Mathematical Biosciences 364 (2023) 109056

Contents lists available at ScienceDirect

Mathematical Biosciences

journal homepage: www.elsevier.com/locate/mbs

Original Research Article

# A topological data analysis study on murine pulmonary arterial trees with pulmonary hypertension

Megan Miller [a,b,1], Natalie Johnston [a,c], Ian Livengood [a], Miya Spinelli [a], Radmila Sazdanovic [a], Mette S. Olufsen [a,*]

[a] North Carolina State University, 2311 Stinson Drive, Raleigh, 27695, NC, USA
[b] Virginia Military Institute, 319 Letcher Avenue, Lexington, 24450, VA, USA
[c] Duke University, 415 Chapel Drive, Durham, 27708, NC, USA

ARTICLE INFO

*Keywords:*
Pulmonary hypertension
Vascular remodeling
Image segmentation
Tree pruning
Strahler order
Persistent homology
Topological data analysis
TDA

ABSTRACT

Pulmonary hypertension (PH), defined by a mean pulmonary arterial blood pressure above 20 mmHg in the main pulmonary artery, is a cardiovascular disease impacting the pulmonary vasculature. PH is accompanied by chronic vascular remodeling, wherein vessels become stiffer, large vessels dilate, and smaller vessels constrict. Some types of PH, including hypoxia-induced PH (HPH), also lead to microvascular rarefaction. This study analyzes the change in pulmonary arterial morphometry in the presence of HPH using novel methods from topological data analysis (TDA). We employ persistent homology to quantify arterial morphometry for control and HPH mice characterizing normalized arterial trees extracted from micro-computed tomography (micro-CT) images. We normalize generated trees using three pruning algorithms before comparing the topology of control and HPH trees.

This proof-of-concept study shows that the pruning method affects the spatial tree statistics and complexity. We find that HPH trees are stiffer than control trees but have more branches and a higher depth. Relative directional complexities are lower in HPH animals in the right, ventral, and posterior directions. For the radius pruned trees, this difference is more significant at lower perfusion pressures enabling analysis of remodeling of larger vessels. At higher pressures, the arterial networks include more distal vessels. Results show that the right, ventral, and posterior relative directional complexities increase in HPH trees, indicating the remodeling of distal vessels in these directions. Strahler order pruning enables us to generate trees of comparable size, and results, at all pressure, show that HPH trees have lower complexity than the control trees.

Our analysis is based on data from 6 animals (3 control and 3 HPH mice), and even though our analysis is performed in a small dataset, this study provides a framework and proof-of-concept for analyzing properties of biological trees using tools from Topological Data Analysis (TDA). Findings derived from this study bring us a step closer to extracting relevant information for quantifying remodeling in HPH.

## 1. Introduction

Cardiovascular diseases are the leading cause of death in the Western world. The World Health Organization (WHO) [1] estimates that in 2019, 17.9 million people died from cardiovascular diseases, approximately 32% of all deaths. While cardiovascular diseases encompass many phenotypes, a common trait is that they are associated with the remodeling of the vasculature and the heart, causing a range of functional problems, the most prominent being high blood pressure. Another important finding is that early detection is essential to improve quality of life and lessen the burden on the healthcare system [1].

Cardiovascular diseases are typically diagnosed by combining imaging measurements, such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), with dynamic (heart rate, blood pressure, and flow), and static measurements. The latter include patient characteristics, physiological and biochemical markers derived from blood tests. Cardiovascular disease can target any part of the vascular system

* Corresponding author.
*E-mail address:* msolufse@ncsu.edu (M.S. Olufsen).
[1] née Chambers.

https://doi.org/10.1016/j.mbs.2023.109056
Received 29 December 2022; Received in revised form 3 July 2023; Accepted 27 July 2023
Available online 6 August 2023

changing blood pressure, blood flow, and heart rate, often caused by tissue remodeling at the local (within specific vessels) and global (network/tree) levels. Tissue remodeling can be viewed from two perspectives: at the vessel level, remodeling causes changes in vessel compliance and diameter, and at the tree level, changes in morphometry, e.g., via vascular rarefaction. Both impact blood pressure and flow. This study uses techniques from topological data analysis to characterize changes in the morphometry of pulmonary arteries in mice with hypoxia-induced pulmonary hypertension (HPH).

Pulmonary hypertension (PH) is characterized by high blood pressure (a mean $\geq$ 20 mmHg) in the main pulmonary artery. Severe increases in pulmonary arterial pressure are typically caused by vascular remodeling and inflammation in the veins or left heart [2]. Symptoms of the disease include shortness of breath, fatigue, dizziness, chest pain, heart palpitations, and swelling of the legs and ankles. These symptoms are common in many illnesses, making PH difficult to diagnose [3]. Moreover, PH is a heterogeneous disease encompassing five subtypes, each with a different pathophysiology [4]. Early diagnosis and targeted treatment can improve quality of life by delaying severe complications, which is essential because all but one type of PH have no cure [5]. This study analyzes the morphometry in murine pulmonary arterial trees excised from healthy and HPH animals. In HPH, vascular remodeling impacts the pulmonary arteries [6,7]. This disease starts in the arterioles, stiffening and constricting the vessels, then migrating to the larger arteries, which stiffen and dilate. The structural modifications in individual vessels are characterized in Hopkins et al. [8] while our study focuses on an open question about the morphometry changes.

**Pulmonary arterial morphometry** characterizes arterial tree properties. Since early contributions by Murray in 1926 [9], many researchers (e.g., [6,7,10–12]) have examined the morphometry of the pulmonary arterial tree. The early studies by Murray [9] and Zamir [13] devise optimality principles, finding that arterial branching minimizes pumping power and lumen volume. These studies assume that the arterial trees bifurcate and that the dimensions of the two daughter vessels can be determined as functions of the parent vessel. Results include a power law defining how vessel radii change across a bifurcation, an asymmetry ratio relating the radii of the two daughter vessels, and an area ratio relating the combined cross-sectional area of daughter vessels to that of the parent vessel. These studies were supplemented by work from Singhal et al. [12], Horsfield [11], and [14,15] incorporating data from lung casts to devise relations between parent and daughter vessels. The casts are generated by injecting liquid resin into the arterial trees. The vessel dimensions are measured using calipers on the hardened resin cast. The data provide geometric information, but each study only examined a single lung, as human cadaver data are not easily obtained. Moreover, the casts are fragile, and there is inherent human error in using calipers to measure the vessel dimensions. More recent studies by Molthen et al. [6], Vanderpool et al. [7], Davidoiu et al. [16], and Chambers et al. [10] use medical imaging to accurately and efficiently extract geometric information from arterial networks.

**Persistent homology** is one of the major techniques of Topological Data Analysis (TDA) using methods from algebraic topology to identify qualitative features of data providing robust numerical descriptors of shape that can further be used as an input for statistical analysis [17]. These techniques have been applied to a range of scientific problems, including biology; see [18] for a concise and focused introduction.

To our knowledge, this study is the first to use persistent homology to characterize pulmonary arterial trees, and it is inspired by ideas by Bendich et al. [19] and Belchi et al. [20]. The study by Bendich et al. [19] analyzes human brain arterial trees generated from a tube-tracking segmentation algorithm on magnetic resonance images (MRI), showing that persistence correlates strongly with age and sex. While Belchi et al. [20] use persistent homology to compare airway CT images from healthy and chronic obstructive pulmonary disease (COPD) patients. They found that persistent homology can distinguish the patient groups and detect inspiration and expiration. Our study introduces new quantitative summaries of spatial trees extracted from micro-CT images based on 0-dimensional persistent homology, using *directional complexity*, to obtain a more detailed characterization than the one from combinatorial invariants, such as tree depth and Strahler order [21,22].

Both numerical descriptors work in synergy to provide a detailed characterization of the pulmonary arteries comparing murine control and HPH trees at four different pressures. The topology-based descriptors capture the spatial tree structure, while the combinatorial features provide insight into the tree size. This study demonstrates the strength of combining these, using the former to generate comparable trees and the latter to differentiate the trees. The former is vital to control differences induced by remodeling. Both trees come from the same species. In HPH, vascular remodeling leads to wall thickening and increased diameter of the large vessels [23] while the medium and small vessels (not visible in the micro-CT) constrict and rarefy [23,24], i.e., HPH trees have the same or fewer branches. As a result, in HPH trees, we can identify more vessels from an image at a finite resolution. However, this is not a feature of the network size. Therefore to extract other differences caused by remodeling, we prune trees according to the radius or their Strahler order and then employ the topological machinery.

## 2. Methods

This study identifies topological and combinatorial features characterizing differences between control and HPH pulmonary arterial trees extracted from murine micro-CT images. The methods include three steps. First, we generate a 3-dimensional (3D) rendered network from the micro-CT image and construct a labeled spatial tree with *edges*, representing vessels, and *vertices*, junctions. Each vertex has coordinates in $\mathbb{R}^3$, and each edge is labeled by its length, spatial orientation, and the radius of its corresponding vessel. Second, we compare the spatial trees obtained from control and HPH animals. To glean biologically important information from these trees, comparing only corresponding parts of the control and HPH trees is essential. In HPH, the diameter of the large vessels increases [10], making more vessels visible in the 3D-rendered trees. Therefore, to compare the trees, we employ pruning algorithms to normalize the trees. Finally, we use persistent homology to characterize the normalized trees. Inspired by Belchi et al. [20], we compute 0-dimensional persistent homology for spatial trees, including height filtration and corresponding directional complexities.

### 2.1. Imaging protocol

The micro-CT images used in this study are made available by Chesler, University of California-Irvine [7]. This study uses images from six C57BL6/J mice, aged 10–12 weeks, and weight 25.8 $\pm$ 2.3 grams. Three of the six mice are controls, and three have PH induced by placing them in a hypoxic environment (FiO2 reduced by half) for ten days. The mice are anesthetized with (52 mg/kg body weight) pentobarbital sodium and euthanized by exsanguination before the lungs are extracted. After extraction, the lungs are imaged following the protocol described in [7]. First, a cannula (PE-90 tubing, 1.27 mm outer and 0.86 mm inner diameter) is positioned in the main pulmonary artery (MPA) well above the first arterial bifurcation. Next, the lungs are ventilated in a gas mixture with (15% $O_2$, 6% $CO_2$, balance nitrogen), rinsed with a physiological salt solution, and perfused with perfluorooctyl bromide (PFOB). The lungs are prepared by adjusting the intravascular pressure from 0 to 25 mmHg multiple times. After preparation, the arterial pressure is kept constant as the lungs are rotated around an X-ray beam at 1° increments to obtain 360 planar images. The lungs are imaged at 6.3, 7.4, 13.0, and 17.4 mmHg pressure measured in the main pulmonary artery (Fig. 2). For each pressure, planar images are constructed using the Feldkamp cone-beam algorithm and converted into a 3D volumetric dataset, saved as a Digital Imaging and Communications in Medicine (DICOM) 3.0 image (Fig. 1(a)).

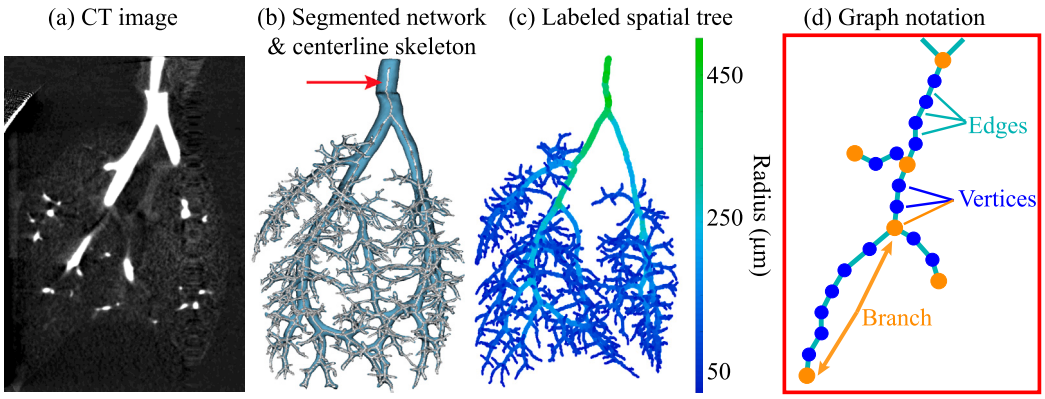


**Fig. 1.** Spatial tree extraction process. (a) Micro-CT image from a control mouse (coronal view). (b) 3D rendering of the segmented arterial tree overlaid with the centerline skeleton (marked with the red arrow). (c) Labeled spatial tree extracted from the skeleton, with branch colors denoting the vessel radii. (d) Detailed representation of edges and vertices in the labeled spatial tree. The edges (teal) connect two vertices (blue). Vertices forming a junction (connected to three or more edges) or creating a root or terminal point (connected to one edge) are marked in orange. The series of vertices and edges between 2 junction points or between a junction point and a terminal point is called a *branch*.

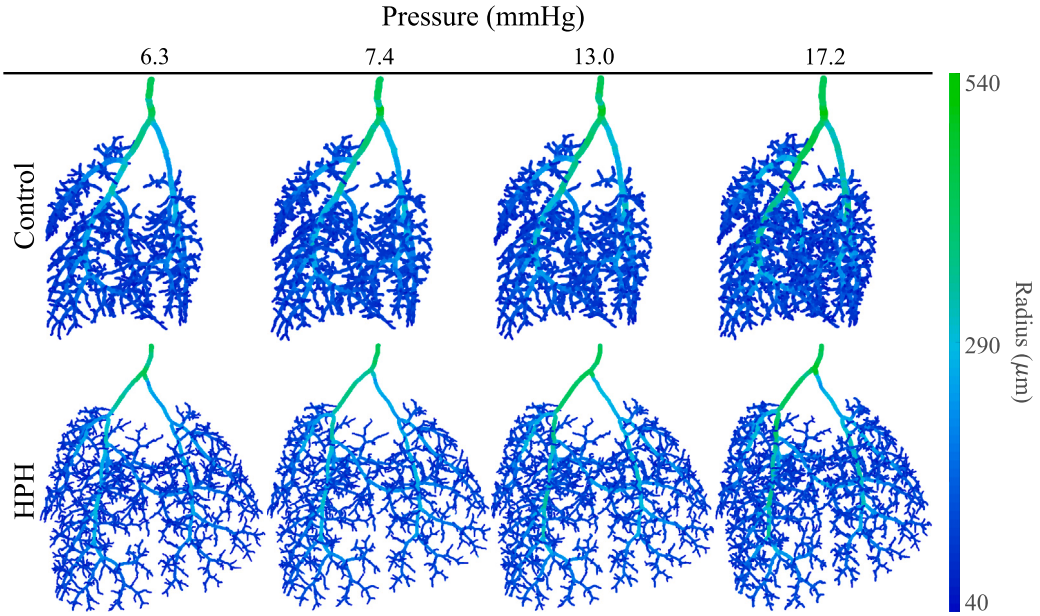


**Fig. 2.** Labeled spatial trees from a control mouse (top row) and an HPH mouse (bottom row) extracted from lungs perfused at four different pressures (6.3, 7.4, 13.0, and 17.4 mmHg) measured in the main pulmonary artery. As the pressure increases, the contrast is transported further, making smaller vessels visible in the image resulting in more branches in the spatial trees.

### 2.2. Segmentation and skeletonization

The micro-CT images are saved in the Digital Imaging and Communications in Medicine (DICOM) 3.0 format. Each image has the dimensions $(497 \times 497 \times 497)$ voxels with a spatial resolution of 30 – 40 μm per voxel. Each voxel $v$ has spatial coordinates $(x_v, y_v, z_v)$ and intensity $I(v) \in [0, 255]$, with 0 representing the intensity of a black voxel and 255 of a white voxel.

**Segmentation** is the process used to partition the voxels into the set of *foreground* voxels, which represent the arteries, and *background* voxels representing the surrounding space or tissue. We segment the images using the open-source program 3D Slicer [25–27], from Kitware, Inc., employing a combination of techniques, including global thresholding, median smoothing, and manual editing as described in detail in our previous study [10]. Global thresholding is used to identify voxels in the foreground, finding voxels with intensities $I(v) \in [\tau_{\min}, \tau_{\max}]$. The pulmonary arteries are the only anatomical structures in the image, so $\tau_{\max} = 255$ for all images. The minimum threshold $\tau_{\min}$ is adjusted ad hoc to ensure that all visible arteries are included. To reduce noise, median smoothing is used, replacing the intensity of all the voxels within a kernel of $(3 \times 3 \times 3)$ voxels with the median of the adjacent voxel intensities. Since the cannula and the hypobaric cylinder can distort the image, manual editing is required to remove voxels from the foreground, which do not represent the actual arterial structure. The result of segmentation is a foreground that can be visualized as a 3D-rendered surface representing the pulmonary arterial tree. The constructed foreground is referred to as the *segmented network* (shown in Fig. 1(b)). Each *segmented artery* is a portion of the segmented network that lies between two junctions (or between a junction and a terminal point).

**Skeletonization:** The *skeleton* is a voxel complex comprising a thinned representation of the segmented arteries where each branch is 1 voxel in width. In the skeleton, the branching structure of the segmented arteries is preserved, i.e., each branch is represented by a chain of adjacent voxels that is centered in the artery. The skeleton is obtained by iteratively removing voxels from the segmented arteries using

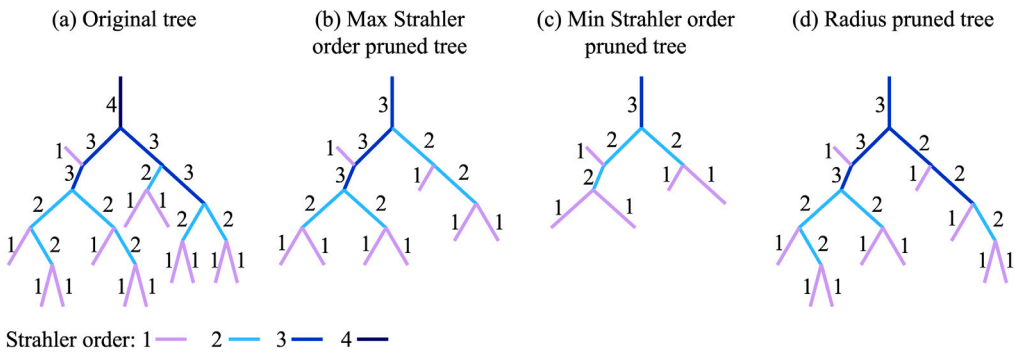


**Fig. 3.** The Strahler ordering system is illustrated on an example tree. (a) Example tree of Strahler order 4. (b) Maximum Strahler order pruning to order 3. This tree removes as few vessels as possible from the original tree to reach Strahler order 3. (c) Minimum Strahler order pruning to order 3. This tree removes as many vessels as possible, still getting a tree of Strahler order 3, i.e., if one more round of pruning is performed, the tree will have Strahler order 2. (d) Radius pruned tree. This tree removes vessels with a radius smaller than a given threshold. Note the pruned tree is Strahler order 3 but has a different structure than any of the Strahler order pruned trees.

Couprie and Bertrand's "Asymmetric Thinning" algorithm. For details see [10,28]. Fig. 1(b) shows an example of a centered skeleton.

**Distance map:** To obtain the dimensions for the segmented arteries, we generate a *distance map*. This is a separate voxel complex with the same spatial dimensions as the original CT image. It encodes distances from the segmented network. Each voxel $v$ in the segmented network has an associated voxel $v_D$ in the distance map at the same spatial coordinates $(x_v, y_v, z_v)$. The intensity of this distance map voxel is $I(v_D) = d(v)$, where $d(v)$ is defined as

$$d(v) = \min_{u \in \text{background}} \|u - v\|_2 , \tag{1}$$

where $\|u - v\|_2$ is the Euclidean distance from $v$ to each background voxel, $u$. In other words, we know the distance from each voxel in the segmented network (AKA image foreground) to the nearest background voxel. Since voxels in the skeleton are approximately in the center of the arteries, their $d(v)$ values approximate the radius of the artery at that point (discussed in the next section). This distance map, the skeleton, and the labeled spatial tree are obtained using the Spatial Graph Extractor (SGEXT) in the Digital Geometry Tools and Algorithms Library (DGtal) available in GitHub [29–31].

### 2.3. Spatial trees

Each skeleton is used to generate a *spatial graph* (also known as an embedded metric graph), a collection of vertices in $\mathbb{R}^3$ connected by edges. Every voxel in the skeleton corresponds to a vertex in the spatial graph. Each vertex, denoted by a numerical ID $A$, has 3D coordinates $(x_A, y_A, z_A)$. Each edge connecting two vertices $A$ and $B$ is denoted $e_{AB}$. If $\deg(A) = 1$, $A$ is either at the inlet to the main pulmonary artery (MPA) (called $A_{\text{root}}$) or a *leaf*, a terminal vertex. If $\deg(A) = 2$, $A$ is a vertex along the length of a segmented artery. If $\deg(A) > 2$, $A$ is a *junction point*, a point where a *parent* artery splits into multiple *daughter* arteries. Typically, junction points have degree 3, as 98%–99% of junctions are bifurcations [10]. All edges are oriented toward blood flow (away from $A_{\text{root}}$). With this orientation in place, we can refer to $e_{AB}$ as an edge from the start vertex $A$ to the end vertex $B$.

Constructing the spatial tree is subject to errors that can be corrected using the protocol introduced in [10]. Errors include edges that do not accurately represent the segmented arteries. These small cycles arise when three adjacent voxels are connected in a loop, duplicated edge points, and duplicated edges that connect the same vertices. False branches must be manually identified and removed from the graph, while small cycles and duplicate edges/points can be removed automatically. Cycles are broken by removing the longest edge. After implementing the corrections, the graph becomes a *spatial tree T* (Fig. 1(c)).

Each vertex in $T$ is labeled with its radius, determined from the distance map. Recall that the skeleton is centered in the segmented arteries. Therefore, for voxel $v$ in the skeleton, the distance $d(v)$ calculated using Eq. (1) gives an estimate for the radius of the artery centered at $(x_v, y_v, z_v)$. Hence, every vertex $A$ in $T$ can be labeled by its radius $r_A = d(v)$, i.e., for the corresponding voxel $v$ $(x_v, y_v, z_v) = (x_A, y_A, z_A)$.

In the 3D rendered surface, every artery is represented by a series of edges connecting consecutive degree 2 vertices between two junctions or a junction and a leaf. We refer to the collection of edges as a *branch* of $T$, i.e., each branch corresponds to an artery. Branches are labeled with an average radius and length in microns (μm). To convert the voxel measurements into μm, all dimensions, $x$, $y$, and $z$, are multiplied by the scaling factor $\lambda_T$, obtained by dividing the cannula's radius (430 μm) by the average of the first five radius measurements of the MPA, the radii of vertices $A_{\text{root}}, A_1, \ldots, A_4$, i.e.,

$$\lambda_T = \frac{5 \times 430 \text{ μm}}{(r_{A_{root}} + r_{A_1} + r_{A_2} + r_{A_3} + r_{A_4}) \text{ voxels}}. \tag{2}$$

The radius $r_{AB}$ of a branch from a leaf/junction point $A$ to a leaf/junction point $B$ is represented by the interquartile mean (IQM) of the radius measurements from $A$ to $B$, and the degree 2 vertices $A_1, \ldots, A_m$ along the branch:

$$r_{AB} = \lambda_T \ (IQM \left(r_A, r_{A_1}, r_{A_2} \ldots, r_{A_m}, r_B\right)). \tag{3}$$

The length $L_{AB}$ of a branch is the scaled sum of the Euclidean distances between consecutive vertices along that branch,

$$L_{AB} = \lambda_T \left( \|A - A_1\|_2 + \sum_{i=2}^{m} \|A_{i-1} - A_i\|_2 + \|A_m - B\|_2 \right). \tag{4}$$

### 2.4. Spatial trees: combinatorial statistics

The number of branches, leaves, tree depth, and the Strahler order (number) are used as features of the spatial trees. The *tree depth* is defined as the number of branches in the longest direct path from $A_{\text{root}}$ to any leaf. The *Strahler order* ($SO_i$) of a branch $i$ is an indicator of its level within the tree [32]. The computation of the Strahler order is illustrated in Fig. 3(a). Terminal branches all have Strahler order one. If two joining daughter vessels have equal Strahler order $SO_{d1} = SO_{d2}$, then the parent's edge Strahler is $SO_p = SO_{d1} + 1$. Otherwise, the Strahler order of the parent's edge is $SO_p = \max\{SO_{d1}, SO_{d2}\}$. The Strahler order of a tree $SO(T)$ refers to the maximum Strahler order of any branch within that tree. Note that the Strahler order and tree depth coincide in symmetric trees.

### 2.5. Spatial trees: pruning

The trees from HPH mice have significantly more branches than those from the control animals, mainly because the corresponding vessels are wider and, therefore, detectable in the micro-CT images. Thus, we use pruning techniques to obtain comparable trees. As mentioned above, if pulmonary trees were symmetric, pruning trees according to the tree depth would be sufficient. However, pulmonary arterial trees are asymmetrical [9,10,13,15]. We examine three pruning methods for the systematical removal of terminal branches. Two pruning techniques rely on computing the minimal and maximal Strahler order, and the third method uses the vessel radius.

**Maximum Strahler order pruning:** This method removes pairs of sister leaves at the lowest Strahler order, i.e., $SO_{d1} = SO_{d2} = 1$. Once all branches of this type have been removed, the branches' Strahler orders are updated and pruning repeats until the desired Strahler order is obtained. This pruning method is illustrated in Fig. 3. Panel (a) shows the original tree and (b) the tree after maximum Strahler order pruning, removing all sister vessels with Strahler order one from the original tree. The Strahler order of the pruned tree in (b) is one lower than that of the original tree.

**Minimum Strahler order pruning:** Since the trees are asymmetric, they vary significantly in size at a given Strahler order. To mitigate this variation, we perform minimum Strahler order pruning. This pruning method removes all pairs of terminating sister branches for which $SO_{d1} = SO_{d2} = 1$. In addition, this algorithm removes any branches that will not impact the final Strahler order. For each set of removed daughter vessels, the Strahler order is recomputed. The process continues until the next iteration of branch removal would decrease the tree's Strahler order. This method is illustrated in Fig. 3(c).

**Radius pruning:** The radius pruning technique removes pairs of sister leaves with a radius less than or equal to a given radius threshold $\tau_r$. This threshold is chosen so that any tree group has the desired number of branches. After radius pruning, all bifurcations with terminal daughter branches have at least one daughter with $r \geq \tau_r$. This method is illustrated in Fig. 3(d).

### 2.6. Spatial trees: topological statistics

This section describes how Topological Data Analysis (TDA) can provide a new numerical descriptor of spatial trees called *directional complexity*. TDA provides an extensive toolbox of data analysis methods for learning data shape. This study uses persistent homology [33–35] to analyze control and HPH spatial pulmonary arterial trees. Specifically, we compute the 0-dimensional persistent homology of height filtrations in biologically relevant directions. We use this to create a barcode and compute a topological marker known as directional complexity. These techniques are robust to noise [18,36], making them useful for biological applications such as the one studied here.

In the remainder of this section, we provide a brief background for TDA methods illustrated by examples from this study. Computing persistent homology involves: (a) building a sequence of topological spaces upon a data set, (b) computing the corresponding homology for each, and (c) tracking the changes in topological summaries that describe the shape of the data through filtration.

**Simplicial homology:** Spatial trees are 1-dimensional simplicial complexes in $\mathbb{R}^3$. We are using simplicial homology in computing persistence. In this section, we provide a brief introduction to simplicial homology. A $k$-dimensional simplex, for $k \geq 0$, is a convex hull of $k+1$ vertices [37]. For example, 0-simplex is a point, a 1-simplex is an edge, a 2-simplex is a triangular face, and a 3-simplex is a filled tetrahedron. A *simplicial complex* is the formal sum of building blocks called *simplices*. Each labeled spatial tree in this study consists of a set of points/vertices, which are 0-simplices, connected by a set of edges which are 1-simplices. Note that all simplices come with orientation which is consistent with the fact that the edges of spatial trees examined in this study are directed away from the tree's root to represent the direction of blood flow.

For simplicity, we describe simplicial homology in terms of linear algebra, using only vector spaces and linear maps. For a given simplicial complex $S$, the *chain vector space*, $C_k(S)$ is the vector space whose basis is the set of oriented $k$-simplices in $S$ [37]. The elements $c = \sum_i a_i \sigma_i \in C_k$ are called a *$k$-chains* where each $a_i$ is an integer and each $\sigma_i$ is an oriented $k$-simplex from $S$. The boundary map, $\partial_k : C_k \to C_{k-1}$, maps $k$-dimensional chains to a linear combination of their $(k-1)$-dimensional boundaries.

Let $L_k(S)$ be the kernel of $\partial_k$ and $B_k(S)$ be the image of $\partial_{k+1}$. The $k$th *simplicial homology*, $H_k(S)$, is the quotient

$$H_k(S) = L_k(S)/B_k(S) = \ker(\partial_k)/\mathrm{im}(\partial_{k+1}). \tag{5}$$

The $k$th *Betti number*, $\beta_k$, is the dimension of the k-dimensional homology $H_k$. Intuitively, $\beta_0$ counts the number of clusters (connected components) in a simplicial complex [18]. Property of a simplicial complex counted by $\beta_k$ is often referred to as a *$k$-feature*. Betti numbers $\beta_0$, $\beta_1$, and $\beta_2$ count the number of clusters (disjoint connected components), (unfilled) loops $S^1$, and (hollow) spheres $S^2$, respectively.

**Persistent homology:** Persistent homology combines ideas from algebraic topology and statistics, providing a powerful tool for data analysis. The idea behind persistent homology is to compute the homology of a sequence of nested simplicial complexes built from data points (point cloud) based on their proximity.

Given any positive parameter $\varepsilon$ and a point cloud, we first build a simplicial complex $S_\varepsilon$ by connecting all points that are less than $\varepsilon$ away from each other. Next, given an increasing sequence of positive parameter values $0 < \varepsilon_1 < \varepsilon_2, \ldots < \varepsilon_m$ we apply the construction above to each $\varepsilon_i$ to obtain a set of nested simplicial complexes based on our data $S_{\varepsilon_1} \subset S_{\varepsilon_2} \subset \ldots \subset S_{\varepsilon_m}$. This is called a *filtration* of a point cloud determined by a parameter $\varepsilon$. Finally, *$k$–dimensional persistent homology* analyzes changes in homology groups of each simplex in the filtration: $H_k(S_{\varepsilon_1}), \ldots, H_k(S_{\varepsilon_m})$. In turn, this provides insights into the evolution of the underlying space's topological features.

In practice, we track the changes of $k$th Betti numbers that count $k$-dimensional features describing the data's shape. If a feature appears for the first time in the homology of some $S_{\varepsilon_i}$ but is no longer present for $S_{\varepsilon_j}$, for $j > i$, then the *birth* of that feature is $\varepsilon_B = \varepsilon_i$, and the *death* is $\varepsilon_D = \varepsilon_j$. The *persistence* of a feature is the difference $\varepsilon_D - \varepsilon_B$. More persistent features exist for a large range of $\varepsilon$ values. Note that for a large enough persistence parameter, the corresponding complex has the topology of a point.

**Barcodes:** Persistent homology $H_k(S_{\varepsilon_i})$ can be visualized by creating a barcode $\mathcal{B}_k$ for each dimension $k$, a diagram which contains one bar or interval for each generator. Each bar spans the length from the birth $\varepsilon_B$ to the death $\varepsilon_D$ for that feature; see Fig. 5. Though the interpretations of the lengths of bars in the barcode are application-specific, shorter bars are often considered to represent noise. In our case, see Fig. 5(b) for the 0-dimensional barcode of one of the spatial trees; the longest vertical bar on the far left represents the whole spatial tree since at the end of filtration all of the branches will be connected to the root.

**Height filtration:** We compute the 0-dimensional persistent homology of the spatial trees for a height filtration [33]. The idea behind the height filtration in 0-dimensional persistence is that by analyzing sub-level sets of the height function, we are recording statistics of when new components appear and existing ones merge. For spatial trees, this translates into the appearance of branches pointing toward lower levels. In particular, we analyze six directions: anterior–posterior, which corresponds to the direction in which lungs grow and air flows; dorsal–ventral, in which gravity acts; and left–right, which captures the asymmetry of lungs, partially due to the position of the heart, see Fig. 4.

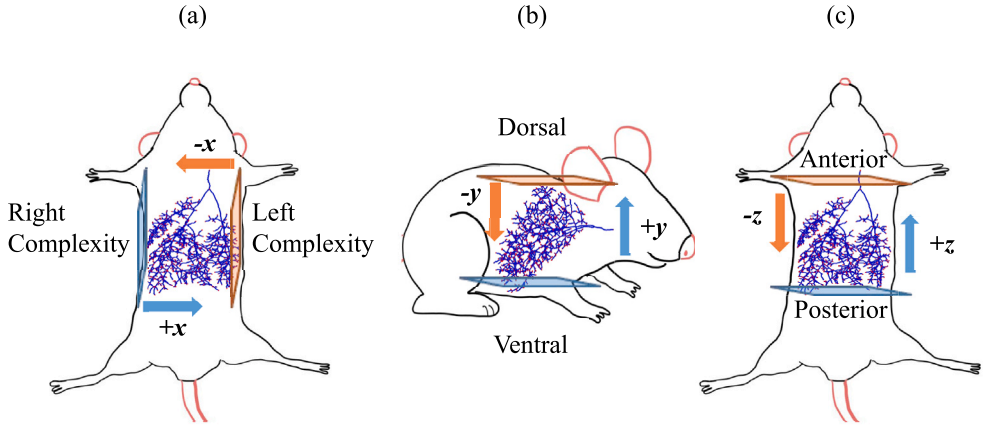


**Fig. 4.** The six height filtration directions are illustrated relative to a mouse's body. The directional complexities are named after the direction of the branches they capture, which will be the reverse of the direction in which the filtration moves.

This filtration was used by Belchi et al. [20] in the case of human lungs and COPD.

Given a spatial tree $T \in \mathbb{R}^3$ and its bounding box placed so that the minimum value of the coordinate corresponding to the direction is equal to zero, consider the distance of any vertex $V$ or a point on some edge $E$ of $T$ to one of the sides of the box along the chosen direction. The height filtration requires that each $S_{\varepsilon_i}$ includes all vertices $v \in V$ and edges $e \in E$ for which the value of the distance function $\varepsilon(\cdot)$ is at most $\varepsilon_i$. The $\varepsilon(\cdot)$ formula depends on the filtration direction examined. As mentioned above, we consider six filtration directions: the positive/negative $x$-directions ($+x/-x$), the positive/negative $y$-directions ($+y/-y$), the positive/negative $z$-directions ($+z/-z$) (shown in Fig. 4). For a given filtration direction the height $\pm\xi$, $\varepsilon$ of the vertex $v \in V$ with spatial coordinates $(x_v, y_v, z_v)$ is defined as

$$\varepsilon(v) = \begin{cases} \xi_v, & \text{if direction} = +\xi \\ \xi^T_{\max} - \xi_v, & \text{if direction} = -\xi, \end{cases} \tag{6}$$

where $\xi^T_{\max}$ is the maximum $\xi$ coordinate of any $v \in V$ for $\xi = x$, $y$, or $z$. For an edge $e \in E$ between two vertices $v_1$ and $v_2$,

$$\varepsilon(e) = \max\left\{\varepsilon(v_1), \varepsilon(v_2)\right\}.$$

For example, a height filtration in the $+z$ direction indicates that the filtration moves along the $z$-axis in the positive direction, and each $S_{\varepsilon_i}$ includes all vertices with $z_v \le \varepsilon_i$ and all edges with $\max\left\{z_{v_1}, z_{v_2}\right\} \le \varepsilon_i$. Conversely, a direction of $-z$ indicates that the filtration moves along the $z$-axis in the negative direction, and each $S_{\varepsilon_i}$ includes all vertices with $z^T_{\max} - z_v \le \varepsilon_i$ and all edges with $\max\left\{z^T_{\max} - z_{v_1}, z^T_{\max} - z_{v_2}\right\} \le \varepsilon_i$. An illustration of the $-z$ height filtration is shown in Fig. 5.

**Directional complexity:** Given a spatial tree $T$ and a filtration direction $\pm\xi$, the 0-dimensional barcode is computed by counting the number of disjoint connected components in the nested set of subgraphs at each $\varepsilon_i$, Fig. 5. For the tree $T$, the directional complexity $DC(T)$ is the total persistence of the bar code associated with this tree, computed by summing the lengths of the bars in the barcode

$$DC(T) = \sum_{k=1}^{\text{no. of bars}} (d_k - b_k). \tag{7}$$

Directional complexities are named relative to murine anatomy and capture the length and occurrence of branches in the designated direction (Fig. 4). For example, the $-z$ filtration yields anterior complexity, referring to branches pointing toward the head of the mouse (Fig. 5).

## 3. Results

Data consists of mouse lung arteries imaged at four contrast pressures: 6.3, 7.4, 13.0, and 17.4 mmHg. Results in Tables 1 and 2 are reported as the mean and standard deviation. Table 1 contains spatial tree statistics of the original and pruned trees, including the number of branches, leaves, tree depth, and Strahler order. Table 2 shows the trees' directional complexities generated from the 0-dimensional persistent homology for the height filtration.

**Table 1**
Spatial tree statistics from micro-CT images at 4 different contrast pressures (6.3, 7.4, 13.0, and 17.2 mmHg). Abbreviations used are "C" for control trees, "HPH" for unpruned HPH trees, "$HPH^R$" for radius-pruned HPH trees, "$HPH^M$" for maximal Strahler-order pruned HPH trees, and "$C^m$" and "$HPH^m$" for the minimal Strahler-order pruned control and HPH trees, respectively. The Strahler order is the same for all trees in each group, except for the $HPH^R$ trees at 7.4 & 13.0 mmHg, where one tree had Strahler order 7, and the other two equal to 6. All other values are reported as mean ± standard deviation.

| Pressure | Type | # Branches | # Leaves | Tree depth | Strahler order |
|---|---|---|---|---|---|
| 6.3 | C | 1018 ± 151 | 513 ± 77 | 25.3 ± 1.5 | 6 |
| | HPH | 2196 ± 778 | 1107 ± 395 | 32.3 ± 4.0 | 7 |
| | $HPH^R$ | 1022 ± 53 | 516 ± 29 | 30.7 ± 3.5 | 6 |
| | $HPH^M$ | 1519 ± 530 | 766 ± 270 | 31.3 ± 4.0 | 6 |
| | $C^m$ | 829 ± 262 | 418 ± 132 | 24.7 ± 1.5 | 6 |
| | $HPH^m$ | 729 ± 178 | 368 ± 91 | 28.7 ± 3.5 | 6 |
| 7.4 | C | 1122 ± 193 | 565 ± 98 | 28.0 ± 5.2 | 6 |
| | HPH | 2362 ± 769 | 1193 ± 390 | 31.7 ± 2.9 | 7 |
| | $HPH^R$ | 1122 ± 82 | 567 ± 42 | 30.0 ± 1.7 | [6 6 7] |
| | $HPH^M$ | 1337 ± 346 | 676 ± 175 | 30.0 ± 2.6 | 6 |
| | $C^m$ | 837 ± 289 | 422 ± 146 | 27.0 ± 5.5 | 6 |
| | $HPH^m$ | 728 ± 106 | 367 ± 53 | 27.7 ± 2.1 | 6 |
| 13.0 | C | 1693 ± 372 | 854 ± 191 | 30.3 ± 1.5 | 6 |
| | HPH | 2730 ± 972 | 1379 ± 495 | 33.0 ± 3.6 | 7 |
| | $HPH^R$ | 1689 ± 206 | 855 ± 108 | 32.0 ± 2.6 | [6 6 7] |
| | $HPH^M$ | 1394 ± 306 | 706 ± 156 | 31.0 ± 3.0 | 6 |
| | $C^m$ | 961 ± 171 | 485 ± 84 | 28.7 ± 2.5 | 6 |
| | $HPH^m$ | 725 ± 84 | 367 ± 43 | 28.3 ± 2.5 | 6 |
| 17.2 | C | 2573 ± 517 | 1301 ± 263 | 32.0 ± 4.4 | 7 |
| | HPH | 3239 ± 1103 | 1639 ± 564 | 35.0 ± 4.0 | 7 |
| | $HPH^R$ | 2592 ± 563 | 1312 ± 289 | 34.7 ± 3.5 | 7 |
| | $HPH^M$[a] | 3239 ± 1103 | 1639 ± 564 | 35.0 ± 4.0 | 7 |
| | $C^m$[a] | 2573 ± 517 | 1301 ± 263 | 32.0 ± 4.4 | 7 |
| | $HPH^m$ | 1955 ± 304 | 989 ± 156 | 33.7 ± 3.5 | 7 |

[a]At 17.2 mmHg, the $HPH^M$ and $C^m$ are the same as their unpruned versions.

### 3.1. Spatial tree statistics

The results in Table 1 indicate that in the unpruned trees, the number of branches, leaves, and tree depth is lower in control than in hypertension (HPH) trees. This can be explained by remodeling increasing the diameter of the larger vessels and, therefore, the number of branches visible in the segmented images. Another observation is that the relative branch count shown in Fig. 6(a) (computed at each

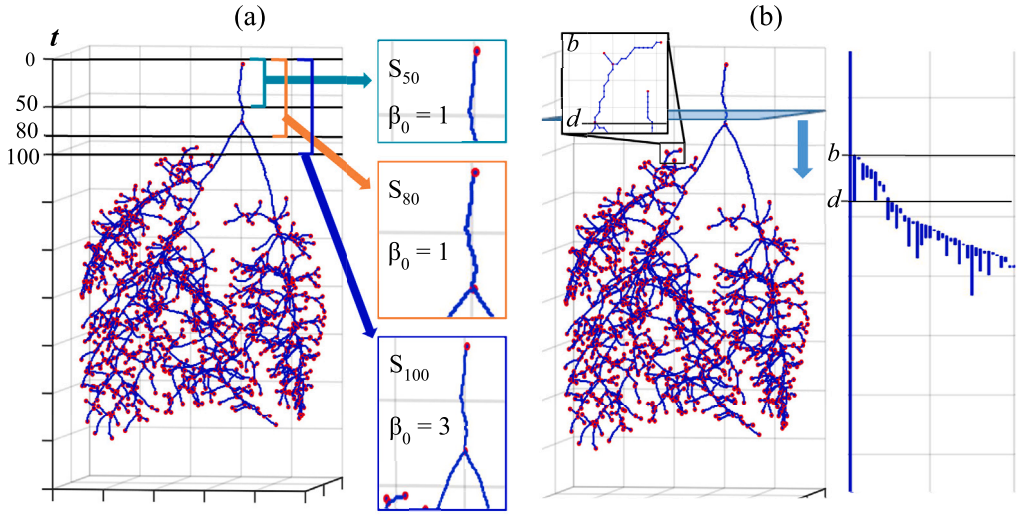


**Fig. 5.** A labeled spatial tree illustrates directional filtration in the anterior ($-z$) direction. Panel (a) shows a spatial tree from a control mouse with several examples of the simplicial complexes $S_t$, where $t$ denotes the distance from the top of the tree, each labeled with their 0-dimensional Betti numbers. $S_{50}$ and $S_{80}$ each contain one connected component, so $\beta_0 = 1$ for both. $S_{100}$ includes 3 connected components, so $\beta_0 = 3$. Panel (b) illustrates the filtration process in the $-z$ direction and shows a portion of the barcode. Each time a new connected component emerges in some $S_t$, a new bar is added to the barcode. For example, the magnified branch would create a new connected component in $S_b$, where $t = b$ is called this component's *birth*. Then, in $S_d$, the branch reconnects to the main tree, no longer being a separate connected component. This means $t = d$ is the *death*, and the bar representing this component is ended. The *persistence* of this feature is the length of its bar, $d - b$. The *directional complexity* (specifically *anterior complexity* for $-z$ filtration) is the sum of the persistence of all these bars (Eq. (7)). The longest and leftmost bar in the barcode begins with the root vessel, indicating that the entire tree comprises one connected component.

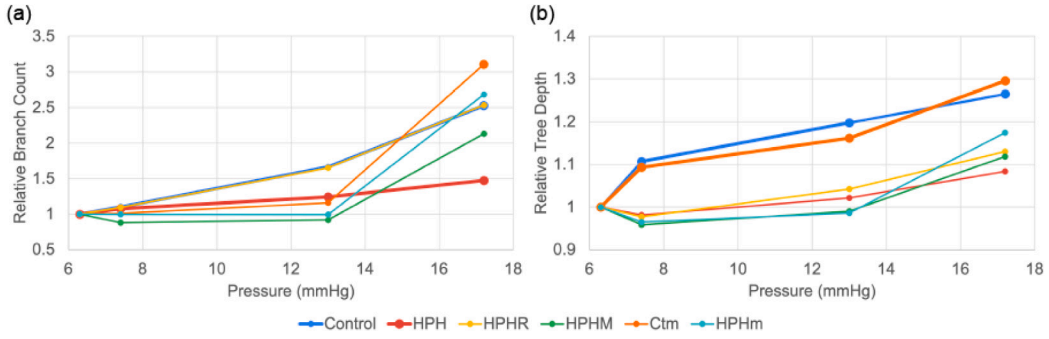


**Fig. 6.** (a) Relative change in the number of branches (or leaves) with pressure for control trees and (b) the relative change in tree depth. All values are plotted as a ratio of the value at that pressure over the value at the lowest pressure (hence, all graphs start at 1).

pressure as a ratio over the lowest pressure, i.e., all graphs start at 1) increases more in control (blue line) than in HPH (red line) trees. This is a consequence of vessels being more compliant in control animals and expanding more with increasing pressure than the stiffer HPH trees. Another critical observation (Fig. 6(a)) is that all the Strahler order pruned trees have approximately the same relative branch count independent of pressure, except for the highest pressure where the Strahler order goes up by 1.

The maximum and minimum Strahler order pruned trees (Fig. 7) all have Strahler order 6, except at the highest pressure (17.2 mmHg), for which all trees have Strahler order 7. The relative increase in the number of branches is approximately constant at the lower pressures. The latter is more consistent for minimum Strahler order than maximum Strahler order pruning. While the number of branches remains approximately constant at the lower pressures, the relative tree depth is lower in the hypertensive trees.

### 3.2. Directional complexity

All directional complexities are higher in the unpruned HPH trees than in the control trees. For the pruned trees, directional complexity can easily be compared for the radius and the minimum Strahler order pruned trees as these have approximately the same number of branches. The radius pruned HPH trees have the same number of branches at each pressure as the control trees, but as shown in Fig. 6 the number of branches increases with pressure. The Strahler order pruned trees have approximately the same number of branches for all pressures except the highest pressure. This feature is most consistent in the minimum trees.

Fig. 8(a) shows that in the radius pruned trees, the right, ventral, and posterior complexity is lower in the HPH animals at all pressures, i.e., these trees have fewer or shorter branches than the control trees. In contrast, hypertensive animals have slightly higher left, ventral, and anterior complexity. At the lower pressures, the control tree mainly contains larger vessels known to exhibit sustained vessel dilation [38].

Since the control trees have compliant vessels, they dilate at higher pressures, allowing us to segment more distal vessels. The ability to capture these additional vessels enables us to quantify the remodeling of the smaller vessels. Results (Fig. 8(a)) show that the right, ventral, and posterior complexity increases (becomes closer to that of the control trees). This agrees with the observation [38] that smaller vessels experience more remodeling than large vessels, possibly generating a tree of similar complexity to the control trees. Like the lower pressure complexity in the left, dorsal and anterior directions increase slightly with increasing pressure (Fig. 8).

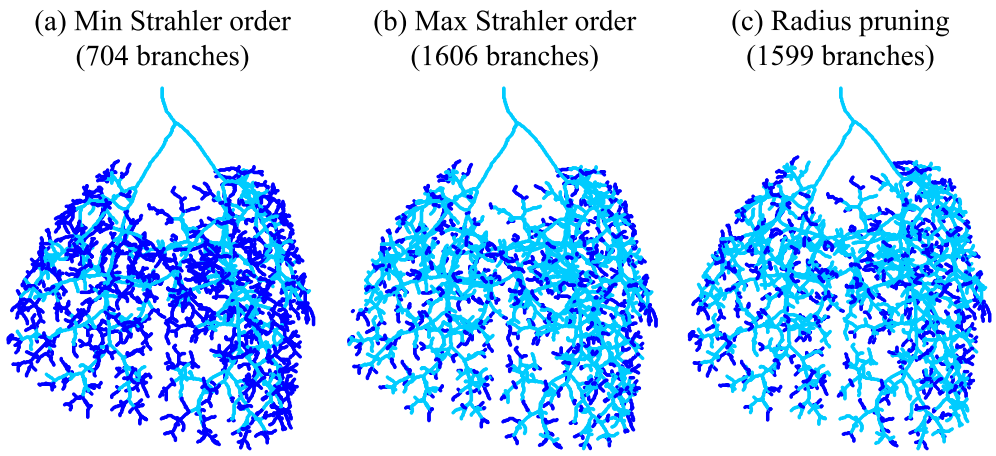


**Fig. 7.** Pruned trees, shown in light blue, overlaid on their original trees in dark blue. The pruned trees are obtained via the three pruning algorithms defined in Section 2.5: (a) minimum Strahler order pruning, (b) maximum Strahler order pruning, and (c) radius pruning.

**Table 2**
Directional complexities from micro-CT images at four different contrast pressures (6.3, 7.4, 13.0, and 17.2 mmHg). Abbreviations used are "C" for control trees, "HPH" for unpruned HPH trees, "$HPH^R$" for radius-pruned HPH trees, "$HPH^M$" for maximal Strahler-order pruned HPH trees, and "$C^m$" and "$HPH^m$" for the minimal Strahler-order pruned control and HPH trees, respectively. All values are reported as mean ± standard deviation.

| Type | Directional complexity | | | | | |
|---|---|---|---|---|---|---|
| | Right | Left | Ventral | Dorsal | Posterior | Anterior |
| Pressure 6.3 mmHg | | | | | | |
| C | 3145 ± 235 | 2679 ± 530 | 2546 ± 456 | 3329 ± 494 | 4502 ± 551 | 1648 ± 423 |
| HPH | 5028 ± 1641 | 5222 ± 1600 | 4460 ± 1341 | 5843 ± 1495 | 6716 ± 1339 | 3659 ± 1369 |
| $HPH^R$ | 2582 ± 338 | 2717 ± 153 | 2160 ± 85 | 3198 ± 25 | 3788 ± 155 | 1678 ± 54 |
| $HPH^M$ | 3745 ± 1249 | 3840 ± 1160 | 3189 ± 925 | 4383 ± 1126 | 5043 ± 980 | 2601 ± 950 |
| $C^m$ | 2627 ± 645 | 2249 ± 820 | 2120 ± 553 | 2801 ± 949 | 3836 ± 830 | 1363 ± 645 |
| $HPH^m$ | 1941 ± 595 | 1968 ± 394 | 1530 ± 246 | 2431 ± 402 | 2930 ± 274 | 1195 ± 293 |
| Pressure 7.4 mmHg | | | | | | |
| C | 3357 ± 299 | 2887 ± 421 | 2784 ± 559 | 3534 ± 264 | 4718 ± 638 | 1810 ± 288 |
| HPH | 5311 ± 1634 | 5549 ± 1418 | 4765 ± 1143 | 6135 ± 1209 | 7070 ± 1262 | 3904 ± 1221 |
| $HPH^R$ | 2814 ± 386 | 2908 ± 61 | 2339 ± 118 | 3435 ± 158 | 4074 ± 159 | 1852 ± 110 |
| $HPH^M$ | 3313 ± 731 | 3427 ± 815 | 2804 ± 636 | 3984 ± 700 | 4593 ± 767 | 2302 ± 613 |
| $C^m$ | 2527 ± 807 | 2244 ± 843 | 2084 ± 753 | 2724 ± 975 | 3742 ± 1021 | 1353 ± 654 |
| $HPH^m$ | 1917 ± 256 | 1991 ± 414 | 1544 ± 313 | 2432 ± 204 | 2980 ± 318 | 1197 ± 207 |
| Pressure 13.0 mmHg | | | | | | |
| C | 4254 ± 706 | 3962 ± 872 | 3771 ± 500 | 4563 ± 1075 | 5821 ± 628 | 2728 ± 906 |
| HPH | 5912 ± 1859 | 6203 ± 1761 | 5368 ± 1397 | 6797 ± 1481 | 7700 ± 1600 | 4445 ± 1426 |
| $HPH^R$ | 4035 ± 530 | 4130 ± 314 | 3454 ± 218 | 4752 ± 156 | 5477 ± 125 | 2828 ± 233 |
| $HPH^M$ | 3397 ± 441 | 3547 ± 840 | 2926 ± 733 | 4059 ± 711 | 4643 ± 793 | 2361 ± 529 |
| $C^m$ | 2702 ± 450 | 2367 ± 349 | 2270 ± 508 | 2879 ± 364 | 3812 ± 619 | 1493 ± 223 |
| $HPH^m$ | 1894 ± 256 | 1973 ± 259 | 1489 ± 190 | 2400 ± 50 | 2886 ± 173 | 1183 ± 75 |
| Pressure 17.2 mmHg | | | | | | |
| C | 5677 ± 796 | 5462 ± 737 | 5197 ± 531 | 6110 ± 1135 | 7400 ± 487 | 4040 ± 955 |
| HPH | 6619 ± 2148 | 7002 ± 2064 | 6134 ± 1655 | 7647 ± 1854 | 8455 ± 1891 | 5191 ± 1659 |
| $HPH^R$ | 5623 ± 1284 | 5797 ± 1091 | 5039 ± 761 | 6476 ± 944 | 7256 ± 922 | 4208 ± 811 |
| $HPH^{M[a]}$ | 6619 ± 2148 | 7002 ± 2064 | 6134 ± 1655 | 7647 ± 1854 | 8455 ± 1891 | 5191 ± 1659 |
| $C^{m[a]}$ | 5677 ± 796 | 5462 ± 737 | 5197 ± 531 | 6110 ± 1135 | 7400 ± 487 | 4040 ± 955 |
| $HPH^m$ | 4432 ± 749 | 4578 ± 704 | 3913 ± 449 | 5232 ± 608 | 5854 ± 505 | 3213 ± 530 |

[a]At 17.2 mmHg, the $HPH^M$ and $C^m$ are the same as their uncropped versions.

The Strahler order pruned trees have approximately the same complexity with increased pressure (except for the highest pressure) (Fig. 6). These trees mostly include larger vessels, which have lower complexity in HPH animals than in control (Fig. 8(b). Again the trees remodel more in the right, ventral, and posterior directions compared to the right, dorsal, and anterior directions.

The maximum Strahler order pruned trees are challenging to compare, likely because the control trees may not be "maximal", i.e., it is possible to add branches without increasing the Strahler order. As a result, the branch count varies significantly for maximum Strahler order trees.

### 3.3. Computational code

Segmented images and computational code used for analysis are provided in the GitHub repositories "mjchambers/Directed_Tree_Extractor" and "mjchambers/TDA _MousePulmonary". To perform the height filtration, we use the repository "ksian/ ML2015FP/3TDATools", created by researchers at Duke University in 2014 [39].

## 4. Discussion

Pulmonary vascular remodeling observed in HPH has been characterized by increased vessel stiffness and chronic dilation of large arteries [10]. This observation was also noted by Strielkov et al. [38], who reported that vessels >500 μm dilate chronically while the small arteries and arterioles (<200 μm) constrict. It has also been shown that the arterioles rarefy [7]. The diameters listed above are measured in unstressed vessels after the animals were sacrificed.

This study uses combinatorial and topological methods to characterize control and HPH arterial tree remodeling, quantifying the number of branches and leaves, tree depth, Strahler order, and directional complexities. For each tree, the minimum radius captured during the segmentation of micro-CT images is approximately 39 μm, with most vessels having a radius greater than 50 μm (97.3% of vessels). As a result, we can quantify the effects of dilation and constriction but not the impact of microvascular rarefaction.

### 4.1. Spatial tree statistics

For the control trees, the tree depth and branch count increase with pressure. The increase is higher than for HPH trees which have stiffer vessels, but the tree depth and branch count are higher in HPH than in control trees. The latter can be explained by chronic dilation making more vessels visible in the CT images. However, the increased number of branches in HPH trees is an artifact of the imaging process and does not translate to the HPH mice having more pulmonary arteries. To eliminate the effects of tree size, we pruned the trees and computed the relative change in tree depth. Results show that the tree depth is higher in control, a phenomenon preserved after pruning.

For the Strahler order pruned trees, the tree depth and branch count are approximately constant following the Strahler order ($SO = 6$). This

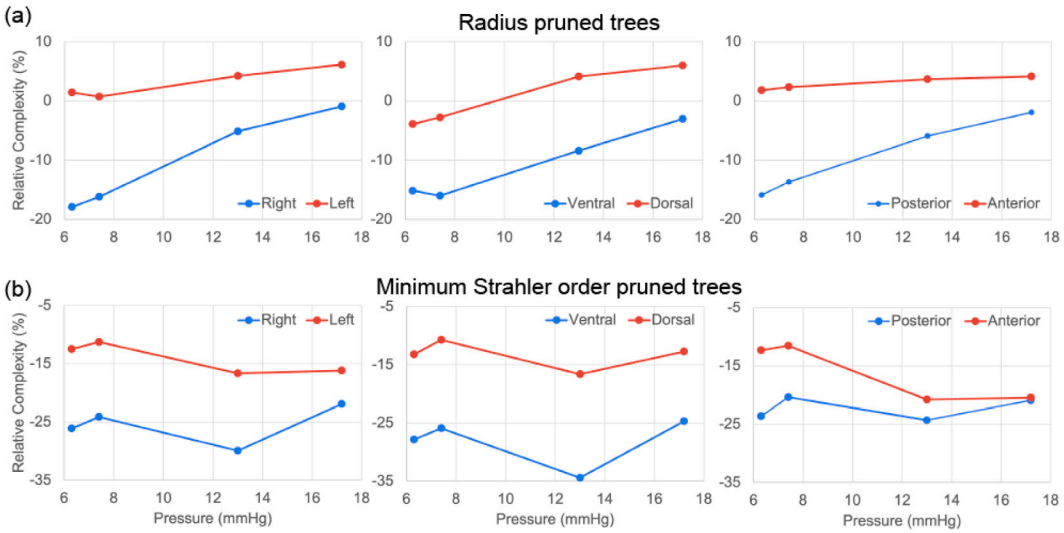


**Fig. 8.** Relative directional complexity of radius pruned trees (a) and minimum Strahler order pruned trees (b). Each is mapped on the mouse in Fig. 4. Each panel shows complexity in two opposing directions (left panel: right/left, center panel: dorsal/ventral, right panel: posterior/anterior) as mapped on the mouse in Fig. 4. The graphs show the relative change in percent relative to the control computed as $100\%\times(\text{HPH}^i\text{-control})/\text{control}$, $i = R, m$ for the radius and minimum Strahler order pruned trees, respectively.

holds for all trees except for the highest pressure, where the Strahler order is 7. The radius pruned trees are constructed to have the same number of branches as the control trees. Therefore, changes in relative tree depth follow that of control. Comparison of radius pruned trees is essential as it allows us to quantify topological remodeling of different vessel sizes, whereas comparison of Strahler order pruned trees enables us to study remodeling in trees not changing with pressure. In the remainder of this discussion, for the radius pruned trees, we refer to vessels visible at the lowest perfusion pressure as the "large vessels" and those visible at the highest perfusion pressure as "small vessels".

### 4.2. Directional complexity

The radius pruned HPH trees have slightly higher left, dorsal, and anterior complexity than the control trees, while the right, ventral, and posterior complexity is lower. Therefore in HPH trees, more or longer vessels point toward the head and the left, while fewer or shorter vessels point in the other directions. This is illustrated in Fig. 8 showing the relative complexity computed as the percent change $100\%\times$(HPH-control)/control. This may indicate that the network is effectively adapting to transport blood to the heart (left) and upper part of the lung.

In general, the directional complexities correlate with the number of branches. By design, the radius-pruning algorithm makes the number of branches in HPH trees close to that of controls. The directional complexity results from these trees are significant, as they allow us to study vessels of different size. With the exception of the highest pressure, the minimum Strahler order pruning is useful as it generates trees with similar vessel counts (Strahler order) independent of pressure. Therefore topological analysis of Strahler order cropped trees are done at all but the highest pressure.

Results of radius and minimum Strahler order pruning reveal that remodeling differentiates between large and small vessels. For the larger vessels (at lower pressures), HPH trees have significantly lower relative complexity than the control trees. This indicates that the network may experience remodeling and/or rarefaction. However, as the networks are perfused at higher pressure, the complexity of the radius pruned HPH tree increases, becoming closer to that of the control tree. This could indicate that remodeling changes in the smaller, more vessels with an unstressed diameter below 200 μm, which are known to constrict. More datasets should be analyzed to differentiate remodeling at different vessel sizes. In particular, since the values reported here reflect the complexity of the whole tree. Finally, the maximum Strahler order pruned trees are harder to analyze. This is likely because the control trees may not "maximal", i.e., adding vessels to these trees may not increase the Strahler order. Therefore, we recommend focusing on radius, and minimum Strahler order pruned trees.

Our use of 0-dimensional persistent homology via height filtration is inspired by Belchi et al. [20], who found that humans with mild to moderate COPD have lower upwards directional complexity of the bronchial network compared to healthy subjects. In this context, "upwards" directional complexity refers to bronchi pointing toward the head. Since HPH is the type of PH most commonly associated with COPD [3] and the pulmonary arteries branch like the airways, we implemented the same filtration as [20]. Comparing our methods to those of Belchi, we need to consider that mice are quadrupeds, their lungs are rotated compared to humans. Thus, dorsal or anterior complexity might be a more apt comparison to human upwards complexity because they are in the direction of the head and countering gravity. Our data show that the dorsal and anterior complexity is higher in the radius-pruned HPH trees. This finding is likely an indication of the differences between humans and mice, and pulmonary arteries vs airways, or it may appear due to the lack of data: we do not have enough animals to generate statistically significant conclusions.

We analyzed the response at several perfusion pressure providing insight into how the large vs. distal vessels remodel. Note that while Belchi et al.'s study analyzed data from more subjects, their trees had fewer generations and lower depth. In contrast, our study examined fewer images, but the generated larger trees had up to 35 generations.

### 4.3. Future work and limitations

This study's primary limitation is the lack of data, we analyzed data from three control and three HPH animals. Lack of data is a common problem in physiological studies as it can be difficult to conduct experiments with many animals or obtain data due to patients' privacy concerns. One way to remedy missing data is using generative machine learning to construct surrogate trees representative of the actual trees, an approach we plan to pursue in future studies. Though before doing this, more work is needed to characterize arterial trees from more animals.

This approach was used in the persistent homology study by Bendich et al. [19] analyzing brain arterial trees. Using machine learning, they generated 98 trees from repeated iterations of a tube-tracking

algorithm. As a first step toward fitting the variables in generative models, we will use simulated trees obtained by attaching self-similar structured trees to principal branches of labeled spatial trees [10] and incorporating angles extracted from our spatial trees to capture the three-dimensional properties of arterial branching. While this approach can generate many networks, more work is needed to ensure that the networks follow the actual topological structure of the arterial trees.

Another challenge is translating the results reported here into the analysis of clinical CT images. Human CT images have lower resolution, containing arteries, airways, and veins, making it challenging to identify the arteries. Yet, the techniques proposed here are promising. In a clinical CT image, it is possible to identify 2-300 vessels compared to the 2000 vessels visible in the micro-CT images analyzed here. Moreover, the mouse images are ideal because the pulmonary arteries were excised, i.e., the arterial trees were the only anatomical objects in the images.

## 5. Conclusion

This study serves as a proof of concept for TDA to identify differences between control and hypoxia-induced pulmonary hypertension (HPH) arterial trees. We lay the foundation and explore the possibilities of identifying features of vascular trees by combining numerical and topological data analysis to study the shape of control and HPH arterial trees. Data consists of spatial trees extracted from micro-CT images in control and HPH mice. The HPH animals experience chronic dilation of the larger vessels, therefore more vessels are visible in the micro CT images. As a result, there is a significant discrepancy between branch counts between the two groups. This result does not reflect that HPH animals have more branches but is a consequence of imaging resolution. To address this limitation, we devised radius and Strahler order pruning algorithms to obtain comparable trees. The latter was done because pulmonary arterial trees branch asymmetrically. In addition to standard graph statistics, we use a numerical summary derived from persistent homology called directional complexity [20] to characterize and compare the pulmonary arterial trees.

Several main findings and major methodological observations arise from this study. The tree depth is higher in HPH trees, indicating that HPH mice experience chronic dilation of the large vessels. A major methodological observation is that we can extract more information by applying different pruning techniques. Analysis of radius pruned trees show that at low perfusion pressure, HPH trees have lower relative directional complexity in 3 of the 6 directions than control, and complexity increases as more vessels are included. The latter agrees with the understanding that large and small pulmonary arterial vessels remodel differently. Analysis of minimum Strahler order pruned, trees allow us to study comparable trees at several pressures. Results agree with the analysis of low-pressure radius pruned trees, namely that HPH complexity is lower than control.

The next steps include overcoming the major limitation in this study, the lack of data. To do so we propose to include more trees and use generative machine learning methods to generate surrogate trees. The most exciting research direction is expanding TDA methods to higher dimensional persistence and analyzing human lungs, the vascular network, and the airways, as their structures and shapes are intertwined and dependent.

## Declaration of competing interest

None.

## Data availability

We have included links to code repository, but the authors do not own the raw images.

## Acknowledgments

This study was initiated during a summer research experience for undergraduates at North Carolina State University, funded by the NSA: H98230-20-1-0259, H98230-21-1-0014, H98230-22-1-0006. Olufsen was supported in part by NSF-DMS 2051010, 1615820 and by the NIH-HLBI R01HL147590. RS was partially supported by the NSF-DMS 1854705. The authors also thank Naomi Chesler, University of California-Irvine, for providing micro-CT images for this study.

## References

[1] World Health Organization 2011. Cardiovascular diseases (CVDs). https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

[2] S. Beshay, A. Guha, S. Sahay, Evaluation, diagnosis and classification of pulmonary hypertension, Methodist Debakey Cardiovasc. J. 17 (2021) 86–91, http://dx.doi.org/10.14797/OCDF4453.

[3] N. Galié, M. Humbert, J.-L. Vachiery, S. Gibbs, I. Lang, A. Torbicki, G. Simonneau, A. Peacock, A. Noordegraaf, M. Beghetti, A. Ghofrani, M. Sanchez, G. Hansmann, W. Klepetko, P. Lancellotti, M. Matucci, T. McDonagh, L. Pierard, P. Trindade, M. Zompatori, M. Hoeper, 2015 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension: The joint task force for the diagnosis and treatment of pulmonary hypertension of the European society of cardiology (ESC) and the European respiratory society (ERS): Endorsed by: Association for European paediatric and congenital cardiology (AEPC), international society for heart and lung transplantation (ISHLT), Eur. Heart J. 37 (2015) 67–119, http://dx.doi.org/10.1093/eurheartj/ehv317.

[4] G. Simonneau, D. Montani, D. Celermajer, C. Denton, M. Gatzoulis, M. Krowka, P. Williams, R. Souza, Haemodynamic definitions and updated clinical classification of pulmonary hypertension, Eur. Respir. J. 53 (2019) 1801913, http://dx.doi.org/10.1183/13993003.01913-2018.

[5] E. Lau, M. Humbert, D. Celemajer, Early detection of pulmonary arterial hypertension, Nat. Rev. Cardiol. 12 (2015) 143–155, http://dx.doi.org/10.1038/nrcardio.2014.191.

[6] R. Molthen, K. Karau, C. Dawson, Quantitative models of the rat pulmonary arterial tree morphometry applied to hypoxia-induced arterial remodeling, J. Appl. Physiol. 97 (2004) 2372–2384, http://dx.doi.org/10.1152/japplphysiol.00454.2004.

[7] R. Vanderpool, A. Kim, R. Molthen, N. Chesler, Effects of acute rho kinase inhibition on chronic hypoxia-induced changes in proximal and distal pulmonary arterial structure and function, J. Appl. Physiol. 110 (2011) 188–198, http://dx.doi.org/10.1152/japplphysiol.00533.2010.

[8] N. Hopkins, P. McLoughlin, The structural basis of pulmonary hypertension in chronic lung disease: remodelling, rarefaction or angiogenesis? J. Anat. 201 (2002) 335–348, http://dx.doi.org/10.1046/j.1469-7580.2002.00096.x.

[9] C. Murray, The physiological principle of minimum work applied to the angle of branching of arteries, J. Gen. Physiol. 9 (1926) 835–841, http://dx.doi.org/10.1085/jgp.9.6.835.

[10] M. Chambers, M. Colebank, M. Qureshi, R. Clipp, M. Olufsen, Structural and hemodynamic properties of murine pulmonary arterial networks under hypoxia-induced pulmonary hypertension, Proc. Inst. Mech. Eng. H: J. Eng. Med. 234 (2020) 1312–1329, http://dx.doi.org/10.1177/0954411920944110.

[11] K. Horsfield, Morphometry of the small pulmonary arteries in man, Circ. Res. 42 (1977) 593–597, http://dx.doi.org/10.1161/01.res.42.5.593.

[12] S. Singhal, R. Henderson, K. Horsfield, K. Harding, G. Cumming, Morphometry of the human pulmonary arterial tree, Circ. Res. 33 (1973) 190–197, http://dx.doi.org/10.1161/01.RES.33.2.190.

[13] M. Zamir, Nonsymmetrical bifurcations in arterial branching, J. Gen. Physiol. 72 (1978) 837–845, http://dx.doi.org/10.1085/jgp.72.6.837.

[14] M. Olufsen, Structured tree outflow condition for blood flow in larger systemic arteries, Am. J. Physiol. 276 (1999) H257–H268, http://dx.doi.org/10.1152/ajpheart.1999.276.1.H257.

[15] M. Olufsen, C. Peskin, W. Kim, E. Pedersen, A. Nadim, J. Larsen, Numerical simulation and experimental validation of blood flow in arteries with structured-tree outflow conditions, Ann. Biomed. Eng. 28 (2000) 1281–1299, http://dx.doi.org/10.1114/1.1326031.

[16] V. Davidoiu, L. Hadjilucas, I. Teh, N. Smith, J. Schneider, J. Lee, Evaluation of noise removal algorithms for imaging and reconstruction of vascular networks using micro-CT, Biomed. Phys. Eng. Exp. 2 (2016) 045015, http://dx.doi.org/10.1088/2057-1976/2/4/045015.

[17] R. Ghrist, Barcodes: the persistent topology of data, Bull. AMS 45 (2008) 61–75.

[18] V. Nanda, R. Sazdanovic, Simplicial models and topological inference in biological systems, in: N. Jonoshka, M. Saito (Eds.), Discrete and Topological Models in Molecular Biology, Springer, Heidelberg, Germany, 2014, pp. 109–141.

[19] P. Bendich, J. Marron, E. Miller, A. Pieloch, S. Skwerer, Persistent homology analysis of brain artery trees, Ann. Appl. Stat. 10 (2016) 198–218, http://dx.doi.org/10.1214/15-AOAS886.

[20] F. Belchi, M. Pirashvili, J. Conway, M. Bennett, R. Djukanovic, J. Brodzki, Lung topology characteristics in patients with chronic obstructive pulmonary disease, Sci. Rep. 8 (2018) 5341, http://dx.doi.org/10.1038/s41598-018-23424-0.

[21] R. Chorley, R. Horton, Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology, GSA Bull. 56 (1945) 275–370, http://dx.doi.org/10.1130/0016-7606(1945)56[275,EDOSAT]2.0.CO;2.

[22] A. Strahler, Hypsometric (area altitude) analysis of erosional topology, Geol. Soc. Am. Bull. 63 (1952) 1117–1142, http://dx.doi.org/10.1130/0016-7606(1952)63[1117,HAAOET]2.0.CO;2.

[23] S. Archer, E. Michelakis, The mechanism(s) of hypoxic pulmonary vasoconstriction: potassium channels, redox o(2) sensors, and controversies, News Physiol. Sci. 17 (2002) 131–137, http://dx.doi.org/10.1152/nips.01388.2002.

[24] F. Kuhr, K. Smith, M. Song, I. Levitan, J. Yuan, New mechanisms of pulmonary arterial hypertension: role of $Ca^2$+ signaling, Am. J. Physiol. 302 (2012) H1546–H1562, http://dx.doi.org/10.1152/ajpheart.00944.2011.

[25] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J. Miller, S. Pieper, R. Kikinis, 3D slicer as an image computing platform for the quantitative imaging network, Mag. Reson. Imaging 30 (2012) 1323–1341, http://dx.doi.org/10.1016/j.mri.2012.05.001.

[26] R. Kikinis, S. Pieper, K. Vosburgh, 3D slicer: a platform for subject-specific image analysis, visualization, and clinical support, in: F. Jolesc (Ed.), Intraoperative Imaging and Image-Guided Therapy, Springer, New York, NY, 2014, pp. 277–289, http://dx.doi.org/10.1007/978-1-4614-7657-3_19.

[27] Kitware Inc., 3D slicer, 2020, https://www.slicer.org/.

[28] M. Couprie, G. Bertrand, Asymmetric parallel 3D thinning scheme and algorithms based on isthmuses, Pattern Recog. Lett. 76 (2016) 22–31, http://dx.doi.org/10.1016/j.patrec.2015.03.014.

[29] DGtal: Digital Geometry Tools and Algorithms Library 2019. DGtal 1.0. http://dgtal.org.

[30] P. Hernandez, Spatial graph extractor (SGEXT), 2018, https://github.com/phcerdan/SGEXT.

[31] P. Hernandez-Cerdan, Biopolymer Networks: Image Analysis, Reconstruction and Modeling PhD Dissertation, Massey University, Manawat#u, New Zealand, 2018.

[32] E. Vanbavel, J. Spaan, Branching patterns in the porcine coronary arterial tree estimation of flow heterogeneity, Pharm. Ther. 71 (1992) 1200–1212, http://dx.doi.org/10.1161/01.RES.71.5.1200.

[33] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, Stability of persistence diagrams, Discret. Comp. Geom. 37 (2007) 103–120, http://dx.doi.org/10.1007/s00454-006-1276-5.

[34] H. Edelsbrunner, E. Mücke, Three-dimensional alpha shapes, ACM Trans. Graph. 13 (1994) 43–72, http://dx.doi.org/10.1145/174462.156635.

[35] V. Robins, Towards computing homology from finite approximations, Topol. Proc. 24 (1999) 503–532.

[36] E. Amezquita, M. Quigley, T. Ophelders, E. Munch, D. Chitwood, The shape of things to come: Topological data analysis and biology, from molecules to organisms, Dev. Dyn. 249 (2020) 816–833, http://dx.doi.org/10.1002/dvdy.175.

[37] A. Hatcher, Algebraic Topology, Cambridge University Press, Cambridge, UK, 2001.

[38] I. Strielkov, N. Krause, N. Sommer, R. Schermuly, H. Ghofrani, F. Grimminger, T. Gudermann, Dietrich, N. Weissmann, Hypoxic pulmonary vasoconstriction in isolated mouse pulmonary arterial vessels, Exp. Physiol. 103 (2018) 1185–1191, http://dx.doi.org/10.1113/EP087117.

[39] J. Harer, R. Bar-On, N. Strawn, C. Tralie, P. Bendich, A. Pieloch, J. Slaczedek, 3Tdatools, 2014, https://github.com/ksian/ML2015FP/tree/master/3TDATools.