preserving network statistics.

Importantly, we note that studying how growing graphs may or may not change after reordering node birth times is not a new concept. Indeed, from a statistical perspective on networks, notable work has been performed in understanding vertex- and edge-exchangeable graphs, which are sequences of graphs whose distributions are invariant with respect to the order of node or edge addition, respectively [21–26]. Results in this area include (but are not limited to) understanding how graphons and point processes relate to exchangeable graphs [27,28] and how one can construct exchangeable graphs [29,30]. Our work differs from the above literature in that we have a specific topological focus which offers us a unique set of mathematical tools with which to understand growing graphs. Still, we suspect interesting links exist between these ideas of exchangeability and reorderability, and we leave the unraveling of such relations for future endeavors.
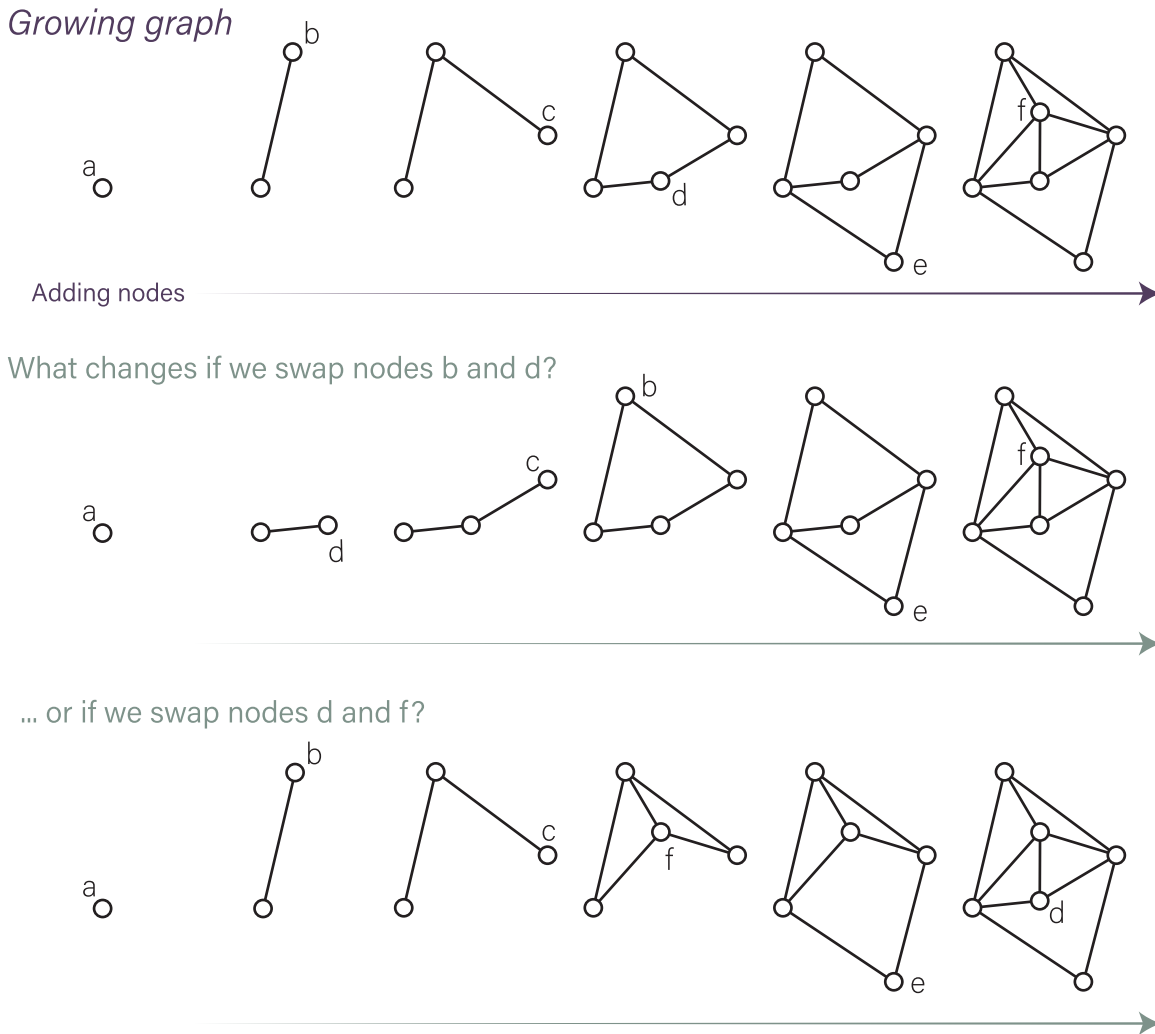
FIG. 1. A growing graph may or may not exhibit differing topology depending on the order of nodes in the growth sequence. Given a growing graph (top row), we ask how the evolving topology changes after we swap nodes b and d (middle row) or after we swap nodes d and f (bottom row) in the node birth order. Curiously, the $b \rightleftarrows d$ swap does not affect the growing graph architecture, while the $d \leftrightarrows f$ swap results in a notable effect.

B. Adding rigor to our intuitions

In order to explicitly study questions of reorderability, we need to first formalize a few of the above concepts. If we assume that we add nodes one at a time (for the sake of simplicity), and that we can only add edges—between the new node and any node already present—at the time of node addition, then we can define a growing graph with only a binary graph B and node order s. Explicitly, we define a growing graph as a pair (B, s) where $B = (V, E)$ is a binary graph with vertex set V and edge set $E \subseteq V \times V$, and s is the map $s : V \to \mathbb{N}$ that defines the order in which we add nodes [Fig. 2(a)]. Note that we could also define s as a map $s : V \to \mathbb{R}$, and then keep only the ordering of the nodes, as has been done in previous work [31]. In the exposition to follow, we will describe the process by which we will extract the topology of a growing graph via persistent homology. For additional intuition on the relevant methods, we direct the interested reader to several excellent sources for further details [9,11,32–34].

1. From graph to simplicial complex

Although we could test the reorderability of any number of graph statistics, in this primary study we aim to understand how the global, holistic graph architecture changes in response to reorderings. Holes or voids in networks are incompressible features that can serve to isolate processes [35] or affect dynamics [36]. To examine the topology of a growing graph, in our case a sequence of binary graphs, with respect to topological cavities or voids, we first must understand how to find topological cavities within one binary graph. To first capture the cavities of all dimensions within a *single* graph, we must translate the graph into a higher relational structure by (abstractly) filling in all cliques with a (higher dimensional) volume of matching dimension, called a simplex. Specifically, an $(n + 1)$-clique is a set of $(n + 1)$ nodes that are all-to-all connected by edges in a graph, and an n-simplex is the convex hull of $n + 1$ affinely positioned nodes [see Fig. 2(b) for examples]. Within a biological system, we often interpret an n-simplex as a set of $n + 1$ nodes that
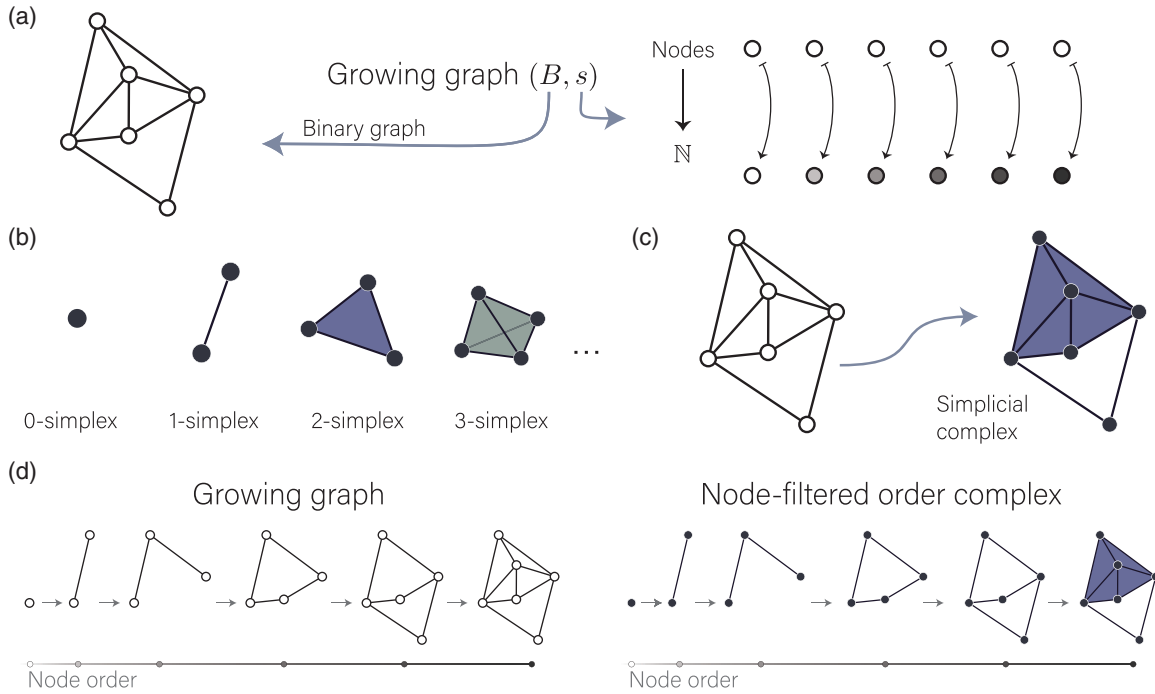
FIG. 2. Illustration of the steps taken to convert a growing graph to a node-filtered order complex. (a) A growing graph is defined as the pair (B, s) where B is a binary graph and s is an ordering of the nodes. (b) Examples of simplices defined using one to four nodes. (c) Filling in all cliques of a binary graph (left) with simplices yields a simplicial complex (right). (d) An example growing graph (left) and the corresponding filtered simplicial complex (right), which we refer to as the node-filtered order complex.

all collectively act together; for example, the set of nodes could act as a collection of brain regions serving the same function [35,37], neurons cofiring [38], or social groups [39]. A collection of simplices nicely glued together is called a simplicial complex, which—more rigorously—is a set of vertices V and a set of simplices K such that if $\sigma \in K$ and $\tau \subseteq \sigma$ then $\tau \in K$. Intuitively, this rule prevents a 2-simplex (a filled triangle) from existing without one of its constituent 1-simplices (edges). We create a simplicial complex from a binary graph by assigning an n-simplex to each $(n + 1)$-clique in the graph as shown in Fig. 2(c). To reiterate, we fill all cliques of a graph with simplices so that we can create a simplicial complex from a binary graph, which then allows us to detect topological cavities of all dimensions.

2. From growing graph to node-filtered order complex

Note that we can perform the conversion from graph to simplicial complex for *every* binary graph in our growing process. That is, after each node addition to the growing graph, we can construct a simplicial complex by filling in cliques with simplices. Furthermore, since we only add elements to our graph as it grows, the maps from the growing graph [Fig. 2(d), left] induce maps between simplicial complexes, now giving us a growing simplicial complex, which we formally refer to as a filtered simplicial complex [Fig. 2(d), right]. Previous studies focusing on filtered simplicial complexes arising from adding edges one at a time and translating all cliques to simplices referred to the construction of these edge-filtered simplicial complexes as the *order complex* of a weighted graph [40]. The rationale for this name is that the

filtered simplicial complex retains the ordering information given by the edge weights (notably this construction is also called the weight rank clique filtration [34,37,41,42] as it preserves the ranks of edge weights). Since in our growing graph case we have an ordering of the nodes (as opposed to an ordering of the edges), we call our above described filtered simplicial complex the *node-filtered order complex* of the growing graph (B, s) [Fig. 2(d)], and we note that this complex can be constructed from either a node-ordered or node-weighted network in which node weights induce the node order [31]. For the sake of brevity, here we shorten this name to the n-order complex of (B, s), which we denote as nord(B, s). We suggest that the node-filtered language described here is potentially a natural description of real network growth.

3. Homology of a simplicial complex

We aim to understand the topology of the n-order complex as defined by cavities or voids within the structure, as detected by homology. To begin the homology computation, for each dimension n, we create the vector space $C_n(K)$ (over $\mathbb{F}_2$, the binary field) in which each basis vector of $C_n(K)$ corresponds directly to a particular n simplex in the simplicial complex K. For example, the simplicial complex in Fig. 2(c) contains 10 edges, and therefore $C_1(K)$ would be $\mathbb{F}_2^{10}$. An example element of $C_1(K)$ is a collection of 1-simplices, called a 1-chain. Naturally we call $C_1(K)$ the first chain group. For simplicity, we conflate the basis vector and the associated n-simplex, and often we will write simply C_n when the simplicial complex under study is obvious. Note that since the physical
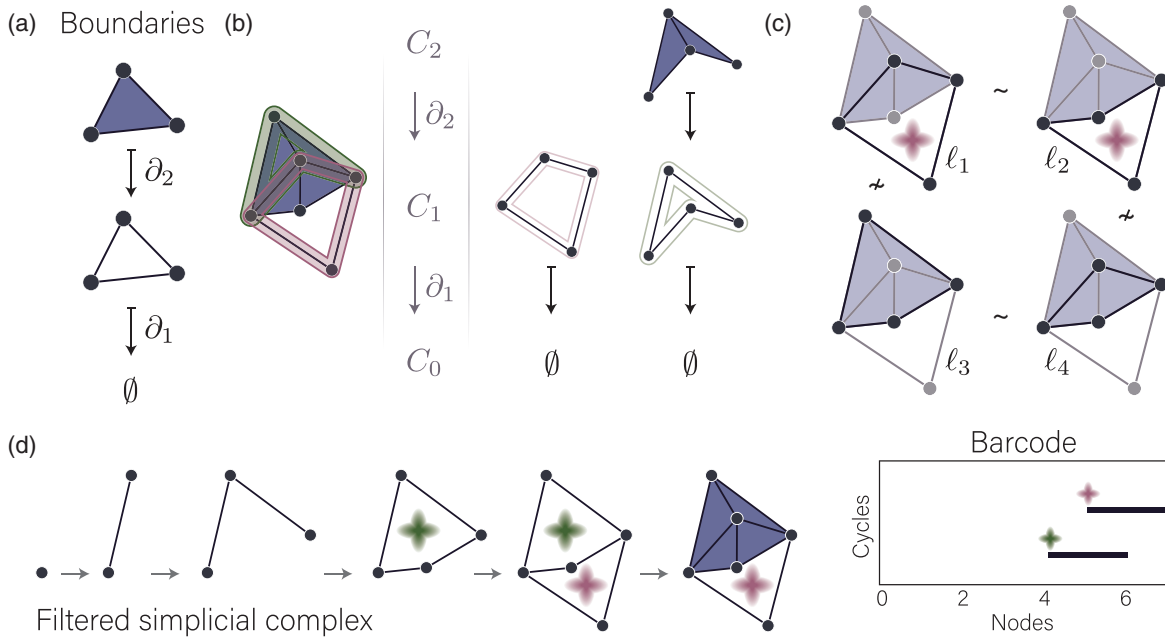
FIG. 3. Concepts in persistent homology. (a) The boundary of a 2-simplex is three 1-simplices, whose boundary in turn is empty. (b) The green and pink 1-cycles both have a zero boundary, but the green cycle is the boundary of a collection of 2-simplices. (c) Loops ℓ_1 and ℓ_2 are equivalent as they both surround the same cavity (pink). Cycles ℓ_3 and ℓ_4 are equivalent as they both surround no cavities. (d) The example filtered simplicial complex contains two persistent cavities of dimension 1 (green and pink stars), which are recorded in the barcode shown on the right.

boundary of an n-simplex is a shell of $(n-1)$-simplices (so a 1-simplex has two nodes as its boundary, a 2-simplex has three 1-simplices as its boundary, and so on), we immediately acquire a relation between simplices of dimension n and $n-1$ corresponding to simplices and their boundaries. This relation allows us to map from $C_n \to C_{n-1}$ via the boundary operator ∂_n, which is explicitly a linear map that sends an n-simplex $\{v_0, v_1, \ldots, v_n\}$ to its boundary $(n-1)$-simplices in C_{n-1} [Fig. 3(a)], mathematically defined as $\partial_n(\{v_0, v_1, \ldots, v_n\}) = \sum_0^n \{v_0, \ldots, v_{i-1}, \hat{v}_i, v_{i+1}, \ldots, v_n\}$, where $\hat{v}_i$ indicates that node v_i is not included in the $(n-1)$-simplex. The collection of chain groups C_* and boundary maps ∂_* forms a chain complex

$$\ldots C_{n+2} \xrightarrow{\partial_{n+2}} C_{n+1} \xrightarrow{\partial_{n+1}} C_n \xrightarrow{\partial_n} C_{n-1} \xrightarrow{\partial_{n-1}} \cdots,$$

which contains all of the assembly information for the simplicial complex.

Homology, generally, counts the voids of a particular dimension within a simplicial complex. Naturally, a void within the complex must be both encapsulated by the complex (being surrounded by a loop of simplices) and have at least some empty space within it (not being filled with higher dimensional simplices). Loops or cycles within a simplicial complex are paths of simplices that have an empty boundary, like higher dimensional analogs of loops within graphs. For example, a 1-cycle (formed by 1-simplices) is a closed path of edges, a 2-cycle is a capsule of 2-simplices, and so on (see Example 1 in the Appendix). All of the topological information about the simplicial complex is housed in the chain complex: The space of cycles of dimension n is $\ker \partial_n =: Z_n$ and the space of boundary cycles—that is, shells

of simplices that are boundaries of higher dimensional simplices—is $\mathrm{im}\,\partial_{n+1} =: B_n$. Note that by definition $B_n \subseteq Z_n \subseteq C_n$ since the boundary of a boundary is null [Fig. 3(b)]. Two cycles surround the same cavity if they differ by a boundary cycle, which induces an equivalence relation on cycles. That is, for $\ell_1, \ell_2 \in Z_n$, $\ell_1 \sim \ell_2 \iff \ell_1 - \ell_2 \in B_n$, so that all cycles of n-simplices that surround the same cavity (or cavities) are equivalent. As an example, in Fig. 3(c) we see $\ell_1 \sim \ell_2$ and that they surround the same cavity, while $\ell_3 \sim \ell_4$ as they both are boundaries of 2-simplices. Intuitively, since any boundary loop (element of B_n) surrounds no cavities, we can add a boundary cycle to any other cycle ℓ and not change the cavities that the cycle ℓ surrounds. We can collect all loops that are equivalent to a particular loop ℓ_0, which together comprise an equivalence class $[\ell_0] = \{\ell \in Z_n | \ell \sim \ell_0\}$. Intuitively, each equivalence class corresponds to a particular set of cavities within the simplicial complex since an equivalence class contains all cycles that surround those particular cavities. Then to pass from the space of cycles Z_n to the space of equivalence classes in which each basis element corresponds to an independent void in the complex, we take the vector space quotient and finally define $H_n(K) := Z_n/B_n$, which is the nth homology group of the simplicial complex K.

4. Persistent homology of a filtered simplicial complex

Moving beyond one simplicial complex, we now consider the computation of persistent homology of a filtered simplicial complex $K^1 \hookrightarrow \cdots \hookrightarrow K^m \hookrightarrow \cdots \hookrightarrow K^M$, where for our purposes the simplicial complex K^m contains m vertices and M is the total number of nodes. The map $K^m \hookrightarrow K^{m+1}$ defines a map between chain complexes $C_*(K^m) \hookrightarrow C_*(K^{m+1})$ in

which basis elements map to basis elements, in a manner consistent with how simplicies map to simplices. Now, the map between chain complexes induces a map between homology groups $H_*(K^m) \to H_*(K^{m+1})$, since cycles map to cycles and boundaries map to boundaries in $C_*(K^m) \hookrightarrow C_*(K^{m+1})$. Having maps between homology groups allows us to track persistent cavities (equivalently called persistent cycles) along the filtration [Fig. 3(d) green and pink stars]. A persistent cycle will be born at some point in the filtration, called the *birth*, will live for some time [i.e., mapped via $H_*(K^m) \to H_*(K^{m+1})$ to a nontrivial equivalence class], and then may finally die [when it maps to the trivial equivalence class, as does the green cavity in Fig. 3(d)] via tessellation by higher dimensional simplices. Those persistent cycles that do not die are formally given a death time of ∞. Then the persistent homology can be summarized as a (multi)set of pairs (b, d), $b \leqslant d$, marking the birth and death time of each persistent cycle. We can visually display these pairs as intervals along the filtration axis, drawing a bar from b to d for each persistent cavity. The resulting plot is called a barcode and illustrated pictorially in Fig. 3(d). Finally, to more easily summarize the persistent homology, we might only record the number of persistent cavities of dimension n alive at each filtration value m with a function called the Betti curve $\beta_n(m)$. We denote the area under $\beta_n(m)$ by $\overline{\beta_n} = \sum d_i - b_i$ with i indexing over persistent cavities of dimension n; the quantity intuitively reflects the total amount of persistent homology seen throughout the growth process. To summarize, we can now take a growing graph model, compute the evolving topology via persistent homology, and record the output as a barcode or Betti curve.

We now have the necessary background to extract the evolving topology from a growing graph (B, s) using persistent homology. With this background, we can re-examine the example in Fig. 1 and see that the persistent homology of the growing graphs in the top and middle rows will be the same, while the persistent homology of the growing graph in the last row will be quite different. Explicitly, the top two rows show two persistent cavities: one born after node 4 and killed with the addition of node 6, and a second born with the addition of node 5 that persists to ∞. Conversely, the bottom row shows only one persistent cavity born with the addition of node 5 that persists to ∞. Note also that this observation suggests that nodes b and d contribute similarly to the persistent homology—both b and d participate in the top persistent cycle, while nodes d and f serve distinct functions as the former usually births a persistent cavity, while the latter usually kills the cavity. The quantification of differences between the persistent homology of two growing graphs will then allow us to make our notions of reorderability more rigorous, and therefore we next focus on the practical consideration of measuring formal distances between topological summary statistics.

C. Distances between barcodes and between Betti curves

Since we wish to compare the persistent homology of growing graphs, we need to define a distance between Betti curves and a distance between barcodes. First, considering the former, we note that Betti curves record the homology of the growing graph slice by slice, and we are therefore motivated

to compare Betti curves in a slice-wise fashion. Specifically, we define the distance between two Betti curves of dimension n, β_n, β'_n to be

$$d_\beta^n(\beta_n, \beta'_n) = \max_m (|\beta_n(m) - \beta'_n(m)|), \tag{1}$$

with m indexing the number of nodes added. Intuitively, this distance records the maximal vertical distance between the Betti curves at any point in the filtration. Second, considering persistence diagrams, we use the bottleneck distance [12] for its simplicity and interpretability, although we recognize that alternatives exist [32,43,44]. Generally, the bottleneck distance measures the maximal left or right difference between matched bars in two barcodes after optimally matching the bars. More specifically, for two barcodes P_1, P_2 of dimension n,

$$d_{BN}^n(P_1, P_2) = \min_\gamma \big(\max_{p \in P_1} (|p_{\text{birth}} - \gamma(p)_{\text{birth}}|,$$
$$|p_{\text{death}} - \gamma(p)_{\text{death}}|) \big), \tag{2}$$

where $p = (p_{\text{birth}}, p_{\text{death}})$ is a bar in P_1 and γ is a matching of bars so that $\gamma(p)$ is a bar of P_2. Note that bars are allowed to match to any point (r, r) for $r \in \mathbb{R}^+$ as well. Importantly, as shown in Ref. [12], the persistent homology with respect to the bottleneck distance is as stable as we could hope for under perturbation. That is, if w_1, w_2 are weight functions on nodes that determine node order, then

$$d_{BN}^n(P_1, P_2) \leqslant ||w_1 - w_2||_\infty, \tag{3}$$

with $|| \cdot ||_\infty$ being the L_∞-norm. Said another way, if we take the ith and jth nodes in the ordering and swap them, the bottleneck distance between the original and swapped persistent homology is bounded above by $|j - i|$. In the Appendix, Sec. 7.1, we provide more detail regarding and intuition for this upper bound.

Specifically in the context of our study, we wish to compare the persistent homology of a growing graph under its original ordering (B, s_0) and under a new ordering $(B, s_{i,j})$ in which we have swapped only the ith and jth nodes, v_i and v_j. Then the bottleneck distance between the persistent homology of these growing graphs, P_0 and $P_{i,j}$, respectively, describes the effective perturbation of the persistent homology caused by the v_i, v_j swap. Furthermore, $d_{BN}^n(P_0, P_{i,j})$ then gives a measure of how similarly v_i and v_j contribute to the persistent homology in dimension n. If $d_{BN}^n(P_0, P_{i,j}) = 0$, then the v_i, v_j swap does not change the persistent homology. If instead $d_{BN}^n(P_0, P_{i,j}) = |j - i|$, then the v_i, v_j node swap causes the maximal effective perturbation to the persistent homology. Combining the bottleneck distance with the theoretical upper bound, we can create a measure of similarity between two nodes in terms of their roles in the persistent homology. We define the *topological similarity* in dimension n as

$$T_n(v_i, v_j) = 1 - \frac{d_{BN}^n(P_0, P_{i,j})}{|j - i|}, \tag{4}$$

so that v_i and v_j have a topological similarity of 1 if their swap in the node ordering does not change the persistent homology, and a topological similarity value of 0 if the persistent homology changes maximally as a result of the node swap. We can more generally discuss the (weighted) topological similarity

of v_i and v_j by averaging over dimensions, that is,

$$T(v_i, v_j) = \frac{1}{D} \sum_{n=1}^{D} T_n(v_i, v_j). \qquad (5)$$

D. Graph summaries and statistics

In this work, we will ask how topological similarity might form motifs described by commonly used graph summaries. First recall that the *degree* of node v_i is the number of edges incident to node v_i and that it is commonly denoted k_i. If we also have edge weight information, we can calculate the *strength* of node v_i as

$$k_i^w = \sum_{i=0, i \neq j}^{N} w_{i,j}, \qquad (6)$$

where $w_{i,j}$ is the weight of edge (v_i, v_j). Additionally, we will ask how the topological similarity between nodes relates to the topological overlap of the nodes in the original graph. We acknowledge that different definitions of topological overlap exist [45,46], but for our purposes we will adapt the definition from Ref. [45] slightly and define topological overlap between nodes v_i and v_j as

$$O(v_i, v_j) = \frac{l_{i,j} + B_{i,j}}{\max(k_i, k_j)} \qquad (7)$$

with $B_{i,j} = 1$ if v_i, v_j are connected by an edge in binary graph B and 0 otherwise, and where $l_{i,j}$ is the number of nodes v_k with $k \neq j, k \neq i$ that are neighbors of both nodes v_i and v_j. Intuitively, the numerator counts the number of shared neighbors between nodes v_i and v_j, and the denominator normalizes by the larger of the two node degrees. Then two nodes who share exactly the same set of neighbors (and no more) will have a topological overlap of 1 and nodes that are not connected and share no neighbors will have a topological overlap of 0. Finally, we ask about the community structure of the topological similarity graph, by applying the Louvain method for modularity maximization [47] as implemented in iGraph [48].

E. Growing graph models

Thus far, we have developed our intuitions on a single growing graph. Yet, in order to understand reorderability at a larger scale, we must work with models that can generate ensembles of growing graphs. Here we will test the reorderability of six growing graph models that span a range of topologies including both embedded and nonembedded models. For each model, we will generate a growing graph (and later an *n*-order complex) by adding nodes $v_1, v_2, \ldots, v_N$ one at a time for a final count of $N = 70$ nodes. For clarity, we refer to the order $v_1, v_2, \ldots, v_N$ as s_0, where the subscript reflects the fact that this order is the original ordering for the growing graph. Note that the current study differs in kind from those that focus on edge-weighted networks, in which both the number of nodes and number of edges are easily held constant. Here in the growing graph case, we seek to compare growing processes across the number of nodes added, and thus we do not restrict our study to graph topologies with

a fixed edge density. We chose 70 nodes in order to keep the computational costs of node swapping manageable; note that in the results, we run persistent homology after swapping every pair of nodes in each replicate. We expect our high-level reorderability classification results to hold if we increase the node number.

The six growing graph models that we study are described in greater detail below.

(1) *Constant probability.* We begin with a random model from Refs. [49,50] that aims to capture random topology with few constraints. At the addition of node v_i, edges are added between node v_i and all previously added nodes v_* with a constant probability $p(v_i, v_*) = c \in [0, 1]$. We report results for $c = 0.4$ in the main text, and note that the Betti curves and barcodes alone for $c = 0.3$ can be found in Ref. [31].

(2) *Proportional probability.* Next we include a random model in which connection probabilities increase throughout growth, instead of being held constant. As in Ref. [31], when adding node v_i we add each edge between v_i and previously added nodes v_* with a probability proportional to the node number added, $p(v_i, v_*) = i/N$.

(3) *Oscillating probability.* Here we examine the effects of a nonmonotonic function determining connection probabilities on the growing topology. In the oscillating probability model, we determine the edge probability using an oscillatory function; that is, we add edges between the new node v_i and previous nodes with probability $p(v_i, v_*) = |\sin(a\pi \frac{i}{N})|$. Here we chose $a = 2$ in order to ensure that we can clearly observe the matching between oscillations in the Betti curves and the resulting high-density edge bands in the adjacency matrix that arise from the periodicity in the growth rule. The η parameter controls the frequency of connection probability oscillations and, as such, larger values of η would be appropriate for larger growing graphs (i.e., those composed of more nodes) so that the effects of each peak can be more easily determined.

(4) *Preferential attachment.* The preferential attachment model famously constructs a graph with a heavy-tailed degree distribution via its "rich-get-richer" growth algorithm. Following Refs. [51,52], we begin with a connected random graph on m_0 nodes. We then add nodes one at a time and at the addition of node v_i, we connect v_i with m previously added nodes with a probability proportional to the degree of each node already present. These rules produce a growing graph in which high-degree nodes are likely to continue receiving more connections as nodes are added, in comparison with low-degree nodes. We show results for $m = m_0 = 4$ in the main text.

(5) *Random geometric.* Next, we move to embedded graphs and begin with a common random graph designed to capture topology arising only from constraints imposed by the embedding. Directly inspired by Refs. [53–55], we sample N points uniformly at random from $[0, 1]^D$. We choose a threshold ϵ and create a binary random geometric graph with edges existing if $d(v_i, v_j) < \epsilon$. To mimic an observer moving through the embedded graph, we order nodes by the value of the first coordinate. For this study, we chose $D = 3$ and $\epsilon = 0.15$, in order to ensure that at least some persistent homology would be present by the end of the filtration [53,56].

(6) *Spatial growth.* Alternatively, an embedded graph may randomly grow out from a particular point, spawning new

nodes that attach most often to nearby nodes. Following Ref. [57], we begin with one node at $(0.5, 0.5)$ in $[0, 1]^2$. We add nodes $v_2, \ldots, v_N$ one at a time, first choosing a location in the unit square uniformly at random and then connecting the new node v_i to previously added nodes v_* with probability $p(v_i, v_*) = \beta e^{-\alpha d(v_i, v_*)}$. Here $d(v_i, v_*)$ is the Euclidean distance between node positions, β is the density parameter, and α controls the spatial range. We show results for $\beta = 1$ and $\alpha = 4$ in the main text, which was shown to produce graphs in the small-world regime [57].

For each growing graph model, we first generate 1000 growing graphs (B_η, s_0), $\eta = 1, \ldots, 1000$ and then we construct the n-order complex as described above by filling in cliques with simplices. Since we can uniquely associate the pair (B, s) with $\mathrm{nord}(B, s)$, we will use the growing graph language for simplicity in the Results section when it does not cause confusion. We compute the persistent homology using Eirene [58] and we use the TDA package in R [59] for bottleneck distances. For the sake of the calculations, any persistent cavity that persists through the end of the filtration (or formally has a death time of ∞) is assigned a death time of $N + 1$. Code for all analyses and growing graph models can be found at Ref. [60].

III. RESULTS

A. Meet the team

Before tackling how the topology of a growing network model might (or might not) be susceptible to node reordering, we first ask how the evolving topology differs between six growing network models chosen to span growth mechanisms relevant for social systems, material systems, spatially embedded systems, and dynamical systems. Because we are interested in the growing global topology of these networks, we use persistent homology, which describes how topological cavities emerge, evolve, and collapse throughout the growth of a network (see the Methods section and Refs. [9,11,33,34] for more details). Persistent homology returns the birth and death times for each persistent cavity of a particular dimension in the form of a barcode, with each bar corresponding to a persistent cavity. Additionally we can summarize the barcode by only recording the number of persistent cavities of a given dimension present after each node addition, and we refer to the data recorded in this way as the Betti curve of dimension n or β_n. Specifically for our growing graph models, we follow the rules of a given generative model to construct a growing graph (B, s_0) with binary graph B and node ordering $s_0 = v_1, v_2, \ldots, v_N$ (Fig. 4, step 1). From this growing graph, we construct the node-filtered order complex (n-order complex) and compute its persistent homology across the entire growth process, displayed as a barcode (Fig. 4, steps 2 and 3). We repeat this process for 1000 instantiations of the generative model, and then we calculate the average Betti curves across instantiations (Fig. 4, step 4). We show the results of this analysis for all growing graph models in Figs. 5 and 6, and in the Appendix we also include expanded figures showing how the average edge density, average degree distributions, and average number of persistent cavities born or killed evolve as the graph matures (Figs. 23 and 24).

1. Constant probability

We first begin with the constant probability model as it is one of the most basic random growing graph models. For this model at each node addition, we add edges with constant probability $c \in [0, 1]$ (here $c = 0.4$). Then if we slice the filtration at node v_i, we will have a random graph on i nodes in which each edge was added with probability c [50]. In the average Betti curve [Fig. 5(a)], we see first a wide peak of persistent cavities of dimension 1 followed by a sudden growth of persistent cavities of dimension 2. Later, as the dimension 1 persistent homology dies, dimension 3 begins to sharply increase. Interestingly, the barcode shows that once persistent cavities in dimensions 2 and 3 are born, they are extremely unlikely to be killed by a later node addition.

2. Proportional probability

We next consider the proportional probability model, which adds edges between the new node v_i and previous nodes with probability i/N. Note that this rule requires that the last node connect to all other nodes in the graph. Consequently, we see no homology can live through the end of the filtration. Still, as shown by the example barcode, there tends to be a wide distribution of lifetimes, particularly in dimension 1. The Betti curves show an increasing pattern of peaks, each taller and narrower than the last. Additionally, the pattern observed in these lower dimensions suggests that the support of each β_n overlaps approximately half of β_{n-1} and β_{n+1}, so that as the β_{n-1} tends to 0, β_{n+1} begins to increase.

3. Oscillating probability

While the previous two functions describing the probability of adding edges emanating from the new node at each node addition have been constant or strictly increasing as the graph grows, we next ask what type of topology we can achieve using a nonmonotonic function. Specifically here we choose an oscillating function, with edges added between a new node v_i and a previously added node with probability $|\sin(2\pi \frac{i}{N})|$. This choice leads to the evolution of two groups of nodes with high degrees: Generally nodes added between nodes 20 and 30 will have a large degree, and those added between nodes 50 and 60 will have an even larger degree, reflecting the periodicity of the edge probability function [shown in Fig. 23(c)]. Although we generally do not see many persistent cycles, on average we see the Betti curves reflect the oscillations in edge density [Fig. 5(c)]. That is, we generally observe a burst of persistent cycle activity (birth and death) while the edge density curve has a positive second derivative.

4. Preferential attachment model

One of the most studied graph models in network science—the preferential attachment model [51,52]—adds a fixed number of edges with each node addition, but connects these edges preferentially to high-degree nodes in the network. For our chosen parameters ($m = m_0 = 4$), we see that the first few nodes end with the highest degrees and that most other nodes have low degrees, as expected [Fig. 24(a)]. We also observe a large abundance of persistent cavities in dimension 1 with very little cavity death [Fig. 6(a)]. The low
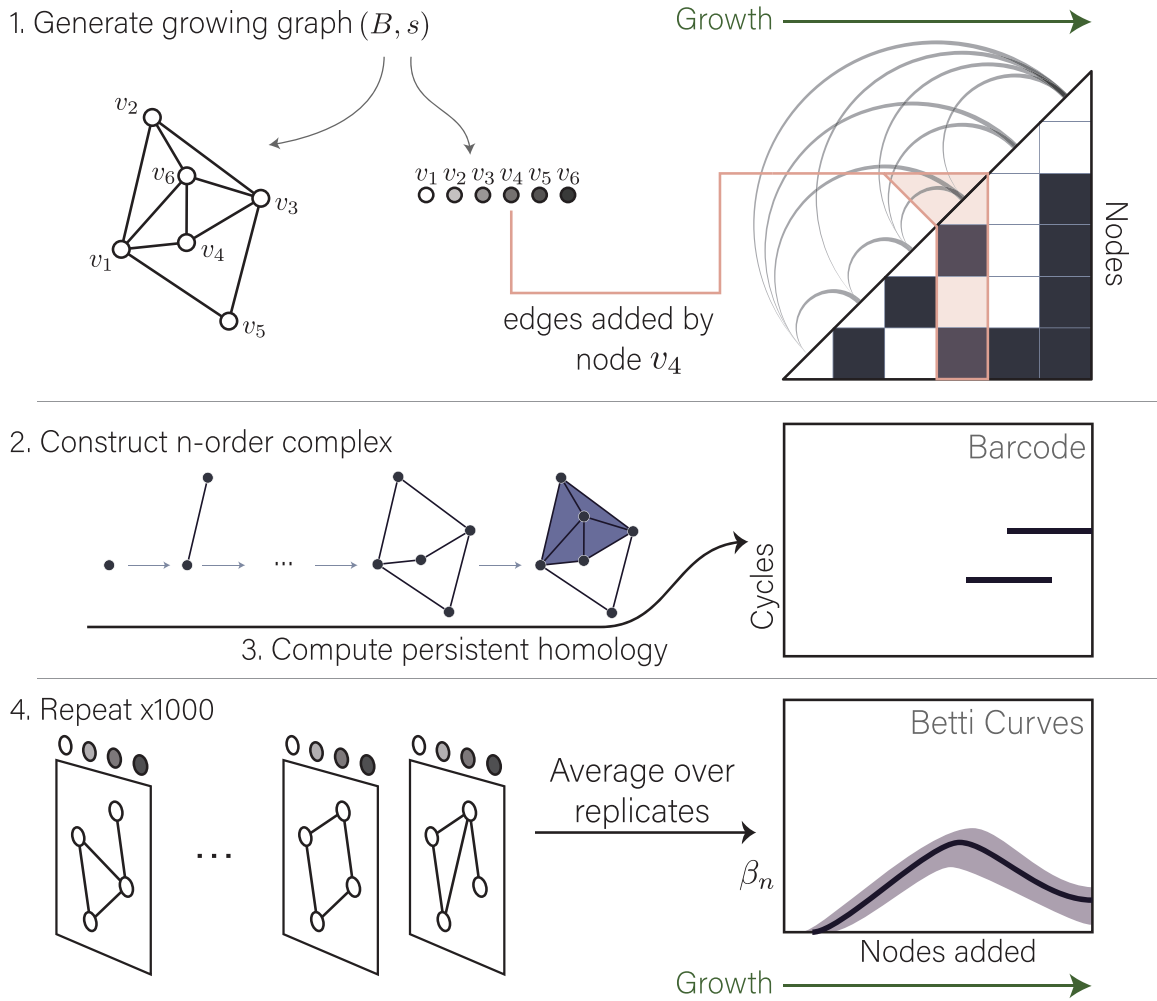
FIG. 4. Outline of the computation of persistent homology for growing graph models. (1) Following the model growth rules, we first construct a growing graph, illustrated by the adjacency matrix and fingerprint graph (right). In the fingerprint graph, we order node positions via ordering s_0 and display edges between nodes in B as gray arcs, so that at the addition of v_4, for example, we see two arcs reaching leftward to connect to nodes v_1 and v_3. We then (2) translate the growing graph to an n-order complex and (3) compute the persistent homology to recover the barcode (right). (4) We repeat this process over 1000 growing graphs constructed from the same generative model, and we present the average Betti curves (solid line) and one standard deviation (shaded).

frequency of persistent cavity death is unsurprising, since to kill a persistent cavity an added node must connect to all nodes of a generating cycle, and a cycle likely involves both high-degree and low-degree nodes.

5. *Random geometric model*

All models that we have considered thus far are not embedded into any particular Euclidean (or non-Euclidean) space. Yet, many real-world systems including the brain, transportation networks, and granular materials exist within and are often constrained by some embedded space such as $\mathbb{R}^3$. Our next model—the random geometric model—incorporates information about how the growing graph lives in the embedded space. To construct this growing graph, we randomly choose locations in $[0, 1]^3$, connect all nodes with distance less than 0.15, and then filter by moving along the x direction through the embedded graph. We imagine this process as walking through a point cloud in which nearby nodes are connected. We find only a small number of persistent cavities emerge

[Fig. 6(b)], and often we only observe persistent homology in dimension 1 (as expected for these parameters [53]). Curiously, this growing graph model produces the only (near) linear average Betti curves for these parameters out of all the models tested. Importantly, we do not include periodic boundary conditions, which is reflected in the heat map of node degree evolution [Fig. 24(b)].

6. *Spatial growth model*

Finally, we investigate a spatial growth model that simulates a growing process in $[0, 1]^2$. In this model, we spawn nodes randomly in the unit square, and we add edges with a probability based on the Euclidean distance between nodes, so that nodes placed near each other will connect with higher probability. In contrast to the random geometric model, we see large amounts of persistent homology in dimensions 1 and 2, and more S-shaped Betti curves that are qualitatively more similar to the nonembedded constant probability model than to the random geometric model. Additionally, the increase of
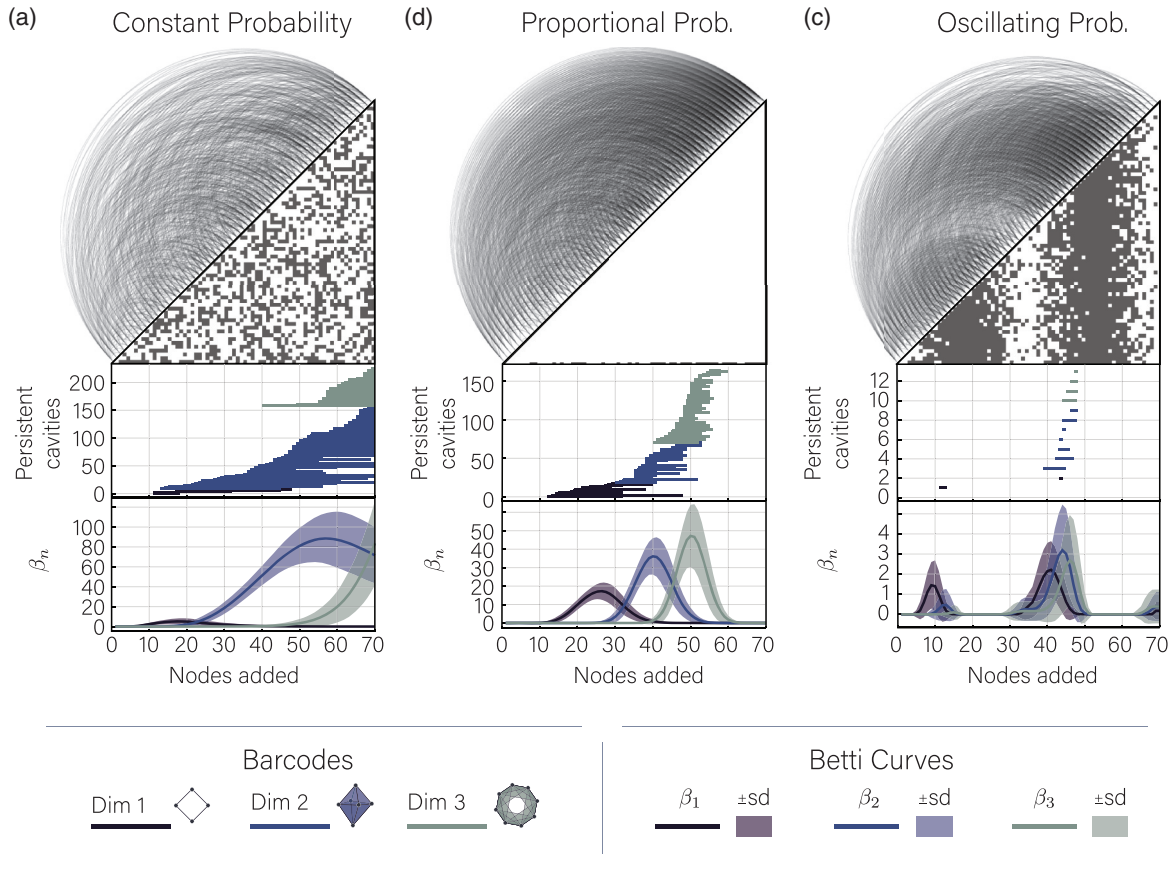
FIG. 5. Growing graph models display a wide range of evolving topologies. Persistent homology of the (a) constant probability, (b) proportional probability, and (c) oscillating probability models. Each panel shows one example adjacency matrix and fingerprint graph (top), the corresponding barcode (middle), and averaged Betti curves across replicates (bottom) with standard deviation (shaded).

persistent homology in dimension 2 aligns with the beginning of the slow decrease of persistent homology in dimension 1, which suggests a predictable transition from a lower dimensional to a higher dimensional persistent homology regime similar to that seen in the proportional probability or constant probability model.

To summarize, the six growing graph models that we consider here display a wide range of persistent homology signatures arising from the differences in growth processes.

B. Random permutations and global reorderability

While the above section considered the evolving topology of the originally ordered growing graphs, here we ask how resilient the topology is to any random ordering of node addition. In real systems, such random ordering would be observed if a group of neurons randomly fired instead of properly propagating a signal [61], or if diseases were to randomly spawn instead of diffuse predictably along a network [62–65]. In each of these cases, we do not generate a new set of connections based on the reordering, but instead keep the connections of the system constant and only alter the node addition order, which in turn impacts the edge addition order. If indeed a growing process was to produce the same final binary graph but with a random node addition order, as discussed in the Methods section, then there are two bounding

possibilities: Either the persistent homology of the growing graph could be independent of the order in which nodes are added or it could be highly dependent on that order. We call the resiliency of the growing graph's persistent homology to random node addition order the *global reorderability* of the growing graph, and in the following experiments, we examine the extent to which each of the above growing graph models is globally reorderable.

In order to examine the global reorderability of each growing graph model, we first generate 100 growing graphs (B_η, s_0) for each $\eta = 1, \ldots, 100$ following the model rules (Fig. 7, green shaded). Next, for each generated growing graph B_η, we create reordered growing graphs in which we add nodes uniformly at random but use the same edges as in B_η (Fig. 7, pink shaded). More specifically, we construct (B_η, s_0) for $\eta = 1, \ldots, 100$, and then for each η we randomly permute s_0 to give a new node order s_{r^m} for $m = 1, \ldots, 100$. We then see the reordered growing graph (B_η, s_{r^m}) as the sequence of induced subgraphs of B_η including nodes in the order of s_{r^m}. All random permutations of s_0 are generated anew for each η. We then compute the persistent homology of each generated and reordered growing graph, and we show the average Betti curves of the originally generated and reordered growing graphs [Fig. 7(c), solid and dashed lines].

Now before directly comparing the persistent homology of the reordered growing graphs to that of the originally gen-
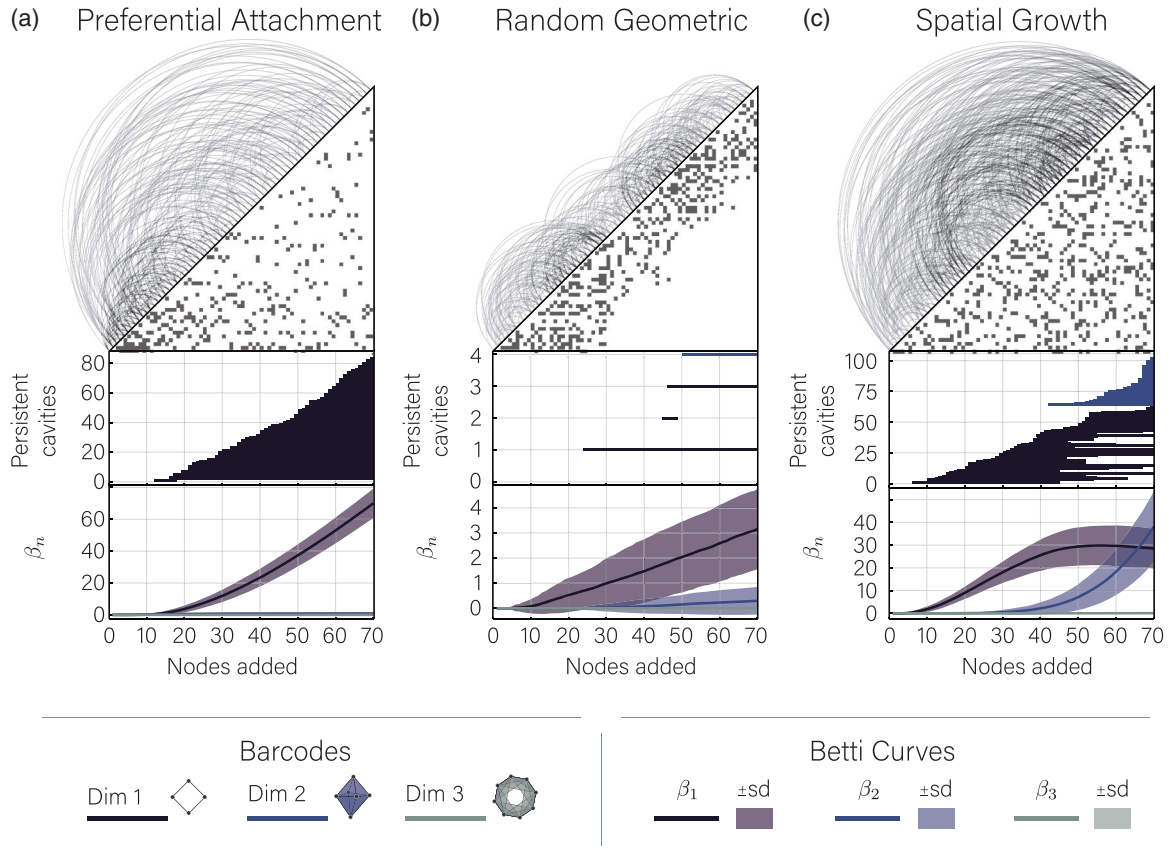
FIG. 6. Growing graph models display a wide range of evolving topologies. Persistent homology of the (a) preferential attachment, (b) random geometric, and (c) spatial growth models. Each panel shows one example adjacency matrix and fingerprint graph (top), the corresponding barcode (middle), and averaged Betti curves across replicates (bottom) with standard deviation (shaded).

erated growing graphs, we must first understand the intrinsic variability of the persistent homology with respect to the range of growing topologies created by the generative model. For example, if we later find that the reordered graphs show a very different persistent homology than that of the originally ordered graphs, could this be simply due to the fact that the underlying graph model produces such a wide range of topologies? In order to answer this question, we sample ten thousand pairs of generated growing graphs with node order s_0 and compute distances between their persistent homology summaries [Fig. 7(d), dark box]. Then, when we calculate the distance between the persistent homology summaries (barcodes or Betti curves) of the originally ordered graph and each of its reorderings for all generated growing graphs [Fig. 7(d), light box, orange lines], we can determine if these distances are smaller (or larger) than expected given the variability of the growing graph model topology. If the distances between the persistent homology of growing graphs and that of their reorderings is smaller than the distances between the persistent homology of randomly sampled generated graphs, then we say that the growing graph model is globally reorderable, since the distance after reordering is smaller than the distance between generated graph topologies.

Next, for illustrative purposes, let us imagine that the persistent homology output of the original and random ordered growing graphs are points in some high dimensional space. It may be that the original ordering s_0 results in an evolving

topology that sits within the masses of persistent homology generated by random orders; on the other extreme, it may be that the persistent homology from s_0 is extremely different from that generated from any random ordering. If the former is true, then we might conclude that the s_0 ordering offers no distinct difference in terms of the evolving topology than any random ordering. But if the latter is true, then we can conclude that the particular s_0 ordering does hold a special significance in terms of the persistent homology for the growing graph model, as the originally ordered persistent homology is far different than that from careless orderings. To assess this behavior, for each set of growing graphs with the same binary graph (of which there are 101), we randomly sample 100 pairs of these growing graphs and compute distances between their topological summaries [Fig. 7(d), blue lines, striped box]. If the distribution of random orderings within a topology (striped box) is larger than or equal to the distribution of distances between the original and reordered growing graphs (light box), s_0 would not appear to generate a persistent homology signature distinct from any random node order permutation. Conversely, if we see that distances between persistent homology outputs within a topology (patterned box) are generally smaller than distances between persistent homology from s_0 and that from a random reordering (light box), then the evolving topology from s_0 must be quite distinct from the evolving topology generated by any random ordering.
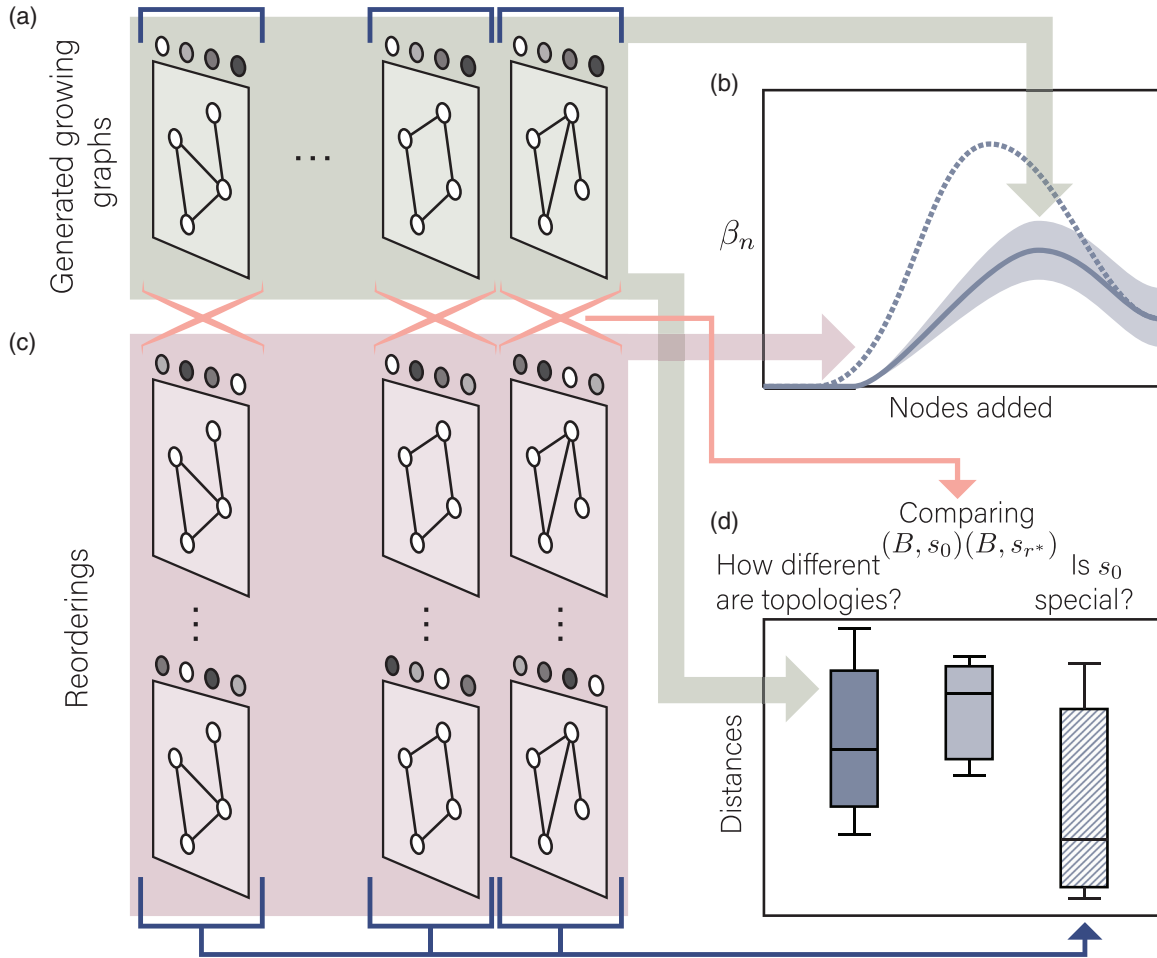
FIG. 7. Global reordering process and comparisons. (a) Following a growing graph model, we generate 100 growing graphs and (b) calculate their persistent homology and average Betti curves (solid line). (c) For each growing graph in panel (a) we randomly permute the node order and then compute the persistent homology and average Betti curve [dashed line in panel (b)] to compare with that of the original growing graphs. (d) To determine reorderability, we plot distributions of the distances between persistent homology outputs of randomly sampled generated graphs (green arrow, dark box), distances between persistent homology outputs of the generated graph and that of their reordered growing graphs (orange lines, light box), and distances between any pair of node orderings on the same binary graph (blue lines, patterned box). For this example, since the lighter box generally sits higher than the dark box, we would say that this growing graph model is not globally reorderable. Additionally, since distances between the persistent homology of randomly sampled orderings on the same topology (striped, columns of growing graphs) are smaller than the distances between the persistent homology of (B, s_0) and the persistent homology of the same graph grown with randomized orderings (B, s_{r*}), we can conclude that the s_0 ordering has a special significance.

1. Differences in Betti curves after reordering

We aim to determine if the Betti curve changes as a result of reordering the nodes in our growing graphs. As a first look, we show the average Betti curves for the original growing graph models (B_*, s_0) and for all reorderings (B_*, s_{r*}) in the left column of Figs. 8 and 9 (solid and dashed lines, respectively). We find little difference in the average Betti curves before and after node reordering for the constant probability, preferential attachment, and spatial growth models. Conversely, the Betti curves of the random geometric and oscillating probability models show marked differences in shape between the original and globally reordered growing graphs. More drastic still, on average we observe that all of the persistent homology seen in the original proportional probability model is killed when we randomly reorder the node addition sequence.

In order to quantify these results, we can compute a distance between Betti curves in each dimension from any pair of growing graphs. For simplicity, we define the distance between two Betti curves as the maximum vertical difference between the curves at any point [see Eq. (1)]. With this Betti curve distance definition, we now show the distribution of Betti curve distances between randomly sampled pairs of generated graphs (Figs. 8 and 9, middle column, solid boxes), between different orderings while holding the binary graph constant (Figs. 8 and 9, middle column, striped boxes), and between the original growing graph and each of its reorderings (Figs. 8 and 9, middle column, light boxes) as discussed above (see Fig. 7). If the distances between the original and reordered graphs are significantly smaller than that of any sampled pair of generated graphs (comparing solid and light boxes), we call the model globally reorderable with respect
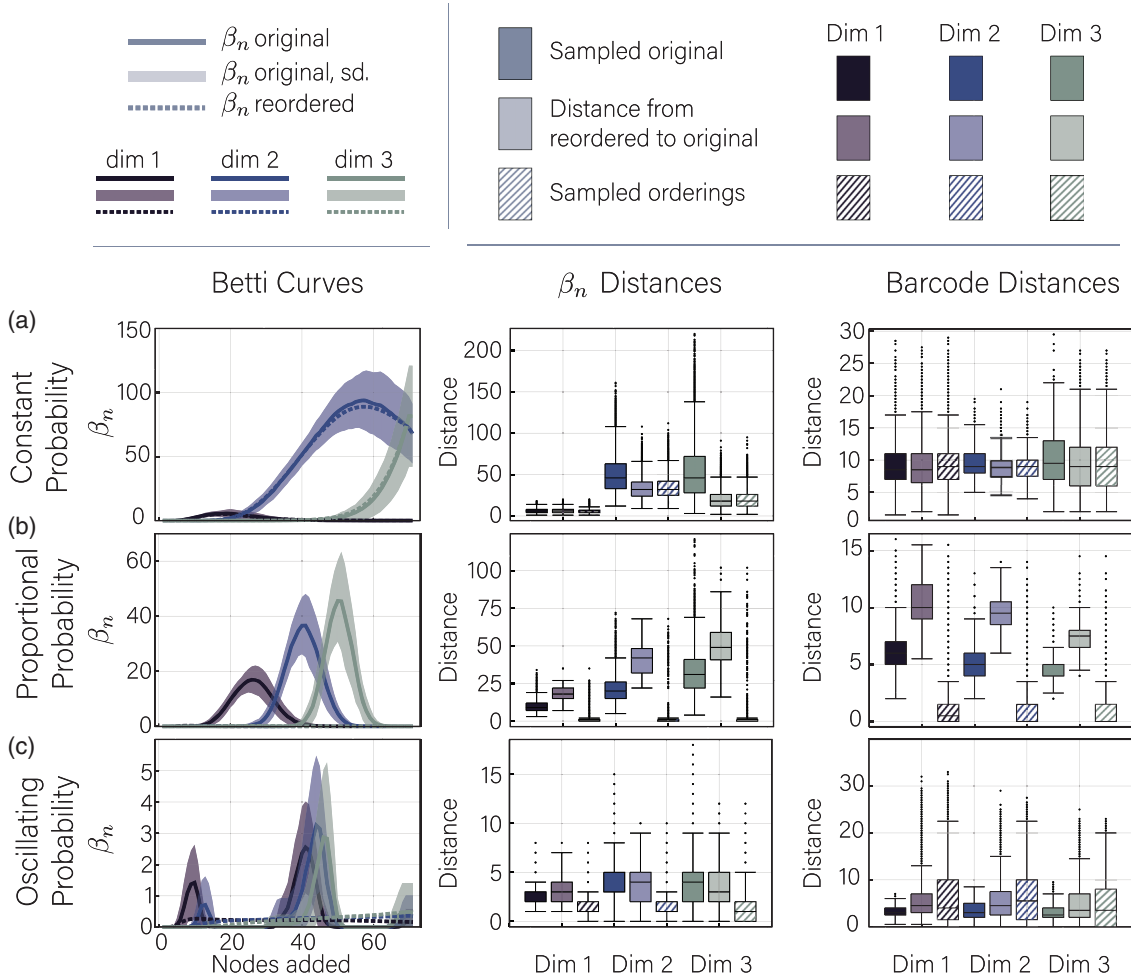
FIG. 8. Global reorderability varies across growing graph models. Results of the random reordering analyses for the (a) constant probability, (b) proportional probability, and (c) oscillating probability models. Within each panel, we show the average Betti curves (left) produced by the originally ordered (solid line, standard deviation shaded) and randomly reordered (dashed line) growing graphs, as well as box plots for Betti curves (middle) and barcodes (right) showing the within-model distribution of distances (solid), distances between the topological summary of the original and randomly reordered growing graphs (lighter shade), and the distribution of differences between randomly sampled pairs of growing graphs ending in the same binary network (striped).

to the Betti curves. We see that the constant probability, preferential attachment, and spatial growth models are clearly globally reorderable with respect to the Betti curves in all three dimensions tested. Additionally, most models show a smaller or near-equal distance between the original and reordered models in comparison to any sampled order of a particular graph B (comparing light to patterned boxes), suggesting that a perturbation from the original growth order will result in a smaller or similar change to the evolving topology than a perturbation to any random order. The proportional probability (and to a lesser extent the oscillating probability) model, however, exhibits notably different behavior, as the Betti curves from any pair of random orderings on a binary graph are likely very similar but the distance between the Betti curves of the original ordering and any reordered growing graph can be an order of magnitude larger.

2. Effect of global reordering on barcodes

The Betti curves still do not capture all of the details of a growing graph's persistent homology. Indeed, multiple barcodes could produce the same Betti curve, and additionally there is no guarantee that the distance (as defined here) between Betti curves will be small as a result of a small perturbation. So we next ask how the barcodes change as a result of global reordering (Figs. 8 and 9, third column). Recall that the bottleneck distance between two barcodes P_1, P_2 is colloquially given by the maximum distance that either end of a bar in P_1 has to move in order to align with its match in P_2 [see Eq. (2)]. As with the Betti curves, we first compute the distributions of barcode distances between the persistent homology of randomly sampled pairs of growing graphs generated by the same model (Figs. 8 and 9, right column, solid boxes). Interestingly, although the random
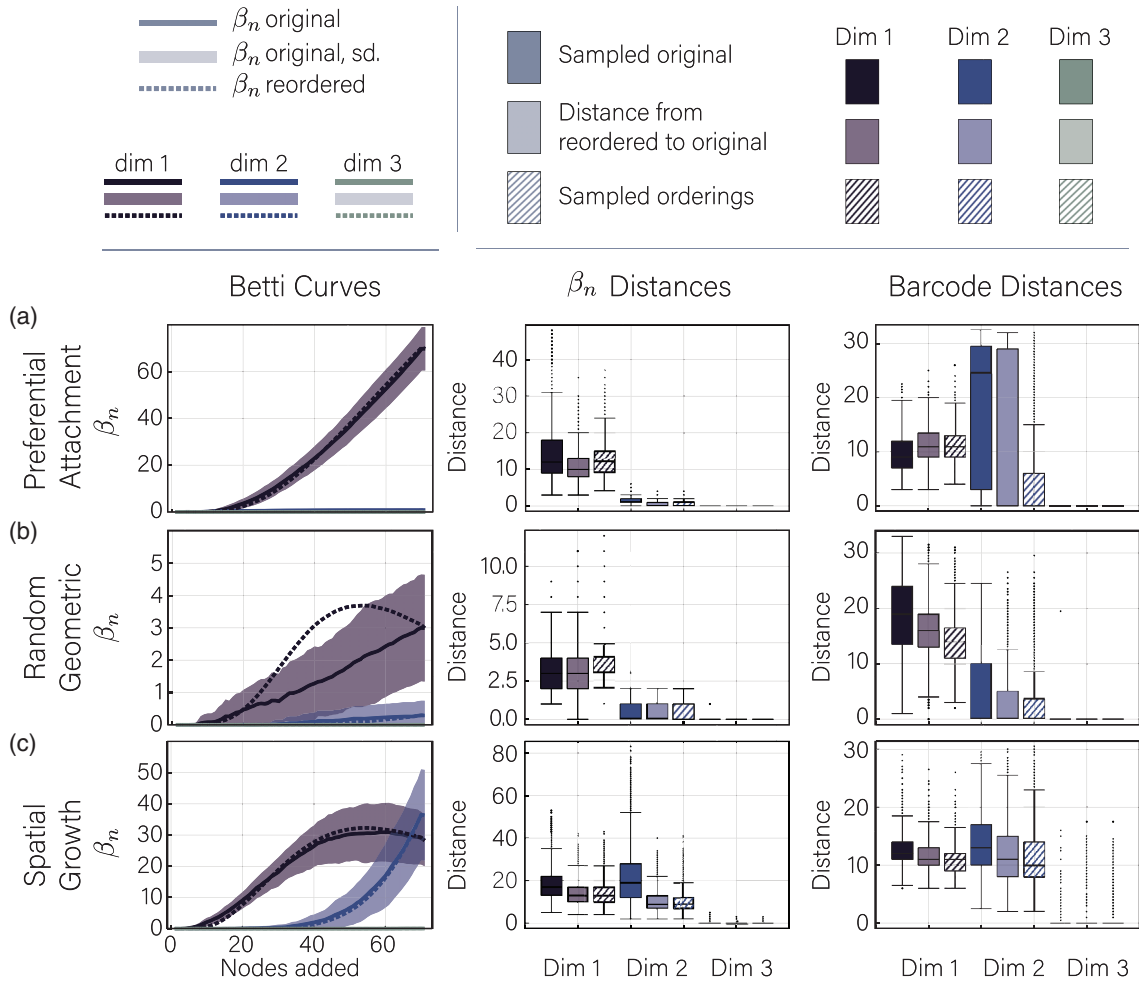
FIG. 9. Global reorderability varies across growing graph models. Results of the random reordering analyses for the (a) preferential attachment, (b) random geometric, and (c) spatial growth models. Within each panel, we show the average Betti curves (left) produced by the originally ordered (solid line, standard deviation shaded) and randomly reordered (dashed line) growing graphs, as well as box plots for Betti curves (middle) and barcodes (right) showing the within-model distribution of distances (solid), distances between the topological summary of the original and randomly reordered growing graphs (lighter shade), and the distribution of differences between randomly sampled pairs of growing graphs ending in the same binary network (striped).

geometric model shows only small differences in within-model Betti curve distances, the barcode distances are some of the highest of all six models, suggesting a large change to only a few highly persistent cavities. In contrast, the oscillating probability model shows low variability in both the Betti curves and barcodes by our measures.

Using the same intuition as for the Betti curves, we can extract the global reorderability of the growing graph models with respect to barcode distances by comparing the barcode distances from random samples of original graphs to the barcode distances from reorderings of the original binary graph (Fig. 7). We see that the preferential attachment model does not exhibit reorderability in dimension 1 and that the random geometric model appears more reorderable in both dimensions 1 and 2 than when we considered Betti curves as the persistent homology summary. Still, we find that the proportional probability model shows extreme sensitivity to reordering and to a lesser extent so does the oscillating

probability model. We observe that the constant probability model, the random geometric, and the spatial growth models display global reorderability with respect to the barcode distance. Finally, we note that for both distances the proportional probability model shows a marked preference for the original s_0 ordering. Then, for this model, if we had let s_0 be any random ordering and performed the same analysis, based on the box plots in Fig. 8(b) we might expect the growing graph to be reorderable.

But why is it that we see such shifts in Betti curves and barcodes after reordering? For example, we may see smaller Betti curves if we increase persistent cycle birth times on average. Or perhaps we might see the same shift in Betti curves but decreased average persistent cycle death times. In Fig. 25, we show distributions of differences between average persistent cycle birth times, average death times, and $\overline{\beta_n}$ values between the original and randomly reordered growing graphs. In particular, for the proportional probability model, we see
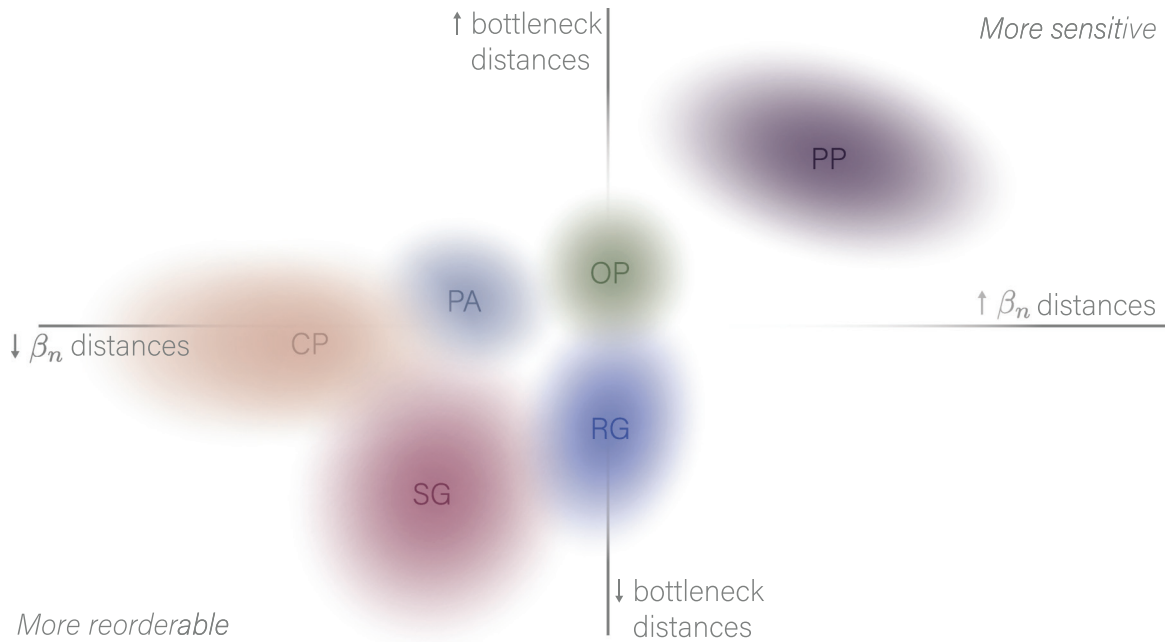
FIG. 10. Illustration of the space of global reorderability. We imagine this two-dimensional space as being parameterized by the distances between Betti curves β_n and barcodes between the reordered and original growing graphs. That is, a model in which the bottleneck distances between the reordered and original growing graphs were higher than expected by random sampling of original graphs would be on the top half of the plane, suggesting that the model is not globally reorderable with respect to this measure. Similarly, a growing graph model in which the Betti curve distances between reorderings and original growing graphs were less than expected by random sampling of original growing graphs would lie on the left half of the plane, suggesting that the model is reorderable with respect to Betti curves. Based on our definitions, we describe global reorderability with respect to both of the distances that we tested as the lower left quadrant. Shapes and locations were chosen based on the distances between persistent homology outputs of the original and reordered growing graphs; we use a fuzzy boarder to indicate that the occupied regions of space may change based on alternative model parameters. Abbreviations: CP, constant probability; PP, proportional probability; OP, oscillating probability; PA, preferential attachment; RG, random geometric; and SG, spatial growth.

that the large barcode distances between global reorderings and the original ordering is likely due to both a decrease in average birth time and a larger decrease in average death time (Fig. 25) so that any persistent cavity that forms does so early in the filtration and lives for only a short time. Additionally, after reordering the proportional probability model, we see that $\overline{\beta_n}$ is generally lower, indicating that we simply see less homology in total, and that when we do see persistent homology it is shorter lived, likely due to the larger probability of early edge addition. From the differences in $\overline{\beta_n}$, we also expect that shifts in the reordered random geometric persistent homology are likely caused by changes in the overall amount of homology and that on average the persistent cavities in the reordered growing graphs die earlier.

When we compare the barcode distances between generated growing graph pairs and their reordered networks, we gain a different perspective on which growing graph models are globally reorderable. Those models that exhibit reorderability with respect to Betti curves but not barcodes are, in a way, similar only at a per-slice level. That is, if we consider any of the models at a given node number i, then the topology in terms of the numbers of cavities in each dimension will be similar between the original and reordered growing graphs. However, if we look across time and take the fruits of the growing process as a collective, the evolving topology has changed drastically. These models illustrate the difference between slicing the growing graph at each node addition

and comparing topology at time points, versus comparing the collective, evolving topology that emerges over the course of graph growth. Conversely, a model that is reorderable by the barcode description but not by Betti curves (though less likely) will show a similar evolving topology throughout the entire growth process but may not show the same homology when comparing slices. In Fig. 10, we show a schematic illustrating where our six studied models with our chosen set of parameters fall on the reorderability axes based on our results. We note that all four quadrants of this graph are accessible and that creating models that optimize one reorderability scheme versus another is an interesting direction for future research.

C. Pairwise swaps and local reorderability

While above we considered randomly swapping all nodes in the ordering, at the other end of the spectrum we can also investigate the effect on the evolving topology of swapping only one pair of nodes in the growing graph. Instead of a growth process progressing randomly, this process would instead be akin to a pair of neurons swapping firing order in a neuronal population [66] or a student switching the order in which they learn a pair of vocabulary words [31]. Here, performing node swaps in our models will offer a measure of the resiliency of a growing graph to a small perturbation of only two nodes as opposed to a complete reordering. As we will see, we gain not only an understanding of each
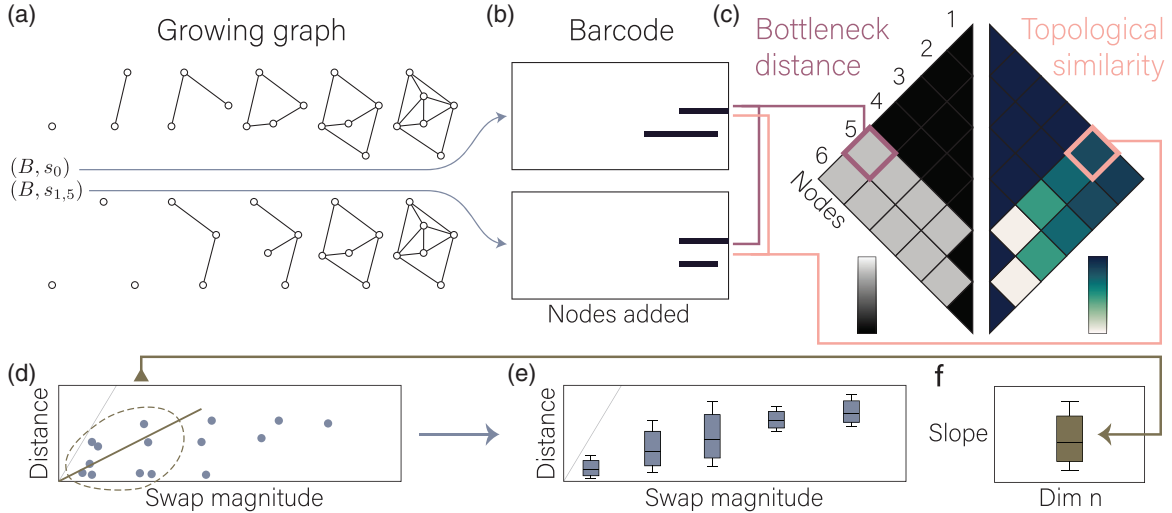
FIG. 11. Analysis flow for node swap studies. (a) Filtration of two growing graphs (B, s_0) and $(B, s_{1,5})$, the latter of which being constructed by swapping nodes v_1 and v_5 in the ordering. (b) The barcode calculated from (B, s_0) and $(B, s_{1,5})$. (c) Calculating the bottleneck distance between the persistent homology of (B, s_0) and (B, s_{v_i, v_j}) for all i, j gives a bottleneck distance matrix (left) and, after dividing by the swap magnitude $|j - i|$ and subtracting from 1, the topological similarity matrix (right). (d) We can investigate the relation between swap magnitude and bottleneck distance by (e) creating box plots of average bottleneck distance against swap magnitude and (f) determining the slope of a line of best fit to points with small swap magnitude and the resulting distribution of such slopes over replicates.

growing graph model's local reorderability but also a deeper appreciation for node participation in the persistent homology.

In order to test how swapping a pair of nodes in the node addition order will alter the persistent homology, we first generate a growing graph (B, s_0) and a node-swapped version $(B, s_{i,j})$ in which the binary graph B is kept unchanged but we swap the ith and jth node in the ordering [Fig. 11(a)]. Recall that since our original ordering is $s_0 = v_1, v_2, \ldots, v_N$ then the ith node in this sequence is v_i. As shown in Fig. 11(b), we then compute the persistent homology and recover the barcodes from each growing graph (B, s_0) and $(B, s_{i,j})$, and next we calculate the bottleneck distance between these barcodes. Repeating this process for every node pair, we can construct a matrix of distances in which the (i, j) entry is the bottleneck distance in dimension n between barcodes arising from (B, s_0) and $(B, s_{i,j})$ [Fig. 11(c), left for dimension 1]. Now, importantly, recall that the bottleneck distance between barcodes recovered from (B, s_0) and $(B, s_{i,j})$ is bounded above by the largest distance any node has moved between the orderings. In our case, this upper bound then is the magnitude of the node swap, or $|j - i|$. Then, finally to calculate topological similarity between nodes, we normalize the bottleneck distance by this swap magnitude and subtract from 1. Explicitly, we define the topological similarity in dimension n of a node pair v_i, v_j as $T_n(v_i, v_j) = 1 - \frac{d_{BN}^n(P_0, P_{i,j})}{|j-i|}$ and the average topological similarity over dimensions as $T(v_i, v_j) = \frac{1}{D} \sum_{n=1}^{D} T_n(v_i, v_j)$ where here our maximum dimension $D = 3$. If a pair of nodes (v_i, v_j) has topological similarity close to 1, then the bottleneck distance between barcodes from (B, s_0) and $(B, s_{i,j})$ must be small compared to the swap magnitude. Said another way, the swap could have caused a large change in persistent homology but did not. Again we can record the topological similarity of each node pair in a matrix as shown in the right of Fig. 11(c). Note that for both sides of the split matrix, a

lighter color indicates node pairs that when swapped will alter the persistent homology more drastically while darker colors indicate node pairs that when swapped have little effect on the persistent homology. Though computed and described here, we will make the impact of the topological similarity matrix apparent even further in the last portion of the Results section. In Fig. 11, we show our process for one growing graph, but for growing graph models we generate 20 replicates and then average bottleneck distances and topological similarity values across replicates and dimensions unless otherwise specified.

Once we have calculated bottleneck distances and topological similarity for each node pair, we next ask how the bottleneck distance varies with the swap magnitude. We can plot the bottleneck distance calculated from each node pair, averaged over dimensions, in one growing graph against the swap magnitude as in Fig. 11(d). Then, we average across dimensions and collate across 20 replicates from a growing graph model and show the results as a sequence of box plots as in Fig. 11(e). Note that neither the box plots in Fig. 11(d) nor the box plots in Fig. 11(e) can surpass the $y = x$ line due to the upper bound on bottleneck distance; thus, for reference we plot the $y = x$ line in gray. Finally, recall that the goal of this section is to determine the robustness of each growing graph model to small perturbations. In order to quantify this robustness, we use the bottleneck distances from all swaps with a magnitude in the smallest 20% of possible swap magnitudes (magnitudes $\leqslant 14$) for each of the 20 replicates to calculate a linear best fit and record the distribution of slopes in a box plot for each dimension [illustrated in Fig. 11(f)]. Smaller slope values indicate that temporally local perturbations to the node addition order have only a small effect on the persistent homology, while slope values close to 1 suggest that nodes temporally close in spawning order contribute very differently to the persistent homology of the growing graph. We include
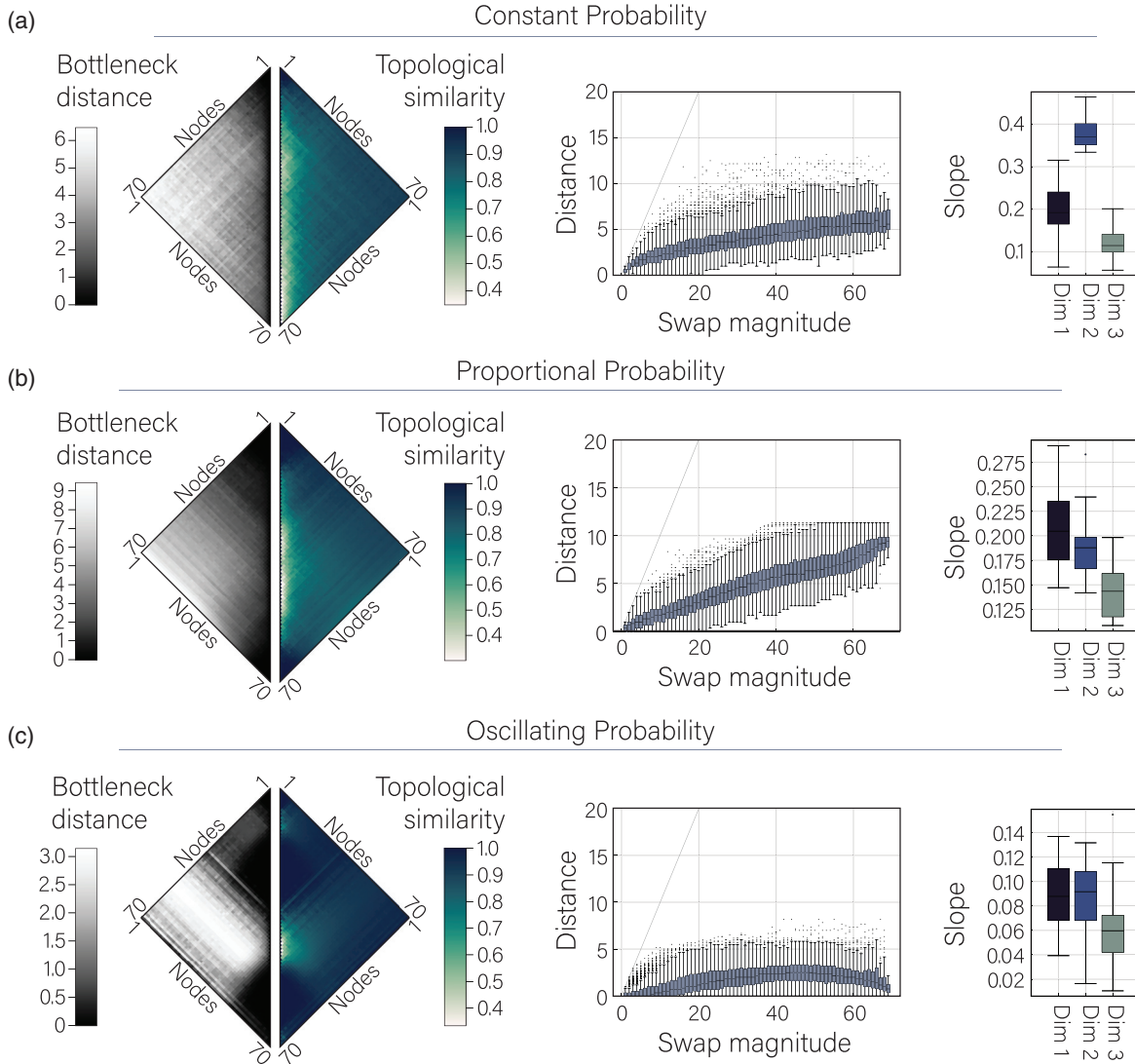
FIG. 12. Node swaps generate a range of effects on persistent homology: Part I. Results of node swap analyses for the (a) constant probability, (b) proportional probability, and (c) oscillating probability models. (Left) The average bottleneck distances across dimensions 1–3 and across replicates (gray color map), and the average topological similarity across dimensions 1–3 and across replicates (green color map). (Middle) Box plots showing distributions of averaged bottleneck distances across dimensions in relation to swap magnitude for all replicates, and distributions of slope values calculated from bottleneck distance and swap magnitude ($\leqslant 14$) pairs from each replicate and for each dimension.

results for slope calculations using swap magnitude cutoffs of 10 and 18 nodes in the Appendix.

1. Node swap analyses on growing graph models

When we use the above analyses to study our six growing graph models, we find again a diverse range of profiles from the models but we also observe that node properties can vary widely. For the constant probability model we see in Fig. 12(a) that the bottleneck distance increases sharply with swap distance for small swaps (magnitude $\approx < 7$) but levels out at the largest swap magnitudes ($\approx > 45$). Interestingly, when broken down by dimension, the distributions of slopes calculated from small swap magnitude ($\leqslant 14$) versus bottleneck distance vary between an average of 0.379 (dimension 2) and an average of 0.122 (dimension 3) across dimensions. In contrast,

the proportional probability model shown in Fig. 12(b) shows a strong linear increase of bottleneck distance with swap magnitudes as expected from the definition of the model. The latest points added will connect to nearly every node; thus, if shifted earlier, this behavior will prevent nontrivial homology from forming. Slope distributions for each dimension have relatively similar averages, reflecting the fact that the increase in bottleneck distance associated with the interval of swaps is similar for each dimension. In Fig. 12(c), we observe that the oscillating probability model has two large regions of zero bottleneck distance between swaps and that these zeroed regions align with periods of no homology in the originally ordered network [Fig. 5(c)]. This phenomenon contrasts with the globally reordered oscillating probability model in which we observed homology through nearly the entire growth process [Fig. 8(c)]. We only observe bottleneck distances <10 in
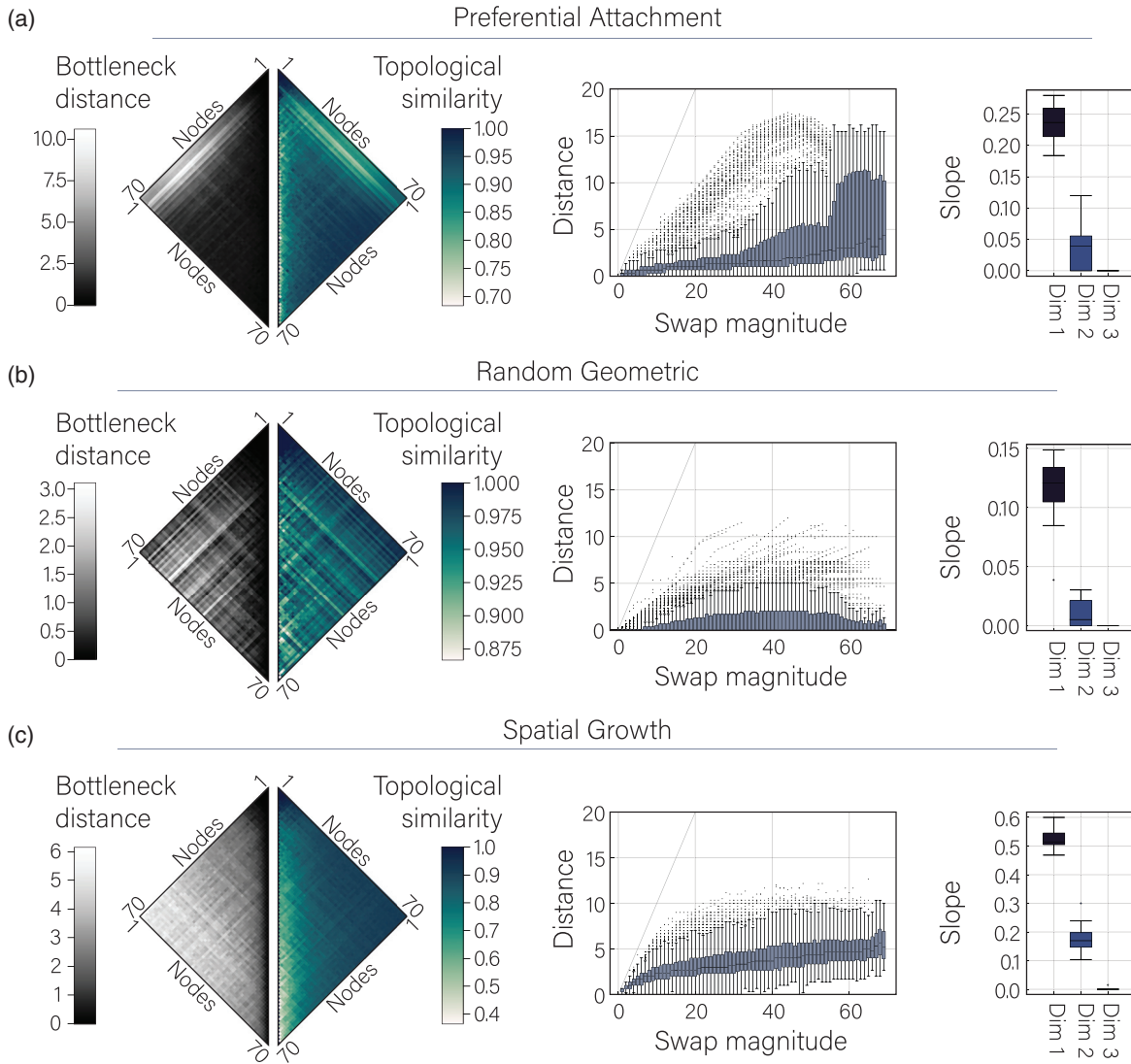
FIG. 13. Node swaps generate a range of effects on persistent homology: Part II. Results of node swap analyses for the (a) preferential attachment, (b) random geometric, and (c) spatial growth models. See Fig. 12 for complete caption.

the locally reordered oscillating probability model, and as a result we see slopes of the lines fit to these data are small.

The other three models show generally less uniform results across nodes. Figure 13(a) shows that the preferential attachment model exhibits a large shift in bottleneck distances between small and large node swaps. As evident in the heat maps in the left of Fig. 13(a), we find that the earliest nodes cause a large shift in bottleneck distances when swapped. Since these early nodes become the few high-degree nodes in the system, we observe by eye (at least for this model) that swapping high-degree nodes will often cause a large change in the persistent homology in comparison to the remaining, low-degree nodes. In Fig. 13(b), we see that the random geometric model shows a striped pattern in the averaged topological similarity and averaged bottleneck distance matrices, suggesting marked inhomogeneity in the graph architecture. We mainly see homology in dimension 1 for this model, resulting in generally low average bottleneck distances. Still, when we restrict ourselves to dimension 1, we see that the average slope of the line fitting the relationship between small swaps

and their associated bottleneck distances is 0.116, which is larger than that of the oscillating probability model. Finally, the spatial growth model displays a surprisingly homogeneous topological similarity matrix [Fig. 13(c)]. However, the box plots of average bottleneck distance display a sharp rise in bottleneck distance for small swaps, a slow rise for mid-range swaps, and finally a quick upturn again for the largest swaps. Indeed, in dimension 1 the slopes of the line fit to these data have an average >0.5, which is the highest average slope seen across all graph models.

We can also use these results from the local reordering analyses to uncover finer details of growing graph architecture. First, if a growing graph model experiences a period of no homology and if or when we swap the nodes that are added within this period we still see no homology, then the earlier node could not cone a cavity that forms before the later node. If the earlier node did cone such a cavity, then moving this node to a later position would allow the cavity to be born, thus making the homology of the resulting growing graph after swapping nonzero. We observe this phenomenon most

strikingly in the oscillating probability model and to a lesser extent in the beginning and end of the proportional probability model. Second, we note that the proportional probability and preferential attachment models are loosely inverses of each other; that is, in the proportional probability model we add the highest degree nodes last, but in the preferential attachment model we add the highest degree nodes first. Indeed, we observe that the bottleneck distance and topological similarity matrices match when flipped vertically by eye. Third and last, we only examine the behavior of lines fit to results from the smallest swaps. The extent to which the bottleneck distance of large swaps plateaus suggests a degree of homogeneity in neighborhood structure between nodes that are temporally distant. For example, the constant probability model plateaus near an average bottleneck distance of 5, which is unsurprising given the random nature of this model; in contrast, the proportional probability model shows steady increases in bottleneck distance across all swap magnitudes.

Additionally, we note that in the global reordering experiments we discover that the spatial growth and constant probability models exhibited similar behavior. When considering the averaged bottleneck distance and averaged topological similarity matrices, this again looks to be the case—at least initially—as the matrices are nearly indistinguishable. However, we find a drastic difference in the values of calculated slopes after fitting the swap distance versus bottleneck distance points. The spatial growth model has the steepest slope in dimension 1 but then is shallower in dimension 2, while the constant probability model shows the opposite effect. Still, these two models are the only growing graph models examined here that display average slopes above 0.3, suggesting that nodes added closely in time contribute very differently to the persistent homology. This property contrasts most with the small slopes observed in the preferential attachment, random geometric, and oscillating probability models, suggesting that nodes added within small time spans form and kill similar cavities.

We have discussed that node swaps with high topological similarity suggest node-neighborhood similarities, but it is interesting to ask whether the topological similarity between nodes could be driven mostly by simply the neighborhood overlap. Certainly if two nodes have exactly the same neighborhoods in B, then swapping these nodes in the ordering could not change the persistent homology. But more generally, could topological similarity be explained by topological overlap? Similarly, we intuitively expect that nodes of high degree should produce a larger effect on the growing topology when swapped than nodes with small degrees (as we generally saw by eye in the preferential attachment model), but is this quantitatively true? To answer the first question, we examine the relationship between the average topological similarity between nodes and the average topological overlap between the same nodes in the original binary graphs B_η, for $\eta = 1, \ldots 20$ (Fig. 26). We see that for all but the proportional probability model, most growing graphs show a weak relation between these two variables, suggesting that the topological overlap alone does not predict the topological similarity between nodes. To answer the latter question, we examine the relation between summed node topological similarity and average degree in the original binary graphs (Fig. 27). Again, we find

that for most models (preferential attachment excluded), the average degree poorly predicts the topological similarity of a node to all others. These results imply that topological similarity can capture information about nodal roles in the larger architecture that is distinct from the graph metric properties tested here.

Together, these node swap experiments suggest a spectrum of local reorderability for growing graph models. Of the models that we examine, we observe that the random geometric and oscillating probability are most resilient to small reorderings in all dimensions tested, followed by the proportional probability and preferential attachment models, and finally the spatial growth and constant probability models, which exhibit the largest topological changes after temporally local reorderings. Furthermore, we find that in most cases the topological similarity cannot be well described by either topological overlap or node degree.

D. Influence of local reorderability and pairwise swaps on global reorderability

After investigating topological resilience to large perturbations (global random reordering) and small perturbations (swapping one node pair), we naturally wonder: Are these two properties related? On even just one growing graph, if *every* node pair was topologically similar (swapping does not change the persistent homology), then could we infer that any possible ordering of nodes would result in the same growing topology? To answer this question and similar questions, we need to dive a little deeper into the results of the node swap analysis. Let us first restrict ourselves to a binary form of topological similarity and say that two nodes are topologically similar if their swap has no effect on the persistent homology, and two nodes are topologically dissimilar otherwise. Then as shown in the Appendix, Sec. 7.2, topological similarity between nodes is a dependency relation: symmetric and reflexive but not transitive (see Counterexample 1 in the Appendix). Thus, we can define the (binary) topological similarity graph of dimension n, denoted T_n, as the graph with the same nodes as our original graph B but with edges between v_i, v_j if v_i, v_j are topologically similar in dimension n. The topological similarity graph allows us to condense the information generated from the $\binom{N}{2}$ node swaps and further unravel the growing graph structure at the mesoscale level. We next can ask questions about the derived topological similarity graph that may offer insights into the growing process. For example, do communities suggesting groups of similar nodes exist in the topological similarity graph? Do we find that all nodes possess a similar capability of swapping without changing persistent homology, or will variation emerge? We will show that from small motifs in T_n we can glean a deeper understanding of the global reorderability of a growing graph, but that these two concepts remain individually manipulable.

1. Intuition on one growing graph

To begin, let us consider one growing graph (B, s_0) on N nodes with topological similarity graph T_n of dimension n and reduce our above questions accordingly. Are the edges in T_n the totality of the information contained in T_n, or does

the structure imply more about the reorderability of (B, s_0)? First, we know that if our graph is totally reorderable (see the Appendix, Sec. 7.2, for examples), then T_n will be a complete graph by definition. So if we can identify that the binary graph has this reorderable property, then we know information about T_n.

Can we also go the other direction and infer information about large node order permutations from the T_n structure? In possibly the simplest case, does finding an m-clique in T_n indicate that all swaps on those nodes admit the same persistent homology? Let us examine what happens when nodes v_i, v_j, and v_k with $i < j < k$ all connect in T_n. By definition, the barcodes (and consequently the Betti curves) produced from the s_0, $s_{i,j}$, $s_{i,k}$, and $s_{j,k}$ orderings on B are all equal in dimension n. If we write out the graph filtration arising from each of these orderings with the same binary graph B, we have

$$G_{v_1} \rightarrow \cdots \rightarrow G_{v_{i-1}} \rightarrow G_{v_i} \rightarrow G_{v_{i+1}} \rightarrow \cdots \rightarrow G_{v_{j-1}}$$
$$\rightarrow G_{v_j} \rightarrow G_{v_{j+1}} \rightarrow \cdots G_{v_{k-1}} \rightarrow G_{v_k} \rightarrow G_{v_{k+1}}$$
$$\rightarrow \cdots \rightarrow G_N,$$

$$G_{v_1} \rightarrow \cdots \rightarrow G_{v_{i-1}} \rightarrow G_{v_j} \rightarrow G_{v_{i+1}} \rightarrow \cdots \rightarrow G_{v_{j-1}}$$
$$\rightarrow G_{v_i} \rightarrow G_{v_{j+1}} \rightarrow \cdots G_{v_{k-1}} \rightarrow G_{v_k} \rightarrow G_{v_{k+1}}$$
$$\rightarrow \cdots \rightarrow G_N,$$

$$G_{v_1} \rightarrow \cdots \rightarrow G_{v_{i-1}} \rightarrow G_{v_k} \rightarrow G_{v_{i+1}} \rightarrow \cdots \rightarrow G_{v_{j-1}}$$
$$\rightarrow G_{v_j} \rightarrow G_{v_{j+1}} \rightarrow \cdots \rightarrow G_{v_{k-1}} \rightarrow G_{v_i} \rightarrow G_{v_{k+1}}$$
$$\rightarrow \cdots \rightarrow G_N,$$

$$G_{v_1} \rightarrow \cdots \rightarrow G_{v_{i-1}} \rightarrow G_{v_i} \rightarrow G_{v_{i+1}} \rightarrow \cdots \rightarrow G_{v_{j-1}}$$
$$\rightarrow G_{v_k} \rightarrow G_{v_{j+1}} \rightarrow \cdots \rightarrow G_{v_{k-1}} \rightarrow G_{v_j} \rightarrow G_{v_{k+1}}$$
$$\rightarrow \cdots \rightarrow G_N,$$

in which each G_{v_*} is the binary graph after the addition of node v_*. More specifically, G_{v_*} is the induced subgraph of B on all nodes added up to and including v_*. This observation is useful, but if we dive deeper we find that we in fact gain more information. Let us examine the graph filtration constructed by swapping nodes v_i, v_j, v_k all together to form the ordering $s_{i,k,j}$ such that now our graph filtration reads as follows:

$$G_{v_1} \rightarrow \cdots \rightarrow G_{v_{i-1}} \rightarrow G_{v_j} \rightarrow G_{v_{i+1}} \rightarrow \cdots \rightarrow G_{v_{j-1}}$$
$$\rightarrow G_{v_k} \rightarrow G_{v_{j+1}} \rightarrow \cdots \rightarrow G_{v_{k-1}} \rightarrow G_{v_i} \rightarrow G_{v_{k+1}}$$
$$\rightarrow \cdots \rightarrow G_N.$$

Notice that here we add the first $j - 1$ nodes in exactly the same order as row 2 above (the $s_{i,j}$ row). After we have added node v_k (the jth node in this $s_{i,k,j}$ sequence), we have added the same subset of nodes as in row 3 above (the $s_{i,k}$ row) at the jth node addition. Then, not only does G_{v_k} from the $s_{i,k,j}$ row equal G_{v_j} in the $s_{i,k}$ row, but we also see that the entire rest of the filtrations are equal to each other. Thus, our new triswap graph filtration is constructed from parts of two pairswap graph filtrations cut and pasted together [67]. Since the graphs are equal at each point (node addition), the homology is also equal at each point, and thus β_n of the triswap must be equal to β_n of the original ordering.

For the above argument, we only used two pairwise swaps to show that the triswap had to produce the same Betti curve. Indeed, there are only two triswaps on three letters and the other swap (adding v_k, v_i, then v_j) can similarly be shown to have the same Betti curve as the original ordering, using only the topological similarity of node pairs v_k, v_i, and v_k, v_j. Then having a 3-clique of nodes v_i, v_j, and v_k in T_n guarantees that any triswap of v_i, v_j, and v_k will also share the same β_n. Does the intuition gained regarding 3-cliques in T_n generalize to higher cliques? In Counterexample 2 in the Appendix, we construct a growing graph in which four nodes are all topologically similar but we observe different Betti curves when we swap all four nodes together. Thus, an arbitrary clique of size 4 or more in T_n does not imply that *all* swaps of the participating nodes will yield the same Betti curves. Still, we can use the above cut-and-paste method with the pairwise-swapped filtrations to construct rules for which permutations of nodes involved in a clique in T_n will produce the same Betti curves. We briefly discuss and explore these rules in Example 3 of the Appendix. Finally, we ask whether the intuitions gained here for Betti curves translate to the bottleneck distance between barcodes. In Counterexample 3 in the Appendix, we construct a growing graph in which v_i, v_j, and v_k are each pairwise topologically similar, but neither triswap produces the same barcode. This finding is perhaps unsurprising, as the barcodes record much more detailed information about the evolving topology of the growing graph than do the Betti curves.

To summarize, we find that the local and global reorderability, though related, are distinct properties of a growing system. A growing graph may be both globally and locally reorderable, locally but not globally reorderable, or not reorderable in either sense. Additionally, our results describing information gained from cliques in T_n re-emphasize the difference in perceived reorderability based on slices (Betti curves) versus longevity (barcodes).

2. For growing graph models

Now we return to our six growing graph models and ask our questions surrounding what the structure of the averaged topological similarity graph tells us about the original growing system. We begin by averaging the topological similarity of node pairs across replicates and across dimensions [Fig. 14(a); see also Figs. 12 and 13]. Intuitively, if a node pair has a resulting high average topological similarity, then generally when constructing a growing graph from the particular model, those two nodes will contribute similarly to the persistent homology. We then ask if there indeed exist strongly connected subgroups of the average topological similarity graph using community detection by modularity maximization [Fig. 14(b); see the Methods section for details]. If a group of nodes forms a community in this averaged topological similarity graph, then often (but not necessarily, as shown above) we expect their pairwise swaps to result in small changes in persistent homology compared to swap magnitude. We represent each of the communities in the detected partition using a different color on the fingerprint graph of an exemplary binary graph B [see Figs. 14(b) and 14(c)] for each of the six studied growing graph models in Fig. 15. We observe that the proportional
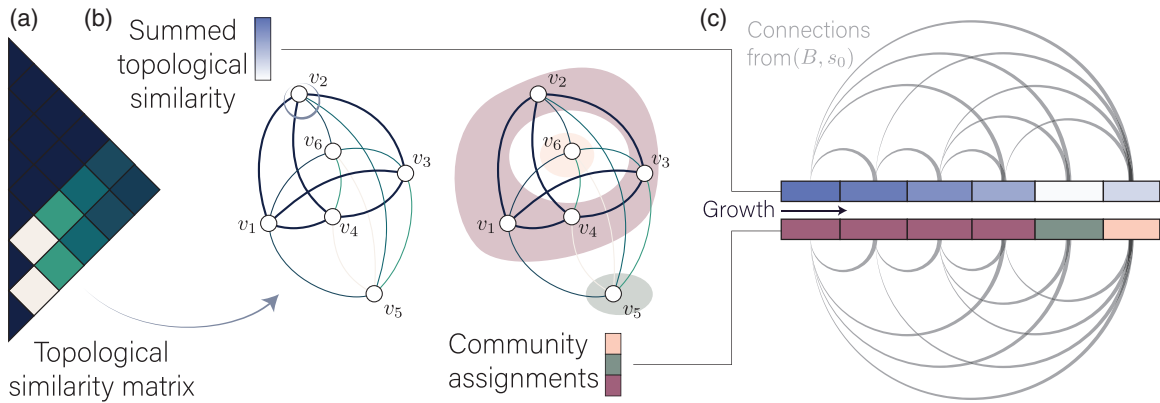
FIG. 14. Analyzing average topological similarity weighted graphs formed from pairwise node order swaps. (a) The topological similarity matrix forms a weighted graph. (b) On the weighted network, node strength corresponds to the summed topological similarity, and community assignments describe sets of nodes that can be pairwise swapped with small consequences on the persistent homology. (c) The summed topological similarity (top) and the community assignments (bottom) shown as vectors paired with the fingerprint graphs from (B, s_0).

probability, oscillating probability, and preferential attachment networks show at least one relatively large, temporally clustered community in their average topological similarity networks. A temporally clustered community found in the averaged topological similarity graph suggests that as the graph develops, the evolving topology will likely change little in response to permutation in node order within the interval in which that community spawns.

Naturally, we next ask when those most or least swappable nodes emerge in the growing graph model. That is, do nodes with the highest summed topological similarity (averaged across dimensions) occur at the beginning of the growth process or at the end? For example, based on the connection patterns of the growing *C. elegans* cellular nervous system, we might expect that neurons born prior to hatching will often have lower summed topological similarity values than neurons
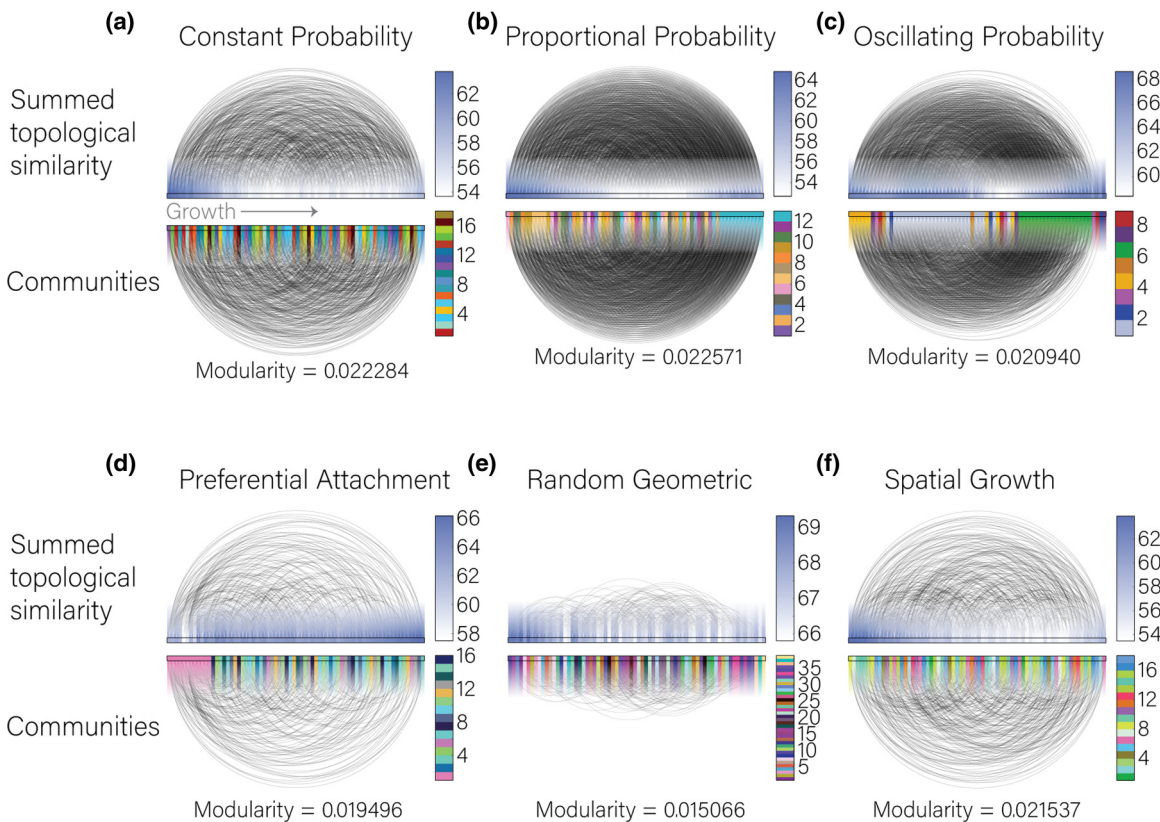
FIG. 15. Community detection and summed topological similarity overlaid on original binary networks. We show the value of the summed topological similarity for each node with the blue color map (top) and calculated community assignment (bottom) overlaid on a representative fingerprint graph for the (a) constant probability, (b) proportional probability, (c) oscillating probability, (d) preferential attachment, (e) random geometric, and (f) spatial growth models.

born after hatching [68]. We calculate the strength (weighted degree of T, or summed topological similarity) of each node within the average topological similarity graphs [Figs. 14(b) and 14(c)] and we depict the summed topological similarity using node color within the fingerprint graphs shown in Fig. 15. We see that for most growing graph models, the earliest added nodes show a relatively high topological similarity, with the exception of the preferential attachment model. Interestingly, the proportional probability, oscillating probability, and to a small extent the spatial growth model show a U-shaped curve of summed topological similarity values. That is, the earliest and latest nodes can be swapped with most other nodes without changing the persistent homology much, but those nodes added in the middle of the growth process are less topologically similar to all other nodes. Then for these growing graph models, the nodes added in the middle are most sensitive to perturbation, suggesting that they exhibit distinct roles in the persistent homology of the growing network.

To summarize, our results suggest that the notion of local reorderability as well as our general analysis approach to studying growing graph processes via node swaps can provide insight into sections of the graph maturation course that may be more topologically robust to node perturbations than other sections. Additionally, these analyses can also expose nodes that are relatively essential for the proper topological evolution of the graph by revealing which nodes on average often create relatively large changes in the evolving topology when swapped.

IV. DISCUSSION

In this paper, we have explicitly defined a convenient framework for studying the evolving topology of growing graphs and the robustness of that topology in six growing graph models. We perturbed each growing process by exchanging the node addition order and tested the ability of each growing graph to retain the same evolving topology after perturbation. We found that the examined graph models display a wide range of resiliency in their topology, here termed *reorderability*, as a result of local and global perturbations. Finally, we found that reorderability at the pairwise level exerts limited restrictions on the global reorderability of a growing network and suggest analyses to understand this influence in the graph model case. In the remainder of this discussion section, we consider the implications of our work first in terms of theoretical matters and then in terms of applications to real-world systems, and we will finally leave the reader with a series of unanswered questions.

A. Persistent homology of growing graph models and their connections to edge filtered complexes

Most commonly in network science, persistent homology has been applied to study edge-weighted networks [41,56]. The edge-weighted network models that have been most well studied from the persistent homology perspective are the independent and identically distributed (i.i.d) uniform random edge weight model [69] and the (edge-filtered) random geometric model [53]. Here we intentionally use similar ideas to construct the constant probability model and the

(node-filtered) random geometric model. Indeed, at the end of these growing processes, we have constructed a binary graph B that could be instead grown from adding edges to a set of N vertices. Specifically, the homology at the end of the filtrations generated by the constant probability model or random geometric model, must match, on average, the homology of their edge-weighted counterparts at that same edge density and number of nodes. Importantly, the connection between these two models is complicated by the variability of edge densities after each node addition in our growing graph models. Still, we strongly suspect that a formal connection between node-filtered and edge-filtered models does exist. If true, a particularly interesting direction for future work is to address the question of whether one could solve the intricacies of an edge-filtered model and apply that solution directly to the node-filtered case, or *vice versa*.

Additionally, as noted in Ref. [31], while our definition of the constant probability model matches the nature of the random i.i.d. edge-filtered complex, the Betti curves of the proportional probability model instead qualitatively match those of the random i.i.d. edge-filtered case [70]. We leave the exploration of the connection between these models for future work, but we speculate that it may be related to the similar progression of randomly attaching higher dimensional simplices. In the i.i.d. edge-filtered model, each new edge added randomly completes larger (higher dimensional) simplices on average, while in the dual of the proportional probability model we randomly add higher dimensional simplices at each step on average. Additionally, both models end with the coning off of all persistent homology, tying at least the homology of the filtration end points to one another.

B. Extending to models of growing simplicial complexes

In this work, we restricted our attention to node-filtered simplicial complexes constructed from clique complexes of graphs. Importantly, many other growing simplicial complex models exist, including those that add new 2-simplices at each filtration step [71], growing simplicial complexes that display hyperbolic geometry [72], weighted growing simplicial complexes [73], point processes [54,55], and more [74]. Some growing simplicial complexes may fit directly into our node-filtered perspective, but others may add exclusively larger simplices at each step. Regardless of simplex size (or distribution of sizes) added at each step, one can ask similar reorderability questions to those asked in this work. For example, how might the persistent homology change if we reordered the growth process to add the smallest simplices first, or instead the largest simplices first? Which generative simplicial complex models exhibit local or global reorderability? Can we determine more reorderable motifs to allow us to simplify or reduce the computation of local reorderability? Given two different growing simplicial complexes, does there exist a reordering of nodes that produces matching persistent homology? Or perhaps can we induce constraints on a developing network that select for reorderability, as with flow networks [75]? These questions on growing simplicial complexes and others raised in the present work additionally build upon previous investigations of the effect of removing vertices or shifting simplex weights on the persistent homology [76].

C. Applications of node-filtered systems

While the reorderability of node-filtered order complexes raises many interesting theoretical questions, we also emphasize the prevalence of node-ordered systems in applications. Certainly many biological systems grow and spread including vasculature networks [77], fungal networks [78,79], and the developing connectome [80,81], but we also see growing processes in nonbiological systems such as evolving mass transportation networks in cities [82] and the evolution of ideas in research journals [83,84]. Moving further still, we can encode signal transduction with the *n*-order complex formalism, such as occurs in the propagation of a focal seizure along the connectome [85] and phosphorylation cascades in response to an extracellular signal [86]. Finally, in graph learning, one traverses a network of topics or ideas and relations between them, making studies involving the learnability of graphs [87] and semantic memory [88] interesting applications for reorderability investigations. Finally and perhaps most interestingly, in certain applications one may be able to experimentally test and manipulate reorderability of complex systems. For example, one might activate genes in different sequences or introduce topics in diverse orders to test the impact on learnability of the material. Or instead if the growing topology of a seizure spread suggests severity, one could use the presented analyses in order to determine which nodes when stimulated would most drastically disrupt the unfavorable topology. In general, we speculate that the reorderability of a complex system will reveal inherently new information about how the system responds to change.

D. Reorderablity derived from Betti curves versus barcodes in applications

As discussed in the previous sections, reorderability from the Betti curve perspective captures a different level of information about the system than reorderability from the barcode perspective. If a graph displays reorderability based on the Betti curves, then slice-by-slice the global reorderings must be similar in their topology. In contrast, if a graph displays reorderability based on the barcodes, then the entire topological evolution of each reordered filtration must be similar. Differences in perceived reorderability based on Betti curves versus barcodes mimic differences seen in cross-sectional versus longitudinal studies [89,90]. While we might tend toward desiring the latter, ofttimes in experiments we might find ourselves constrained to the former [91]. For example, if studying brain development we may perform a cross-sectional study instead of a longitudinal study due to the availability of participants or technology. With cross-sectional slices, one would generally be confined to studying reorderability from Betti curves. However, if we instead performed a longitudinal study, the presence of natural maps from one time point into the next would readily support a comparison of topology based on the barcode. Overall, our work underscores the importance of longitudinal investigations in unearthing information about a growing system.

E. Reorderability and dynamics

Loops in networks can play many roles including that of a pacemaker, limit cycle, and noise buffer, depending on the specific system dynamics. Then if the underlying structural connectivity graph of the system displays reorderability, we would expect the dynamics to evolve in a relatively stable manner despite perturbation. As discussed in detail elsewhere [92], one type of meaningful noise in biology arises from random fluctuations in time delays between nodes. Interestingly, the authors of Ref. [92] found that the feed-forward cycle motif itself offers a particular robustness to this type of noise that would result in small perturbations to node activity order. The authors of Ref. [93] similarly found that feed-forward motifs can decrease randomness, while Ref. [94] showed that loops can act as dynamic buffers against noise. Furthermore, Ref. [36] discusses the role of cycles as pacemakers in dynamics and particularly shows how long cycles contribute to sustained activity. If we restrict our system to a simple case which includes only one pacemaker loop, for example, then a reorderable graph would activate the pacemaker at the same time regardless of noise in the system, whereas a very nonreorderable graph could begin the pacemaker very early on, or very late, in response to noise that affects the node activation order. Although many previous studies involving dynamics on networks have investigated, in our topological terms, loops in dimension 1, we would expect the local reorderability of a graph to correlate strongly with the reliability of the network dynamics in response to small node order perturbations, such as those analyzed in Ref. [92]. From the perspective of synchronizability, previous studies have linked network homogeneity with synchronizability [95,96], but it remains to be shown how topological regularity based on the stability of a graph's persistent homology may also contribute to optimal synchronizability. In truth, the full relationship between graph reorderability and network dynamics remains to be understood and we propose the idea as an interesting avenue for future research.

F. Why might we design a reorderable system?

The results from the local-to-global studies showed that the local reorderability and global reorderability are surprisingly distinct properties of a growing system. This result raises the question of in what contexts we might wish to design a system that exhibits, for example, local but not global reorderability. One such context might be cell proliferation. Here, marked stochasticity exists at the gene transcription level, but only a large fluctuation in particular gene sets will alter a phenotype [97,98]. It is intuitively plausible that cell regulatory networks exhibit local reorderability to account for the small variations in gene expression but do not demonstrate global reorderability if the cell is differentiated, making cells difficult to reprogram [99]. Indeed, testing the reorderability of such systems is a potentially fruitful avenue for future research. On the engineering side, if we seek to design a growing system that performs a particular function from scratch and that also needs to display robustness to perturbation, the reorderability framework presented here may offer a plausible test of proper network function. Additionally further investigations into reorderability and specifically motifs that support or prevent reorderability may lead to better predictions of system responses to perturbation.

G. Open questions not addressed in this paper

Often research projects hatch more questions than they answer. We take this last opportunity to highlight a few additional open questions not addressed in the current work. First, the constant probability model, as discussed, grows into a random i.i.d. network that has a particular Betti curve signature [69]. In Fig. 5, we see the constant probability model Betti curves may follow a similar yet stretched pattern that is truncated by the number of nodes. One might ask whether one would continue to see increasing peaks of increasing dimensions as the number of nodes grows (see Ref. [70]), and generally if the persistent homology signatures seen here scale with graph size. Second, our examples and intuition suggest that many biological networks may be locally but not globally reorderable. We leave open the question of which, if any, biological or developmental networks indeed exhibit such properties. Third, in the current work we only consider the topological similarity graph, although we in fact recover a topological similarity simplicial complex in which k nodes form a simplex if any permutation of those k nodes does not alter the persistent homology. What more can one learn about the growing graph architecture from more thoroughly studying the simplicial complex formed by the topological similarity relation? Is it even true that permutations of nodes within a clique of T_n alter the persistent homology to a smaller degree than permutations, including nodes within and outside of the clique? Finally, we note that while we studied reorderability of growing graphs with respect to one node ordering, removing the base ordering and investigating the variability of the persistent homology generated from one binary graph and any random node order would reveal the topological regularity of a binary graph. That is, given a binary graph the persistent homology may vary greatly or insignificantly across random growth orders, suggesting a sameness in the topology across the binary graph as observed previously in a semantic feature network [31].

V. CONCLUSION

In this work, we explore the evolving topology of growing graph models and test the robustness of this evolving topology with respect to perturbations of node order. We find that both spatially embedded and nonembedded models can exhibit reorderability at a global level and that globally reorderable growing graphs may not be the most locally reorderable. We determine that reorderability at the pairwise level does not necessarily imply global reorderability but still can be used to determine more information about the growing graph architecture. Finally, we proffer suggestions for deepening the theory behind node-filtered networks and speculate that real-world systems will be found to display a range of reorderability properties based on system function.

APPENDIX

We organize this Appendix into two main sections. First we provide additional useful examples and supporting arguments for statements made in the main text. Second, we report the results from any additional experiments run on the growing graph models.

1. Intuition for the stability of barcodes in the node-filtered case

The original stability theorem for persistence diagrams (equivalently barcodes) [12] was expertly crafted in more general terms of weight functions on filtrations. In this work, we restrict ourselves to a very specific type of filtered simplicial complex, which is one generated by filling in cliques as simplices within a growing graph. This special case that we consider here lends itself to a simple interpretation of the stability theorem for node-filtered order complexes, which we include below. We wish to emphasize that the following section contains no novel results and only interprets the stability theorem from Ref. [12] from the perspective of node-filtered order complexes.

In the main text, we state that the stability theorem gives us an upper bound for the bottleneck distance between P_0 and $P_{i,j}$ of $|j - i|$. How do we see this upper bound in our growing graph models? If we grow the same binary network B, then recall that we only need to know the ordering s to specify the growing graph. In Fig. 16, we assume the same binary graph B and show only the original ordering of nodes, s_0 (top). Then when we swap nodes v_i and v_j, $i < j$, we obtain the $s_{i,j}$ ordering shown in the bottom row of Fig. 16. Notice that from the addition of the first node through node v_{i-1}, we do not change the nodes added or the order in which nodes are added, so the filtration from node v_1 through v_{i-1} must be equal between (B, s_0) and $(B, s_{i,j})$. Additionally, note that if we have added a subset of nodes $v \subseteq V$, the order in which nodes of v were added cannot change the subgraph of B induced by v. Therefore, the binary graph after the addition of v_i in $s_{i,j}$ must be equal to the binary graph after the addition of v_j in s_0 since at this point we have added all nodes $v_1, \ldots, v_j$, and furthermore the filtrations must match from the addition
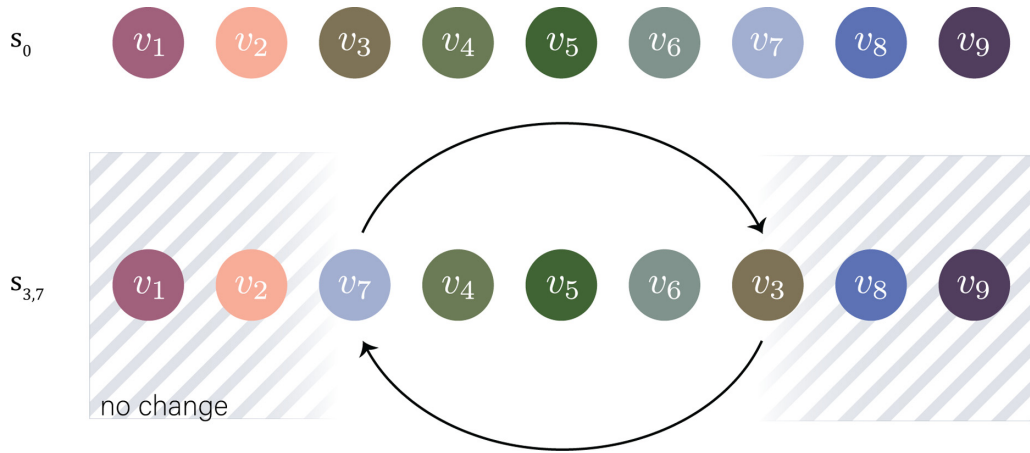
FIG. 16. A node swap defines the maximum extent to which the resulting persistent homology can change. When swapping the original order (top) of nodes v_3 and v_7 (bottom), the filtration from nodes v_1 to v_2 and the filtration from the addition of the seventh node to the end do not change. Thus, outside the swap bubble the persistent homology cannot be affected. It is only inside the swap bubble that the persistent homology could change in response to the node swap shown.

of the jth node on. We illustrate these regions of the filtration in which no changes to the filtration—and consequently to the persistent homology—could have occurred as gray-shaded regions of the orderings in Fig. 16.

Now, we can imagine a situation in which the addition of node v_i begins a cavity and that all nodes $v_{i+1}, \ldots, v_j$ do not participate in the cavity formation. So when we swap nodes v_i and v_j, then we move a persistent cavity birth time from i to j, and if no other persistent cavities emerge, then $j - i$ will be the bottleneck distance between barcodes generated by s_0 and $s_{i,j}$. Thus, we could swap nodes v_i and v_j and produce a barcode distance of $j - i$, but no more. To see these points directly from the definition given in Ref. [12], here our weight function on simplices is defined by the node order so that the weight of a node is the number of that node in the sequence, and the norm of the weight function on the simplicial complex is the same as the norm of the function

on the vertices. From this definition, the interpretation above follows.

2. Examples and counterexamples

Example 1: Cycles. While cycles composed of edges are often familiar, cycles, boundaries, and cavities of other dimensions are often less common. To expand upon our description of homology in the main text, here we present an example of one cavity-enclosing and one boundary n-cycle for $n = 0, 1, 2, 3$. Shown in Fig. 17, we see first an example of a 0-cycle that is not a boundary (two disconnected nodes) and a 0-cycle that is also a 0-boundary (two connected nodes). Next, we see the familiar cavity-enclosing and boundary 1-cycles in which the 1-simplices either surround a two-dimensional void or form a boundary of two 2-simplices. The octahedron is an example cavity-enclosing 2-cycle, but filling the interior with 3-simplices then offers an example of a boundary
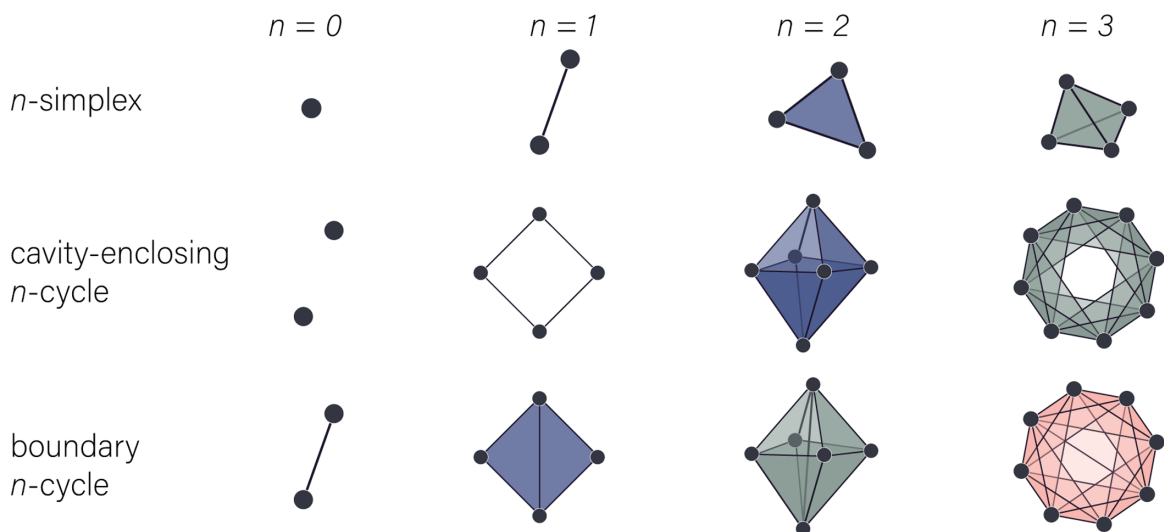

FIG. 17. Examples of n-cycles. Simplices (top row), example of a cavity-enclosing n-cycle (middle row), and an example of a boundary n-cycle (bottom row) for small values of n.
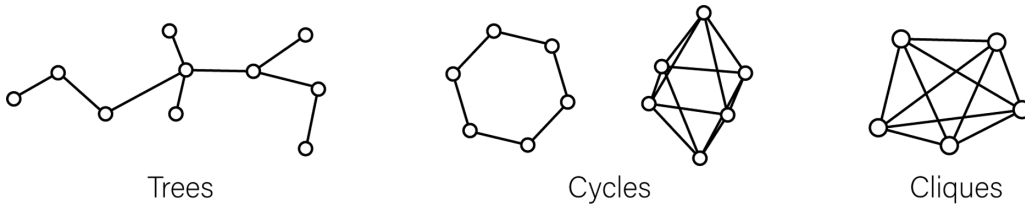
FIG. 18. Examples of completely reorderable graphs. If we assume that we will form a filtered simplicial complex via the n-order complex construction, then the tree graph (left) is completely reorderable in dimensions $\geqslant 1$, a minimal n-cycle (middle) is completely reorderable in dimensions $\geqslant n$, and cliques (right) are totally reorderable in all dimensions.

2-cycle. Finally, we illustrate a projection of a cavity-enclosing 3-cycle which surrounds a four-dimensional void with a shell of 3-simplices, and a 3-cycle that is also a boundary in which the shell of 3-simplices surrounds a collection of 4-simplices (peach).

Example 2: Totally reorderable graphs. In Fig. 1, we see the transposition of nodes d and f change the persistent homology of the growing graph, and in the main text we see that the reorderability of a growing graph depends on the topology of the binary graph B. In fact, some special binary graphs exist that will yield the same persistent homology in a particular dimension for *any* ordering of nodes. First, if B is a tree (Fig. 18, left) equipped with some node ordering $s : V \rightarrow \mathbb{N}$ then there will be no homology in dimension n for $n \geqslant 1$ along the entire filtration. Indeed, any node addition order will yield the same persistent homology (here this means no persistent cavities emerge) in dimension n for $n \geqslant 1$. Second, any minimal n-cycle (specifically an $(n + 1)$-cross polytope; see Ref. [40]) such as the minimal 1-cycle or minimal 2-cycle shown in Fig. 18 (we assume that we always create the clique complex in this case) in which all nodes and n-simplices are required for the cycle will produce the same persistent homology regardless of node order for dimensions $\geqslant n$. Finally, an n-clique in B as shown on the right of Fig. 18 (again assuming that we take the clique complex) will produce the same persistent homology in all dimensions regardless of node order. These examples help us to understand the sort of topological homogeneity necessary for a binary network B to exhibit complete reorderability.

The above totally reorderable graphs also offer examples of how network statistics may or may not be preserved across the growing process in totally reorderable graphs. For example, if we grow the tree graph with a node order assigned left to right, and then we calculated the average degree after each node addition (creating an average degree curve), we could find an alternative node ordering in which the average degree curve would differ. Similarly, we can grow the octahedron in Fig. 18 in node orders which would change the clustering coefficient curve. These examples demonstrate the differences between preserving the topology of growing processes and preserving network statistic curves.

Example 3: Rules of permuting nodes within a k-clique in T_n. In the main text, we stated that while finding a 3-clique in T_n implies that any permutation of the nodes involved in the clique would produce the same Betti curves in dimension n as the original ordering, and additionally that this fact does not hold for all permutations of nodes within a 4-clique of T_n. Here we illustrate a rule for permutations on

m nodes of an m-clique in T_n that will guarantee conservation of β_n.

Let $\{v_{i_1}, \ldots, v_{i_m}\}$ be the nodes of an m-clique found in T_n with $i_j < i_{j+1}$. Note that in the original ordering s_0, $v_{i_{j+1}}$ may not follow directly after v_{i_j}, but for the following proof we can ignore any nodes added between each $v_{i_j}, v_{i_{j+1}}$ as they do not affect the outcome. We only require that $i_{j+1} > i_j$ so that in the original ordering s_0, $v_{i_{j+1}}$ comes after v_{i_j}. We can also ignore any nodes added before v_{i_1} or after v_{i_m} in s_0 as they will not cause a change in the Betti curves (see Fig. 16). For the sake of clarity, we can rewrite our nodes of the m-clique as $v_1, v_2, \ldots, v_m$ and only focus on the following portion of the graph filtration $\cdots \rightarrow G_{v_1} \rightarrow \cdots \rightarrow G_{v_m} \rightarrow \cdots$ since each step is an induced subset of the ending binary graph B on all nodes added up to and including that step. We call a permutation σ of nodes $v_1, \ldots, v_m$ *admissible* if the permutation s_σ of these nodes with respect to the original ordering s_0 results in a growing graph (B, s_σ) with the equal β_n to that of (B, s_0). Based on our cut-and-paste strategy illustrated in the main text, we propose the following rule for permutations σ that if met, will show that the permutation σ of nodes in an m-clique in T_n is admissible.

Rule 1. Let $v_1, \ldots, v_m$ form an m-clique in T_n and v_k denote the first k nodes in a reordering of $v_1, \ldots, v_m$. Then a reordering of $v_1, \ldots, v_m$ is admissible if for at least $(k - 1)$ nodes $v_i \in v_k, i \leqslant k$, for all $k = 2, \ldots, m$.

Informally, this rule states that if we take the first two nodes in the reordering, check that at least v_1 or v_2 is present, then move to the first three nodes and check that at least two of v_1, v_2, and v_3 are present, and so on, then the permutation will be admissible.

In order to prove that an ordering following Rule 1 is indeed admissible, we fix the number of nodes m and induct on k, so that for each k we prove the modified rule:

Rule$_k$. Let $v_1, \ldots, v_m$ form an m-clique in T_n and v_r denote the first r nodes in the reordering. The reordering is admissible up to the kth node if for at least $(r - 1)$ nodes we have $v_i \in v_r, i \leqslant r$, for all $r = 2, \ldots, k$.

First, we show that the rule holds for $k = 2$. Let $2 < q \leqslant m$. Then if $v_2 = (v_1, v_2)$, our filtration through the addition of v_2 is equal to that from s_0. If $v_2 = (v_1, v_q)$ or $v_2 = (v_q, v_2)$, then since v_1, v_2, v_q are connected in T_n, β_n must remain unchanged. Finally, if $v_2 = (v_q, v_1)$ or $v_2 = (v_2, v_q)$, then using the cut-and-paste method we can show up through the second node added that the reordering will not change β_n. We show the $v_2 = (v_q, v_1)$ case below in Fig. 19(a) to illustrate that by using the topological similarity of the v_1, v_q pair and the v_2, v_q pair, we can piecewise construct the relevant parts of
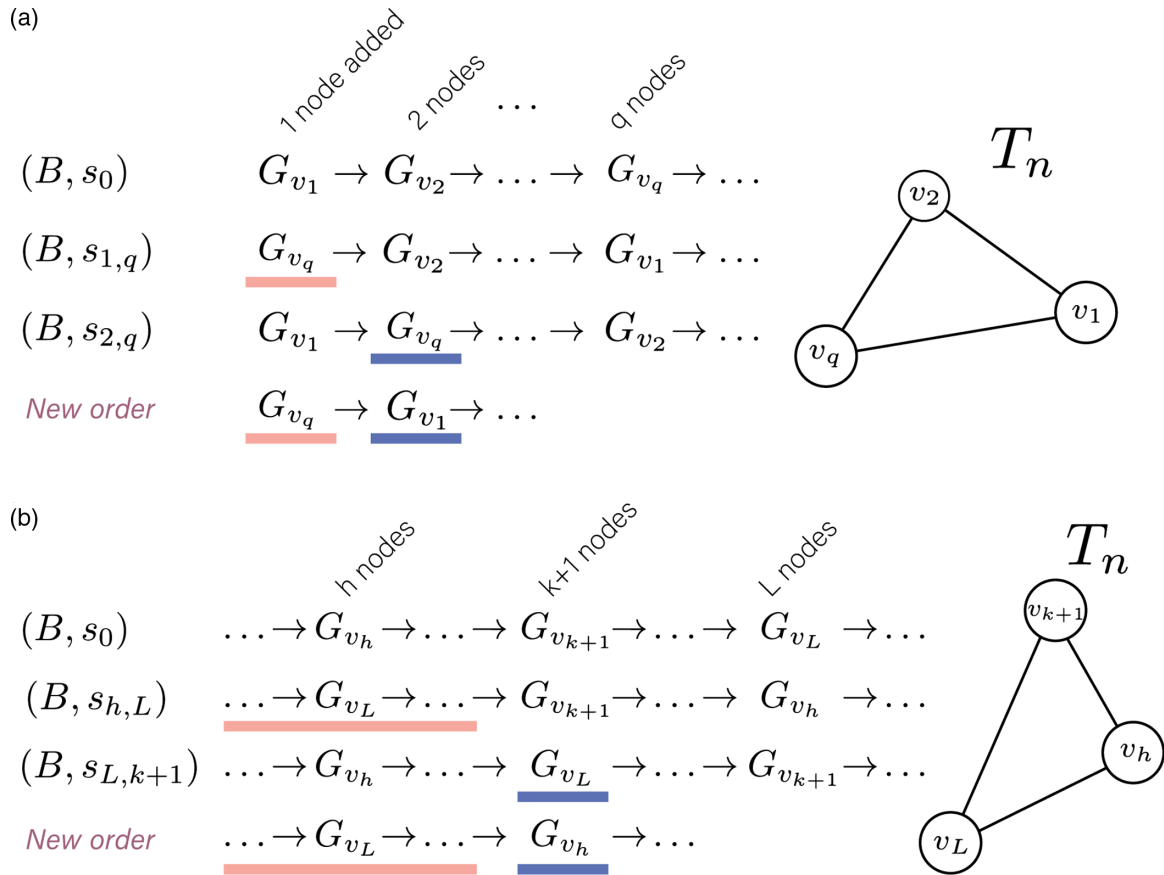
(a)

(b)

FIG. 19. Examining node swaps that conserve Betti curves. (a) Illustration of using the cut-and-paste process to show up through the first two nodes of the reordering that β_n will be unchanged. (b) Illustration detailing checking that β_n is conserved through the first $k+1$ nodes as described in Case 2(a).

the new graph filtration. In particular, the filtration up through adding v_q is the same as from $(B, s_{1,q})$ (peach) and adding v_1 in the new ordering will create the same binary graph as in $(B, s_{2,q})$ (blue) and thus must have conserved the Betti curve for at least the first two nodes added. We leave the $v_2 = (v_2, v_q)$ for the reader.

Generally we see that for an admissible ordering, we will need to use no swaps, one swap, or two swaps with the cut-and-paste process.

Now we need to show that given a reordering in which the first k nodes of the reordering satisfy Rule$_k$; if Rule$_{k+1}$ is satisfied then the reordering up through the first $(k+1)$ nodes

will be admissible. Assume the first k nodes in the reordering follow Rule$_k$ so that for $v_i \in v_r$, we have $i \leqslant r$ for at least $(r-1)$ of the nodes in v_r, for all $r = 2, \ldots, k$. We need to consider the possible nodes that might come next, at the $k+1$ spot. We break this portion into two cases. As for the $k=2$ case, we will need to use the topological similarity of at most two node pairs.

Case 1. For all k nodes $v_i \in v_k$, $i \leqslant k$. Then at the addition of the kth node in the reordering, we have the graph induced by the subset of nodes in v_k, which is equal to the graph after the kth node is added in the original ordering s_0. Let the next node added be v_L, and then $L \geqslant k+1$. Since v_L is
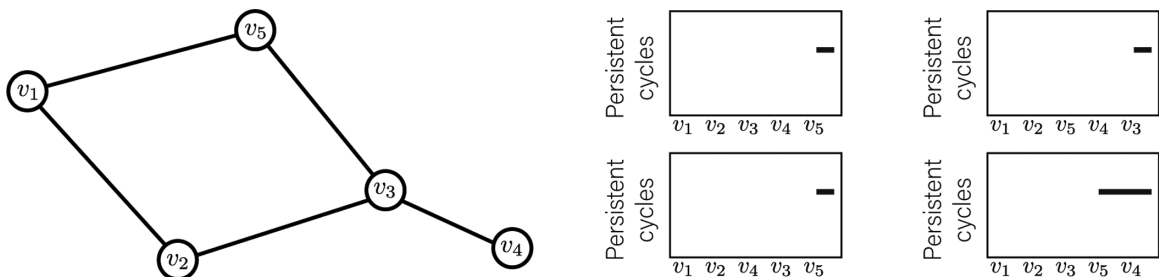

FIG. 20. Topological similarity is not a transitive relation. The v_3 node is topologically similar to nodes v_4 and v_5, but v_4 and v_5 are not topologically similar as illustrated by the barcodes on the right.
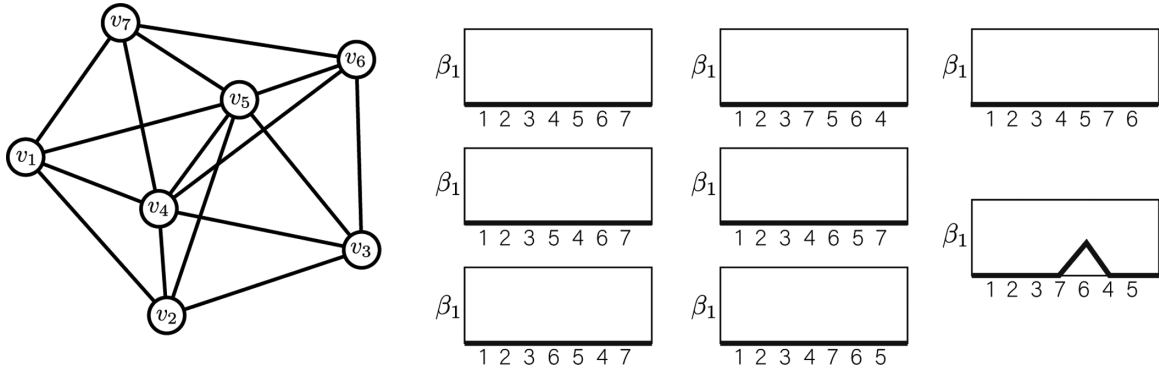
FIG. 21. Swapping nodes from a 4-clique in T_n does not necessarily conserve Betti curves. Nodes v_4, v_5, v_6, and v_7 are all pairwise topologically similar and all swaps of no more than three of these four nodes will not alter the Betti curves (right). However, the final ordering in which nodes v_7 and v_6 precede nodes v_4 and v_5 produces a different β_1 than the original ordering.

topologically similar to v_{k+1} by the assumption, the ordering $s_{L,k+1}$ yields the same β_n, and we are done.

Case 2. Exactly one of the $v_i \in v_k$ has $i \geqslant k + 1$; call this node v_L. Then exactly one node v_i with $i \leqslant k$ will be added later (after the kth node); call this node v_h.

Case 2(a). Let $L \geqslant k + 2$. Then Rule$_{k+1}$ requires the $k + 1$ node to be v_h. As illustrated in Fig. 19(b), we use the topological similarity of the v_L, v_h pair and the v_{k+1}, v_L pair to construct the binary graphs in the reordered filtration. Therefore, our reordered growing graph has the same β_n as the original up through the first $(k + 1)$ nodes.

Case 2(b). Let $L = k + 1$. Then the $k + 1$ node could be any node, as we will still have that for k of the $v_i \in v_{k+1}$, $i \leqslant k + 1$. Denote the $k + 1$ node by v_M. If $M = h$, then we are done as we only need to map into the original filtration. If $M \neq h$, then the reordered filtration will have equal β_n up

through the $(k + 1)$ node to that emerging from the original ordering due to the topological similarity of the v_h, v_L pair and the v_h, v_M pair.

This completes the induction step and therefore we can conclude that if the reordering of the m nodes of an m-clique in T_n adheres to Rule 1, then β_n will remain unchanged.

Counterexample 1: Nontransitivity of the topological similarity relation. Here we provide details proving that the binary topological similarity relation between node pairs is a dependency relation as opposed to an equivalence relation. Recall that we defined two nodes v_i and v_j in a single growing graph (B, s_0) as topologically similar if $d_{BN}^n(P_0, P_{i,j}) = 0$, and for this section we write $v_i \equiv_r v_j$ if this is so. First, this definition of topological similarity is trivially reflexive and symmetric. To show that topological similarity is not transitive, consider the growing graph in Fig. 20 in which $v_3 \equiv_r v_4$ and $v_3 \equiv_r v_5$.
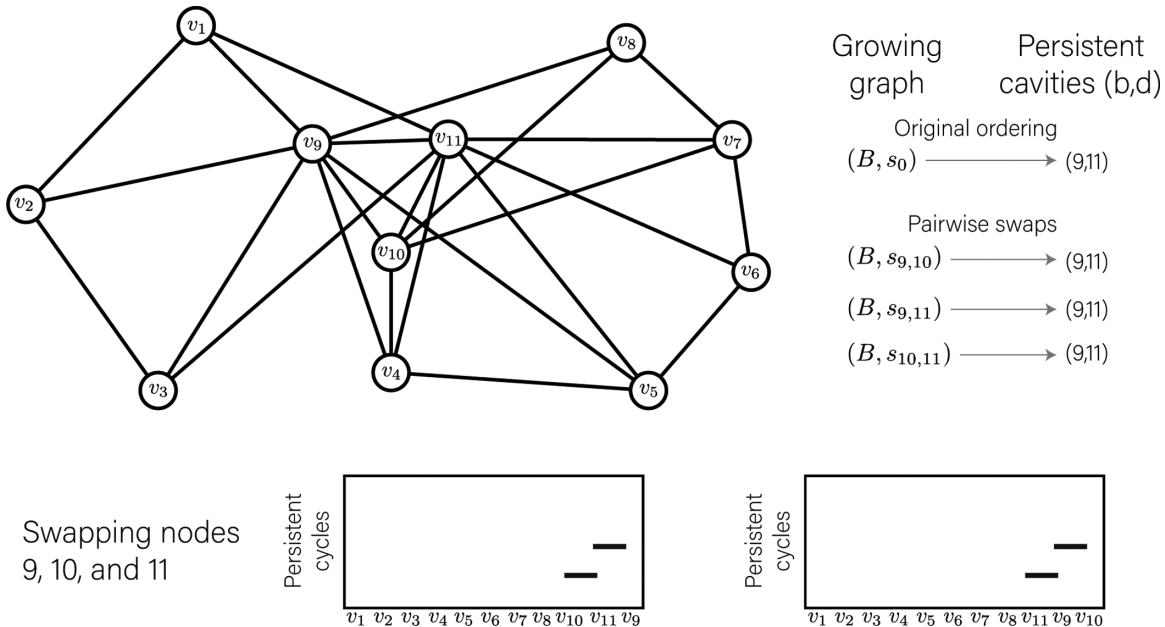

FIG. 22. Not all permutations of nodes within a 3-clique of T_n conserve the persistent homology. Nodes v_9, v_{10}, and v_{11} are all pairwise topologically similar as shown by the persistent homology description (right). However, the permutations of these three nodes (shown in the bottom barcodes) produces different barcodes than those produced from the original ordering.
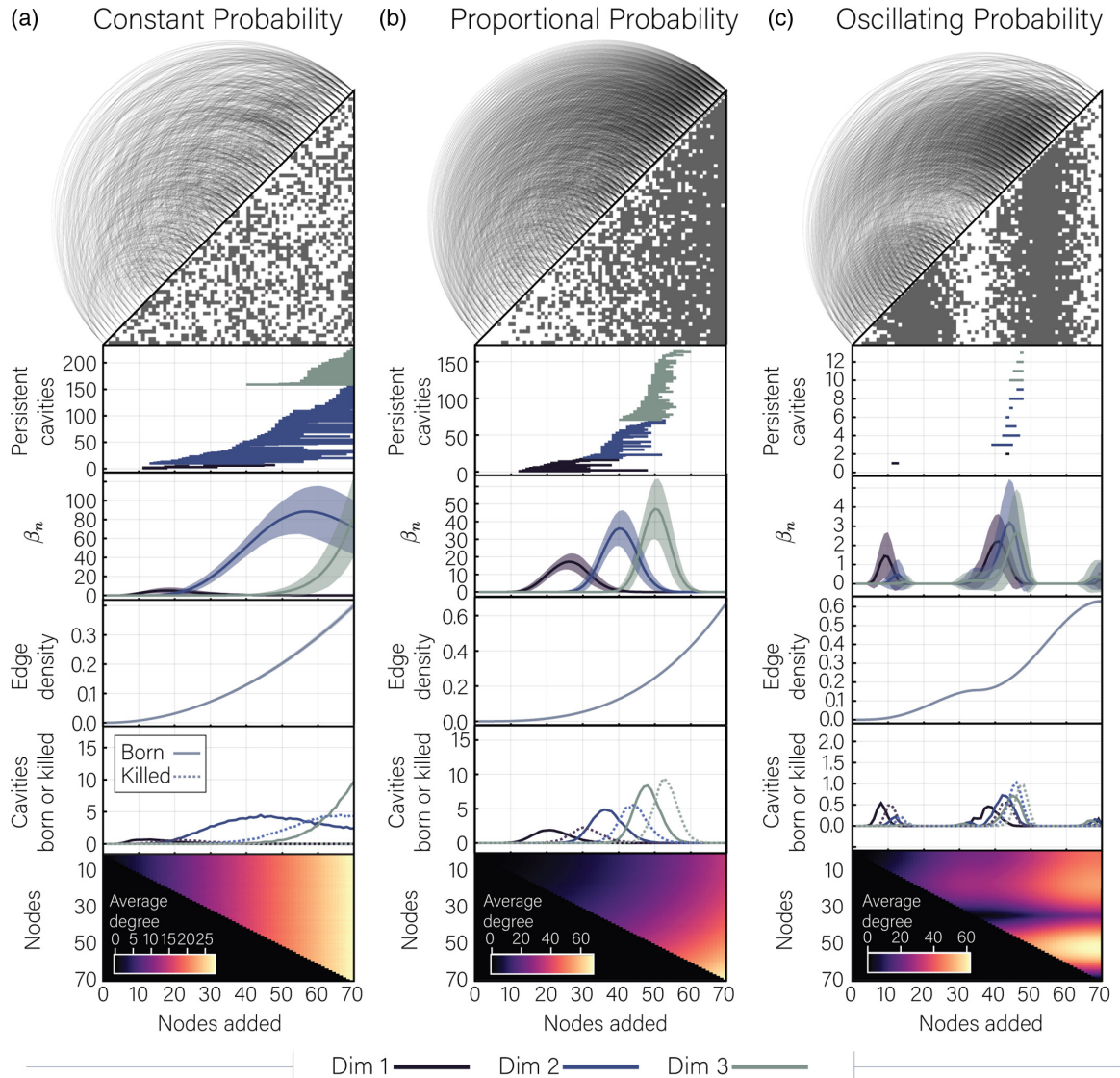
FIG. 23. Extended descriptions of the growing graph models: Part I. Here we provide extended descriptions for the (a) constant probability, (b) proportional probability, and (c) oscillating probability growing graph models. From top to bottom: fingerprint graph and binary adjacency matrix, barcode plot of one example growing graph, average Betti curves, edge density of the growing graph with respect to the final node count, plot showing the number of cavities born or killed at each node addition, and heat map of evolving average degree for each node across the filtration. Results for dimension 1 shown in purple, dimension 2 shown in blue, and dimension 3 shown in green.

Indeed, we see that the $v_4 \leftrightarrows v_5$ swap does not yield the same persistent homology as the original ordering, and so $v_4 \not\equiv_r v_5$. Thus, topological similarity is not transitive, and therefore is also not an equivalence relation.

Counterexample 2: Not all swaps between four topologically similar nodes yield the same Betti curves. In the main text, we consider how permutations of the k nodes found in a k-clique in T_n would affect the persistent homology. Here we provide a growing graph in which four nodes are all pairwise topologically similar, but a permutation of all four nodes does not result in the same persistent homology. Specifically, in the graph above (Fig. 21), one can show that nodes v_4, v_5, v_6, and v_7 form a 4-clique in T_1. As a consequence, all pairwise swaps of these nodes yield the same barcode as the original binary graph (Fig. 21, right), which shows no persistent cavities. However, we see that if we perform a reordering such that

we add these four nodes in order v_7, v_6, v_4, and then v_5, β_1 has changed as now we see a peak at the addition of the fifth node (v_6).

Counterexample 3: Not all swaps between three topologically similar nodes yield the same persistent homology. Continuing our investigation into how a clique in T_n may determine the persistent homology after permutation of nodes within the clique, we now turn to observing the effect of such a permutation on the barcode generated by a growing graph. We see in the above 11-node graph (Fig. 22, left) that v_9, v_{10}, and v_{11} are pairwise topologically similar, as swapping any pair results in the same barcode (Fig. 22, right). However, if these three nodes emerge in the network either in order v_{10}, v_{11}, v_9 or in order v_{11}, v_9, v_{10}, then we observe two bars of dimension 1 in the corresponding barcode (Fig. 22, bottom). Thus, we find that permutations of nodes' participation in the
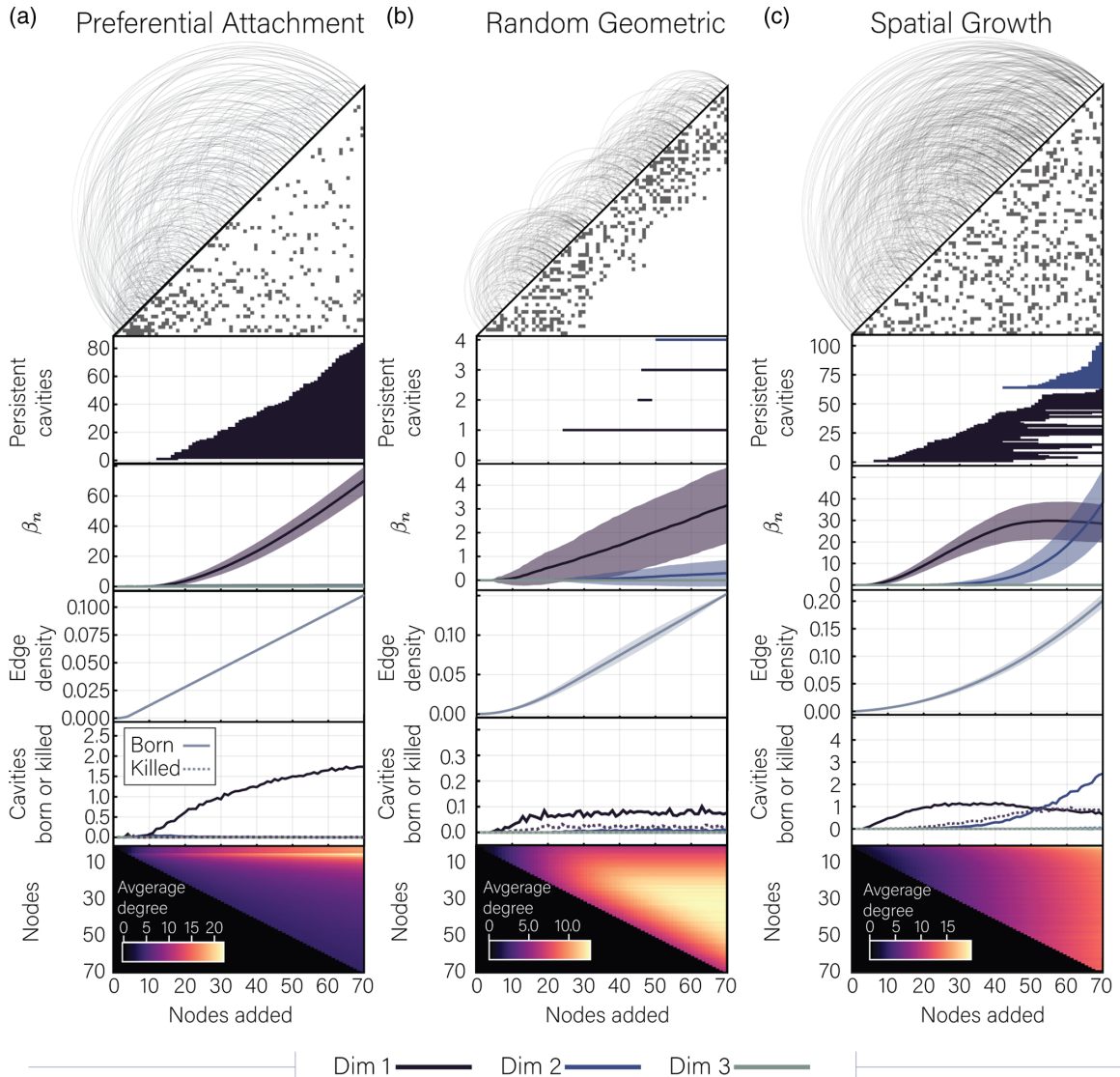
FIG. 24. Extended descriptions of the growing graph models: Part II. Here we provide extended descriptions for the (a) preferential attachment, (b) random geometric, and (c) spatial growth growing graph models. From top to bottom: fingerprint graph and binary adjacency matrix, barcode plot of one example growing graph, average Betti curves, edge density of the growing graph with respect to the final node count, plot showing the number of cavities born or killed at each node addition, and heat map of evolving average degree for each node across the filtration. Results for dimension 1 shown in purple, dimension 2 shown in blue, and dimension 3 shown in green.

same clique of three nodes or more in T_n does not necessarily preserve the persistent homology.

3. Meet the team, extended

In this section, we include additional information about each of the six growing graph models defined and studied in the main text. We show results for the constant probability, proportional probability, and oscillating probability models in Fig. 23 and for the preferential attachment, random geometric, and spatial growth models in Fig. 24. We include the fingerprint graphs, adjacency matrix images, barcodes, and Betti curves from Fig. 5, and we also add three more panels for each model. The first added panel records the average edge density of the growing graph after each node addition. Importantly, we compute the edge density at each step relative to the size of the final graph. That is, the edge density after the addition of

v_i is $\rho(i) =$ edges added up through $v_i / \binom{N}{2}$ with $N = 70$. The second additional panel shows the number of persistent cavities born (solid lines) or killed (dashed lines) after each node addition averaged across replicates. The final panel (last row in Figs. 23 and 24) shows a heat map of the evolving average degree of each node as the graph grows. More specifically, for each node v_i we calculate the mean of the degree of v_i after the addition of node v_j across replicates. Repeating for all $i, j = 1, \ldots, 70$, we recover the presented heat map. For all plots, the horizontal axis follows graph growth.

4. Additional experiments for investigating global and local reorderability

After determining the reorderability of growing graph models, we naturally ask what types of changes to the barcodes or Betti curves lead to the distribution of distances
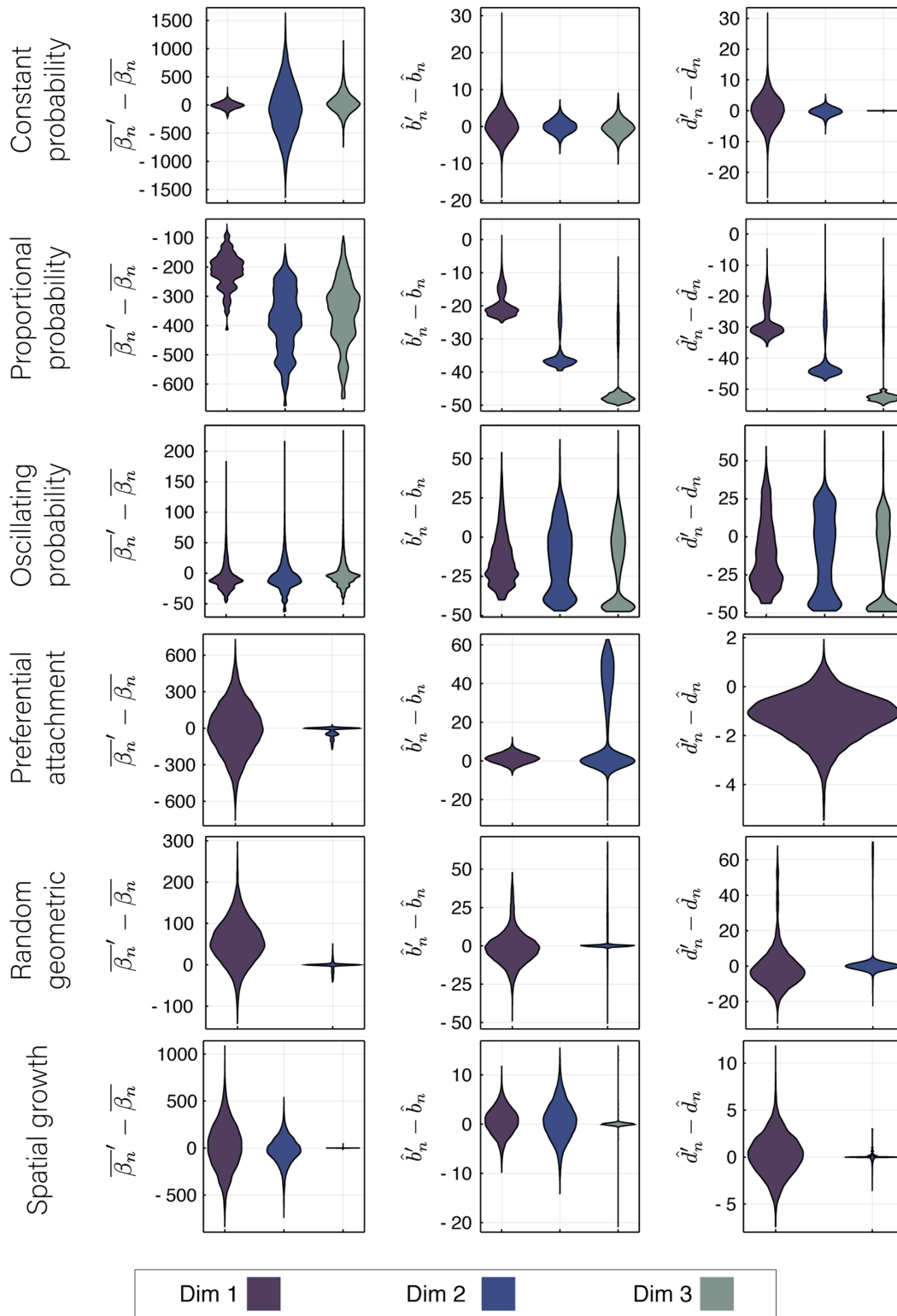
FIG. 25. Differences among Betti bar values, average birth, and average death times between the reordered and original growing graph persistent homology. (Left) Distributions of $\overline{\beta_n}' - \overline{\beta_n}$ values, (middle) distributions of $\hat{b}_n' - \hat{b}_n$ values, and (right) distributions of $\hat{d}_n' - \hat{d}_n$ values for each of the six growing graph models.

shown in Figs. 8 and 9. Recall that $\overline{\beta_n}$ is the sum of all persistent cavity lifetimes. We first ask how $\overline{\beta_n}$ differs between the originally generated growing graphs and their associated reordered growing graphs. We show these distributions in the

left column of Fig. 25. Specifically if $\overline{\beta_n}'$ is the Betti bar value for a reordered growing graph and $\overline{\beta_n}$ is for the original, we plot the distribution of $\overline{\beta_n}' - \overline{\beta_n}$ values. These plots help us to understand that the change in Betti curves of, for example, the
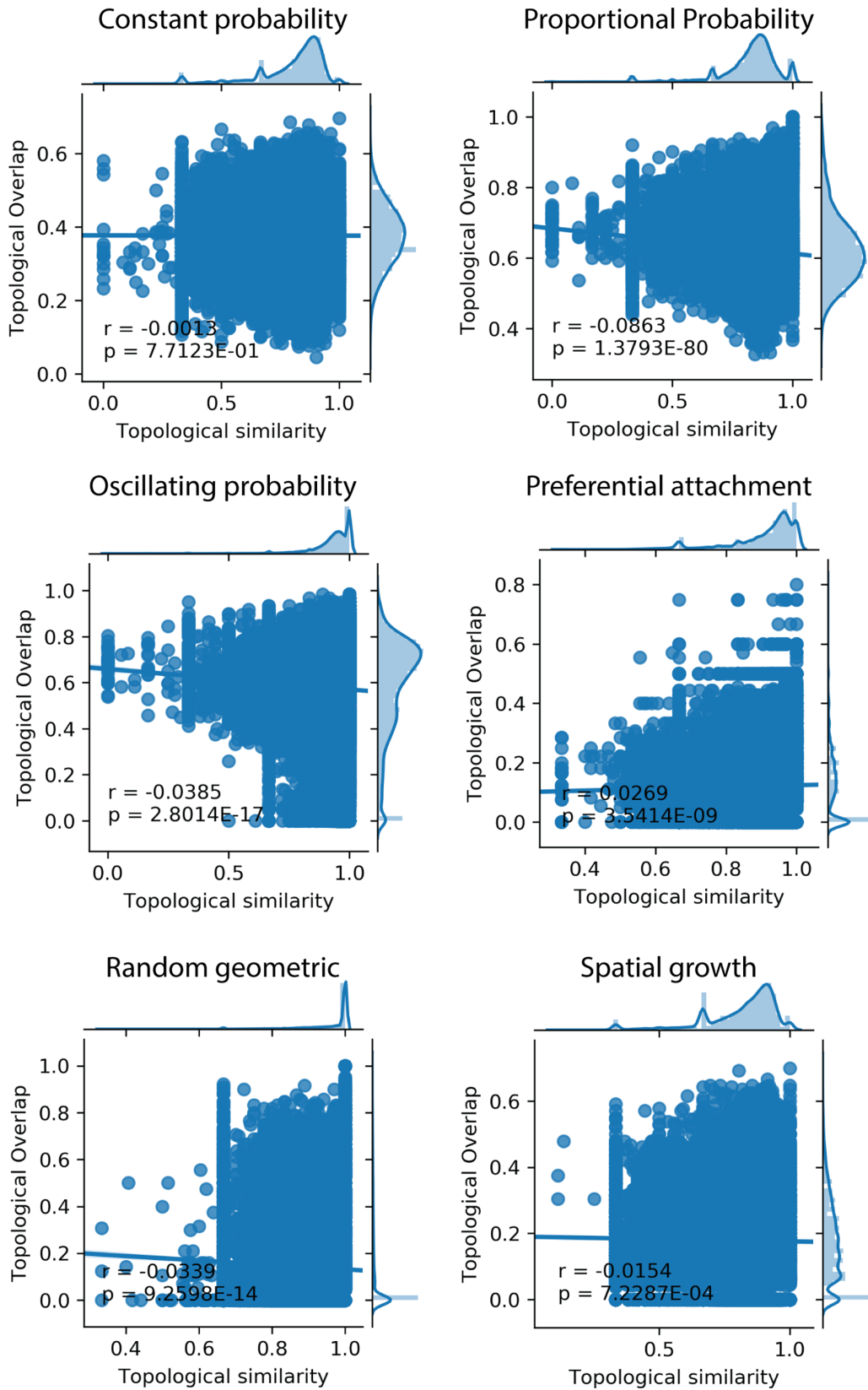
FIG. 26. Scatterplots and distributions of the averaged topological similarity and averaged topological overlap between node pairs. Plots of topological similarity averaged over dimension between node pairs vs their computed topological overlap, calculated for each node pair in each growing graph model replicate. Line of best fit overlaid and Pearson correlation coefficient (r) and associated p-value p displayed.
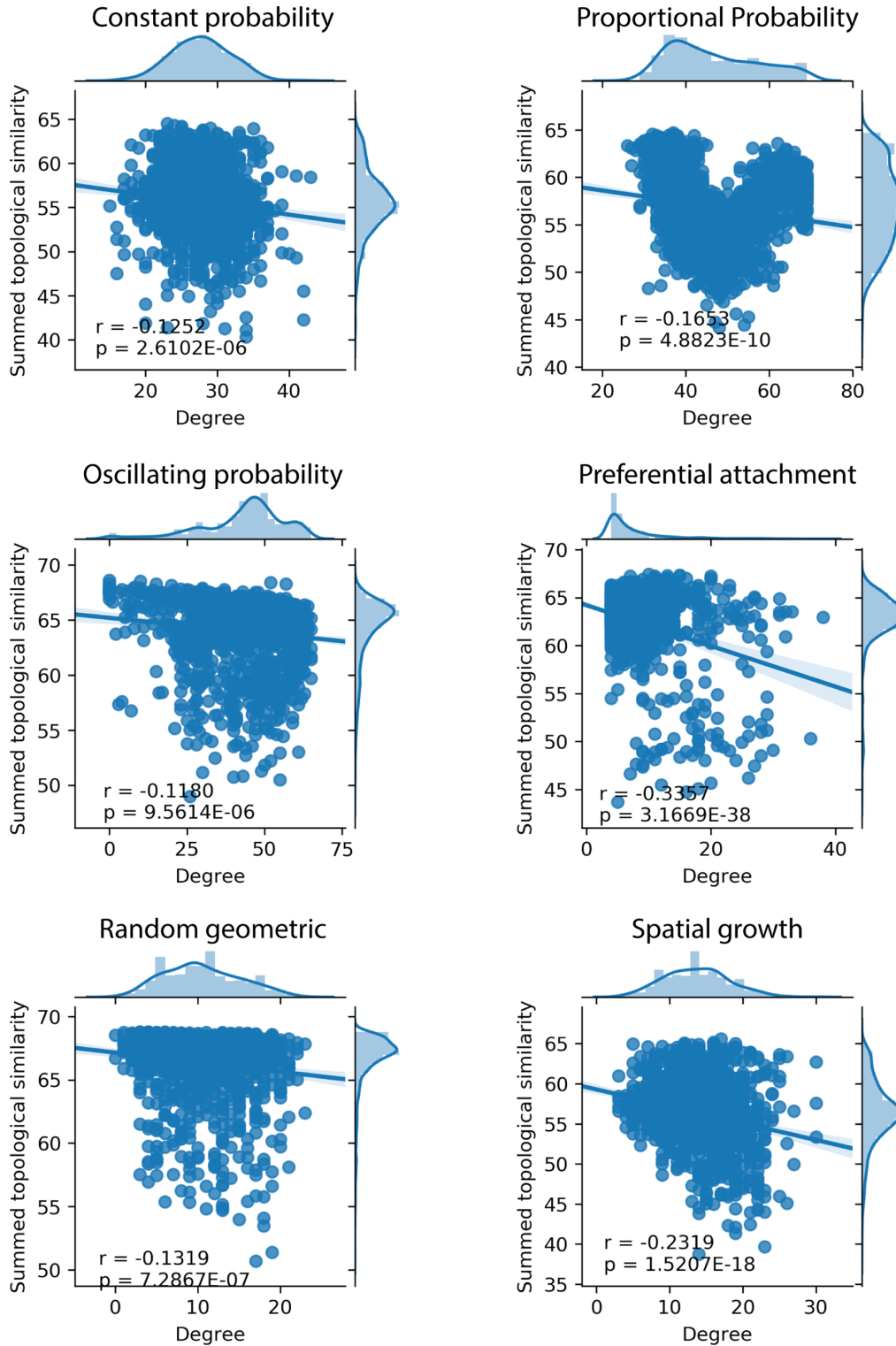
FIG. 27. Scatterplots and distributions of degree and summed node averaged topological similarity. Plots of node degree vs summed topological similarity with line of best fit overlayed and marginal distributions for each of the six growing graph models. Additionally, the Pearson correlation coefficient r and associated p value are included.

globally reordered proportional probability graphs occurs in a manner that also decreases the value of $\overline{\beta_n}$ on average.

Additionally, we can perform the same calculation but for the average birth and death times of persistent cavities. Specifically, we calculate the average birth (death) time of all

persistent cavities of dimension n for the originally ordered growing graph $\hat{b}_n$ ($\hat{d}_n$), and we subtract this value from the average birth (death) time calculated from a random reordering of that growing graph $\hat{b}'_n$ ($\hat{d}'_n$). We repeat for all replicates and all reorderings and show the distributions as violin plots
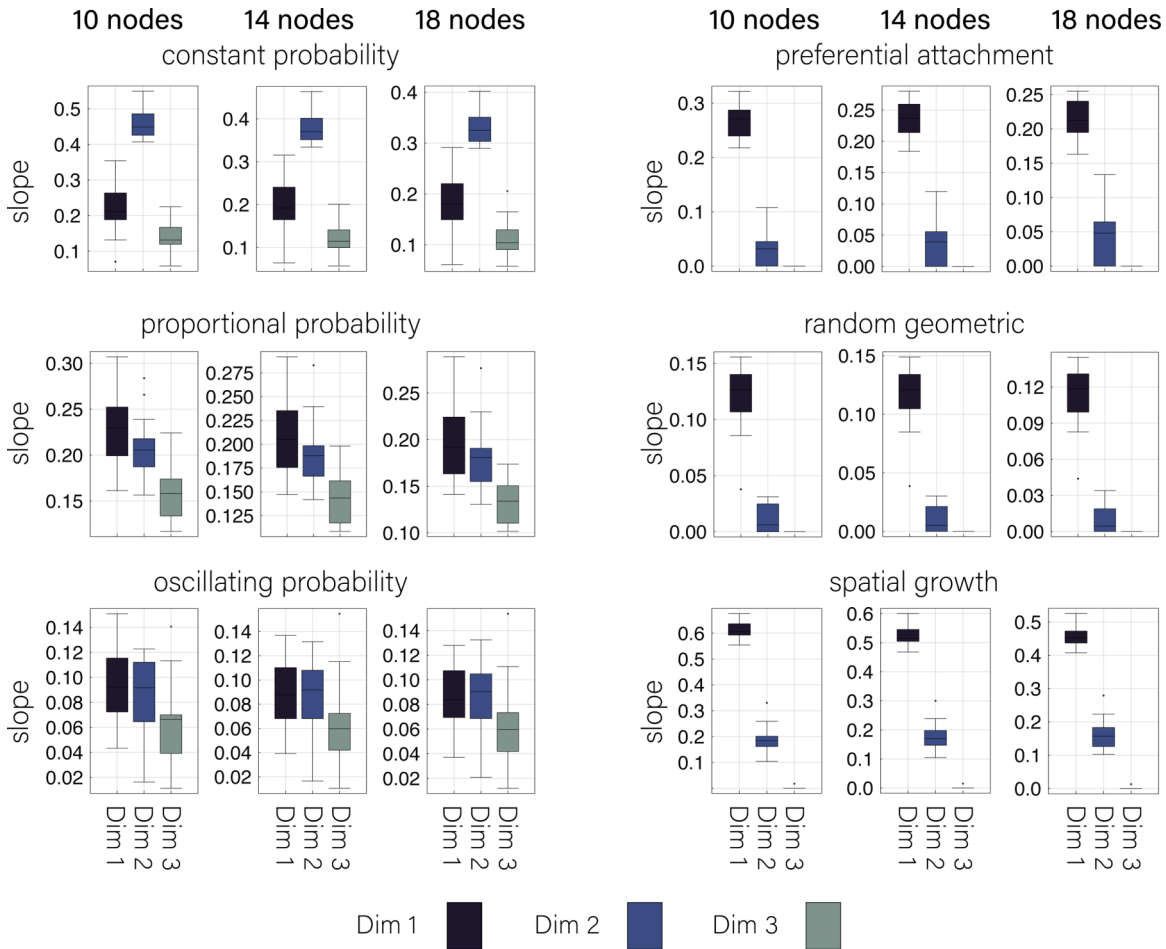
FIG. 28. Effect of changing the threshold in local reorderability calculations. Each box plot shows the distribution of calculated slopes as described in Fig. 11, colored by dimension. For each graph model, the left column shows results for swap magnitudes ⩽10 nodes, the middle column shows results for swap magnitudes ⩽14 nodes (repeated from Figs. 12 and 13), and the right column shows results for swap magnitudes ⩽18 nodes.

in the middle and right columns of Fig. 25. As an example, for the preferential attachment model in dimension 1, these plots reveal that after reordering on average we decrease the average death time of persistent cavities and often increase average birth times.

When investigating how the persistent homology of growing graph models changes in response to node swaps, we intuitively expect node pairs with similar connection patterns to be the most topologically similar. In order to test this intuition, we compute the average topological overlap $O(v_i, v_j)$ of each node pair across binary networks generated from the growing graph model, and we plot this value against topological similarity averaged across dimensions in Fig. 26. For each growing graph model, we also show the calculated Pearson correlation r and associated p value along with the line of best fit; on the top and right axes we plot the marginal distributions of each variable. We see in the topological overlap versus topological similarity scatterplots (Fig. 26) that for most growing graph models, topological overlap is not a strong predictor of topological similarity.

We note that for both the preferential attachment and proportional probability models, in Fig. 15 we observe by eye that the degree of a node in B looks to affect the overall

topological similarity of that node. In order to investigate this potential relation further, we calculate the degree of each v_i in each B_η and plot this value against the summed topological similarity (averaged over dimensions) of node v_i for that replicate (Fig. 27). Note that this statistic is equivalent to the strength or weighted degree of node v_i in the averaged topological similarity graph constructed for each of the 20 replicates. For each growing graph model, we also include the calculated Pearson correlation r and p value along with the line of best fit; on the top and right axes, we plot the marginal distributions of each variable. We find that for most graph models, the average degree of a node poorly predicts the ability of that node to swap with all other nodes while preserving the persistent homology. Interestingly, the proportional probability model shows a clear U-shaped distribution of points suggesting that the lowest degree nodes and highest degree nodes can often swap with other nodes without changing the persistent homology much (relative to all nodes), while it is the midrange degree nodes that are the least swappable.

Finally, we examine the robustness of our local reorderability slope calculations to our choice of 20% as a threshold (Fig. 28). Recall that for Figs. 12 and 13, we calculated the bottleneck distance between the barcodes generated from the

original and node i, j-swapped orderings. For each replicate, we calculate a line of best fit for points involving swaps with magnitude $\leqslant 14$ and show the distribution of the resulting slopes in box plots in the right column of Figs. 12 and 13. We chose a swap magnitude ceiling of 14 (in other words a 20% threshold) by examining the data and observing that most models exhibit a near-linear increase between the two

variables for swap magnitudes between 0 and 14. Still, it remains important to determine how our calculated slopes might change if we alter the upper limit to 18 or lower it to 10. Below we see that the calculated slopes change little, suggesting that the specific choice of threshold for calculating the line of best fit has little effect on the local reorderability conclusions presented in the main text.

[1] A. Arenas, A. Diaz-Guilera, and C. J. Perez-Vicente, Phys. Rev. Lett. **96**, 114102 (2006).

[2] S. Karthika, T. Radhakrishnan, and P. Kalaichelvi, Cryst. Growth Des. **16**, 6663 (2016).

[3] C. Rodriguez-Navarro, E. Ruiz-Agudo, J. Harris, and S. E. Wolf, J. Struct. Biol. **196**, 260 (2016).

[4] A. N. Khambhati, K. A. Davis, B. S. Oommen, S. H. Chen, T. H. Lucas, B. Litt, and D. S. Bassett, PLoS Comput. Biol. **11**, e1004608 (2015).

[5] M. X. Henderson, E. J. Cornblath, B. Z. Adam Darwich, H. Brown, R. J. Gathagan, R. M. Sandler, D. S. Bassett, J. Q. Trojanowski, and V. M. Lee, Nat. Neurosci. **22**, 1248 (2019).

[6] A. Raj, A. Kuceyeski, and M. Weiner, Neuron **73**, 1204 (2012).

[7] P. M. Chaikin, T. C. Lubensky, and T. A. Witten, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, UK, 1995), Vol. 1.

[8] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse, Nat. Phys. **6**, 888 (2010).

[9] G. Carlsson, Bull. Am. Math. Soc. **46**, 255 (2009).

[10] G. Carlsson, G. Singh, and A. Zomorodian, in *Proceedings of the International Symposium on Algorithms and Computation* (Springer, Berlin, 2009), pp. 730–739.

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

[12] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, Discrete Comput. Geom. **37**, 103 (2007).

[13] M. Berry, A. Rogers, and J. Eayrs, Nature (London) **203**, 591 (1964).

[14] H.-H. Yu, C.-F. Kao, Y. He, P. Ding, J.-C. Kao, and T. Lee, PLoS Biol. **8**, e1000461 (2010).

[15] F. L. Gomes, G. Zhang, F. Carbonell, J. A. Correa, W. A. Harris, B. D. Simons, and M. Cayouette, Development **138**, 227 (2011).

[16] S. Kienle, N. Gritti, J. R. Kroll, A. Kriselj, Y. Goos, and J. S. van Zon, bioRxiv, 588418 (2019).

[17] J. He, G. Zhang, A. D. Almeida, M. Cayouette, B. D. Simons, and W. A. Harris, Neuron **75**, 786 (2012).

[18] B. J. Pearson and C. Q. Doe, Annu. Rev. Cell Dev. Biol. **20**, 619 (2004).

[19] J. Jacob, C. Maurange, and A. P. Gould, Development **135**, 3481 (2008).

[20] S. Varier and M. Kaiser, PLoS Comput. Biol. **7**, e1001044 (2011).

[21] D. J. Aldous, J. Multivariate Anal. **11**, 581 (1981).

[22] D. J. Aldous, in *École d'Été de Probabilités de Saint-Flour XIII-1983* (Springer, Berlin, 1985), pp. 1–198.

[23] H. Crane and W. Dempsey, arXiv:1603.04571.

[24] D. N. Hoover, Institute for Advanced Study, Princeton, preprint, 1979 (unpublished).

[25] O. Kallenberg, J. Theor. Prob. **3**, 81 (1990).

[26] P. Orbanz and D. M. Roy, IEEE Trans. Pattern Anal. Mach. Intell. **37**, 437 (2014).

[27] C. Borgs, J. T. Chayes, H. Cohn, and N. Holden, J. Mach. Learn. Res. **18**, 7740 (2017).

[28] L. Lovász and B. Szegedy, J. Comb. Theory, Ser. B **96**, 933 (2006).

[29] D. Cai, T. Campbell, and T. Broderick, in *Advances in Neural Information Processing Systems 29 (NIPS)* (Curran Associates, Inc., 2016), pp. 4249–4257.

[30] J. Lloyd, P. Orbanz, Z. Ghahramani, and D. M. Roy, in *Advances in Neural Information Processing Systems 25 (NIPS)* (Curran Associates, Inc., 2012), pp. 998–1006.

[31] A. E. Sizemore, E. A. Karuza, C. Giusti, and D. S. Bassett, Nature Human Behav. **2**, 682 (2018).

[32] H. Edelsbrunner and D. Morozov, Persistent homology: Theory and practice, Tech. Rep., Lawrence Berkeley National Laboratory (LBNL), Berkeley, CA, 2012 (unpublished).

[33] R. Ghrist, Math. Data **25**, 273 (2018).

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

[35] A. E. Sizemore, C. Giusti, A. Kahn, J. M. Vettel, R. F. Betzel, and D. S. Bassett, J. Comput. Neurosci. **44**, 115 (2018).

[36] G. C. Garcia, A. Lesne, C. C. Hilgetag, and M.-T. Hütt, Phys. Rev. E **90**, 052805 (2014).

[37] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. Hellyer, and F. Vaccarino, J. R. Soc., Interface **11**, 20140873 (2014).

[38] M. W. Reimann, M. Nolte, M. Scolamiero, K. Turner, R. Perin, G. Chindemi, P. Dłotko, R. Levi, K. Hess, and H. Markram, Front. Comput. Neurosci. **11**, 48 (2017).

[39] I. Iacopini, G. Petri, A. Barrat, and V. Latora, Nat. Commun. **10**, 2485 (2019).

[40] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, Proc. Natl. Acad. Sci. USA **112**, 13455 (2015).

[41] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino, PLoS ONE **8**, e66506 (2013).

[42] B. Stolz, Master's thesis, University of Oxford, 2014.

[43] P. Bubenik, J. Mach. Learn. Res. **16**, 77 (2015).

[44] I. Chevyrev, V. Nanda, and H. Oberhauser, IEEE Trans. Pattern Anal. Mach. Intell. **42**, 192 (2018).

[45] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A.-L. Barabási, Science **297**, 1551 (2002).

[46] A. M. Yip and S. Horvath, BMC Bioinf. **8**, 22 (2007).

[47] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, J. Stat. Mech. (2008) P10008.

[48] G. Csardi and T. Nepusz, Intl. J. Complex Syst. **1695**, 1 (2006).

[49] P. Erdös and A. Rényi, Publicat. Math. (Debrecen) **6**, 290 (1959).

[50] E. N. Gilbert, Ann. Math. Stat. **30**, 1141 (1959).

[51] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[52] D. De Solla Price, J. Am. Soc. Inf. Sci. **27**, 292 (1976).

[53] M. Kahle, Discrete Comput. Geom. **45**, 553 (2011).

[54] T. Owada and R. J. Adler, Ann. Probab. **45**, 2004 (2017).

[55] D. Yogeshwaran and R. J. Adler, Ann. Appl. Probab. **25**, 3338 (2015).

[56] A. Sizemore, C. Giusti, and D. S. Bassett, J. Complex Netw. **5**, 245 (2017).

[57] M. Kaiser and C. C. Hilgetag, Phys. Rev. E **69**, 036103 (2004).

[58] G. Henselman and R. Ghrist, arXiv:1606.00199.

[59] https://CRAN.R-project.org/package=TDA.

[60] https://github.com/BassettLab/Reorderability_scripts.

[61] H. Ju, J. Z. Kim, and D. S. Bassett, arXiv:1812.09361.

[62] C. L. Barrett, K. R. Bisset, S. G. Eubank, X. Feng, and M. V. Marathe, in *Proceedings of the 2008 ACM/IEEE Conference on Supercomputing: SC'08* (IEEE, Piscataway, NJ, 2008), pp. 1–112.

[63] M. Granovetter and R. Soong, J. Math. Sociol. **9**, 165 (1983).

[64] M. E. J. Newman, Phys. Rev. E **66**, 016128 (2002).

[65] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[66] A. K. Lee and M. A. Wilson, J. Neurophysiol. **92**, 2555 (2004).

[67] Excluding the map from the first to second part, which will cause trouble for us later.

[68] V. Nicosia, P. E. Vértes, W. R. Schafer, V. Latora, and E. T. Bullmore, Proc. Natl. Acad. Sci. USA **110**, 7880 (2013).

[69] M. Kahle, Discrete Math. **309**, 1658 (2009).

[70] M. Kahle, AMS Contemp. Math. **620**, 201 (2014).

[71] Z. Wu, G. Menichetti, C. Rahmede, and G. Bianconi, Sci. Rep. **5**, 10073 (2015).

[72] G. Bianconi and C. Rahmede, Sci. Rep. **7**, 41974 (2017).

[73] O. T. Courtney and G. Bianconi, Phys. Rev. E **95**, 062301 (2017).

[74] C. F. Fowler, Discrete Comput. Geom. **62**, 87 (2019).

[75] J. W. Rocks, A. J. Liu, and E. Katifori, arXiv:1901.00822.

[76] F. Mémoli and O. B. Okutan, Discrete Comput. Geom. (2018), doi: 10.1007/s00454-019-00104-y

[77] T. Gavrilchenko and E. Katifori, Phys. Rev. E **99**, 012321 (2019).

[78] L. Heaton, B. Obara, V. Grau, N. Jones, T. Nakagaki, L. Boddy, and M. D. Fricker, Fungal Biol. Rev. **26**, 12 (2012).

[79] L. Papadopoulos, P. Blinder, H. Ronellenfitsch, F. Klimm, E. Katifori, D. Kleinfeld, and D. S. Bassett, PLoS Comput. Biol. **14**, e1006428 (2018).

[80] A. A. Mansour, J. T. Gonçalves, C. W. Bloyd, H. Li, S. Fernandes, D. Quang, S. Johnston, S. L. Parylak, X. Jin, and F. H. Gage, Nat. Biotechnol. **36**, 432 (2018).

[81] G. Quadrato, T. Nguyen, E. Z. Macosko, J. L. Sherwood, S. M. Yang, D. R. Berger, N. Maria, J. Scholvin, M. Goldman, J. P. Kinney *et al.*, Nature **545**, 48 (2017).

[82] M. M. Sperry, Q. K. Telesford, F. Klimm, and D. S. Bassett, J. Complex Netw. **5**, 199 (2016).

[83] J. D. Dworkin, R. T. Shinohara, and D. S. Bassett, NeuroImage **183**, 872 (2018).

[84] J. D. Dworkin, R. T. Shinohara, and D. S. Bassett, PLoS ONE **14**, e0216146 (2019).

[85] M. A. Kramer, U. T. Eden, E. D. Kolaczyk, R. Zepeda, E. N. Eskandar, and S. S. Cash, J. Neurosci. **30**, 10076 (2010).

[86] D. Bray, R. B. Bourret, and M. I. Simon, Mol. Biol. Cell **4**, 469 (1993).

[87] C. W. Lynn, A. E. Kahn, N. Nyema, and D. S. Bassett, arXiv:1805.12491 [Nat. Comm. (to be published)].

[88] T. T. Hills, M. N. Jones, and P. M. Todd, Psychol. Rev. **119**, 431 (2012).

[89] J. N. Giedd, J. Blumenthal, N. O. Jeffries, F. X. Castellanos, H. Liu, A. Zijdenbos, T. Paus, A. C. Evans, and J. L. Rapoport, Nat. Neurosci. **2**, 861 (1999).

[90] C. Pantelis, D. Velakoulis, P. D. McGorry, S. J. Wood, J. Suckling, L. J. Phillips, A. R. Yung, E. T. Bullmore, W. Brewer, B. Soulsby, D. Patricia, and K. M. Philip, Lancet **361**, 281 (2003).

[91] K. Robinson, T. Schmidt, and D. M. Teti, in *Handbook of Research Methods in Developmental Science*, edited by D. M. Teti (Blackwell Publishing Ltd, 2008), pp. 1–20

[92] K. Klemm and S. Bornholdt, Proc. Natl. Acad. Sci. USA **102**, 18414 (2005).

[93] C. Oosawa, K. Takemoto, and M. A. Savageau, arXiv:0711.2730.

[94] H. Zhang, Y. Chen, and Y. Chen, PloS ONE **7**, e51840 (2012).

[95] D. Shi, G. Chen, W. W. K. Thong, and X. Yan, IEEE Circuits Syst. Mag. **13**, 66 (2013).

[96] D. Shi, L. Lü, and G. Chen, Natl. Sci. Rev. **6**, 962 (2019).

[97] S. M. Shaffer, M. C. Dunagin, S. R. Torborg, E. A. Torre, B. Emert, C. Krepler, M. Beqiri, K. Sproesser, P. A. Brafford, M. Xiao *et al.*, Nature (London) **546**, 431 (2017).

[98] S. M. Shaffer, B. L. Emert, A. E. Sizemore, R. Gupte, E. Torre, D. S. Bassett, and A. Raj, BioRxiv, 379016 (2018).

[99] M. Ieda, J.-D. Fu, P. Delgado-Olguin, V. Vedantham, Y. Hayashi, B. G. Bruneau, and D. Srivastava, Cell **142**, 375 (2010).