ABSTRACT

CHAMBERS, MEGAN JEANNE. Morphological Properties of Pulmonary Arteries in Control and Hypertensive Mice. (Under the direction of Mette Olufsen.)

In computed tomography (CT) scans, one can observe that the pulmonary arteries form a tree-like network. This structure is necessary for distributing deoxygenated blood to all areas of the lung, where it enters the alveolar capillaries and receives oxygen. In patients with pulmonary hypertension (PH), morphological changes in the pulmonary arteries have been observed, a phenomenon known as vascular remodeling. Quantifying these morphological changes is a key step to understanding the effect of PH on the vasculature and may help in early detection of the disease.

In collaboration with Kitware, Inc. we use 3D Slicer, an open source image analysis software, to extract representative spatial trees from CT scans of control and pulmonary hypertensive arterial networks in mice. Such trees can contain hundreds or thousands of vessels. Pressure and flow in these vessels can be determined by solving a 1D fluid dynamics model, but this is computationally expensive to solve explicitly in every vessel. Instead, we model the larger vessels in terms of their geometry and organize the smaller vessels into self-similar structured trees defined by parameters such as length-to-radius ratio, radius scaling factors, and 3D branching angles.

This dissertation uses a data-driven approach to analyze the effects of PH on the morphological properties of the pulmonary arterial network. We find that with PH, arteries become stiffer and have an increased length-to-radius ratio, while radius scaling factors do not change. We also find that our PH trees have more branches than trees extracted from control specimens, which is in conflict with previous findings and is believed to be an artifact of the imaging process. Using systematic pruning algorithms to minimize the effect of branch count on other properties of interest, we determine that tree depth was higher with PH in nearly all cases. We also use persistent homology, a topological data analysis technique which can characterize the shape of data, to capture the network complexity in different directions. Ultimately, more data is needed to obtain strong topological results; simulated data can be useful for this since raw data is not plentiful. To generate realistic simulated data, we extract arterial branching angles from our trees and implement ABC analysis to infer which angle values provide simulated data that closely resembles the raw data we have observed.
Morphological Properties of Pulmonary Arteries in Control and Hypertensive Mice

by
Megan Jeanne Chambers

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina
2022

APPROVED BY:

________________________   _________________________
Radmila Sazdanovic           Sharon Lubkin

________________________   _________________________
Srijan Sengupta              Rachel Clipp

________________________
Mette Olufsen
Chair of Advisory Committee
DEDICATION

To my mom, dad, and Andrew
Megan grew up in Boardman, Ohio, where she attended school until graduating in 2011. While in school, she enjoyed and excelled in her mathematics classes, which was largely due to her exceptional math educators. In her senior year, her AP Calculus class attended the annual Mathfest event at Youngstown State University (YSU), which involved various mathematical workshops, speakers, and a problem solving competition. This experience connected Megan with professors in their math department, who inspired her to want to study mathematics in college. Upon graduating from high school, Megan enrolled in YSU as a mathematics major. She became heavily involved in the Pi Mu Epsilon Ohio Xi Chapter and the Association for Women in Mathematics. While pursuing her degree, her professors advised her on independent research projects, which allowed her to travel to conferences and present her work. She also completed summer REU programs at Carleton College in Northfield, Minnesota and the University of Hawai’i at Hilo. These experiences caused her love of mathematics to grow. After receiving her B.S. in Applied Mathematics in 2015, she moved to Raleigh and began graduate school at North Carolina State University (NCSU).

Megan entered her Ph.D. program as a pure mathematics major. She soon switched to applied mathematics but maintained an interest on how pure and applied math could be combined to solve problems. In her third year at NCSU, she took a course in fluid dynamics with Dr. Mette Olufsen and became more interested in biological applications of mathematics. With the encouragement of Dr. Olufsen and her postdoctoral scholar, Dr. Umar Qureshi, Megan became involved in research with the Cardiovascular Dynamics Group. This gave her the opportunity to spend the summer at Kitware, Inc., a nearby company working in image analysis software. She learned to use their programs to collect data for her thesis project on network analysis of mouse pulmonary arteries. In addition to this collaboration with Kitware, Megan has received invaluable assistance from the SofTMech program at University of Glasgow and her undergraduate mentees during NCSU’s virtual REU in summer 2020. After her defense, she will continue to work at Virginia Military Institute (VMI), where she has been working as a mathematics lecturer since August 2021.
ACKNOWLEDGEMENTS

While I am listed as the author of this dissertation, it would not exist without the many individuals who supported me for many years. To begin, I want to thank God for blessing me with a mathematical mind and keeping me going throughout this journey.

I express my deepest gratitude to my advisor Mette Olufsen for her encouragement, insight, and patience through this process. I am grateful to her for taking me on as an advisee and constantly reassuring me of my abilities, especially when I frequently doubted them myself. I also wish to thank Radmila Sazdanovic for providing instrumental feedback on the topological side of my work, and reminding me to be kind to myself during this difficult process. Thank you to Rachel Clipp, Pablo Hernandez, and the entire team at Kitware for teaching me all about their software and giving me the tools I needed to obtain the data which made this dissertation possible. I am also thankful for Sharon Lubkin and Srijan Sengupta for graciously offering their time to serve as my committee members and for the comments they have provided on my work.

I am grateful for the additional collaborators who have contributed to my research in various ways over the years. Thank you to Naomi Chesler for providing the images analyzed in this dissertation. Thank you to Umar Qureshi and Paweł Dłotko for helping me in the early stages of the graph analysis process. Thank you to Nick Hill and Dirk Husmeier for teaching me about 3D branching angles and ABC analysis, as well as the entire SofTMech group for your input during our weekly discussion meetings. Thank you to Natalie Johnston, Ian Livengood, Miya Spinella, Mariam Kharbat, and Robert Sternquist for their hard work and during the 2020 REU and for their continued dedication as we work to publish our findings. Thank you to Mitchel Colebank for being a valued collaborator and friend, and to the entire CDG research group for giving me opportunities to discuss my research among peers. I also am thankful for the funding I have received from the NSF-DMS grant #1615820, NSA Contract H98230-20-1-0259, and the AWM NSF Workshop grant.

For fostering my love of higher level mathematics, I would like to thank my high school and undergraduate math educators, including Carrie Davis, Nathan Ritchey, Angela Spalsbury, and Alicia Prieto. You each taught me what it means to be a mathematician and prepared me to further my education in graduate school. I am also grateful to Troy Siemers and the VMI Applied Mathematics department for supporting me in my first faculty position.

Thank you to my family for always encouraging me to follow my passion and to never give up on myself. Thank you Mom, Dad, Paul, Nana & Papa, and Oma & Opa, for all your love and support. To my friends, thank you for being a source of light and joy during this grueling process and for always being my biggest cheerleaders. Thank you to Sarah, Blain, Nikki, Emily, Amanda, and Ivy for all the encouragement and advice you have given me.

Finally, I wish to thank my fiancé, Andrew Miller, for his unending love and encouragement. Thank you for making me tea and food during massive writing sessions, cracking me up with our favorite TV references and dad jokes, and enduring many late nights and tearful breakdowns. You are my best friend, and I am so excited for the life we have ahead of us.
# TABLE OF CONTENTS

**LIST OF TABLES** ................................................................. viii

**LIST OF FIGURES** ............................................................. x

**Chapter 1** Introduction .................................................. 1

**Chapter 2** Cardiovascular Physiology ................................. 4
  2.1 The cardiovascular system ............................................. 4
  2.2 Pulmonary hypertension ............................................... 7

**Chapter 3** Morphometry of the Pulmonary Vasculature ............ 9
  3.1 Vessel dimensions ..................................................... 9
  3.1.1 Radius ................................................................... 10
  3.1.2 Length ................................................................... 14
  3.1.3 Wall thickness ....................................................... 15
  3.2 Branching angles ....................................................... 15

**Chapter 4** Processing Computed Tomography Images ............. 19
  4.1 Micro-CT images ......................................................... 21
  4.2 Segmented arteries ...................................................... 21
  4.3 Distance map ............................................................. 25
  4.4 Centerline skeleton ..................................................... 25
    4.4.1 Critical cliques .................................................. 26
    4.4.2 Centerlines ......................................................... 27
    4.4.3 Vessel radii ......................................................... 28
  4.5 Labeled spatial tree .................................................... 29
    4.5.1 Graph theory ..................................................... 29
    4.5.2 Spatial graph ..................................................... 30
    4.5.3 Labeled spatial tree ............................................. 31
  4.6 Multiscale geometric model .......................................... 32
    4.6.1 Principal pathway ............................................... 32
    4.6.2 Small vessels .................................................... 33
  4.7 Summary ................................................................. 34

**Chapter 5** Persistent Homology .......................................... 36
  5.1 Persistent homology background .................................... 37
    5.1.1 Homology basics ................................................. 37
    5.1.2 Persistent homology ............................................ 39
  5.2 TDA for biological data analysis ..................................... 42

**Chapter 6** Structural and Hemodynamic Properties of Murine Pulmonary Arterial Networks under HYpoxia-induced Pulmonary Hypertension ......................................................... 43
  6.1 Abstract ................................................................. 43
  6.2 Introduction .............................................................. 44
    6.2.1 Geometry of pulmonary arterial vasculature ............... 44
    6.2.2 Pulmonary arterial hemodynamics ............................ 46
LIST OF TABLES

Table 2.1 Total blood volume and mean blood pressure of components of the circulatory system [76]. Abbreviations used are P (pulmonary), S (systemic), A (arteries), V (veins) and C (capillaries), as defined in Figure 2.1. 6

Table 4.1 Examples of segmented images generated by different lower thresholds $\tau_L$. Notice that $\tau_L = 13$ is too low, as noise in the image was labeled as part of the arterial network. By contrast, $\tau_L = 65$ is too high, as voxels which clearly belong to arteries were not included in the foreground. For the images in this dissertation, we found $\tau_L = 18$ to be a good starting point for the lower threshold and adjusted this in an ad hoc method for each image. 22

Table 4.2 Examples of segmented images generated using lower threshold $\tau_L = 18$ and median smoothing with different sized smoothing kernels. In our study, we applied median smoothing with a kernel size of $3 \times 3 \times 3$ voxels. This kernel size allowed us to appropriately smooth out any noisy surfaces without removing many distal branches. When smoothing is applied, small islands often appear; some examples of these are circled in red in the images. 23

Table 4.3 Summary of methods for labeled spatial graph extraction (part 1). 34

Table 4.4 Summary of methods for labeled spatial graph extraction (part 2). 35

Table 6.1 Graph terminology. 52

Table 6.2 Network information. The $\ell_{rr}$ constants $C_1$ and $C_2$ are reported as coefficients $\pm$ margin of error for a 95% confidence interval. All other values are mean $\pm$ std. deviation. 61

Table 6.3 Stiffness parameters used in model simulations. PP-Principal pathway; ST Structured tree. 63

Table 7.1 Spatial tree statistics from micro-CT images at 4 different contrast pressures. 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 $SO_{MPA} = 7$, and the other two had $SO_{MPA} = 6$. All other values are reported as mean $\pm$ standard deviation. Data for HPH$^M$ at pressure 17.2 mmHg is missing because the control and original HPH trees at this pressure already have the same Strahler order. 86

Table 7.2 Directional complexities of trees from micro-CT images at 4 different contrast pressures in 6 different directions. 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 $\pm$ standard deviation. Data for HPH$^M$ at pressure 17.2 mmHg is missing because the control and original HPH trees at this pressure already have the same Strahler order. 87
Table 8.1  Breakdown of vessel vectors with $n = 0, 1, 2$ change points for 3 control and 3 PH mice (imaged at contrast pressure 6.3 mmHg). Note that the number of total vessel vectors does not equal the total number of branches because a branch may be a daughter of one junction and parent of another, in which case its vessel vector is computed twice for different hub nodes.
LIST OF FIGURES

Figure 2.1  A) The cardiovascular system. Arrows indicate the direction of blood flow. Oxygenated blood flows from the left heart into the SA to be distributed throughout the body. Oxygen and nutrients are exchanged for CO\textsubscript{2} and waste products in the SC. The deoxgenated blood then passes in to the SV. From the SV, it enters the right heart and then the PA. In the PC, blood is re-oxygenated and then enters the PV to be returned to the left heart. Reproduced and modified from [59] with permission. B) The pulmonary circulation. Notice that the PA form a rapidly branching structure, the morphometry of which is the focus of this thesis. Reproduced and modified from [60] with permission.

Figure 2.2  Blood pressure in different components of the circulatory system. When pressure is oscillating, the upper threshold reached is the systolic pressure; the lower bound is the diastolic pressure. Reproduced from [54] with permission.

Figure 2.3  The 5 types of pulmonary hypertension (PH) and their affected areas. I: Pulmonary arterial hypertension (PAH) initially starts in the pulmonary arterioles and progresses to the larger pulmonary arteries. II: PH due to left heart disease, also called pulmonary venous hypertension (PVH), starts in the left atrium and progresses to the pulmonary veins and, potentially, the venules. III: PH due to hypoxia initially affects the pulmonary arterioles and capillaries and progresses to the arteries and venules. IV: Chronic thromboembolic PH (CTEPH) initiates in the pulmonary arteries and progresses down to the arterioles. Finally, V: idiopathic PH refers to PH due to unknown or multifactor causes.

Figure 3.1  A) In humans, the right lung has 3 lobes (superior, middle, and inferior), and the left lung has 2 (superior and inferior) [60]. Images obtained for free personal use from http://clipart-library.com/, modified. B) In mice, the right lung has 4 lobes (superior, middle, inferior, and post-caval), and the left lung is a single lobe [5].

Figure 3.2  A) A section of an artery which we model as a cylinder with radius \( r \), length \( \ell \), and wall thickness \( h \). B) A typical bifurcation with parent/daughter radii, lengths, and 2D branching angles labeled. C) A typical bifurcation with 3D branching angles.

Figure 4.1  Semi-automated algorithm for labeled spatial tree extraction. Solid gray arrows represent automated processes; striped arrows represent processes with a manual component. B) The distance map encodes each voxel with its \((x, y, z)\) coordinates and its distance from the nearest background voxel. (E) In the final labeled spatial tree, each branch is labeled with the corresponding artery’s radius \( r \), length \( \ell \), length-to-radius-ratio \( \ell / r \), and orientation \( o \). Each junction vertex is labeled with the radius scaling parameters \( \alpha \) and \( \beta \) and 3D branching angles \( \psi_{1,2} \) and \( \theta_{1,2} \).
Figure 4.2  A) The hypobaric chamber which contains the arteries during the imaging process was captured in the CT scan. We remove this by selecting the arteries as the island to keep. B) Glares in the image caused many small islands to appear toward the bottom of the network. We remove them by keeping the largest island, which we know to be the arterial network.

Figure 4.3  Segmented network with centerlines before and after manual segmentation was applied to the MPA. Notice that the cannula inserted for imaging distorted the appearance of the MPA, causing a “clump” to form in the centerline skeleton. To improve the skeleton, the distortion must be manually reduced to a more uniform cylindrical shape. Care is taken to remove only erroneous protrusions and not to reduce the actual size of the MPA.

Figure 4.4  A) Segmented human aorta. When the space origin and directions of the original (bottom) are reset to \((0,0,0)\) and \((1,0,0),(0,1,0),(0,0,1)\), the adjusted aorta (top) is shifted and visibly skewed. B) Segmented murine pulmonary arteries. The space directions of the original (left) are already \((1,0,0),(0,1,0),(0,0,1)\). When the space origin is reset to \((0,0,0)\), the adjusted arteries (right) are shifted, but not skewed as the space directions did not change.

Figure 4.5  A) A voxel complex made up of voxels \(\{v_1, ..., v_9\}\). B) \(\mathcal{C}_1\) is a 2-clique. Its voxels, highlighted in blue, share the 2-face \(\chi_2\). \(\mathcal{C}_1\) is essential because it is the largest clique whose intersection is \(\chi_2\). Also, \(\mathcal{C}_1\) is regular because \(N^*(\mathcal{C}_1)\), highlighted in orange, is reducible. C) \(\mathcal{C}_2\) is a 1-clique. Its voxels share the 1-face \(\chi_1\). \(\mathcal{C}_2\) is not essential because it is not the largest clique whose intersection is \(\chi_1\). D) \(\mathcal{C}_3\) is the largest 1-clique whose intersection is \(\chi_1\), so it is essential. It is also critical because \(N^*(\mathcal{C}_3)\), highlighted in orange, is not reducible.

Figure 4.6  An example of the Select function at work. \(S\) is a voxel complex which is a work-in-progress skeleton of a vessel junction (shaded in gray). \(\mathcal{C}\) (in red) is a critical clique of \(S\). Any one of the voxels \(v_1, v_2, \text{ or } v_3\) can be kept, and the others removed, to preserve this critical clique in the skeleton. Our Select(\(\mathcal{C}\)) function chooses \(v_2\) because it has the largest distance map value. Notice how it is also the more centered of the 3 voxels.

Figure 4.7  A) A graph with vertices \(a - h\) and 2 connected components. B) A undirected and directed tree, with the root labeled on the directed tree. C) A spatial graph generated from one of the murine pulmonary artery CT scans. This graph sits in \(\mathbb{R}^3\); each node and edge-point has \((x, y, z)\) coordinates.

Figure 4.8  Sample network shown with the principal pathways (shown in dark blue on top of the light blue network) obtained from different choices of \(\tau_r\). We found \(\tau_r = 0.4\) to be a suitable choice for our networks. If \(\tau_r < 4\), too many distal branches were picked up in the principal pathway. If \(\tau_r > 4\), every lobe was not reached.

Figure 5.1  Example of a 0-, 1-, 2- , and 3- simplex, and a simplicial complex. The simplicial complex \(\mathcal{S}\) comprises five 0-simplices, six 1-simplices, and one 2-simplex.

Figure 5.2  Examples of oriented \(k\)-simplices for \(k = 1, 2, 3\), along with their boundaries.

Figure 5.3  A simplicial complex \(\mathcal{S}\), along with some of its chains and their boundaries. Note from examining \(\mathcal{C}_2\) that a chain does not need to be a connected simplicial complex. Also, note that the chain \(c_3\) forms a loop of 1-simplices. Since all vertices in a loop both begin and end an edge, we find that \(\partial_1(c_3) = 0\).
Figure 5.4 Čech complex filtration applied to a point cloud. A) Visual representation of the filtration, showing the Čech complexes at various \( \epsilon \) values. B) Barcode tracking the evolution of the 0- and 1-features of the complexes. C) Persistence diagram, showing points plotted at coordinates \((\epsilon_B, \epsilon_D)\) for each feature.

Figure 5.5 Height filtration applied to the graph of a pulmonary arterial network. For various \( \epsilon \) values, we have shown the simplicial complex \( S_{\epsilon} \) and stated the degree-0 Betti number of each.

Figure 6.1 (a) Micro-CT image from a control mouse. (b) 3D segmentation and skeletonization (one-voxel wide) networks from representative control and hypertensive mouse. (c) Main pulmonary artery flow (ml/s) and pressure (mmHg) waveforms from 7 control and 5 hypertensive animals.

Figure 6.2 Workflow for the segmentation and graph extraction process. Solid lines represent processes that are fully automated, and dashed lines represent processes requiring user interaction.

Figure 6.3 Example spatial graph before and after merging. Voxels are labeled with coordinates, degree, and node ID. The red edges in the left graph represent a small cycle in \( G \), which is broken in the merged graph \( G^* \).

Figure 6.4 Schematics of the four types of errors that can occur during skeletonization: (a) false branches (FB), (b) double edges (DE), (c) duplicate points (DP), and (d) small cycles (SC). These figures are not drawn to scale but represent the connectivity observed in each type of error. Any gray edges or white nodes with dashed borders are removed to correct the error. Black nodes with white borders are merged into edge points, while white nodes with black borders remain nodes. An exception is the red edge in SC4. To correct this error, this edge must be inserted manually.

Figure 6.5 Example structured tree. A) Radii at bifurcations are determined by multiplying the radius of the parent vessel by \( \alpha \) (for the daughter of larger radius) and \( \beta \). In our fluids simulations, structured trees are attached at all terminal nodes of the principal pathway, shown on the right in blue. B) Projection of the structured tree in 2D noting the \( \alpha \)-branch and the \( \beta \)-branch.

Figure 6.6 Structured tree parameters plotted as a function of the vessel radius. In all graphs, the predictions are grouped into 20 bins, and results include estimates from the 3 control and 3 hypertensive animals. (a) depicts \( \alpha \) and \( \beta \), the horizontal lines denote the mean value for each group. (b) depicts \( \ell_{rr} \) along with a decreasing exponential curve.

Figure 6.7 Pressure predictions in the MPA (marked with red, cyan and blue lines) for each network along with measurements from 5 control and 7 hypertensive animals (marked with gray lines). The shaded gray region denotes the range between the lowest and highest measurements. Results show that model predictions agree well with data and that by adapting vessel stiffness using separate ranges for control and hypertensive animals, we can predict pressure in both control and hypertension.
Figure 6.8  Pressure and flow predictions along the principal pathway and $\alpha/\beta$ structured tree branches in a control and a hypertensive mouse. Mean pressure and flow predictions in the structured tree model are also provided. Note that the pressure drop along the principal pathway is more sudden compared to the pressure drop in the structured tree and that the hypertensive animal has a lower cardiac output. Principal pathway predictions run from the MPA (lightest grey) to the terminal vessel (darkest grey) proximal to the structured tree. Structured tree predictions run through the $\alpha$ side (in red) and the $\beta$ side (in blue), with higher generations denoted by lower magnitude pressure and flow values, as well as a darker color.

Figure 6.9  Mean pressure predictions in the three control and hypertensive mice. For mouse three we show predictions both along the principal pathway and the structured trees, whereas for mice 1 and 2 we only show predictions along the principal pathway. Below each 3D graph we show the mean pressure drop down a representative pathway, with a variation of $\pm2.3$ mmHg in the control mice and $\pm8.5$ mmHg in the hypertensive, accounting for measurement uncertainty. To illustrate results the structured trees were rendered in 3D, though computations are all done in 1D as described in the method section.

Figure 7.1  (a)-(c) The labeled spatial tree extraction process. (a) An example of a micro-CT image of a control mouse's pulmonary arteries (the coronal view). (b) The segmented arteries (a result of segmentation) overlaid with the centerline skeleton (a result of skeletonization). (c) The spatial tree extracted from the skeleton, with branches colored to show vessel radii. Branches are also encoded with other labels, such as orientation and Strahler order. (d) An example of the segmented arteries and skeleton from a control mouse. (e) An example of the segmented arteries and skeleton from a HPH mouse. (f) The structure of the labeled spatial tree in (c) shown in more detail. The tree is comprised of vertices connected by edges. Vertices labeled in blue are points along a segmented artery's length, while orange vertices are branching or terminal points. A series of vertices and edges between 2 branching points or between a branching point and a terminal point is called a branch.

Figure 7.2  The labeled spatial trees from a control mouse (top row) and a HPH mouse (bottom row) extracted from images taken at the 4 different arterial pressures. As pressure increases, the number of branches increases, since a higher arterial pressure perfuses the imaging contrast further down the arterial tree.

Figure 7.3  The Strahler ordering system illustrated on a small tree. All terminal branches have $SO = 1$. Whenever 2 branches of the same Strahler order meet, their parent has Strahler order which is 1 higher than that of its daughters. If 2 branches of different Strahler order meet, their parent is labeled with the maximum Strahler order of its daughters.

Figure 7.4  Examples of pruned trees for each of our 3 pruning algorithms. Pruned trees are shown in light blue, overlaid on their original trees in dark blue. Any dark blue branches that are visible represent branches that are removed from the original tree during pruning. The pruned trees are obtained via the 3 pruning algorithms defined in Section 7.3.5: (a) minimum Strahler order pruning, (b) maximum Strahler order pruning, & (c) radius pruning.
Figure 7.5 The 6 height filtration directions, 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.

Figure 7.6 Directional filtration in the $z_-$ direction illustrated on a labeled spatial tree. (a) shows the tree with several examples of the subtrees $S$ for different values of $t$, which is the distance in the $z_-$ direction from the top of the tree. These subtrees are labeled with their 0 degree Betti numbers $\beta_0$. At $t = 50$, $S_{50}$ is a portion of the root branch of the tree, which is 1 connected component and thus, $\beta_0 = 1$. $S_{100}$ includes the root branch along with the first bifurcation. This is still 1 connected component and thus, $\beta_0 = 1$. $S_{100}$ includes 3 connected components, labeled in the bottom box. Therefore, $\beta_0 = 3$ for $S_{100}$.

Figure 8.1 Example bifurcation, with parent branch ($D_0$) and two daughter branches ($D_1$ & $D_2$) shown overlapping the segmented junction. The hub node ($H$) and other branch nodes ($N_{D_0}$, $N_{D_1}$, $N_{D_2}$) are labeled, along with edge points $e_1,...,e_{10}$ for $D_0$.

Figure 8.2 Vessel vectors $v_0$, $v_1$, $v_2$ generated for an example bifurcation when at most A) 0, B) 1, and C) 2 change points are allowed.

Figure 8.3 Example bifurcation at hub node $H$ with parent vessel vectors $v_0$ and daughters $v_1$ and $v_2$. A coordinate frame $\tilde{x}$, $\tilde{y}$, $\tilde{z}$ is shown in blue, where the origin is the hub node $H$. The $\tilde{z}$ axis is aligned in the reverse direction of the parent vector $v_0$. The $\tilde{x}$-axis is defined as the vector perpendicular to $\tilde{z}$ in the plane defined by $\tilde{z}$ and $A$ (the “aunt” vector, the other daughter vector from the previous hub node). The $\tilde{y}$ is the cross product of $\tilde{z}$ and $\tilde{x}$. The branching angles $\psi_i$ are computed as the arccosine of the normalized dot product of $\tilde{z}$ and $v_i$, as defined in equation (8.3). The angles $\theta_i$ are computed using the arctangent of the ratio $(v_i \cdot \tilde{y})/(v_i \cdot \tilde{x})$, as defined in equation (8.4).

Figure 8.4 Histograms (blue: control, orange: HPH) and KDE distributions (green lines) of angle measures for control and PH mice. A) $\psi$ angles, B) $\theta$ angles, C) absolute value of $|\theta|$ angles.

Figure 8.5 Trees generated for control (C) and hypertensive (H) mice. The red branches belong to the principal pathway, and in gray are the original branches of the spatial trees. The generated trees are in blue and were created by pulling $\psi_1, \psi_2, \theta_1$, and $\theta_2$ from the KDE distributions in Figure 8.4, as well as using the structured tree parameters from [102]. $\alpha = 0.86$, $\beta = 0.67$, control $\ell_{rr} = 13.4e^{-0.0077}r$, and hypertensive $\ell_{rr} = 10.9e^{-0.0079}r$.

Figure 8.6 A) 2D tree example, generated from parameters $\alpha = 0.86$, $\beta = 0.67$, $\ell_{rr} = 3.59$, $\psi_1 = \pi/3$, $\psi_2 = -\pi/3$, $\theta_1, \theta_2 = 0$. B) Prior and posterior distributions for $\psi_{1,2}$ for different acceptance thresholds $\epsilon$. The acceptance rate, $\alpha_r$, is the % of trees generated whose parameters were accepted by the ABC rejection sampler, which decreases with $\epsilon$. C) Two representative trees for each $\epsilon$. The sample tree is shown in blue. The simulated trees share the same root branch with the example, and their subsequent branches are shown in red.
CHAPTER 1

INTRODUCTION

Pulmonary hypertension (PH) is a cardiovascular condition defined as blood pressure in the main pulmonary artery (MPA) above 20 mmHg [97]. PH is classified into 5 types, with the second most prevalent type being Group III PH: PH associated with lung disease and/or hypoxemia [97, 57]. This type of PH is often associated with conditions such as chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), and sleep apnea, as well as with long-term exposure to high altitudes [44, 99, 36]. Symptoms of PH include shortness of breath, fatigue, and exercise limitations [44], and despite the availability of some treatments, PH is typically progressive and fatal [99].

PH is only diagnosable via right-heart catheterization, an invasive procedure where a catheter is inserted into the jugular vein and threaded through the right atrium, right ventricle, and into the MPA to measure blood pressure [49]. However, the effects of PH can also be observed through new advancements in medical imaging, as PH has long been associated with changes in the structure of the pulmonary arteries [99, 11, 35]. Analyzing pulmonary arterial networks in images can increase our understanding of their morphometry and how it is affected by PH. In this study, we analyze spatial trees extracted from medical images of control and pulmonary hypertensive mice. These trees are obtained by a semi-automated graph extraction algorithm (detailed in Chapters 4 and 6) which we developed with the help of our colleagues at Kitware, Inc.

We used these trees to inform a multi-scale 1D fluid dynamics model, extracting larger vessels to form a principal pathway and representing smaller vessels with self-similar structured trees attached to the terminals of the principal pathway. The dimensions of the structured trees are governed by radius scaling parameters $\alpha$ and $\beta$, as well as the length-to-radius ratio $\ell_{rr}$. These parameters inform the fluids model, and are extracted directly from data to produce pressure predictions in the expected range.
In addition to the aforementioned parameters, information about branching angles of the trees is obtained and used to generate structured trees in 3D in order to best fit the actual data. Previous studies [31, 85] of the structured tree model have stayed in 2D or used angle predictions from theoretical optimality principles. We instead extract actual angle values from our data and generate structured trees in 3D to model the small arteries. Angles are computed at each bifurcation, which requires the definition of a parent and 2 daughter vectors modeling the vessel branching at that bifurcation. These vectors are not trivially defined, as some vessels exhibit considerable bending along their length. To account for this, linear regression with change points is applied to the points along each vessel to define a set of vectors capturing that vessel’s trajectory. The three vectors nearest to a bifurcation point are used to define the angles $\psi_1$ and $\psi_2$ (in-plane branching angles of the 2 daughters) and $\theta_1$ and $\theta_2$ (out-of-plane angles). New 3D structured trees were generated with these angle values and fitted to data by examining a set of summary statistics.

Another technique that has been used to analyze biological networks is topological data analysis (TDA), which includes techniques for analyzing the shape of data. In particular, persistent homology has proven useful in the past for gaining insight about biological data [79, 70]. During an REU program in summer 2020, we worked with our undergraduate mentees to analyze the complexity of the spatial trees. We did this by computing degree-0 persistent homology via a height filtration, which tracked the emergence of new branches in different directions.

Overview of dissertation

The work in this dissertation has led to three studies in various stages of progress. These studies are included in Chapters 6 & 7 as full manuscripts, with preliminary results from the third included in Chapter 8.

- Chapter 2 includes an overview of cardiovascular physiology and previous work relating to arterial branching properties.
- Chapter 3 provides a literature review of the previous work that has been done toward characterizing arterial morphology.
- Chapter 4 discusses our medical image analysis algorithm, detailing the image segmentation, skeletonization, and post-processing steps to obtain spatial trees representing physiological networks.
- Chapter 5 provides topological background about persistent homology and how it has been used previously on biological networks.
- Chapter 6 includes our paper entitled “Structural and hemodynamic properties of murine pulmonary arterial networks under hypoxia-induced pulmonary hypertension”, published in the July 2020 [102]. This paper was a collaboration with our colleagues at Kitware, Inc. and
details the tree extraction protocol, 1D fluids model, and the multi-scale approach of modeling large and small arteries. We also present calculations for the structured tree parameters from the trees in our data.

- Chapter 7 includes our paper entitled “Topological Data Analysis of Murine Pulmonary Arterial Networks Under Hypoxic Pulmonary Hypertension”, currently in progress. This paper is a collaboration with our REU students from Summer 2020 and applies persistent homology to the trees extracted from control and pulmonary hypertensive mice. In particular, degree-0 persistent homology of spatial trees is computed with respect to a height filtration to compute directional complexities. This analysis is applied to the original control and pulmonary hypertensive trees, and also to systematically pruned trees in order to compare corresponding vessels in control and pulmonary hypertensive mice.

- Chapter 8 includes a description of our methods and preliminary results for our in-progress study on the use of ABC analysis to generate 3D structured trees using arterial branching angles. This paper is a collaboration with our colleagues at SoftMech, Inc. and extends the work of Chapter 6 to include branching angles in the structured-tree model. The goal of this study is to extract angles which generate the best structured-trees compared to data and to predict perfusion in control and pulmonary hypertensive mice.

- Chapter 9 provides a conclusive summary of the results found in this work and future research directions.
2.1 The cardiovascular system

The cardiovascular system (shown in Figure 2.1) is responsible for distributing oxygenated blood to all parts of the body. It is comprised of the heart and the systemic and pulmonary circulations. The heart has 4 chambers: the upper chambers are called atria and the lower chambers are called ventricles. The atria are separated from the ventricles by valves that control blood flow into the ventricles. The mitral valve separates the left atrium and ventricle and the tricuspid valve the right atrium and ventricle. We refer to the left ventricle and atrium as the left heart, and the right ventricle and atrium as the right heart. The circulations are comprised of arteries, veins, and capillaries. Arteries carry blood away from the heart and veins toward the heart. Capillaries are tiny blood vessels which connect the arteries and veins and are the site of gas and nutrient exchange.

Oxygenated blood is pumped by the left heart into the aorta through the aortic valve, entering the systemic arteries. The systemic arteries transport blood throughout the body, perfusing the upper and lower body and organs. They form a network starting at the aorta and ending at the systemic capillaries. This network is mostly bifurcating, and at each bifurcation the cross-sectional area of each new branch is smaller than that of its parent. Despite this size reduction, the total cross-sectional area of the network expands as the network branches. At the distal ends of the network, the arteries give rise to the arterioles, which are small blood vessels about 5-100 µm in diameter [56]. The arterioles connect to the systemic capillaries, which are 5-10 µm in diameter. In the systemic capillaries, blood supplies the body tissues with oxygen and nutrients and absorbs carbon dioxide and waste products from the tissues. Unlike the bifurcating network of the systemic arteries, the systemic capillaries form a mesh.
Figure 2.1 A) The cardiovascular system. Arrows indicate the direction of blood flow. Oxygenated blood flows from the left heart into the SA to be distributed throughout the body. Oxygen and nutrients are exchanged for CO₂ and waste products in the SC. The deoxygenated blood then passes into the SV. From the SV, it enters the right heart and then the PA. In the PC, blood is re-oxygenated and then enters the PV to be returned to the left heart. Reproduced and modified from [59] with permission. Access for free at https://openstax.org/books/college-physics/pages/1-introduction-to-science-and-the-realm-of-physics-physical-quantities-and-units. B) The pulmonary circulation. Notice that the PA form a rapidly branching structure, the morphometry of which is the focus of this thesis. Reproduced and modified from [60] with permission. Access for free at https://openstax.org/books/anatomy-and-physiology/pages/1-introduction.
From the systemic capillaries, the deoxygenated blood then enters the systemic venules and veins, which return it to the right heart. From the right heart, deoxygenated blood is pumped into the pulmonary circulation via the main pulmonary artery (MPA) through the pulmonary valve. The pulmonary arteries, the focus of this dissertation, distribute the deoxygenated blood to all areas of the lung via a rapidly branching network of highly compliant vessels. The pulmonary arteries contain about 7.2% of the blood volume, which is low compared to the 14% held in the systemic arteries. Even though CO is approximately equal from the left and right heart, the mean pressure in the pulmonary arteries is much lower (15 mmHg) than in the systemic arteries (95 mmHg) due to the higher resistance in the systemic arteries (see Figure 2.2 and Table 2.1 for more information on blood volume and pressure in the circulatory system). At the end of the pulmonary arteries are capillaries surrounding the alveoli. Here, dissolved oxygen enters the blood and carbon dioxide is removed. The blood (now oxygenated) enters the pulmonary veins to return to the left heart and begin the cardiac cycle again [76].

Table 2.1 Total blood volume and mean blood pressure of components of the circulatory system [76]. Abbreviations used are P (pulmonary), S (systemic), A (arteries), V (veins) and C (capillaries), as defined in Figure 2.1.

<table>
<thead>
<tr>
<th>Component</th>
<th>Blood volume (mL)</th>
<th>Relative blood volume (%)</th>
<th>Mean blood pressure (mmHg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>700</td>
<td>14</td>
<td>95 (60 in arterioles)</td>
</tr>
<tr>
<td>SC</td>
<td>300</td>
<td>6.0</td>
<td>25</td>
</tr>
<tr>
<td>SV</td>
<td>3200</td>
<td>64</td>
<td>3 (15 in venules)</td>
</tr>
<tr>
<td>Heart</td>
<td>360</td>
<td>7.2</td>
<td>0-120 (lower in diastole, up to 120 in the ventricles during systole)</td>
</tr>
<tr>
<td>PA</td>
<td>130</td>
<td>2.6</td>
<td>15</td>
</tr>
<tr>
<td>PC</td>
<td>110</td>
<td>2.2</td>
<td>10</td>
</tr>
<tr>
<td>PV</td>
<td>200</td>
<td>4.0</td>
<td>5</td>
</tr>
</tbody>
</table>

During the cardiac cycle, the heart rhythmically contracts and relaxes to pump blood into the system. When the heart contracts, the system is said to be in **systole**, and in **diastole** when the heart relaxes. The amount of blood ejected by the heart during a single heartbeat is known as the **stroke volume** (SV), measured in liters. Humans typically have approximately 5 L of blood in the body. **Cardiac output** (CO) is the mean blood flow in the circulation, measured in L/min. CO is given by

\[
CO = SV \cdot HR
\]  

(2.1)

where HR is heart rate in beats/min. At rest, CO is about 5 L/min in a 70-kg human, but it increases dramatically during exercise. The right and left heart have approximately the same CO, since blood is not added or lost during the cardiac cycle.
2.2 Pulmonary hypertension

Pulmonary hypertension (PH), defined as mean pulmonary arterial pressure mPAP $> 20$ mmHg [97], is classified into 5 types which affect various parts of the pulmonary circulation (Figure 2.3).

![Figure 2.2](image1.png) **Figure 2.2** Blood pressure in different components of the circulatory system. When pressure is oscillating, the upper threshold reached is the systolic pressure; the lower bound is the diastolic pressure. Reproduced from [54] with permission.

![Figure 2.3](image2.png) **Figure 2.3** The 5 types of pulmonary hypertension (PH) and their affected areas. I: Pulmonary arterial hypertension (PAH) initially starts in the pulmonary arterioles and progresses to the larger pulmonary arteries. II: PH due to left heart disease, also called pulmonary venous hypertension (PVH), starts in the left atrium and progresses to the pulmonary veins and, potentially, the venules. III: PH due to hypoxia initially affects the pulmonary arterioles and capillaries and progresses to the arteries and venules. IV: Chronic thromboembolic PH (CTEPH) initiates in the pulmonary arteries and progresses down to the arterioles. Finally, V: idiopathic PH refers to PH due to unknown or multifactor causes.
Group I PH is pulmonary arterial hypertension (PAH), which initially affects the pulmonary arterioles and progresses to the larger pulmonary arteries. In PAH, the walls of the pulmonary arteries and arterioles become thick and stiff, limiting blood flow and increasing vascular resistance. As PAH advances, lesions and blood clots may form in the affected vessels [101]. PAH can be sporadic or hereditary and has been associated with vascular disease, HIV, certain drugs and toxins, and other causes [36].

Group II is PH associated with left heart disease, which is the most common form of PH [101]. It is also known as pulmonary venous hypertension (PVH), as it starts in the left heart and progressively affects the pulmonary veins and venules [36]. In PVH, the increased mPAP is caused by a weak, stiffened left heart, which cannot contract or relax properly to pump the blood returning from the pulmonary circulation back into the systemic side. This causes a backup of blood in the pulmonary veins, causing the higher pressure [101].

Group III is hypoxia-related PH (HPH). HPH is often associated with diseases which hinder lung function, such as chronic obstructive pulmonary disease (COPD) and interstitial lung disease (ILD), as well as exposure to high altitudes [36]. HPH initially affects the pulmonary capillaries and arterioles. In HPH patients, some alveoli are collapsed due to the patient’s lung disease or high altitude environment, which causes the pulmonary capillaries and arterioles to constrict and force blood into more well-ventilated areas of the lung, increasing mPAP [101]. Progressive HPH can affect the pulmonary arteries and veins.

Group IV is chronic thromboembolic pulmonary hypertension (CTEPH). CTEPH is typically caused by clots, embolisms, and obstructions in the pulmonary arteries. Damage due to CTEPH can progress to the arterioles over time, causing scarring in the pulmonary vasculature. Clots and scarring block the flow of blood, resulting in an elevated mPAP [101].

The final type, group V, is PH associated with unknown or multifaceted causes, collectively called idiopathic PH. A few causes which fall into this category include inflammatory diseases (e.g. sarcoidosis, schistosomiasis), certain anemias (e.g. sickle cell, chronic hemolytic), splenectomy, glycogen storage disease, and thyroid disorders [36, 101].

Studies indicate that HPH is the second commonest form of PH [57, 90], with one study suggesting it may be the most deadly over time [90]. This dissertation focuses on the pulmonary arterial morphometry of mice induced with HPH.
CHAPTER

3

MORPHOMETRY OF THE PULMONARY VASCULATURE

The pulmonary circulation facilitates gas exchange with the airways in the lungs. The lungs are organized into several lobes separated by fissures and connective tissue [60]. Air is supplied to the alveoli in each lobe by subtrees of airways branching off of the trachea. Blood is supplied to each lobe via a different subtree of pulmonary arteries off of the main vessel pathway. The number and names of lobes vary between species (Figure 3.1).

In each lobe, the pulmonary arteries form rapidly branching trees that have often been studied for their fractal properties [1, 2, 14, 12, 15, 10, 4, 8, 13, 31, 39]. Specific attention has been devoted to the dimensions of the pulmonary arteries, relationships across arterial junctions, and the angles of arterial branching. In this chapter, we discuss work that has been done to characterize the vessel morphometry and note the novel contributions we have made to this field.

3.1 Vessel dimensions

In the rapidly branching pulmonary arterial network, we refer to a vessel section between junctions as a single artery. These sections are modeled as cylinders, each with radius $r$, length $\ell$, and wall thickness $h$ (Figure 3.2A). Some models allow for tapering to occur along the length of a vessel; for simplicity we assume the radius and thickness are uniform along the length of an artery. Additionally, a blood vessel’s walls stretch and relax periodically throughout the cardiac cycle. Note that whenever we reference vessel radius, we are referring to the unstressed radius, which is the in vivo radius at a pressure of 0 mmHg.
3.1.1 Radius

Murray’s Law: The first scientist to characterize radius relations in the pulmonary arteries was Cecil Murray, who studied the efficiency of the pulmonary arterial system [2]. His work led to Murray’s Law, a cubic power law relating the vessel radius to flow. This law was derived from energy considerations, assuming that the net amount of work $E$ (in ergs/second) required to pump blood through a section
of artery with radius \( r \) and length \( \ell \) is given by

\[
E = (\Delta p)q + b\ell \pi r^2, \quad (3.1)
\]

where \( \Delta p \) is the difference in pressure from the inlet to the outlet of the artery in dyn/cm\(^2\), \( q \) is the blood flow in cm\(^3\)/s, and \( b \) the cost of maintaining blood volume in the body in erg/(s cm\(^3\)) (assumed constant). The first term \((\Delta p)q\) represents the work to pump the blood through an artery that is attributed to friction, while the second term \(b\ell \pi r^2\) is the work attributed to maintaining blood in the body. According to Poiseuille, the pressure difference \( \Delta p \) for steady flow is

\[
\Delta p = \frac{8q\ell \mu}{\pi r^4},
\]

where \( \mu \) the blood viscosity (assumed constant) in (dyn s)/cm\(^2\) (unit called a “poise”). This means equation (3.1) can be rewritten as

\[
E = \frac{8q^2\ell \mu}{\pi r^4} + b\ell \pi r^2. \quad (3.2)
\]

Murray posited that the arterial system is most efficient when \( E \) is at a minimum, i.e.,

\[
\frac{dE}{dr} = -\frac{32q^2\ell \mu}{\pi r^5} + 2b\ell \pi r = 0
\]

\[
\Rightarrow b = \frac{16q^2\mu}{\pi^2 r^6}. \quad (3.3)
\]

Rewriting equation (3.3) yields

\[
q^2 = \frac{\pi^2 r^6 b}{16\mu} \Rightarrow q = r^3 \sqrt{\frac{\pi^2 b}{16\mu}}. \quad (3.4)
\]

This equation states that the flow in any section of artery is proportional to the cube of the radius, with a proportionality constant \( c = \sqrt{\frac{\pi^2 b}{16\mu}} \).

The junction of a parent vessel, denoted with subscript 0, splitting into 2 daughter vessels, denoted with subscripts 1,2, is called a bifurcation (see Figure 3.2). Conservation of blood flow across bifurcations, dictates

\[
q_0 = q_1 + q_2. \quad (3.5)
\]

By combining equations (3.4) and (3.5), Murray’s Law defines the relationship between the radius of a parent vessel \( r_0 \) and the radii of its daughter vessels \( r_1, r_2 \) in arteries with steady flow to be

\[
c r_0^3 = c r_1^3 + c r_2^3 \Rightarrow r_0^3 = r_1^3 + r_2^3. \quad (3.6)
\]

As a convention, daughter 1 is always labeled such that \( r_1 \geq r_2 \).
Uylings' generalization of Murray's Law: While Murray's work assumed steady flow [2], in general, blood flow in arteries is not steady. This observation inspired Uylings in 1977 to develop a generalized equation relating radii across bifurcations under multiple flow regimes [14].

As previously mentioned, the first term \((\Delta p)q\) in (3.1) represents the work due to friction of pumping the blood. Uylings expressed \(\Delta p\) as

\[
\Delta p = \frac{\lambda \rho q^2}{4\pi^2 r^5} \ell,
\]

where \(\rho\) is the blood density in g/cm³ (assumed constant) and \(\lambda\) is a dimensionless friction coefficient of the form

\[
\lambda = g \left( \frac{\mu}{2\rho \ell} \right)^k,
\]

where \(g\) and \(k\) are values related to the "roughness" of the vessel wall. It is stated [14] that \(k\) is such that \(0 \leq k \leq 1\), where \(k = 0\) for turbulent flow and \(k = 1\) for laminar flow.

Substituting (3.7) and (3.8) into (3.1) yields

\[
E = g \left( \frac{\mu}{2\rho \ell} \right)^k \frac{\rho q^3}{4\pi^2 r^5} \ell + b \ell \pi r^2
\]

\[
\Rightarrow E = g \left( \frac{\mu}{2\rho} \right)^k \left( \frac{\rho}{4\pi^2} \right) q^{3-k} r^{k-5} \ell + b \ell \pi r^2
\]

\[
\Rightarrow E = C q^{j-2} r^j - b \ell \pi r^2,
\]

where \(C = g \left( \frac{\mu}{2\rho} \right)^k \left( \frac{\rho}{4\pi^2} \right)\) and \(j = 5 - k\). Hence, \(4 \leq j \leq 5\) with \(j = 4\) for laminar flow and \(j = 5\) for turbulent flow.

Uylings also used the optimality principle of minimum work, stating that efficiency is maximized when (3.9) is minimized, i.e.,

\[
\frac{dE}{dr} = -jC q^{j-2} \frac{r^{j+1} \ell}{r^j + 2b \ell \pi r} = 0
\]

\[
\Rightarrow -jC q^{j-2} \ell + 2b \ell \pi r^{j+2} = 0
\]

\[
\Rightarrow q^{j-2} = \frac{2b \pi}{jC} r^{j+2}
\]

\[
\Rightarrow q = \left( \frac{2b \pi}{jC} \right)^{1/(j-2)} r^{(j+2)/(j-2)}.
\]

This equation gives a different power law than (3.4), stating that the flow in any section of artery is proportional to the \(r^{(j+2)/(j-2)}\) with a proportionality constant \(K_j = \left( \frac{2b \pi}{jC} \right)^{1/(j-2)}\).

Using 3.5, Uylings obtained

\[
K_j r_0^{(j+2)/(j-2)} = K_j r_1^{(j+2)/(j-2)} + K_j r_2^{(j+2)/(j-2)} \Rightarrow r_0^\xi = r_1^\xi + r_2^\xi
\]
where $\xi = (j + 2)/(j - 2)$. Thus, $\xi = 2.33$ for turbulent flow and $\xi = 3.00$ for laminar flow (and steady flow, as Murray suggested). For most blood flow applications, it is assumed that $\xi = 2.76$ [14, 18, 31]. For the pulmonary arterial networks we studied, we found $\xi = 2.61$ based on a simple average of $\xi$ values over all bifurcations in each network.

**Zamir’s asymmetry and area ratios:** In 1978, Zamir characterized asymmetrical arterial bifurcations. The majority of arterial bifurcations in the body are asymmetrical [15]. He defined 2 additional relations to describe vessel dimensions at bifurcations: an asymmetry ratio

$$\gamma = \left(\frac{r_2}{r_1}\right)^2 \tag{3.11}$$

and an area ratio

$$\eta = \frac{r_1^2 + r_2^2}{r_0^2}. \tag{3.12}$$

It is clear that $0 < \gamma \leq 1$, since $r_1 \geq r_2$. An asymmetry ratio $\gamma = 1$ would indicate a symmetrical bifurcation ($r_1 = r_2$), while a $\gamma$ near 0 would indicate a highly asymmetrical bifurcation.

By combining equations (3.10), (3.11), and (3.12), Zamir related $\eta$, $\gamma$, and $\xi$ by writing

$$\eta = \frac{1 + \gamma}{(1 + \gamma \xi^2)^{2/\xi}}. \tag{3.13}$$

Equation (3.13) indicates that $\eta$ would be near 1 for a highly asymmetrical bifurcation and $\eta = 2^{1 - 2/\xi}$ for a symmetrical bifurcation. This means in a network with turbulent flow, $1 < \eta \leq 1.10$, and in a network with steady flow, $1 < \eta \leq 1.26$. For the networks examined in this dissertation, we found an average asymmetry ratio of $\gamma = 0.66$ and an average area ratio of $\eta = 1.25$.

Zamir observed that in optimal asymmetrical bifurcations, $r_1$ was approximately equal to $r_0$. He examined the ratios of each daughter’s radius to parent radius and found that, for $\gamma = 0.5$ and $\xi = 3.00$, $r_1/r_0 \approx 0.9$ and $r_2/r_0 \approx 0.65$. This shows that in an asymmetrical bifurcation, $r_1$ can still be as much as 90% of $r_0$.

**Olufsen’s radius scaling factors:** In 2000, Olufsen et al. [31] used parameters $\eta$, $\gamma$, and $\xi$ to characterize the geometry of arterial trees in a 1D fluid dynamics model. They modeled the arterial network as a *structured tree*—a self-similar fractal tree in which branches’ dimensions are found by scaling the parent branch’s dimensions. This allowed them to solve the Navier-Stokes equations in these trees semi-analytically, and reduce the number of operations required to compute impedance in the entire tree. The scaling factors, $\alpha$ and $\beta$, were such that

$$r_1 = \alpha r_0, \quad \text{and} \quad r_2 = \beta r_0. \tag{3.14}$$

By combining equations (3.10), (3.11), and (3.14), they derived the following formulas for $\alpha$ and $\beta$ in
terms of $\xi$ and $\gamma$,

$$\alpha = \left(1 + \gamma^{\xi/2}\right)^{-1/\xi}, \quad \text{and} \quad \beta = \alpha \sqrt{\gamma}. \quad (3.15)$$

Based on a review of several experimental studies, Olufsen et al. found $\xi = 2.76$, $\eta = 1.16$, and $\gamma = 0.41$ to be appropriate for modeling blood flow in structured trees. These parameter values in equation (3.15) yielded $\alpha = 0.9$ and $\beta = 0.6$. These agree with the values of $r_1 / r_0 \approx 0.9$ and $r_2 / r_0 \approx 0.65$ found in [15]. In our study of arterial networks, we extracted $\alpha$ and $\beta$ values from medical images and found $\alpha = 0.87$ and $\beta = 0.68$ [102] (further discussed in Chapters 4 and 6).

All previous work has relied on determining $\xi$ and $\eta$ to compute the values of the other parameters. This has been accomplished via optimality principles [2, 15] or physical measurements of arteries from specimens [13, 10], which can be difficult to obtain and are subject to human error. In this dissertation, we use data analysis to extract values of these parameters from medical images and discuss the validity of previously obtained values.

### 3.1.2 Length

The length of an artery varies widely depending on its location in the body. For example, the brachial arteries run down the length of the arms without branching, making them very long. By comparison, pulmonary arteries are much shorter, branching rapidly to fill the lobes of the lung. Since we refer to a single artery as a segment between junctions, this rapid branching pattern makes for much shorter arterial lengths.

With numerous equations relating vessel radii at a bifurcation (equations (3.10)-(3.14)), it is natural to investigate what relationship, if any, exists between vessel radius and length. In 1963, Suwa et al. [4] used the fractal properties of arterial networks in resin casts to estimate intravascular pressure. In their casts, they found that arterial radius and length were related by

$$\ell = m r^i, \quad (3.16)$$

where $m$ and $i$ are organ specific constants. Suwa et al. found that for various organs, $i$ ranged between 0.76 and 1.21, with $i = 1.16$ in the lungs. Since the exponent $i$ in (3.16) was around 1 in the lung [4], it can be concluded that the length-to-radius ratio,

$$\ell_{rr} = \ell / r, \quad (3.17)$$

in the pulmonary arteries tends to be constant. If the length and radius have the same unit, $\ell_{rr}$ is non-dimensional. Iberall [6] made this assertion in 1967, and by combining data from Suwa et al. [4] and Patel et al. [7], found that $\ell_{rr} \approx 50$ was reasonable for modeling blood flow. In 1999, Dawson et al. [28] also defined length along the power law (3.16). They examined several studies and found exponent values of $0.84 \leq i \leq 1.16$ in the pulmonary arteries, again concluding that $\ell_{rr}$ in the lung is roughly constant. By averaging these studies’ data, Dawson et al. found that $\ell_{rr} \approx 5$ [28].
Iberall’s [6] value of $\ell_{rr} \approx 50$ was used by Olufsen et al. [31] for their 1D fluid dynamics model in the arterial structured trees. With a constant $\ell_{rr}$ value, the daughters’ lengths $\ell_1$ and $\ell_2$ at a bifurcation can be found using the same scaling factors $\alpha$ ad $\beta$ used for finding the radii, i.e.,

$$r_1 = \alpha r_0 \implies \frac{\ell_1}{\ell_{rr}} = \alpha \frac{\ell_0}{\ell_{rr}} \implies \ell_1 = \alpha \ell_0, \text{ and}$$

$$r_2 = \beta r_0 \implies \frac{\ell_2}{\ell_{rr}} = \beta \frac{\ell_0}{\ell_{rr}} \implies \ell_2 = \beta \ell_0.$$

In the trees examined in this dissertation, we found great variation in $\ell_{rr}$ values. We found that $\ell_{rr}$ can be modeled as a decaying exponential function of $r$,

$$\ell_{rr}(r) = c_1 e^{-c_2 r}, \quad (3.18)$$

where $c_1$ is nondimensional and $c_2$ is in $\mu m^{-1}$. In control mice, $\ell_{rr}$ was found to be higher ($c_1 = 13.4 \pm 1.2$ and $c_2 = 0.00771 \pm 0.00083$) compared to pulmonary hypertensive mice ($c_1 = 10.9 \pm 1.1$ and $c_2 = 0.00797 \pm 0.00010$) [102]. These values are reported $\pm$ their margin of error for a 95% confidence interval.

### 3.1.3 Wall thickness

Another dimension of interest when modeling pulmonary arteries is the thickness $h$ of the vessel wall. In 2012, Olufsen et al. determined that in pulmonary arteries, the wall thickness $h$ is

$$h = \frac{34.5}{E} r,$$

where $E$ is Young’s modulus and $r$ is the vessel radius (unstressed, as stated before)[55]. Therefore, the wall thicknesses $h_1$ and $h_2$ of the daughter vessels at a bifurcation can be determined by scaling the parent’s wall thickness $h_0$ by $\alpha$ and $\beta$.

$$r_1 = \alpha r_0 \implies \frac{34.5}{E} r_1 = \alpha \frac{34.5}{E} r_0 \implies h_1 = \alpha h_0, \text{ and}$$

$$r_2 = \beta r_0 \implies \frac{34.5}{E} r_2 = \beta \frac{34.5}{E} r_0 \implies h_2 = \beta h_0.$$

### 3.2 Branching angles

**Murray’s 2D branching angles:** In 1926, Murray determined formulas for optimal 2D branching angles $\psi_1$ and $\psi_2$ of daughter arteries at a bifurcation [1] (Figure 3.2). Substituting equation (3.4)
into equation (3.2) yields

\[ E = \frac{\pi^2 r^6 b}{16\mu} \frac{8\ell \mu}{\pi r^4} + b \ell \pi r^2 \]

\[ \Rightarrow E = \frac{\pi^2 r^6 b}{2(4\ell \mu)} \frac{8\ell \mu}{\pi r^4} + b \ell \pi r^2 \]

\[ \Rightarrow E = \frac{1}{2} b \ell \pi r^2 + b \ell \pi r^2 \]

\[ \Rightarrow E = \frac{3}{2} b \ell \pi r^2. \]

This can be rewritten as

\[ kE = \ell r^2, \text{ where } k = \frac{2}{3b\pi}. \] (3.19)

The value \( kE \) is still a measure of work. Note that while Murray reports \( k = \frac{1}{3b\pi} \) [1], we have found this to be in error.

Consider if the length \( \ell_0 \) of the parent is increased by an infinitesimal amount \( d\ell_0 \). Then the work \( kE \) in the parent will increase by \( d\ell_0 r_0^2 \), and the work in the daughters will decrease by \( \cos(\psi_1)d\ell_0 r_1^2 \) and \( \cos(\psi_2)d\ell_0 r_2^2 \). If daughter 1’s length \( \ell_1 \) is increased by an infinitesimal \( d\ell_1 \), the work \( kE \) in daughter 1 will increase by \( d\ell_1 r_1^2 \), the work in the parent will decrease by \( \cos(\psi_1)d\ell_1 r_0^2 \), and the work in daughter 2 would be decreased by \( -\cos(\psi_1 + \psi_2)d\ell_1 r_1^2 \). Likewise, if daughter 2’s length \( \ell_2 \) is increased by an infinitesimal \( d\ell_2 \), the work \( kE \) in daughter 2 will increase by \( d\ell_2 r_2^2 \), the work in the parent will decrease by \( \cos(\psi_2)d\ell_2 r_1^2 \), and the work in daughter 1 will decrease by \( -\cos(\psi_1 + \psi_2)d\ell_2 r_2^2 \). The principle of virtual work in mechanics states that when work is minimum, an infinitesimal change in the configuration of a system results in no change in the total work [2]. Thus, the changes in work described above should be in balance, yielding

\[ d\ell_0 r_0^2 = \cos(\psi_1)d\ell_0 r_1^2 + \cos(\psi_2)d\ell_0 r_2^2 \]
\[ d\ell_1 r_1^2 = -\cos(\psi_1 + \psi_2)d\ell_1 r_2^2 + \cos(\psi_1)d\ell_1 r_0^2 \]
\[ d\ell_2 r_2^2 = -\cos(\psi_1 + \psi_2)d\ell_2 r_1^2 + \cos(\psi_2)d\ell_2 r_0^2. \]

By simplifying and combining the above equations, Murray arrived at the following formulas for branching angles \( \psi_1 \) and \( \psi_2 \) which are optimal for minimum work,

\[ \cos(\psi_1) = \frac{r_0^4 + r_1^4 - r_2^4}{2r_0^2 r_1^2}, \quad \text{and} \quad \cos(\psi_2) = \frac{r_0^4 + r_2^4 - r_1^4}{2r_0^2 r_2^2}. \] (3.20)

**Zamir’s 2D branching angles:** In 1976, Zamir [15] computed angles \( \psi_1 \) and \( \psi_2 \) for arteries under different optimality principles, examining minimum lumen volume, minimum lumen surface area, minimum pumping power, and minimum total drag force. Using equation (3.4), he determined that
for both minimum lumen volume and minimum pumping power, the angles were such that

\[
\cos(\psi_1) = \frac{r_0^4 + r_1^4 - r_2^4}{2r_0^2 r_1^2}, \quad \text{and} \quad \cos(\psi_2) = \frac{r_0^4 + r_2^4 - r_1^4}{2r_0^2 r_2^2},
\]

which are identical to Murray’s angle definitions (3.20). For both minimum lumen surface and minimum drag force, the angles were such that

\[
\cos(\psi_1) = \frac{r_0^2 + r_1^2 - r_2^2}{2r_0 r_1}, \quad \text{and} \quad \cos(\psi_2) = \frac{r_0^2 + r_2^2 - r_1^2}{2r_0 r_2}.
\]

Both Murray and Zamir’s aforementioned studies only computed 2D branching angles between the daughter vessels and the parent vessel direction. In 1983, Zamir et al. investigated the importance of including an out-of-plane branching angle to account for 3D branching [17]. Using a resin cast of the arteries of a rat, they determined that parent and daughter vessels mostly lie in the same plane.

**Recent work on 3D branching angles:** While it was found that most arterial bifurcations are 2D [17], Tawhai and Hunter et al. [37] in 2004 modeled the bronchial tree with 3D branching angles. They defined angles \(\psi_{1,2}\) to be the branching angle between each daughter and the parent branch within their common plane. Angles \(\theta_{1,2}\) were the out-of-plane rotational angles between each daughter and the plane between the parent vessel and its sibling, referred to as the parent-aunt bronchial plane (Figure 3.2C). Angle values were extracted from CT images of human and sheep bronchi.

In 2005, Burrowes, Hunter, and Tawhai [38] extended this work [37] by applying it to pulmonary arteries, which branch similarly to bronchi. They used the same angle definitions used in [37], but added a set of supernumerary branches to represent the microvasculature off of the main arterial pathways. These supernumerary branches were added to large arteries following a space-filling algorithm. They exhibited a 2D branching angle of around 90\(^\circ\) from their parent.

In 2018, Lee et al. modeled branching of the microvasculature with the same 3D branching angles defined in [37, 38]. The in-plane branching angles \(\psi_{1,2}\) were taken to be the same as what Murray defined as optimal angles for minimum work (equation 3.20), while the out-of-plane branching angles \(\theta_{1,2}\) were taken from a uniform distribution from 0 to \(\pi\).

In this dissertation, we define our branching angles in the same way as [37, 38, 85], with the in-plane and out-of-plane angle values extracted from pulmonary arterial networks obtained from micro-CT images of mice. These values are then used to define distributions from which angle values can be pulled to generate structured trees which are similar to imaged networks.

**Vessel directions:** Defining angles also involves determining the direction of vessels at a junction. Zamir did this by only considering the local region around the junction and defining the parent and daughter vectors to be tangent to the centerlines at the junction point [12]. Other studies [75, 85] define vessel direction as the vector from the junction point to the vessel endpoint. Tawhai et al. [37] and Burrowes et al. [38] define vessel direction in a similar way to [75, 85], but use the
forward direction vector of the parent and daughters (vector from start point to junction point for parent vessel, vectors from junction point to end point for daughters). A recent paper by Dong et al.\cite{103} computes the 2D branching angles from arterial data, but fails to explain the directions of the vessels. In this dissertation, vessel directions are defined by fitting vectors to points along the vessels’ centerlines, allowing for up to 2 change points to capture vessel bending.
To characterize the morphometry of the pulmonary arteries, we need an efficient algorithm to extract the geometry from medical images. The pulmonary arteries form a complex network with the following physiological and anatomical properties.

- The pulmonary arterial network begins at the inlet of the main pulmonary artery (MPA).
- Blood flows from the right ventricle and enters the network via the MPA.
- Pulmonary arteries branch rapidly and share the 3D space of the lung with airways and pulmonary veins. The pulmonary arteries transport blood to the alveoli, following a similar branching pattern as the airways.
- Each artery connects two junctions and is characterized by its radius and length.
- More than 98% of junctions are bifurcations, i.e., a parent artery splits into exactly daughters.
- No cycles (loops) exist in the network\(^1\).

To satisfy these properties and our goal of extracting parameters for 1D fluid mechanics computations in these arteries, it is natural to represent the pulmonary arterial network as a directed tree. In

\(^1\)Theoretically, only 5 loops should exist in the entire arterial system: one in each hand and foot, and one in the brain. Elsewhere, maximum efficiency of blood flow dictates that arteries should not form loops. Any loops in the pulmonary arteries observed in the images are classified as image errors.
this tree, all branches extend in three dimensions, but are locally one-dimensional. The tree is rooted at the start of the MPA with all branches directed away from the root and with arterial information encoded in the corresponding branches and junctions. This structure is called a **labeled spatial tree**, a directed tree in \( \mathbb{R}^3 \) with a single root and all branches directed away from the root, labeled with geometric information about their representative vessels\(^2\).

In collaboration with Kitware, Inc., we developed a semi-automated algorithm for creating labeled spatial trees to represent the pulmonary arterial network. This algorithm is illustrated in Figure 4.1 and has been described in [102] (Chapter 6). It should be noted that while the methods described in this section were developed on micro-computed tomography (micro-CT) images of murine arteries, they have a much wider range of applications. These methods can be applied to CT or magnetic resonance images (MRI), other biological networks (bronchial, venous, etc.), and other species. We have had some success using these methods to extract graphs of human arterial and bronchial networks, though this data can be difficult to obtain.

In this chapter, we provide details about each component of our algorithm in Figure 4.1, including an overview of common segmentation techniques, details about the distance map, skeletonization, and graph extraction algorithm [72, 84], and instructions for using our MATLAB code to extract the connectivity and geometry information of the vessels (i.e., to label the geometric tree).

---

\(^2\)A directed tree with a single root and all branches directed away from the root is also sometimes called an **arborescence** [47], as tree typically refers to an undirected acyclic graph. Since all the trees discussed in this dissertation represent pulmonary arteries, they all have their branches oriented in the direction of blood flow, so we assume any tree mentioned in this section is directed.
4.1 Micro-CT images

This dissertation analyzes micro-CT images from male C57BL6/J mice, aged 10-12 weeks. These images were obtained from studies conducted at the University of Wisconsin, Madison, and made available by Naomi Chesler, University of California Irvine. Details of the imaging protocol can be found in [51, 102]. We analyze images from 3 control and 3 pulmonary hypertensive (PH) mice, induced by keeping the mice in an hypoxic environment (FiO₂ reduced by half, to 10%) for 10 days. The mice were euthanized by exsanguination, their lungs were extracted, and a cannula (PE-90 tubing, 1.27 mm outer and 0.86 mm inner diameter) was placed in the opening of the MPA³. The pulmonary arteries were perfused with perfluorooctyl bromide at a pressure of 7.4 mmHg and placed in a micro-CT scanner. Lungs were rotated 360° in the imaging chamber, and planar images were obtained at 1° increments, resulting in 360 planar images. The planar images were reconstructed using the Feldkamp cone-beam algorithm and converted into 3D volumetric datasets and stored as Digital Imaging and Communications in Medicine (DICOM) 3.0 images. The perfusion pressure was then decreased to 6.3 mmHg, and increased to 13.0 and 17.2 mmHg, with new images obtained at each pressure.

The result is 24 micro-CT scans providing 3D images, saved as Digital Imaging and Communications in Medicine (DICOM) files. There were 6 mice in this study (3 control, 3 PH) imaged at each of the 4 perfusion pressures. Each 3D image forms a voxel complex, a finite set of voxels which are 3D units of image data analogous to the “pixels” in a 2D image [84]. The dimensions of each image are (497 × 497 × 497) voxels, and the resolutions are between 30-40 µm due to slight variation in the size of the mice. Each voxel v in the image has 3D coordinates (x_v, y_v, z_v) and image intensity I(v) ∈ [0, 255]. The higher the voxel intensity, the lighter the voxel appears.

4.2 Segmented arteries

Image segmentation involves classifying the voxels in an image as belonging to certain structures of interest. It is an invaluable tool for analyzing medical images, with applications in morphological studies, diagnostics, surgery planning, and medical device development. Outside of medicine, segmentation has extensive applications in computer vision, including facial/object recognition and pedestrian/traffic detection.

There are many programs for performing segmentation, each with its benefits and drawbacks. Some have expensive licenses (Mimics, OsirIX, Aivia) or can take a long time to segment (ITKSnap). There are also segmentation programs made specifically for analyzing images of the brain (FSL, FreeSurfer, BIC Toolbox). For this dissertation, we collaborated with Kitware, Inc. and used 3D Slicer [53, 63, 104], an open-source image analysis program which is user-friendly and well-documented.

In the micro-CT images of mice, the pulmonary arterial networks were the structures of interest. By segmenting the images, we separated the voxels into foreground voxels (belonging to the arteries)

³The cannula was placed well above the first bifurcation so as to preserve its morphometry. Still, measures sometimes needed to be taken to adjust the image for errors caused by the presence of the cannula.
and background voxels (belonging to non-arterial components of the image). Since the pulmonary arteries are extracted from the mice before imaging, they were the only anatomical structures in the images, making segmentation less complicated than it typically is for human images. Several common segmentation techniques were implemented in 3D Slicer to isolate the foreground.

1. **Thresholding** defines an intensity range from a lower intensity threshold $\tau_L$ to an upper threshold $\tau_U$ such that all voxels with $I(v) \in [\tau_L, \tau_U]$ are included in the foreground of the image. Structures captured in the image will have higher (lighter) voxel intensities than the image background and noise. We set $\tau_U = 255$ for the images in this study, since the arteries typically include the lightest voxels. We choose $\tau_L$ in an ad hoc method, typically starting at $\tau_L = 18$ and adjusting from there until all arteries in the image are captured (Table 4.1). The final $\tau_L$ value varies with each image.

   **Table 4.1** Examples of segmented images generated by different lower thresholds $\tau_L$. Notice that $\tau_L = 13$ is too low, as noise in the image was labeled as part of the arterial network. By contrast, $\tau_L = 65$ is too high, as voxels which clearly belong to arteries were not included in the foreground. For the images in this dissertation, we found $\tau_L = 18$ to be a good starting point for the lower threshold and adjusted this in an ad hoc method for each image.

<table>
<thead>
<tr>
<th>Lower threshold</th>
<th>$\tau_L = 13$</th>
<th>$\tau_L = 18$</th>
<th>$\tau_L = 65$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmented micro-CT scan</td>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
<tr>
<td>Segmented 3D rendering</td>
<td><img src="image4.png" alt="Image" /></td>
<td><img src="image5.png" alt="Image" /></td>
<td><img src="image6.png" alt="Image" /></td>
</tr>
</tbody>
</table>

2. **Smoothing** removes noise from the image foreground. Each foreground voxel’s intensity is replaced by some function of the intensities of a group of adjacent voxels, called the smoothing kernel. This is used to make the foreground to appear smoother and removes noise. In this study, we applied median smoothing with a smoothing kernel size of $(3 \times 3 \times 3)$ voxels to each image (Table 4.2).
**Table 4.2** Examples of segmented images generated using lower threshold $\tau_L = 18$ and median smoothing with different sized smoothing kernels. In our study, we applied median smoothing with a kernel size of $3 \times 3 \times 3$ voxels. This kernel size allowed us to appropriately smooth out any noisy surfaces without removing many distal branches. When smoothing is applied, small islands often appear; some examples of these are circled in red in the images.

<table>
<thead>
<tr>
<th>Median smoothing kernel size</th>
<th>None</th>
<th>$3 \times 3 \times 3$ voxels</th>
<th>$5 \times 5 \times 5$ voxels</th>
<th>$7 \times 7 \times 7$ voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Segmented micro-CT scan</td>
<td><img src="image" alt="Segmented Image" /></td>
<td><img src="image" alt="Segmented Image" /></td>
<td><img src="image" alt="Segmented Image" /></td>
<td><img src="image" alt="Segmented Image" /></td>
</tr>
</tbody>
</table>

3. **Islands** are groups of voxels in the image foreground that are disconnected from each other. Islands can represent individual structures of interest in an image, or noise, glares, and other non-biological components. In our images, the pulmonary arterial network should form the only island included in the foreground. Sometimes the imaging chamber is also captured, which can be removed using the “Keep selected island” function in 3D Slicer (Figure 4.2A). Additionally, small islands can appear as a result of smoothing (Table 4.2) or from glares on the bottom of the chamber (Figure 4.2B). In these instances, the “Keep largest island” function can easily remove these islands without needing to select the island to keep.

![Figure 4.2](image)

**Figure 4.2** A) The hypobaric chamber which contains the arteries during the imaging process was captured in the CT scan. We remove this by selecting the arteries as the island to keep. B) Glares in the image caused many small islands to appear toward the bottom of the network. We remove them by keeping the largest island, which we know to be the arterial network.
4. **Manual editing** is applied sparingly to correct errors in the image which are not corrected with other methods. 3D Slicer has several functions for manual segmentation, including “Cut”, “Erase”, “Paint”, and “Draw”. Manual segmentation is done to smooth out the appearance of the MPA, which can be affected by the cannula, or to remove obvious image errors such as glares or noise that cannot be removed by other methods. Manual segmentation must be performed carefully so as to not alter the true arterial morphometry (Figure 4.3).

![Figure 4.3](image)

**Figure 4.3** Segmented network with centerlines before and after manual segmentation was applied to the MPA. Notice that the cannula inserted for imaging distorted the appearance of the MPA, causing a “clump” to form in the centerline skeleton. To improve the skeleton, the distortion must be manually reduced to a more uniform cylindrical shape. Care is taken to remove only erroneous protrusions and not to reduce the actual size of the MPA.

The result of segmentation is a voxel complex which depicts a 3D volumetric representation of the image foreground (the pulmonary arterial network). We refer to this 3D depiction as the **segmented network**, and refer to an individual segment of the network between vessel junctions (or between a junction and a vessel endpoint) as a **segmented artery**. The segmentation output is saved as a “nearly raw raster data” (.nrrd) file, hereby referred to “seg.nrrd”.

The file *seg.nrrd* (as in any .nrrd file containing a voxel complex) has 2 important properties: space directions and space origin. The **space directions** denote the $x$, $y$, and $z$ scale of the segmented network. Space directions of $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ refer to cubic voxels, i.e., the voxel faces have the same length in every direction, while space directions of $(2,0,0)$, $(0,1,0)$, $(0,0,1)$ refer to a voxel where the $x$ scale is twice that of the $y$, and $z$ scales. The **space origin** is the coordinate-triple $(x, y, z)$ in the lower-left-front corner of *seg.nrrd*. These properties can be easily found by opening *seg.nrrd* with a text editor. In all the mouse images we analyzed, the voxels were cubes (space directions were $(1,0,0)$, $(0,1,0)$, $(0,0,1)$).
4.3 Distance map

The **distance map** is a voxel complex associated with `seg.nrrd` which has the same spatial dimensions and encodes a distance measurement in each voxel. For each voxel \( v \) in the foreground of a segmented network, we define the function \( d(v) \) to be the Euclidean distance from \( v \) to the nearest background voxel \( u \),

\[
    d(v) = \min_{u \in \text{background}} ||u - v||_2. \tag{4.1}
\]

If \( v \) is a voxel in the center of a segmented artery, \( d(v) \) provides an estimate for that artery’s radius at that point. Once the segmented network is skeletonized (described in the next section), the distance map will be used to obtain the pulmonary arterial radii.

The distance map is obtained by using the DGTal program “create_distance_map” [80], using the `seg.nrrd` file as input. The output is the distance map, stored as a separate file `dmap.nrrd`. While working with this software, we noticed a bug in the code that can sometimes distort the distances. Before the distance map is computed, the “create_distance_map” function changes the space directions and space origin of `seg.nrrd` to \((1, 0, 0)\), \((0, 1, 0)\), \((0, 0, 1)\) and \((0, 0, 0)\), respectively. In cases where `seg.nrrd` had space directions of magnitude other than 1, the segmented network was skewed and produced values in the distance map that are not reflective of the actual distances from the background. For example, if `seg.nrrd` has space directions \((2, 0, 0)\), \((0, 1, 0)\), \((0, 0, 1)\), “create_distance_map” changes these to \((1, 0, 0)\), \((0, 1, 0)\), \((0, 0, 1)\) before computing the distance map. This means the distance map is being computed on a version of the segmented network that, in the \(x\)-direction, is half the size of the true segmented network (Figure 4.4).

This is a bug which needs to be addressed to use these methods on human/other images in the future. Fortunately for this work, the segmented networks have space directions \((1, 0, 0)\), \((0, 1, 0)\), \((0, 0, 1)\) initially, so the distance map values are not distorted. There is, however, a shift in space origin. The segmented networks have space origins of, for example, \((212, 57, 2)\), \((205, 89, 4)\), etc., while “create_distance_map” changes the space origin to \((0, 0, 0)\). In the next step, we will discuss how the centerline skeleton is adjusted to account for this shift.

4.4 Centerline skeleton

The **centerlines** of a segmented network form an associated network in which branches are only 1 pixel wide and run down the center of each of the segmented arteries. While convenient, centerlines are not necessary in order to obtain a graph representation of a segmented structure. Without using centerlines, a representative graph would look like a 3D mesh of the outside of the segmented structure. However, we wish to run 1D fluid dynamics simulations in the pulmonary arteries, so each branch must be viewed as a 1D blood conduit and is modeled with parameters used in the 1D code. Therefore, we obtain centerlines to build our labeled spatial tree.

There are various algorithms and softwares for extracting centerlines. The Vascular Modeling Toolkit (VMTK) [43], used in [96, 92, 93], requires the user to manually place target points at the
ends of terminal arteries, and the algorithm creates centerlines between the target points along the centers of maximally inscribed spheres in the segmented arteries. Belchi et al. [79] used custom written MATLAB code to extract their centerlines, though the details of the algorithm used for this are omitted from [79]. In this dissertation, we obtained the centerlines through a process called skeletonization, wherein voxels from the segmented network are iteratively removed until a skeleton which preserves the branching pattern of the network is obtained. This method follows the Asymmetric Thinning Algorithm [72, 84], which iteratively removes voxels while preserving critical cliques. Specific definitions regarding critical cliques can be found in [72, 84]; select definitions are described below for clarity.

4.4.1 Critical cliques

A simple voxel is one that can be removed from a voxel complex without changing the topology of the complex (i.e. creating a hole, splitting up connected components, etc.). A voxel complex is called reducible if it can be transformed into a single voxel by iteratively removing simple voxels. Two voxels are adjacent if they share a face (more specifically a k-face), which can be a 0-face (corner), 1-face (edge), 2-face (square 2D side) or 3-face (entire voxel). A clique (more specifically, a k-clique) is a group of voxels which all share a common k-face $\chi_k$. A clique is essential if it is the largest clique in the complex that shares the face $\chi_k$ (Figure 4.5).
Within a voxel complex, there are many essential cliques, and they can be categorized as regular or critical. For a clique \( C \), the **deleted neighborhood** of \( C \), \( N^*(C) \), is the set of all voxels which are adjacent to some voxel in \( C \) minus voxels in \( C \). An essential clique \( C \) is **regular** if \( N(C) \) is reducible, and it is **critical** if it is not regular. The Asymmetric Thinning Algorithm [72] iteratively removes voxels from a voxel complex such that after each iteration, at least 1 voxel from each critical clique remains. After several iterations of the algorithm, no more voxels can be removed without removing a critical clique. The result is a **skeleton** of the original voxel complex, a thinned structure which captures the topology of the original.

### 4.4.2 Centerlines

Skeletons generated by the Asymmetric Thinning Algorithm are not inherently centered. During skeletonization, there is often more than one choice of which voxel to keep from a critical clique \( C \). The algorithm uses a **selection function** to determine which voxel \( v' \in C \) to keep [72, 84]. There are different choices for selection criteria, a natural one being to keep the voxel whose coordinates come first in lexicographical order. In this study, we use the selection function

\[
\text{Select}(C) = v' \text{ for which } d(v') = \max_{v \in C} d(v),
\]
Figure 4.6 An example of the Select function at work. $S$ is a voxel complex which is a work-in-progress skeleton of a vessel junction (shaded in gray). $C$ (in red) is a critical clique of $S$. Any one of the voxels $v_1$, $v_2$, or $v_3$ can be kept, and the others removed, to preserve this critical clique in the skeleton. Our Select($C$) function chooses $v_2$ because it has the largest distance map value. Notice how it is also the more centered of the 3 voxels.

meaning we keep the voxel $v' \in C$ for which the distance map value (equation 6.1) is maximum, which tends to be the most centered voxel (Figure 4.6). Using this selection function ensures that the skeleton is centered in the segmented network; we refer to such a structure as the centerline skeleton, or the centerlines, of the network.

We use the DGTal program “thin” [80] to obtain the centerlines. The inputs for this function are the seg.nrrd and dmap.nrrd files, and we set the options “select=dmax” to ensure that the maximum distance map value is used as the selection criterion and “skel=1isthmus” to ensure we get a skeleton with 1D branches. The output is a new file skel.nrrd, which contains the centerlines and can be loaded and viewed in 3D Slicer.

4.4.3 Vessel radii

Since any voxel in the centerline skeleton is situated in the center of its corresponding vessels in the segmented network, its distance map value serves as an estimate of the vessel’s radius at that point. This definition of vessel radius works well with one caveat. As previously mentioned, the “create_distance_map” function resets the space origin of seg.nrrd to $(1, 0, 0), (0, 1, 0), (0, 0, 1)$ and $(0, 0, 0)$ respectively before generating dmap.nrrd. The centerlines, however, preserve the initial space directions and space origin of seg.nrrd. The result is a centerline skeleton that can be viewed in 3D Slicer as precisely overlapping the segmented network (Figure 4.4). However, when trying to pull
distance map values for voxels in the skeleton, their coordinates will not match up with the voxels in \textit{dmap.nrrd} due to the shift in space origin. To fix this, we manually change the space origin of the centerlines to \((0, 0, 0)\) by opening \textit{skel.nrrd} in a text editor\(^4\).

With this small adjustment, we can obtain the radii of the pulmonary arteries in the segmented network. To clarify, each voxel \(v\) in the centerline skeleton is situated in the center of some segmented artery. The distance map value \(d(v)\) gives the distance from that centered voxel to the nearest “background voxel”, i.e., distance from that centered voxel to the vessel wall. Therefore, \(d(v)\) provides an estimate \(r(v)\) of that artery’s radius at the point where \(v\) is the center voxel.

The centerline skeleton has a tree-like appearance (Figure 4.4). This appearance, coupled with the properties of the pulmonary arteries mentioned at the beginning of this chapter, make it natural to model the pulmonary arteries with a graph, specifically a labeled spatial tree. Details of this extraction are included in Chapter 6, with relevant graph theory background \([47]\) and explanation provided in the next section.

### 4.5 Labeled spatial tree

The arteries’ geometric and connectivity information are needed to conduct 1D fluid simulations in the pulmonary network. We obtain this by analyzing the skeleton as a graph, processing and correcting the graph to obtain a directed tree, and summarizing the geometric and connectivity information into matrices.

#### 4.5.1 Graph theory

Beginning with some graph theory background, a graph \(G = (V, E)\) is a mathematical object consisting of a set of vertices, \(V\), connected by a set of edges, \(E\) (Figure 4.7). If an edge connects two vertices \(a, b \in V\), we denote that edge \((a, b) \in E\) \([47]\).

Unless otherwise stated, a graph is assumed to be \textbf{undirected}, meaning no order is imposed on the vertices of an edge (edge \((a, b)\) may also be denoted \((b, a)\)). In a \textbf{directed} graph, each edge is assigned a direction, so an edge \((a, b)\) is said to run “from” vertex \(a\) “to” vertex \(b\), and an edge \((b, a)\) (which may or may not exist in the graph) would be directed from vertex \(b\) to vertex \(a\). The degree of a vertex \(a\), denoted \(\text{deg}(a)\), is the number of edges attached to \(a\). For example in Figure 4.7A, \(\text{deg}(a) = 3\). In a directed graph, vertices also have an \textbf{in-degree} and \textbf{out-degree}, indicating how many edges are directed toward or away from the vertex, respectively \([47]\). A \textbf{path}, is a collection of edges which connect to vertices in sequence. We denote a path between vertices \(a_1\) and \(a_n\)

\[
P[a_1, a_n] = \{(a_1, a_2), (a_2, a_3), ..., (a_{n-1}, a_n)\}.
\]

\(^4\)Recall that even if the space origin is adjusted to match the distance map, the radius values may still be skewed due to the reset of space directions. This is a bug that needs to be fixed to expand the utility of the code, but for the images used in this dissertation, this was not a problem due to the fact that the original segmented network has space directions of \((1, 0, 0), (0, 1, 0), (0, 0, 1)\).
Figure 4.7 A) A graph with vertices \( a-h \) and 2 connected components. B) A undirected and directed tree, with the root labeled on the directed tree. C) A spatial graph generated from one of the murine pulmonary artery CT scans. This graph sits in \( \mathbb{R}^3 \); each node and edge-point has \((x, y, z)\) coordinates.

For example, in Figure 4.7A, two paths which connect vertices \( b \) and \( d \) are \( P_1[b, d] = \{(b, a), (a, d)\} \) and \( P_2[b, d] = \{(b, a), (a, c), (c, d)\} \). \( P_1 \) is easily seen to be the shortest path, but for more complex networks (such as Figure 4.7C), algorithms like Dijkstra's Shortest Path algorithm [3, 41] exist to find the shortest path between vertices.

A graph may have more than one connected component, which is a portion of the graph where a path exists connecting every pair of vertices. For example, the graph in Figure 4.7A has 2 connected components, one containing vertices \{\( a, b, c, d, e \)\} and one with vertices \{\( f, g, h \)\}. Additionally, a graph may contain a cycle of edges, which is a path that exists from one vertex to itself (such as \{\( (a, c), (c, d), (d, a) \)\} in Figure 4.7A). A tree is a graph with one connected component and no cycles. A tree is typically undirected, but all trees referred to in this dissertation will be directed trees, as branches are oriented in the direction of blood flow. Our trees also each have a root, which is a unique vertex which all edges are directed away from (labeled in 4.7B).

4.5.2 Spatial graph

The skeleton is comprised of thin segments made up of sequences of adjacent voxels. The details of this network can be seen in a spatial graph, which is a graph where each voxel in the skeleton generates a vertex that is connected by edges to all adjacent voxels’ corresponding vertices. Every vertex has the same \((x, y, z)\) coordinates as its corresponding voxel in the skeleton.
The degree of a vertex in a spatial graph classifies it into one of 3 categories: terminal points (deg = 1), edge points (deg = 2), and branching points (deg ≥ 3). In a spatial graph, we refer to terminal and branching points as nodes. A branch is a path \( P[a, b] \) between two nodes \( a \) and \( b \) where the only vertices in the path other than \( a \) and \( b \) are edge points. An edge in the spatial graph represents two voxels in the skeleton being adjacent, while a branch represents the trajectory of a particular segmented artery in the pulmonary arterial network. Hence, branches and nodes are more interesting and relevant to study in a spatial graph than edges and vertices.

In the same way that a typical graph is denoted \( G = (V, E) \) as a collection of vertices \( (V) \) and edges \( (E) \), we denote a spatial graph as \( G_s = (N, B) \), a collection of nodes \( (N) \) connected by branches \( (B) \) that bend in 3D space. We denote a branch \( B_{ab} \in B \) between nodes \( a, b \in N \) as

\[
B_{ab} = (a, b : e_1, e_2, ..., e_k),
\]

where \( e_1, e_2, ..., e_k \) are edge points along the branch.

To extract a spatial graph from the centerline skeleton \( skel.nrrd \) (where space origin has already been adjusted), we run the DGTal program “analyze_graph”, using \( skel.nrrd \) as the input. This program performs a Depth First Search (DFS), where it analyzes every node and follows adjacent edge points until another node is reached, logging the nodes and edges in the graph [84] (further detail in Chapter 6). When running this program, we enable 3 options: reduce graph (“-r”) to ensure that degree 2 vertices are classified as edge points, merge nodes (“-m”) so that any 3 nodes connected in a cycle are merged into 1, remove extra edges (“-c”) to clear necessary edges as a result of merging.

Two files are outputs of this program, \textit{graph.txt} and \textit{data.txt}. The file \textit{graph.txt} contains the list of the nodes, labeled with numerical IDs and their \((x, y, z)\) coordinates, followed by the list of branches between nodes with their edge points’ coordinates listed. The file \textit{data.txt} contains some additional information about the graph, such as node degrees, branch lengths, etc. We do not use the \textit{data.txt} file for anything because the information provided in this file is not well-labeled and was originally only included for statistical purposes.

### 4.5.3 Labeled spatial tree

There are several reasons why the spatial graph provided in \textit{graph.txt} is not a sufficient final product to analyze if we want to extract 1D fluid dynamics parameters from our networks. The information is saved in a text file, which is not conducive to performing calculations. Also, the vessel radii are not included, and we have found that some common errors arise that do not reflect the biological nature of the network. To this end, we have written custom MATLAB programs for processing the raw graph output of “analyze_graph” and published them in the Github Repository TBN.

The function “formatData(‘graph.txt’, ‘dmap.nrrd’)”, generates 2 important structures: “nodes” and “arcs”. Nodes is a matrix with a row for each node and 5 columns which contain the numerical

---

5The program actually saves this as a .dot file, but simply changing the name to \textit{graph.txt} will convert it to a text file which makes it easier to view and read in the MATLAB code which follows.
node IDs (column 1), \((x, y, z)\) coordinates (columns 2-4), and node degrees (column 5). Arcs is a cell structure with a cell for each branch that contains the node IDs of the 2 nodes of the branch and the \((x, y, z)\) coordinates and radius of each node and edge point along the branch. The spatial graph can be visualized by running “plotSlicerData(arcs, nodes)”, which generates a 3D plot of the graph. This program also outputs the root node’s ID, taken to be the node with the highest \(z\)-coordinate. One should verify that this ID is in fact that of the root node by looking at the graph.

We noticed that the raw graphs typically contained some errors, such as small branches which are not reflective of the segmented network or small cycles created by 3 nearby voxels in the skeleton which are not merged in the graph extraction (more details in Chapter 6). Thus, we use the “correctionEngine(arcs, nodes, root_node_ID)” program to correct these errors. When all errors are corrected, there will be no remaining cycles because in a healthy pulmonary arterial network, no natural cycles occur so any that are shown in the graph are categorized as image error. Therefore, the spatial graph can now be called a spatial tree.

To label the tree’s branches with radii, lengths, and orientation, we run the program “directedGraphInfo(arcs, nodes, path)”. From this we get 2 important structures: a “network” matrix and a “connectivity” matrix. The network matrix has a row for each branch and 6 columns which contain each branch’s numerical branch ID, top node ID, bottom node ID, branch radius (taken as the interquartile mean of the node and edge point radii), branch length, and standard deviation of the node and edge point radii. The connection matrix has a row for each node and either 6, 7, or 8 columns. Each row contains the numerical node ID (column 1), out degree (column 2), in degree (column 3), daughter vessel IDs (for up to 4 daughters, columns 4, 5, 6 if there is a trifurcation present, and 7 if there is a quadfurcation present), and parent vessel ID (last column).

4.6 Multiscale geometric model

The pulmonary arterial network consists of vessels of varying sizes. Larger vessels are typically responsible for transporting blood to the various lobes of the lung, while smaller vessels distribute the blood within each lobe. Since these vessels have different roles, it is natural that we separate the large and small vessels and model them differently.

4.6.1 Principal pathway

The principal pathway of a pulmonary arterial network is the subtree of large vessels which transport blood to each lobe of the lung. Determining the principal pathway in the pulmonary arterial network is a non-trivial task. In the systemic arteries, the principal pathway is formed by named vessels (e.g., aorta, left/right coronary artery, etc.), distributing blood to all regions in the body. In the pulmonary vasculature, only the main, right, and left pulmonary arteries are named. These will of course be the first part of the principal pathway, but it is not sufficient to only include these arteries as they do not reach each lobe of the lung. To ensure all lobes are reached, we considered several criteria for deciding which branches of the labeled spatial tree to include in the principal pathway.
Figure 4.8 Sample network shown with the principal pathways (shown in dark blue on top of the light blue network) obtained from different choices of $\tau_r$. We found $\tau_r = 0.4$ to be a suitable choice for our networks. If $\tau_r < 4$, too many distal branches were picked up in the principal pathway. If $\tau_r > 4$, every lobe was not reached.

Alternatively, a left and right branch of the principal pathway can be defined based on criteria at junctions. With this method, we started with including the main, right, and left pulmonary arteries, and then building the rest of the pathway by choosing a single daughter to follow at each subsequent junction until a terminal vertex was reached. We examined several criteria for choosing which daughter to follow, such as larger radius, larger downstream volume, and larger number of vessels downstream. Ultimately, this method had the inherent flaw of only providing a single branch down the left and right, which was not enough to get into every lobe of the lung.

We ultimately decided to define our principal pathway based on the radius of the MPA, $r_{MPA}$. We defined a radius ratio threshold $\tau_r$, took the principal pathway to be the largest subtree of vessels whose radii are all greater than or equal to $\tau_r r_{MPA}$. Figure 4.8 shows the principal pathways generated for several values of $\tau_r$; we found $\tau_r = 0.4$ to be a suitable threshold value for determining the principal pathways in this study. In other words, the principal pathway is the largest subtree of vessels whose radii are all 40% of the MPA’s radius.

4.6.2 Small vessels

When examining the junctions of vessels off of the principal pathways, we found that over 98% of them were bifurcations. Therefore, we decided to model these vessels with self-similar bifurcating structured trees defined by the parameters discussed in [12, 31, 85] and Chapters 3 and 6,

\[
\alpha = \text{scaling factor for daughter 1 radius},
\beta = \text{scaling factor for daughter 2 radius},
\ell_{rr} = \text{length-to-radius ratio},
\theta_1, \theta_2 = 2D \text{ in-plane branching angles of daughters 1, 2},
\psi_1, \psi_2 = \text{out-of-plane rotational angles of daughters 1, 2}.
\]

Using these parameters, 1D fluid dynamic predictions can be computed in the small vessels.
### 4.7 Summary

The methods for labeled spatial tree extraction discussed in this section are summarized in Tables 4.3 and 4.4. Suppose that we start with a DICOM image called *myImage.dcm* of some network of interest (pulmonary arteries, bronchi, etc.). Assume that *myImage.dcm* has cubic voxels, meaning the space directions of the segmented/skeletonized image are \((1, 0, 0), (0, 1, 0), (0, 0, 1)\). The steps in Tables 4.3 and 4.4 can be followed to extract a labeled spatial tree representation of the network.

**Table 4.3** Summary of methods for labeled spatial graph extraction (part 1).

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Software</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Segmentation</td>
<td>Load <em>myImage.dcm</em> into 3D Slicer. Use the Segment Editor module to isolate the network of interest in the image. Save the result as <em>seg.nrrd</em>.</td>
<td>3D Slicer</td>
<td><em>myImage.dcm</em></td>
<td><em>seg.nrrd</em></td>
</tr>
<tr>
<td>2. Distance map</td>
<td>Navigate to the “DGTal-&gt;build-&gt;cpp-scripts” folder in the terminal. Run <code>.create_distance_map --input=path to *seg.nrrd* --outputFolder=path to destination folder for *dmap.nrrd*</code>.</td>
<td>DGTal</td>
<td><em>seg.nrrd</em></td>
<td><em>dmap.nrrd</em></td>
</tr>
<tr>
<td>3. Skeletonization</td>
<td>Still in “cpp-scripts” folder in the terminal, run <code>.thin --input=path to *seg.nrrd* --select=dmax --skel=1isthmus --inputDistanceMapImageFilename=path to *dmap.nrrd* --foreground=white --exportImage=path to destination folder for *skel.nrrd*</code>.</td>
<td>DGTal</td>
<td><em>seg.nrrd</em>, <em>dmap.nrrd</em></td>
<td><em>skel.nrrd</em></td>
</tr>
<tr>
<td>4. Adjust space origin</td>
<td>Open <em>skel.nrrd</em> with a text editor. Change the space origin to ((0,0,0)) and save the adjusted skeleton as <em>skel_adj.nrrd</em>.</td>
<td>Text editor</td>
<td><em>skel.nrrd</em></td>
<td><em>skel_adj.nrrd</em></td>
</tr>
<tr>
<td>5. Raw graph</td>
<td>Still in “cpp-scripts” folder in the terminal, run <code>.analyze_graph --input=path to *skel_adj.nrrd* -r -c -m --exportReducedGraph=path to destination folder for *graph.txt* --exportData=path to destination folder for *data.txt*</code>.</td>
<td>DGTal</td>
<td><em>skel_adj.nrrd</em></td>
<td><em>graph.txt</em>, <em>data.txt</em></td>
</tr>
<tr>
<td>6. Graph processing</td>
<td>Run MATLAB program ([\text{arcs}, \text{nodes}] = \text{formatData('graph.txt', 'dmap.nrrd')},) and then run ([\text{rootNode}]=\text{plotSlicerData(arcs, nodes)}.) Verify rootNode visually.</td>
<td>Github</td>
<td><em>graph.txt</em>, <em>dmap.nrrd</em></td>
<td><em>arcs</em>, <em>nodes</em>, <em>rootNode</em></td>
</tr>
</tbody>
</table>
Table 4.4 Summary of methods for labeled spatial graph extraction (part 2).

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Software</th>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Tree</td>
<td>Examine the figure made from 'plotSlicerData' and note any branches which need manual removal. Run ([\text{path}, \text{arcs}, \text{nodes}] = \text{correctionEngine(}	ext{arcs, nodes, rootNode})) and ([\text{orientation}, \text{newNetwork}, \text{connectivity, nodeDetails}] = \text{directedGraphInfo(}	ext{arcs, nodes, path})).</td>
<td>Github</td>
<td>(\text{arcs, nodes, rootNode})</td>
<td>arcs, nodes, path, orientation, newNetwork, connectivity, nodeDetails.</td>
</tr>
<tr>
<td>8. Principal</td>
<td>Run ([\text{arcs}<em>{pp}, \text{nodes}</em>{pp}, \text{path}<em>{pp}, \text{newNetwork}</em>{pp}, \text{Connection}<em>{pp}, \text{arcs}</em>{offpp}, \text{nodes}<em>{offpp}, \text{newNetwork}</em>{offpp}, \text{Connection}_{offpp}] = \text{ExtractPP(}	ext{arcs, nodes, newNetwork, MPA} _\text{ID})).</td>
<td>Github</td>
<td>(\text{arcs, nodes, newNetwork, MPA} _\text{ID})</td>
<td>arcs_pp, nodes_pp, path_pp, newNetwork_pp, Connection_pp, arcs_offpp, nodes_offpp, newNetwork_offpp, Connection_offpp</td>
</tr>
<tr>
<td>9. Structured</td>
<td>Run ([\text{alphas}<em>{offpp}, \text{betas}</em>{offpp}, \text{lrr}<em>{offpp}, \text{Angles}</em>{offpp}] = \text{ST_Parameters(}	ext{arcs}<em>{offpp}, \text{newNetwork}</em>{offpp}, \text{Connection}_{offpp})).</td>
<td>Github</td>
<td>(\text{arcs}<em>{offpp}, \text{newNetwork} _\text{offpp, Connection}</em>{offpp})</td>
<td>(\text{alphas}<em>{offpp}, \text{betas}</em>{offpp}, \text{lrr}<em>{offpp}, \text{Angles}</em>{offpp})</td>
</tr>
</tbody>
</table>
CHAPTER 5

PERSISTENT HOMOLOGY

Topological data analysis (TDA) is a field of mathematics which involves using topological markers in data to learn more about its shape or behavior [105]. Since topological properties are those which are preserved under continuous transformations [100], TDA has proven to be a useful tool for analyzing large data sets as it is robust to minor perturbations/noise in data [40, 64]. Typically, TDA methods involve using a point cloud from data as the basis for building a shape or set of shapes whose topology can then be analyzed to obtain new features and descriptors for the underlying structure of the data [105]. A particularly useful concept in the TDA realm, persistent homology, involves examining the evolution of the homology of nested complexes built upon data in search of topological properties that “persist” through several of the complexes.

The concept of persistent homology was initially developed by Frosini et al. [19], Edelsbrunner et al. [16, 23, 26], and Robins [30] during the 1980’s and 1990’s. This early work focused on “size functions” and “alpha shapes”. Since its early development, persistent homology has grown into an invaluable tool for studying the shape of data, particularly in biology [107, 106, 70, 100, 89, 91, 79]. The work by Belchi et al. [79] on classifying bronchial networks of healthy/COPD patients was of particular interest for our work since bronchial networks branch similarly to pulmonary arterial networks. Since the aim of this dissertation is to learn about the morphometry (geometric and topological properties) of the pulmonary arteries, we implemented a method inspired by [79] to see if persistent homology could distinguish PH networks from control networks. Before the method can be described, some basic topology definitions are provided.
5.1 Persistent homology background

To understand the biological applications of persistent homology, it is necessary to present some basic topological definitions. In summary, persistent homology involves building a sequence of topological spaces upon a point cloud of data, computing their homology, and tracking a topological summary that describes the shape of the data. Homology is used to detect features such as loops, connected components, and 3D cavities in data. The information in this section follows Sizemore’s TDA Primer at the Broad Institute’s 2016 Models, Inference and Algorithms meeting [74], Hatcher’s *Algebraic Topology* [32], and Nanda & Sazdanovic’s chapter “Simplicial Models and Topological Inference in Biological Systems” in *Discrete and Topological Models in Molecular Biology* [64].

5.1.1 Homology basics

A $k$-simplex is a convex hull of $k + 1$ vertices. In particular, a 0-simplex is a vertex, a 1-simplex is an edge, a 2-simplex is a triangular face, and a 3-simplex is a filled tetrahedron. The formal sum of simplices is called a simplicial complex $S$ (Figure 5.1).

![Figure 5.1 Example of a 0-, 1-, 2-, and 3- simplex, and a simplicial complex. The simplicial complex $S$ comprises five 0-simplices, six 1-simplices, and one 2-simplex.](image)

An oriented $k$-simplex is a $k$-simplex with an ordering $(v_0, v_1, ..., v_k)$ on its vertices (Figure 5.2). We denote an oriented $k$-simplex $\sigma_{0,...,k}$. Reversing the orientation of an oriented $k$-simplex is denoted with a negative sign. For example, the oriented 1-simplex $\sigma_{01}$ in Figure 5.2 is an edge from $v_0$ to $v_1$. With its orientation reversed, it is an edge from $v_1$ to $v_0$ and would be denoted $\sigma_{10} = -\sigma_{01}$.

An orientated 0-simplex is a vertex with a positive or negative sign (see $\partial_1(\sigma_{01})$ in Figure 5.2).

For $k \geq 1$, the boundary map, $\partial_k(\cdot)$, sends $\sigma_{0,...,k}$ to its boundary. The boundary $\partial_k(\sigma_{0,...,k})$ is the alternating formal sum of unique $k - 1$-simplices it contains. More specifically,

$$\partial_k(\sigma_{0,...,k}) = \sum_{i=0}^{k} (-1)^i \sigma_{0,...,\hat{i},...k}, \quad (5.1)$$

where $\sigma_{0,...,\hat{i},...k}$ denotes a simplex on all vertices of $\sigma_{0,...,k}$ except the vertex $\hat{i}$, which has been removed (Figure 5.2).
Figure 5.2 Examples of oriented $k$-simplices for $k = 1, 2, 3$, along with their boundaries.

For a simplicial complex $S$, the **chain vector space**, $C_k(S)$ is the vector space whose basis is the set of oriented $k$-simplices in $S$. Each element of $C_k$ is called a $k$-**chain**, $c$, such that

$$c = \sum_i a_i \sigma_i,$$

where each $a_i$ is an integer and each $\sigma_i$ is an oriented $k$-simplex from $S$. The boundary of a chain, $\partial_k(c)$, is the sum of the boundaries of its simplices. Hence, the boundary map is defined as $\partial_k : C_k \to C_{k-1}$ since it sends $k$-dimensional chains to their boundaries, which are $k-1$-dimensional chains. Some examples of chains and their boundaries are illustrated in Figure 5.3.

For geometric applications, it only makes sense to consider chain vector spaces $C_k$ for $k = 0, 1, 2, 3$, since they are generated from chains of vertices, edges, triangles, and tetrahedra, respectively. The boundary maps between these vector spaces are

$$\cdots \rightarrow C_3 \xrightarrow{\partial_3} C_2 \xrightarrow{\partial_2} C_1 \xrightarrow{\partial_1} C_0.$$

Let $L_k$ denote the kernel of $\partial_k$, i.e., everything in $C_k$ with a boundary $\partial_k = 0$. Let $B_k$ be the image of $\partial_{k+1}$, i.e., everything in $C_k$ which is the boundary of something in $C_{k+1}$. The $k$th **simplicial homology**, $H_k$, is the quotient of $L_k$ with $B_k$, i.e.

$$H_k = L_k / B_k = \ker(\partial_k) / \text{im}(\partial_{k+1}).$$

The $k$th **Betti number**$^{12}$ $\beta_k$ is the dimension of the homology $H_k$.

As we can see from chain $c_3$ in Figure 5.3, a 1-chain with a boundary of 0 forms a **loop**, a

---

$^1$Named for Italian mathematician Enrico Betti.

$^2$Note that this $\beta_k$ differs from the $\beta$ value described in Chapters 3, 4, and 6, which is a radius scaling factor across bifurcations. When a subscript is included, we take $\beta_k$ to mean the $k$th Betti number.
Figure 5.3 A simplicial complex $\mathcal{S}$, along with some of its chains and their boundaries. Note from examining $c_3$ that a chain does not need to be a connected simplicial complex. Also, note that the chain $c_3$ forms a loop of 1-simplices. Since all vertices in a loop both begin and end an edge, we find that $\partial_1(c_3) = 0$.

1-dimensional hole, in $\mathcal{S}$. This notion can be extended to say that any chain $c \in L_k$ forms a $k$-dimensional “hole”. Therefore, the Betti number $\beta_k$ counts the number of $k$-dimensional holes in a simplicial complex. In particular, $\beta_1$ is the number of loops and $\beta_2$ is the number of spheres $S^2$. The $0^{th}$ Betti number, $\beta_0$, is the number of disjoint clusters in a simplicial complex. Going forward, we will define a $k$-feature to be a property of a simplicial complex which is counted by $\beta_k$ (0-features are connected components, 1-features are loops, 2-features are spheres $S^2$).

5.1.2 Persistent homology

The methods of persistent homology can reveal information about the underlying shape of a data set. Often, our data set is a graph or point cloud in some $\mathbb{R}^n$. The $k^{th}$ homology $H_k$ of our data, and its dimension $\beta_k$ for $k = 0, 1, 2$, determine the number of $k$-features in the data. The working hypothesis is that this information about the shape can reflect other properties of interest in the data. For example, if $\beta_0 = 2$ for our data, then the data consists of 2 disjoint clusters which may reflect points with different values for some other property.

A point cloud itself is a simplicial complex made up of many separate clusters (individual 0-simplices), but it does not reveal relations between the points. The idea behind persistent homology is to iteratively add edges between the points based on their proximity and examine the topology of the intermittent shapes this creates. This process is called a filtration, whereby simplicial complexes are built upon a point cloud, with their formation governed by a distance function $\epsilon$. For an increasing sequence of $\epsilon$ values $\epsilon_1, ..., \epsilon_m$, a set of nested simplicial complexes $\mathcal{S}_{\epsilon_1} \subset \mathcal{S}_{\epsilon_2} \subset ... \subset \mathcal{S}_{\epsilon_m}$ is built upon
the data. **Degree-\(k\) persistent homology** is found by computing \(H_k(S_{\varepsilon_1}), H_k(S_{\varepsilon_2}), \ldots, H_k(S_{\varepsilon_m})\) and examining the evolution of their \(k\)th Betti numbers. For any \(i = 1, \ldots, m\), the Betti number \(\beta_k(S_{\varepsilon_i})\) counts the number of \(k\)-features in \(S_{\varepsilon_i}\). If a feature appears for the first time in a certain \(S_{\varepsilon_i}\), but is no longer present in \(S_{\varepsilon_j}\) for some \(j > i\), 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 \(P = \varepsilon_D - \varepsilon_B\). More “persistent” features are ones present for a large range of \(\varepsilon\) values. The last complex created in a filtration will have the topology of a point.

To decide on a filtration to use for computing persistent homology, a rule needs to be established for how to construct the set of nested simplicial complexes \(S_{\varepsilon_1} \subset S_{\varepsilon_2} \subset \ldots \subset S_{\varepsilon_m}\). A common filtration is the **čech complex filtration**, illustrated in Figure 5.4. Suppose our data is a point cloud \(S\). In the čech complex filtration, the distance function \(\varepsilon\) is the diameter of \(\varepsilon\)-balls centered at the points. Balls \(B_{\varepsilon_i}(s)\) of diameter \(\varepsilon_i\) are centered at each point \(s \in S\). Whenever \(k\) balls intersect, a \(k-1\) simplex is inserted between the centers of the \(k\) balls, resulting in the simplicial complex \(S_{\varepsilon_i}\). A similar filtration is the **\(\alpha\)-complex filtration**, which involves the Voronoi cells of \(S\). A **Voronoi cell** \(V_s\) is the region around \(s \in S\) such that every point \(x \in V_s\) is closer to \(s\) than it is to any other point \(s' \in S\). In the \(\alpha\)-complex filtration, balls \(B_{\varepsilon_i}(s)\) are constructed around each point \(s \in S\) and intersected with \(V_s\), forming a region \(R_s = B_{\varepsilon_i}(s) \cap V_s\). Whenever \(k\) regions \(R_s\) meet, a \(k-1\) simplex is inserted between their centers, resulting in the simplicial complex \(S_{\varepsilon_i}\). Other filtrations that are commonly used on point clouds are the Vietoris-Rips and Delaunay filtrations.

![Figure 5.4 Čech complex filtration applied to a point cloud. A) Visual representation of the filtration, showing the Čech complexes at various \(\varepsilon\) values. B) Barcode tracking the evolution of the 0- and 1- features of the complexes. C) Persistence diagram, showing points plotted at coordinates \((\varepsilon_B, \varepsilon_D)\) for each feature.](image)

If the data being examined is a graph \(G\), persistent homology can be computed via a **height filtration** (sometimes also called the “lower/upper star filtration”). A box is imagined that fully contains \(G\), and the distance function \(\varepsilon\) is the distance from the top of the box as we pan down the whole graph. At a given height \(\varepsilon_i\) from the top of the box, the simplicial complex \(S_{\varepsilon_i}\) is taken to be the portion of the graph that lies less than \(\varepsilon_i\) units from the top of the graph (Figure 5.5).

**Barcodes** and **persistence diagrams** are two ways of visualizing persistent homology. Barcodes contain 1 bar for each new \(k\)-feature detected in a filtration. Each bar spans the length from \(\varepsilon_B\) to...
Figure 5.5 Height filtration applied to the graph of a pulmonary arterial network. For various $\varepsilon$ values, we have shown the simplicial complex $S_\varepsilon$ and stated the degree-0 Betti number of each.

$\varepsilon_D$ of that feature (Figure 5.4B). Longer bars represent features that persist longer in the chosen filtration. Features with shorter bars are often interpreted to be noise, though this depends on the application. In persistence diagrams, this same information can be presented as points plotted in the unit square (Figure 5.4C). There is 1 point for each feature detected, plotted at the $(x, y)$ coordinates $(\varepsilon_B, \varepsilon_D)$. Clearly, all points will lie above the diagonal line $y = x$ since $\varepsilon_D \geq \varepsilon_B$. Points further from the line $y = x$ correspond to longer bars in the barcode.

The bottleneck distance, $d_B(X, Y)$ is used to compare two persistence diagrams. Given two persistence diagrams $X = \{x := (x_b, x_d)\}$ and $Y = \{y := (y_b, y_d)\}$, the bottleneck distance is

$$d_B(X, Y) = \inf_{\gamma: X \rightarrow Y} \sup_{x \in X} ||x - \gamma(x)||_\infty$$

where $\gamma$ ranges over bijections from $X$ to $Y$. Such bijections, sometimes referred to as “matchings”, also consider points on the diagonal in the event that $X$ and $Y$ have different cardinality [40, 109].

Persistence diagrams are known to be stable under small perturbations in the data [40, 64]. Consider a point cloud $P$ in which each point is perturbed by less than some fixed distance $h > 0$, generating another point cloud $Q$. For a given filtration, the degree-$k$ persistence diagrams $X_P$ and $X_Q$ are such that the bottleneck distance $d_B(X, Y)$ is bounded above by $h$. In other words, if the point clouds $P$ and $Q$ are close, their persistence diagrams will be equally close or closer in terms of bottleneck distance. It should be noted that the converse has not been proven true; close persistence diagrams do not necessarily indicate close data sets.
5.2 TDA for biological data analysis

In the last few decades, TDA has emerged as a powerful tool for applying mathematics to biological problems, including analysis of tumor growth [107, 106], brain function [70, 89], and bronchial networks [79]. In [107], persistent homology was used to analyze segmented CT images of lungs from the Cancer Imaging Archive. They defined a filtration that generated cubic complexes based on voxel intensity and tracked the persistence of features. The barcodes were then used to generate topological “feature curves”. They found that the first moments of the distribution in the 0D feature curves correlated inversely with lung cancer survival rate (those with the lowest first moments had significantly better survival rates). In [106], TDA was used in case studies of lung and brain cancer patients to determine how tumor shape relates to survival rates. They also defined a cubic complex filtration based on voxel intensity and tracked the persistence of features. They used this information in the Cox hazards model to determine the predicted survival length of the patients, and found that irregular and heterogeneous shape patterns were correlated with a shorter survival rate.

In [70], persistent homology was used to analyze brain arterial trees from magnetic resonance imaging (MRI) scans of humans of varying ages and sexes. The arterial trees were extracted using a tube-tracking segmentation algorithm. Degree-0 persistent homology was computed via height filtration (moving bottom to top, rather than top to bottom), while degree-1 persistent homology was computed by Vietoris-Rips filtration. Degree-0 persistence was strongly correlated with age, while degree-1 persistence was strongly correlated with sex. In [89], TDA was used to analyze neuron activity. The question in this study was whether or not persistent homology could reveal information about behavioral covariates, such as head position and spatial direction. Degree-1 and -2 persistent homology was computed for order complexes built from examining spikes in neural activity.

The work by Belchi et al. [79] applying TDA to bronchial networks was of particular interest for our study. It inspired us to consider if similar methods could be used to characterize pulmonary arteries, which have a similar branching pattern to bronchi. This work is, to our knowledge, the only study to consider persistent homology on a pulmonary network. Their study used TDA to study lung CT scans of control and mild/moderate chronic obstructive pulmonary disease (COPD) patients. Degree-0 persistent homology was computed for a height filtration, and it was able to distinguish between scans of healthy, mild COPD, and moderate COPD patients. Degree-2 homology was computed for the $\alpha$-complex filtration to distinguish inspiratory and expiratory scans. Chapter 7 includes a paper wherein we present our results from computing degree-0 homology for the height filtration on our pulmonary arterial graphs. This work was completed with students involved in an NCSU math REU in the summer of 2020.
CHAPTER 6

STRUCTURAL AND HEMODYNAMIC PROPERTIES OF MURINE PULMONARY ARTERIAL NETWORKS UNDER HYPOXIA-INDUCED PULMONARY HYPERTENSION

This chapter includes a verbatim reproduction of the manuscript MJ Chambers, MJ Colebank, MU Qureshi, R Clipp, MS Olufsen, “Structural and hemodynamic properties of murine pulmonary arterial networks under hypoxia-induced pulmonary hypertension”. Proc Inst Mech Eng, Part H: J Eng Med, 234(11): 1312–1329, 2020 [102]. I was responsible for designing this study, segmenting and skeletonizing the images, devising error corrections schemes, extracting network information, and computing structured tree parameters.

6.1 Abstract

Detection and monitoring of patients with pulmonary hypertension, defined as a mean blood pressure in the main pulmonary artery above 25 mmHg, requires a combination of imaging and hemodynamic measurements. This study demonstrates how to combine imaging data from micro-computed tomography (micro-CT) images with hemodynamic pressure and flow waveforms from
control and hypertensive mice. Specific attention is devoted to developing a tool that processes CT images, generating subject-specific arterial networks in which 1D fluid dynamics modeling is used to predict blood pressure and flow. Each arterial network is modeled as a directed graph representing vessels along the principal pathway to ensure perfusion of all lobes. The 1D model couples these networks with structured tree boundary conditions representing the small arteries and arterioles. Fluid dynamics equations are solved in this network and compared to measurements of pressure in the main pulmonary artery. Analysis of micro-CT images reveals that the branching ratio is the same in the control and hypertensive animals, but that the vessel length to radius ratio is significantly lower in the hypertensive animals. Fluid dynamics predictions show that in addition to changed network geometry, vessel stiffness is higher in the hypertensive animal models than in the control models.

Keywords: pulmonary hypertension, fractal networks, image segmentation, center line extraction, one-dimensional fluid dynamics, Navier-Stokes equations

6.2 Introduction

The pulmonary vasculature forms a rapidly branching network of highly compliant vessels that, in healthy subjects, conducts blood at low pressure. Blood is ejected from the right ventricle into the main pulmonary artery (MPA) and transported to every lobe within the lung. Since the lungs are situated deep in the body, it is not possible to measure blood pressure noninvasively. However, such measurements are essential to diagnose and assess the progression of pulmonary hypertension (PH), defined as a mean pressure above 25 mmHg in the MPA [67]. PH is rare, but the incidence rate is increasing [90], and the disease is associated with high morbidity and mortality [71]. A positive diagnosis requires an assessment of chest images and blood pressure measurements. Chest images can be obtained using radiography, ultrasound, magnetic resonance imaging (MRI), or computed tomography (CT) [87, 94], and blood pressure is measured invasively using right heart catheterization (RHC) [94]. To predict outcomes of interventions and to assess disease progression, these data streams should be integrated. One way to do so is by using a computational fluid dynamics model to predict blood pressure and flow in networks extracted from medical images and to compare these predictions with data. In this study, we solve a one-dimensional (1D) fluid dynamics model in pulmonary arterial networks extracted from micro-computed tomography (micro-CT) images from control and hypertensive mice and compare pressure predictions with \textit{in-vivo} pressure and flow measurements.

6.2.1 Geometry of pulmonary arterial vasculature

The morphometry of vascular networks has been the topic of numerous studies [1, 15, 29, 31, 10, 13, 27, 35, 73]. In 1926, Murray proposed a power law based on an optimality principle describing how vessel radii change across bifurcations [1]. Murray's law predicts optimal vessel dimensions, which minimize the total work under steady flow. In 1978, Zamir [15] derived additional optimality
principles, predicting a flow-radius relationship that minimizes the pumping power and lumen volume in cardiovascular networks. Zamir's optimality principles also describe how radii change across bifurcations. He introduced two ratio laws: an asymmetry ratio relating the radii of the two daughter vessels, and an area ratio relating the combined area of the daughter vessels to the area of their parent vessel. Both Murray's and Zamir's laws were inspired by data, but were derived from theoretical principles.

The studies by Murray and Zamir were set up to characterize whole cardiovascular networks, but since they are derived under a steady flow assumption, they are more appropriate for describing the flow in the small vessels, where pulsatility plays a minor role. Olufsen [29, 31] conducted the first study utilizing a multi-scale approach distinguishing large and small vessels. She solved nonlinear 1D fluid dynamics equations in the large vessels and used Murray's and Zamir's optimality principles to predict pressure and flow in the small vessels. The large vessels were represented by their radius, length, and connectivity within the network, while the small vessels were represented by a self-similar, asymmetric structured tree, relating the radii and length of daughter vessels to their parent. Rather than using an exponent of 3, proposed by Murray, Olufsen used an exponent of 2.76 obtained from analysis of arterial casts [20, 4, 10, 13, 27]. These casts were generated by injecting liquid resin into the arterial network, and the vessel dimensions were measured using calipers on the hardened resin cast. While these data provide essential geometric information, they were all obtained in a single lung, and therefore do not capture the variation between individuals [27]. Moreover, the casts were fragile, some pieces of the casts broke, and there is an inherent human error in measuring the dimensions with a physical tool. As noted above, Olufsen's original structured tree was informed by the cast data, i.e., it is not subject-specific. However, as discussed by Colebank et al. [92], inter-individual variation in vessel diameters, length, and connectivity significantly impact flow predictions, highlighting the importance of generating networks encoding subject-specific geometry.

In recent years, medical imaging technologies have emerged as valuable and efficient for obtaining high fidelity measurements of vascular geometries [35, 73]. Image-based analyses of pulmonary vascular networks have been done in a range of species using a variety of technologies. These studies provide a detailed description of the network geometry but were not used to investigate if the geometry satisfies aforementioned optimality principles. To our knowledge, only two studies (Burrowes et al. [38] and Clark et al. [86]) have generated large pulmonary subject-specific network models. These studies used imaging data to generate large arterial and venous space-filling networks. This method ensured that supernumerary vessels, small vessels that emerge at nearly 90° angles from the large arteries, were included in the model domain. In this study, we expand on these results by using data to generate self-similar networks informed by the branching structure in healthy control and hypertensive arterial networks from mice.
6.2.2 Pulmonary arterial hemodynamics

The pulmonary arteries form a rapidly bifurcating network with more than 20 generations, depending on the species [61]. Conducting nonlinear fluid dynamics simulations in subject-specific geometric networks of this size is not feasible, as the network would include more than $2^{20}$ vessels. One way to avoid this is to represent the vasculature by an artificially generated network that encodes the subject-specific branching structure.

Several pulmonary studies have used fluid dynamics to predict flow and pressure in the large vessels. Yang et al. [98] used a 3D subject-specific model to characterize the time-averaged wall shear stress in the MPA in pediatric pulmonary hypertension, and Kheyfets et al. [68] used a 3D model subject-specific model to compute spatially averaged wall shear stress under PH conditions. While 3D models provide high fidelity predictions, they require vast computing power making it difficult to use them for clinical analysis of large datasets. At the other end of the computational spectrum are 0D models that predict blood flow and pressure using an electrical circuit analogy. As noted in the review by Tawhai et al. [50], 0D models can predict hemodynamics at a low computational cost, but they require a plethora of parameters that are not uniquely identifiable [88]. Besides, 0D models cannot predict wave propagation. PH is associated with stiffening of large and small arteries, as well as microvascular rarefaction [36]. These morphological changes increase wave-propagation and, eventually, the load on the heart (i.e., ventricular afterload). As shown in our previous study [96], wave propagation can be predicted effectively using 1D models, achieving a higher fidelity prediction than the 0D model at a computational cost that is significantly lower than 3D models.

1D fluid dynamics models have been used to predict hemodynamics (flow, pressure, and cross-sectional area) in both arterial and venous networks. This model type connects large vessels, characterized by their length and radius, in a network informed by data. At the network inlet, an inflow profile, driving the 1D model, is specified from data or computed by a heart model. At the outlet, terminal vessels are coupled to the micro-circulation via boundary conditions formulated using: a) a Windkessel model (an electrical circuit with two resistors and a capacitor); b) a lumped parameter model linking the arterial network model to a closed-loop circuit; or c) a multiscale model coupling the larger arteries to a small-vessel geometric model.

The most commonly used 1D models use boundary conditions of type a). Studies by Fossan et al. [81] and Epstein et al. [66] found that in the systemic circulation, flow and pressure in the large arteries can be predicted from a network including the aorta and one branch off this vessel. Colebank et al. [93] drew similar conclusions in the pulmonary circulation, showing that the model sensitivity to boundary conditions decreases with network size. These results suggest that it is not necessary to explicitly represent all vessels in the vasculature and that it is possible to separate the network into large vessels (modeled explicitly) and small vessels (represented by boundary conditions). Another result by Colebank et al. [92] showed that changes in connectivity significantly impact flow and pressure predictions in the pulmonary circulation, emphasizing the importance of generating subject-specific models.
In the second model type, b), the arterial network model is linked to a 0D electrical circuit representation forming a closed-loop cardiovascular model. Mynard and Smolich [69] developed the most advanced model of this type. Their model represents the large arteries and veins in 1D, while the heart and small vessels are modeled using an electrical circuit. This model is ideal for studying flow and pressure in generic subjects, but due to its complexity, it is challenging to conduct subject-specific simulations.

Lastly, models of type c) utilize a multiscale approach representing the large vessels explicitly and the small vessels by fractal trees. The advantage of this approach is that it becomes feasible to predict hemodynamics in the large vessels using the full 1D Navier-Stokes equations, while small vessel hemodynamics can be computed using a simple linearized model. Olufsen [29] developed the first model of this type, studying hemodynamics in the systemic arteries. Spilker et al. [42] extended Olufsen’s results deriving a fractal tree model for the pulmonary arteries using data collected by Huang et al. [27]. Spilker’s work successfully estimated pressure, flow, and impedance (magnitude and phase) in pigs, but did not verify if the data in [27] provide a valid representation of porcine pulmonary vasculature.

These early studies [29, 31, 42] introduced multiscale models including both large and small vessels, but the structured trees representing the small arteries and arterioles were only used to provide an impedance boundary condition (as an alternative to the Windkessel model). Recent studies by Olufsen et al. [55] and Qureshi et al. [65] expanded these results developing a multiscale model predicting dynamic flow and pressure in both large and small pulmonary arteries. In these works, the vessel geometry for the large vessels was determined from data, while the fractal network was parameterized using literature data.

An alternative approach is to solve simplified equations (e.g., Poiseuille flow) in all vessels of the pulmonary tree. This approach was used by Borrowes et al. [38] who constructed a full network of the arteries and veins from imaging data. Clark et al. [86] solved the fluid dynamics equations using a linear transmission line model. This recent study added another level to the multiscale model, coupling arteries and veins to an alveolar sheet model and enabling prediction of perfusion. However, by not separating the large and small vessels, these alternative approaches ignored inertial effects in the large vessels.

The studies discussed above use data to guide model predictions but did not fit predictions to data. Our goal is to separate the vasculature into two parts: large vessels transporting blood to every lobe in the lung through a principal pathway, and small vessels perfusing each sub-area effectively. The large vessels will be described explicitly by their length, radius, and connectivity within the network informed by subject-specific CT images, while the small vessels will be represented by subject-specific fractal trees, using parameters extracted from micro-CT images. The latter is of importance as it will enable us to characterize remodeling with PH, which for most PH groups [36] starts in the small vessels, increasing vessel stiffness and decreasing the area. As the disease progresses, remodeling also affects the large vessels [35], described by large vessel parameters.
In summary, this study presents a data-driven approach to predict hemodynamics in pulmonary arterial networks extracted from control and hypertensive mice. Geometric data are obtained from micro-CT images from excised mouse lungs \[11, 51\] segmented using 3D Slicer® \[53, 63, 104\], skeletonized, and used to construct directed graphs. From these graphs, we extract principal pathway vessels and determine subject-specific fractal properties in the non-principal vessels. We conduct 1D fluid dynamics simulations predicting pressure and flow in both the large and small pulmonary arteries, and compare predictions between the control and hypertensive mice.

6.3 Materials and methods

This study includes three components: image analysis, network generation, and hemodynamics modeling. We briefly describe protocols used for data acquisition, followed by a detailed description of the image segmentation process, construction of directed graphs, principal pathway extraction, structured tree parameter calculation, and the 1D fluids model for predicting hemodynamics.

6.3.1 Experimental protocol

This study analyzes micro-CT images from male C57BL6/J mice, aged 10-12 weeks. The images were made available by Naomi Chesler, University of Wisconsin, Madison, and details of the imaging protocol can be found in \[11, 51\]. We analyze images from 3 control and 3 hypertensive mice, induced by keeping the mice in an hypoxic environment (FiO\(_2\) reduced by half, to 10%) for 10 days. The mice were euthanized by exsanguination, their lungs were extracted, and the MPA was cannulated (PE-90 tubing, 1.27 mm outer and 0.86 mm inner diameter) well above the first bifurcation. The pulmonary arteries were perfused with perfluorooctyl bromide at a pressure of 7.4 mmHg and placed in a micro-CT scanner. Lungs were rotated 360° in the imaging chamber, and planar images were obtained at 1° increments, resulting in 360 planar images. The planar images were reconstructed using the Feldkamp cone-beam algorithm and converted into 3D volumetric datasets and stored as Digital Imaging and Communications in Medicine (DICOM) 3.0 images.

Hemodynamic waveforms are extracted in-vivo from 7 control and 5 hypertensive adult male C57BL6/J mice aged 12-13 weeks. Hypertension is induced by placing mice in a hypoxic environment (FiO\(_2\) reduced by half, to 10%) for 21 days, while both control and hypertensive mice were exposed to a 12-hour light-dark cycle. For both groups, blood pressure is measured in the MPA using a 1.0-F pressure-tip catheter (Millar Instruments, Houston, TX) and recorded on a hemodynamic workstation (Cardiovascular Engineering, Norwood, MA) at 5 kHz. Blood flow velocity is obtained by spectral analysis of Doppler ultrasound audio signals (Visualsonics, Toronto, Ontario, CA) measured just distal to the valve angling the probe to obtain maximal possible velocity, as it provides more accurate measurement of the inner diameter, needed to calculate the volumetric flowrate. The resulting signal was inspected for quality and recorded on a hemodynamic workstation (Cardiovascular Engineering, Norwood, MA) at 30 MHz. The flow velocity and pressure signals are measured simultaneously in the MPA distal to the pulmonary valve. This calculation is conducted under the
assumption of a circular lumen. Both the pressure and flow waveforms were averaged over twenty cardiac cycles gated to the ECG. More details on the experimental protocol can be found in [58].

6.3.2 Image segmentation

A 3D volumetric representation (Figure 6.1b) was rendered from micro-CT scans (Figure 6.1a). To construct a vascular network encoded by a directed graph, the 3D representation was skeletonized (Figure 6.1b) using the image analysis process outlined in Figure 6.2.

![Figure 6.1](image)

**Figure 6.1** (a) Micro-CT image from a control mouse. (b) 3D segmentation and skeletonization (one-voxel wide) networks from representative control and hypertensive mouse. (c) Main pulmonary artery flow (ml/s) and pressure (mmHg) waveforms from 7 control and 5 hypertensive animals.

Each micro-CT scan yields a 3D image comprised of cubic cells of image data called voxels,
which are analogous to the square "pixels" in a 2D image. The dimensions of each image are $(497 \times 497 \times 497)$ voxels. Size variation between mice produces slightly varied spatial resolutions for each image, ranging between 30-40 $\mu$m [51].

The image analysis process in Figure 6.2 involves strategically extracting voxels from the image to obtain the basic network structure. We refer to each image as a voxel complex, $V$, defined as a finite set of voxels [72]. Each voxel $v \in V$ has spatial coordinates $(x_v, y_v, z_v)$ and intensity value $I(v) \in [0, 255]$, with 0 and 255 denoting the intensity of a black and a white voxel, respectively. This is the standard range for image intensity; it covers all voxels in the images analyzed in this study. However, it is worth noting that for more complicated images, this range could be expanded as necessary to capture the more varied intensities. In general, voxels representing anatomical features will have higher intensities than other voxels.

Segmentation involves isolation of the pulmonary arterial tree. This is done by partitioning $V$ into a set $F$ of “foreground” voxels (belonging to the tree) and a set $B$ of “background” voxels, $V = F \cup B$. We use the image analysis program 3D Slicer® [53, 63] to segment and construct a 3D representation of the foreground voxels (Figure 6.1b). This program was chosen because it is open-source and produces results quickly with little user input required.

The segmentation is obtained using global thresholding, median smoothing, and some manual editing. Global thresholding requires specification of a lower ($\tau_l$) and an upper ($\tau_u$) intensity. Every
voxel with \( I(v) \in [\tau_l, \tau_u] \) is included in the foreground \( F \). Thresholds are selected \textit{ad hoc} to ensure the entire tree is included in \( F \). Recall that the lung is excised and placed in a cylindrical hypobaric chamber before imaging, and that the arteries are perfused with a contrast, i.e., the high-intensity voxels belong to either the arterial tree or to the cylinder. As a result, we only need to specify the lower threshold \( \tau_l \) keeping the upper threshold at \( \tau_u = 255 \). To isolate the arterial network, we manually remove the voxels representing the imaging chamber from \( F \).

Median smoothing reduces overall noise in the segmentation by replacing all voxels in the image with voxels whose intensities are the median of nearby voxel intensities. A kernel is specified to dictate how many nearby voxels are averaged. A median smoothing kernel of \((3 \times 3 \times 3)\) voxels was applied to the images in this study. Also, the cannula can distort the appearance of the MPA, which can impede skeletonization. To prevent this, we manually smoothed the MPA, using as little manual editing as possible and taking care to not distort the true vessel geometry.

6.3.3 Skeletonization and graph extraction

The result of the segmentation process is a 3D volumetric representation of the foreground \( F \) (representing the arterial tree). The next steps needed to generate a spatial graph include obtaining 1) a distance map and 2) a centered skeletonization. Using these, we design a subject-specific directed graph (including all vessels visible in the image).

The \textbf{distance map} \( D(V) \), or simply \( D \), is a voxel complex with the same spatial dimensions as \( V \), which encodes distance. For a segmentation \( V = F \cup B \), we define the function \( d : V \rightarrow \mathbb{R} \) to be:

\[
d(v) = \min_{u \in B} ||u - v||_2
\]  

where \( ||u - v||_2 \) is the Euclidean distance between \( u \) and \( v \). In other words, \( d(v) \) is the distance from the voxel \( v \) to the nearest background voxel. For each \( v \in V \), there exists a corresponding voxel \( v_D \in D \) with the same spatial coordinates \((x_v, y_v, z_v)\) and intensity \( I(v_D) = d(v) \).

The \textbf{skeletonization} \( S(F, D) \) refers to a thinned version of the foreground voxel complex \( F \), which preserves the connection and branching pattern of \( F \) (Figure 6.1b). We compute \( S \) using Couprie and Bertrand’s “Asymmetric Thinning” algorithm, which iteratively removes voxels from \( F \) following the framework of “critical kernels” \([72, 84, 80]\). This creates a thinned copy of \( F \), in which each branch is one voxel wide. During the “Asymmetric Thinning” algorithm, a choice must periodically be made to chose what voxels, within a group of voxels \( X \subseteq F \), should be included in \( S \). To determine which voxels \( \hat{v} \in X \subseteq F \) should be kept in \( S \), we define the function \text{Select}: \( F \rightarrow F \) given by

\[
\text{Select}(X) = \{ \hat{v} \} \text{ such that } d(\hat{v}) = \max_{v \in X} d(v).
\]
The voxel with the maximum \( d(v) \) value is the most centered. By choosing this voxel, the resulting skeletonized network \( S \) is centered in \( F \) (Figure 6.1b) [84]. We refer to \( S \) as the “skeletonization” or the “centerline network” of \( F \).

To examine vessel branching patterns [80] we extract a spatial graph \( G(S) \), or simply \( G \), from \( S \). Details of the graph extraction are described in [84], and relevant graph terminology is included in Table 6.1. Each voxel in \( S \) corresponds to a terminal node, a junction node, or an edge point in \( G \) and is connected to all adjacent voxels of \( S \). Starting at one of the terminal nodes \( a \), the algorithm records coordinates of adjacent edge points \( e_1, e_2, \ldots \) until another terminal or junction node \( b \) is reached. Next, the nodes \( a \) and \( b \) are added to \( N_G \) and the edge \( E_{(a,b)} = \{[a,b], e_1, e_2, \ldots, e_m \} \) is added to \( E_G \), while marking nodes that have already been visited. This is repeated for all terminal and junction nodes. To ensure that the resulting graph is a tree, small cycles of nodes of degree \( \leq 3 \) are broken by removing the longest edge in the cycle (Figure 6.3). This is known as “merging”.

### 6.3.4 Exception handling

The skeletonization process produces errors in the graphs that do not align with the 3D representation. These errors can be categorized into four types: false branches, double edges, duplicate points, and small cycles. The false branches have to be detected manually, while the other error types can be detected automatically.

**False branches (FB)** arise when the skeletonization has nodes/edges that do not represent actual

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>degree</td>
<td>The degree of a voxel ( v \in S ), denoted ( \text{deg}(v) ), is the number of voxels in ( S ) adjacent to ( v ).</td>
</tr>
<tr>
<td>nodes</td>
<td>A node is a voxel ( v \in S ) with either ( \text{deg}(v) = 1 ) (terminal node) or ( \text{deg}(v) &gt; 2 ) (junction node). Nodes are given numerical node IDs, starting at 0 and denoted ( nID(v) ). The root node of a network is the terminal node at the end of the edge representing the MPA.</td>
</tr>
<tr>
<td>edge points</td>
<td>An edge point is a voxel ( v \in S ) with ( \text{deg}(v) = 2 ) that lies along an edge between two nodes.</td>
</tr>
<tr>
<td>edges</td>
<td>An edge, denoted ( E_{(a,b)} ), is a collection of consecutively adjacent voxels between two nodes ( a, b \in N_G ). Any ( v \in E_{(a,b)} \cap {a,b} ) are edge points. We write the edges in the form ( E_{(a,b)} = {[a,b]: e_1, e_2, \ldots, e_m } \in E_G ), where ( e_i ) is the ( i )th edge point along ( E_{(a,b)} ). An arbitrary ordering is placed on the edges so that each ( E_{(a,b)} ) has a numerical ID ( k ) ranging from 1 to (</td>
</tr>
<tr>
<td>spatial graph</td>
<td>A spatial graph ( G = (N_G, E_G) ) is a collection of nodes ( (N_G) ) connected by a collection of edges ( (E_G) ). For two nodes ( u, v \in S ) with ( nID(u) = a ) and ( nID(v) = b ), we say ( a, b \in N_G ) and denote an edge between them as ( E_{(a,b)} \in E_G ).</td>
</tr>
<tr>
<td>small cycle</td>
<td>A cycle occurs in a spatial graph ( G ) if for three nodes ( a, b, c \in N_G ), we have edges ( E_{(a,b)}, E_{(b,c)}, E_{(c,a)} \in E_G ). A small cycle is a cycle where each of the three edges involved consist of only their nodes and no additional edge points, ie: ( E_{(\text{node1, node2})} \cap {\text{node1, node2}} ) for all ( E_{(\text{node1, node2})} ) in the cycle. These small cycles are discounted as graphical errors rather than anatomical loops since the cycle only involves 3 adjacent voxels.</td>
</tr>
</tbody>
</table>
Figure 6.3 Example spatial graph before and after merging. Voxels are labeled with coordinates, degree, and node ID. The red edges in the left graph represent a small cycle in $G$, which is broken in the merged graph $G^\ast$. 
Figure 6.4 Schematics of the four types of errors that can occur during skeletonization: (a) false branches (FB), (b) double edges (DE), (c) duplicate points (DP), and (d) small cycles (SC). These figures are not drawn to scale but represent the connectivity observed in each type of error. Any gray edges or white nodes with dashed borders are removed to correct the error. Black nodes with white borders are merged into edge points, while white nodes with black borders remain nodes. An exception is the red edge in SC4. To correct this error, this edge must be inserted manually.
branches in the 3D representation. These are corrected manually as the graph data has no record of the original segmentation. The only way to identify a false branch is by visually detecting that the branch is not part of the 3D representation. False branches can be separated into 4 categories FB1-FB4 (Figure 6.4a). FB1 occurs when branches form a “v” at the root node. FB2 occurs when an edge branches off of a straight vessel. FB3 occurs when a vessel further down the tree branches upward close to the first bifurcation, causing a false connection. FB4 occurs when a branch connects the first two daughter vessels, forming a cycle.

**Double edges (DE)** occur when two nodes are connected by two different edges. The longer of the two edges is removed (Figure 6.4b).

**Duplicate points (DP)** occur when an edge \( E(a,b) = \{[a, b]: e_1, e_2, ..., e_m \} \) has \( e_i = e_{i+1} \). In this case we delete the duplicate points \( e_{i+1} \). This error is shown in Figure 6.4c.

**Small cycles (SC)** are associated with nodes of degree 4. The merging protocol (Figure 6.3) is meant to correct all small cycles, but only works for cycles containing nodes of degree \( \leq 3 \). Therefore, some SC remain after the merging protocol. To correct these, we apply Dijkstra's shortest path algorithm [3, 41]. Let \( a_{root} \in N_G \) be the ID of the root node in a graph \( G = (N_G, E_G) \). A path of length \( p \) is an ordered list of nodes \([a_1, a_2, ..., a_p] \subseteq N_G \) such that any 2 nodes \( a_k, a_{k+1} \) are connected via an edge in \( E_G \). For each \( a \in N_G \), Dijkstra's algorithm returns the shortest path from \( a_{root} \) to \( a \). We denote this path \( P_a \) and the set of all these paths for a graph as \( P_G \).

\[
P_G = \{P_a = [a_1, a_2, ..., a_p]\}
\]

where \( a_1 = a_{root} \) and \( a_p = a \). For all edges \( E(a,b) \in E_G \), if no path in \( P_G \) has the nodes \( a \) and \( b \) listed consecutively. Short cycles are broken by removing the edge from \( E_G \).

The small cycles include 4 types, SC1-SC4, shown in Figure 6.4d. SC1s are found in bifurcations where the bottom node of one of the daughter vessels has degree 4. SC2 occurs when a cycle forms between two daughters in a trifurcation. SC3 occurs when 2 cycles form between 2 pairs of daughters in a trifurcation. SC4 occurs at a bifurcation if the bottom nodes of each daughter are the top nodes of edges ending in the same bifurcation node, causing a loop of 4 voxels. In SC1-SC3, the protocol described above automatically corrects the cycles. For SC4, it is necessary first to create an edge manually (shown in red in Figure 6.4(d)), and then use Dijkstra's algorithm to break the cycles.

After error correction is complete, some nodes of degree 2 may exist as a result of removing edges. We convert these to edge points of a newly defined edge connecting the two adjacent nodes (Figure 6.4).

### 6.3.5 Network generation

The result of the graph extraction process is a tree with the same branching pattern as the arterial tree. The edges in this tree are one voxel in diameter, and each edge represents a vessel between junction nodes (or between a junction and a terminal node). The next steps needed to generate an
arterial network are: 1) orienting the edges, 2) obtaining the vessel radii, and 3) obtaining the vessel length.

**Edge orientation:** Since blood flow in the pulmonary arteries has a direction, we represent the network by a directed graph, i.e., every edge must have a “start” node and “end” node indicating the direction of blood flow through that vessel. The graph extraction described above does not distinguish between start and end nodes, i.e., for a given edge \( E(a, b) \in E_G \) it is not necessarily true that the edge starts at node \( a \) and ends at node \( b \). We use the paths \( P_G \) (equation (6.3)) to obtain proper vessel orientations. If \( E(a, b) \in E_G \) with \( k = vID \{ E(a, b) \} \) and \( P_b \in P_G \) has length \( p_b \), then the orientation of \( E(a, b) \) is given by

\[
o(k) = \begin{cases} 
1 & \text{if } P_b(p_b - 1) = a \\
-1 & \text{otherwise}
\end{cases}
\]  

(6.4)

In other words, \( o(\{ E(a, b) \}) \) is 1 if \( a \) is the start node and \(-1\) if \( b \) is the start node.

We use the orientation to correct the order of the nodes for each \( E(a, b) \in E_G \). If \( o(\{ E(a, b) \}) = 1 \), we leave \( E(a, b) \) unchanged. If \( o(\{ E(a, b) \}) = -1 \), we replace \( E(a, b) = \{ [a, b] : e_1, \ldots, e_m \} \) with the edge \( E(b, a) = \{ [b, a] : e_m, \ldots, e_1 \} \). In the resulting network, any edges \( E(i, a) \in E_G \) has the start node ID \( a \) and the end node ID \( b \).

**Vessel dimensions:** Each voxel \( v \in S \) is centered in \( F \), so the value \( d(v) \) from the distance map (equation (6.1)) gives an approximation for the vessel radius at \( v \). For each vessel \( E(i, a) \in E_G \) with \( k = vID \{ E(i, a) \} \), the vessel radius \( r(k) \) is defined as the inter-quartile mean (IQM) averaging \( d(v) \) values for all voxels \( v \in E(i, a) \). The vessel length \( L(k) \) is computed by summing the Euclidean distances between each pair of consecutive points along the vessel.

### 6.3.6 Fluid dynamics domain

To conduct fluid dynamics simulations, we characterize vessels as either large or small. Large arteries transport blood to each lobe within the lung, and the small arteries distribute the blood within each lobe.

**Principal pathway:** The large arteries form a subtree known as the *principal pathway*. In the systemic arteries, the principal pathway is formed by named vessels (e.g., the aorta, left and right coronary artery, etc.) distributing blood to all regions in the body. In the pulmonary vasculature, only the right and left pulmonary artery are named, and the vessel morphometry differs significantly between individuals and can therefore not be identified uniquely. However, it is not sufficient to only include the left and right pulmonary arteries as they do not reach each lobe of the lung. To ensure all lobes are reached, we use a scaling argument to select vessels within the principal pathway.

The MPA and its two daughter vessels are always part of the principal pathway. Every other vessel whose radius is at least 40% of its parent vessel’s radius is designated as a principal pathway vessel. The network of principal pathways is obtained by taking the largest connected component of
Figure 6.5 Example structured tree. A) Radii at bifurcations are determined by multiplying the radius of the parent vessel by $\alpha$ (for the daughter of larger radius) and $\beta$. In our fluids simulations, structured trees are attached at all terminal nodes of the principal pathway, shown on the right in blue. B) Projection of the structured tree in 2D noting the $\alpha$-branch and the $\beta$-branch.

these vessels. We denote the principal pathway by $PP = (N_{PP}, E_{PP})$, where $N_{PP}$ are the nodes and $E_{PP}$ are the edges. The principal pathway is a subtree of the entire network graph $G$, as shown in Figure 6.5a.

**Structured tree:** Vessels off the principal pathway are represented by structured trees in which the daughter vessels' radii are scaled relative to their parent vessel (Figure 6.5(A)) [1, 15, 31]:

$$r(d_1) = \alpha \ r(p), \quad r(d_2) = \beta \ r(p)$$

(6.5)

where $p, d_1, d_2$ denote the vessel IDs of the parent and two daughter vessels respectively, with $r(d_1) \geq r(d_2)$. For each network, we computed $\alpha$ and $\beta$ at every bifurcation and subsequently averaged values to get one $\alpha$ and $\beta$ value for the entire network. Similarly, we compute the length to
radius ratio \( \ell_{rr}(k) \) for each vessel \( k \) off the principal pathway by solving

\[
\ell_{rr}(k) = L(k)/r(k).
\] (6.6)

### 6.3.7 Fluid dynamics

To simulate pulmonary hemodynamics, we use the 1D fluid dynamics model, predicting flow, area, and pressure in the large and small arteries. In the large vessels, comprising the principal pathway, we solve the 1D Navier-Stokes equations, and in the small arteries and arterioles, represented by structured trees, we solve a linearized 1D model.

**Principal pathway:** Similar to our previous studies [96, 92, 55], we predict flow \( q(x, t) \) (mL/s), pressure \( p(x, t) \) (mmHg), and area \( A(x, t) \) (cm\(^2\)) in each large vessel within the principal pathway, ensuring conservation of mass

\[
\frac{\partial A}{\partial t} + \frac{\partial q}{\partial x} = 0,
\] (6.7)

and momentum

\[
\frac{\partial q}{\partial t} + \frac{\partial}{\partial x} \left( \frac{q^2}{A} \right) + \frac{A}{\rho} \frac{\partial p}{\partial x} = -2\pi \nu r \frac{q}{\delta},
\] (6.8)

where \( \rho = 1.057 \) (g/mL) is the blood density, \( \nu \) (cm\(^2\)/s) is the kinematic viscosity, and \( \delta = 0.03 \) (cm) is the boundary layer thickness, approximated by \( \sqrt{\nu T/2\pi} \), where \( T = 0.11 \) (s) is the length of the cardiac cycle. The right hand-side of the momentum equation (6.8) is derived under the assumption of a flat velocity profile with a linearly decreasing boundary layer \( \delta \). To close the above system of equations, we relate pressure and area [31] as

\[
p(x, t) = p_0 + \frac{4}{3} \frac{E h}{r_0} \left( 1 - \sqrt{\frac{A_0}{A}} \right),
\] (6.9)

where \( p_0 \) (mmHg) and \( A_0 = \pi r_0^2 \) (cm\(^2\)) are the reference pressure and area, \( E \) (mmHg) is the Young’s modulus, and \( h \) (cm) is the wall thickness. The vascular stiffness \( E h/r_0 \) is related to the unstressed vessel radius \( r_0 \), and can be approximated using the functional form

\[
\frac{E h}{r_0} = k_1 e^{-k_2 r_0} + k_3
\] (6.10)

where \( k_1 \) (mmHg), \( k_2 \) (cm\(^{-1}\)), and \( k_3 \) (mmHg) are positive constants.

Since the system of PDEs (6.7)-(6.9) is hyperbolic with characteristics of opposite sign, each vessel requires two boundary conditions, one at each end of the vessel. Hence, when combining vessels in a bifurcating network, three types of boundary conditions are needed: 1) at the inlet to the root vessel, 2) at each junction node, and 3) at each terminal node. At the inlet of the root vessel, we prescribe flow informed by data. Hence, at the junctions between vessels, three conditions are needed: a condition at the outlets of the parent vessel and two conditions at the inlet to the daughter
vessels. These are obtained by enforcing continuity of total pressure and conservation of flow, i.e.,
\begin{align*}
p_p(L_p, t) &= p_{d_1}(0, t) = p_{d_1}(0, t), \tag{6.11} \\
q_p(L_p, t) &= q_{d_1}(0, t) + q_{d_2}(0, t) \tag{6.12}
\end{align*}
which holds \( \forall t \in [0, T] \). Finally, the end of each terminal vessel is linked to a structured tree by matching the impedance by relating pressure and flow via the discrete approximation of the convolution integral given by
\[ p(L, t_i) = \Delta t \sum_{k=0}^{M+1} z(0, t_k) q(L, t_{i-k}) \tag{6.13} \]
where \( z(0, t_k) \) is the impedance of the structured tree, \( \Delta t \) is the magnitude of the time step and \( M = \Delta t / T \).

The large artery equations are solved numerically using the two-step Lax-Wendroff finite difference scheme [31].

**Structured tree:** Fluid dynamics in the large arteries of the pulmonary circulation are predominantly inertia driven, whereas viscous forces predominantly influence hemodynamics in the small arteries and arterioles represented by structured trees. In these vessels, we ignore inertia enabling us to solve linearized forms of eqs. (6.7) and (6.8). These equations do not depend on the nonlinear convective terms, and have a Reynolds number that is sufficiently small so that a fully developed flow is present along the length of the small arteries and arterioles. Writing pressure and flow as periodic, frequency domain functions (i.e., \( P(x, \omega) \) and \( Q(x, \omega) \)), we can write linearized mass conservation and momentum balance equations for each frequency \( \omega_k \) as
\begin{align*}
i \omega_k C P + \frac{\partial Q}{\partial x} &= 0, \tag{6.14} \\
i \omega_k Q + \frac{A_0(1 - F_j)}{\rho} \frac{\partial P}{\partial x} &= 0 \tag{6.15}
\end{align*}
respectively, where \( F_j \) denotes the quotient of first and zero order Bessel functions of the first kind, \( C \) is the vessel compliance, and \( w_0 \) is the Womersley number (see Appendix 1 for detailed derivations).

Taking a derivative of equation (6.14) with respect to \( x \) and using equation (6.15) to solve for \( P \) gives the wave equation
\[ \frac{\omega_k^2}{c} Q + \frac{\partial^2 Q}{\partial x^2} = 0, \quad c = \sqrt{\frac{A_0(1 - F_j)}{\rho C}} \tag{6.16} \]
where \( c \) (cm/s) is the pulse wave propagation velocity. Solving equation (6.16) for \( Q \) and using this...
solution in (6.15) gives

\[ Q(x, \omega_k) = a \cos(\omega_k x/c) + b \sin(\omega_k x/c), \]
\[ P(x, \omega_k) = \frac{i}{g} (-a \sin(\omega_k x/c) + b \cos(\omega_k x/c)) \]

(6.17)

where \(a\) and \(b\) are arbitrary integration constants and \(g = c C = \sqrt{CA_0(1-F_J)/\rho} \).

The structured tree equations (6.17) can be solved analytically, taking advantage of their periodic nature. From these equations, we predict vascular impedance in the frequency domain as

\[ Z(x, \omega) = \frac{P(x, \omega)}{Q(x, \omega)} = \frac{i(b \cos(\omega x/c) - a \sin(\omega x/c))}{g(a \cos(\omega x/c) + b \sin(\omega x/c))}, \]

(6.18)
i.e., the impedance at the start and end of each small artery is given by

\[ Z(0, \omega) = \frac{i\ b}{g\ a}, \]
\[ Z(L, \omega) = \frac{i(b \cos(\omega L/c) - a \sin(\omega L/c))}{g(a \cos(\omega L/c) + b \sin(\omega L/c))}. \]

(6.19)

Combining these equations allows us to predict the impedance at \(x = 0\) as a function of \(Z(L, \omega)\)

\[ Z(0, \omega) = \frac{ig^{-1} \sin(\omega L/c) + Z(L, \omega) \cos(\omega L/c)}{\cos(\omega L/c) + ig Z(L, \omega) \sin(\omega L/c)}. \]

(6.20)

The input impedance at the zero frequency, i.e., analogous to the DC component, is given by

\[ \lim_{\omega \to 0} Z(0, \omega) = \frac{8 \mu_{rr}}{\pi r_0^3} + Z(L, 0). \]

(6.21)

As the radius of the small arteries decreases, the effects of blood viscosity become more important. Following \([21, 24]\), we assume that the viscosity in the small arteries and arterioles follows

\[ \mu^*(r_0) = [1 + (\mu_{0.45} - 1) \left( \frac{1 - \text{Hct}}{0.45} \right)^C - 1] \mathcal{G} \]

(6.22)

\[ \mu_{0.45}(r_0) = 6 e^{-0.1770} + 3.2 - 2.44 e^{-0.1270^{0.645}} \]

(6.23)

where \(\text{Hct}\) is the blood hematocrit, \(\mathcal{G} = \left( \frac{2r_0}{\pi r_0^2} \right)^2 \mu_{0.45}(r_0)\) is the relative viscosity at an average hematocrit level of 0.45, and \(C\) is a shape parameter derived from a nonlinear, empirical formula (see \([24]\) for derivation). In this study we use Hct= 0.6 for the control animals and Hct= 0.81 for the hypertensive animals. The above equation was used to match viscosity values in humans, where the viscosity of the large vessels is 3.2 (g/cm s). To adapt this model to the mouse viscosity of 0.049, we scale (6.22) giving \(\mu_{ST} = \mu \cdot \mu^*(r_0)/3.2.\)
Table 6.2 Network information. The \( \ell_{rr} \) constants \( C_1 \) and \( C_2 \) are reported as coefficients ± margin of error for a 95% confidence interval. All other values are mean ± std. deviation.

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Hypertensive</th>
</tr>
</thead>
<tbody>
<tr>
<td># of vessels</td>
<td>2573 ± 517</td>
<td>3239 ± 1103</td>
</tr>
<tr>
<td>% of total vessels in ( PP )</td>
<td>5.63 ± 1.34%</td>
<td>4.49 ± 2.03%</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.883 ± 0.122</td>
<td>0.864 ± 0.135</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.666 ± 0.141</td>
<td>0.655 ± 0.137</td>
</tr>
<tr>
<td>( \ell_{rr} (C_1) )</td>
<td>13.4 ± 1.2</td>
<td>10.9 ± 1.1</td>
</tr>
<tr>
<td>( \ell_{rr} (C_2) )</td>
<td>0.00771 ± 0.00083</td>
<td>0.00797 ± 0.0001</td>
</tr>
</tbody>
</table>

Similar to the large vessels, at each junction we enforce conservation of pressure and flow giving

\[
Z_p(L, \omega)^{-1} = Z_{d_1}(0, \omega)^{-1} + Z_{d_2}(0, \omega)^{-1}.
\] (6.24)

The structured tree is generated using the \( \alpha, \beta, \) and \( \ell_{rr} \) values obtained from the analysis of the data. The structured tree bifurcates until the radius of any vessel is less than a specified critical minimal radius value, \( r_{\text{min}} = 0.005 \) (cm), where the size of the red blood cell is no longer negligible compared to the size of the vessel. At this radius, we require a known terminal impedance, \( Z_{\text{trm}} \), to be able to back calculate all the structured tree impedance values. To match experimental setup conducted ex-vivo, with a zero-external pressure, at the end of the structured trees we enforce zero pressure, giving \( Z_{\text{trm}} = 0 \). The structured tree equations are solved by recursively computing the impedance of each daughter vessel starting at the terminal branches. Due to the self-similarity of the tree, we do not recompute the impedance in vessels scaled by \( \alpha^i \beta^j \) if it has already been computed previously, speeding up computation.

Finally, to compute pressure along vessels within the structured tree, we use the root impedance in a forward algorithm to predict pressure along the \( i^{th} \) branch [55] as

\[
P^i_L = P^i_0 \cos \left( \frac{\omega_k L^i}{c^i} \right) - \frac{i q_c}{\lambda^i p_c} Q^i_0 \sin \left( \frac{\omega_k L^i}{c^i} \right),
\] (6.25)

where \( i = \alpha, \beta, \alpha \beta, ..., \alpha^n \beta^m \), and \( n \) and \( m \) are the maximal \( \alpha \) and \( \beta \) branches obtained before reaching \( r_{\text{min}} \) (see Figure 5b).

### 6.4 Results

#### 6.4.1 Structured tree parameters

Structured tree parameters \( \alpha, \beta, \) and \( \ell_{rr} \) are determined in all vessel off the principal pathway (the small vessels). Table 6.2 reports average values over all small vessels in the networks. Figure 6.6 depicts parameter values from the data plotted against vessel radius. For each parameter, the observations are divided into 20 bins, and bin averages are plotted at the midpoint of the bin. Both
graphs include computations in each of the 3 control and hypertensive animals.

Figure 6.6a shows $\alpha$ and $\beta$ as a function of the parent vessel radius. Reference lines are plotted denoting the average $\alpha$ and $\beta$ values for each group of mice. Results show that neither $\alpha$ nor $\beta$ differ between the control and hypertensive animals, and that the subject-specific values agree with values reported by Olufsen [31] ($\alpha = 0.9$, $\beta = 0.6$).

Figure 6.6b depicts $\ell_{rr}$ as a function of the vessel radius. Using a nonlinear least-squares fit with bisquare robustness, we fitted the decaying exponential curve including all points less than three standard deviations from the mean

$$f(r) = C_1 e^{-C_2 r} \quad (6.26)$$

Results show that the length to radius ratio $\ell_{rr}$ decreases with an increasing radius in both control and hypertensive animals. Values for $C_1$ and $C_2$ are given in Table 6.2. The predicted values are lower in hypertensive animals compared to controls. This agrees with physiological observations that the radius expands in hypertensive animals, while the length is not affected, resulting in a lower $\ell_{rr}$.

**Figure 6.6** Structured tree parameters plotted as a function of the vessel radius. In all graphs, the predictions are grouped into 20 bins, and results include estimates from the 3 control and 3 hypertensive animals. (a) depicts $\alpha$ and $\beta$, the horizontal lines denote the mean value for each group. (b) depicts $\ell_{rr}$ along with a decreasing exponential curve.
Table 6.3 Stiffness parameters used in model simulations. PP-Principal pathway; ST Structured tree.

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Hypertensive</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_{1}^{PP}$ (mmHg)</td>
<td>75</td>
<td>300</td>
</tr>
<tr>
<td>$k_{2}^{PP}$ (cm$^{-1}$)</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>$k_{3}^{PP}$ (mmHg)</td>
<td>270</td>
<td>525</td>
</tr>
<tr>
<td>$k_{1}^{ST}$ (mmHg)</td>
<td>375</td>
<td>$2.25 \times 10^{3}$</td>
</tr>
<tr>
<td>$k_{2}^{ST}$ (cm$^{-1}$)</td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td>$k_{3}^{ST}$ (mmHg)</td>
<td>150</td>
<td>750</td>
</tr>
</tbody>
</table>

6.4.2 Fluid dynamic predictions

In each network, we specify flow at the inlet to the MPA and compute pressure and flow along each vessel in the principal pathway by solving the fluid dynamics equations (6.7) and (6.8). In the structured tree, we solve the linearized fluid dynamics model (6.14) and (6.15) with varying viscosity using Hct values from [58] of 0.60 and 0.81 in control and hypertensive groups, respectively.

To demonstrate the effect of variations in geometry, we use the same flow profile for each of the two groups. The flow profile used for the hypertensive mice had a lower cardiac output ($0.14$ ml/s) than the control mouse flow profile ($0.17$ ml/s), which was consistent in all the profiles available. Simulations were conducted under the constraint that control animals have significantly more compliant vessels than the hypertensive animals, and that stiffness increases in smaller vessels; i.e., the stiffness in the structured trees is larger than in the large vessels. Stiffness values for the principal pathway and structured tree are given in Table 6.3. Stiffness values used in this study were informed by previous studies [93, 96, 65] adapted to the model geometry used in this study.

The Reynolds number is calculated as

$$\text{Re} = \frac{\rho R \bar{q}}{\mu \bar{A}},$$

where $\bar{q}$ and $\bar{A}$ are computed by averaging $q(t)$ and $A(t)$ over the cardiac cycle. $\text{Re} \approx 50$ in the MPA of the control mice, decreasing to $\approx 1$ in the smallest vessels within the principal pathway. In the hypertensive animals, $\text{Re} \approx 40$ in the MPA, also decreasing to $\approx 1$ in the smallest vessels in the principal pathway. In the structured tree (comprising the small arteries and arterioles), the Reynolds number decreases from $\approx 1$ to $\approx 0.5$ in control animals, and to $\approx 0.02$ in the hypertensive animals. It is expected that the hypertensive animals have a smaller cardiac Reynolds number as their cardiac output is smaller compared to the control animals.

Figure 6.7 shows pressure predictions in the MPA (marked with red, cyan and blue lines) for each network along with measurements from 5 control and 7 hypertensive animals (marked with gray lines). Results show that model predictions agree well with data and that by adapting vessel stiffness using separate ranges for control and hypertensive animals, we can predict pressure in both control and hypertension. Figure 6.8 shows changes in the pressure and flow profiles along the principal pathway as well as predictions within the $\alpha$ and $\beta$ branches of the structured tree for one
control and hypertensive mouse. Predictions in the $\alpha$ and $\beta$ branches correspond to vessels with radius $\alpha r_0, \alpha^2 r_0, \ldots, \alpha^n r_0$ on the $\alpha$ side and $\beta r_0, \beta^2 r_0, \ldots, \beta^m r_0$ on the $\beta$ side, where $n$ and $m$ are the number of branches before reaching a radius less than $r_{\text{min}}$. The mean pressure and mean flow drop along the $\alpha$ and $\beta$ branches within the structured tree are also shown in Figure 6.8 as a function of radius. Closer scrutiny of the pulse pressure predictions in the MPA (Figure 6.7) show that the model predictions provide a better agreement with data from hypertensive animals than for the controls. In particular, control predictions overestimate diastolic pressure values, while the hypertensive model underestimate pressure predictions at diastole. The mean pressure drop in both disease states decreases nonlinearly as a function of the structured tree radius and drops quicker in the $\alpha$ branch than the $\beta$ branch.

A benefit of the 1D model is that it is easy to predict pressure in both large and small vessels. Figure 6.9 depicts the mean pressure along the principal pathway and structured trees in the 3D network. In the control animals, the mean pressure drops from approximately 16 mmHg in the MPA to 1 mmHg at the level of the arterioles, while in the hypertensive animals, the mean pressure changes from 22 mmHg to 2 mmHg. To illustrate the impact of measurement uncertainty, we computed the mean discrepancy in mean pressure along a representative principal pathway varying $\alpha$, $\beta$ and $\ell_{rr}$ with $\pm$ one standard deviation, using values reported in Table 6.2. Changes in $\ell_{rr}$ did not affect predictions significantly, while changes in $\alpha$ and $\beta$ lead to a mean MPA pressure change of approximately 2.3 mmHg in the control mice and 8.5 mmHg in the hypertensive mice. These results are shown in Figure 6.9.

**Figure 6.7** Pressure predictions in the MPA (marked with red, cyan and blue lines) for each network along with measurements from 5 control and 7 hypertensive animals (marked with gray lines). The shaded gray region denotes the range between the lowest and highest measurements. Results show that model predictions agree well with data and that by adapting vessel stiffness using separate ranges for control and hypertensive animals, we can predict pressure in both control and hypertension.
Figure 6.8 Pressure and flow predictions along the principal pathway and α/β structured tree branches in a control and a hypertensive mouse. Mean pressure and flow predictions in the structured tree model are also provided. Note that the pressure drop along the principal pathway is more sudden compared to the pressure drop in the structured tree and that the hypertensive animal has a lower cardiac output. Principal pathway predictions run from the MPA (lightest grey) to the terminal vessel (darkest grey) proximal to the structured tree. Structured tree predictions run through the α side (in red) and the β side (in blue), with higher generations denoted by lower magnitude pressure and flow values, as well as a darker color.
Figure 6.9 Mean pressure predictions in the three control and hypertensive mice. For mouse three we show predictions both along the principal pathway and the structured trees, whereas for mice 1 and 2 we only show predictions along the principal pathway. Below each 3D graph we show the mean pressure drop down a representative pathway, with a variation of ±2.3 mmHg in the control mice and ±8.5 mmHg in the hypertensive, accounting for measurement uncertainty. To illustrate results the structured trees were rendered in 3D, though computations are all done in 1D as described in the method section.
6.5 Discussion

6.5.1 Network reconstruction

We have presented a network extraction process that captures the length, radius, and connectivity of vascular networks using an accurate, semi-automated method. Our investigation into exception handling for skeletonized arterial trees is, to our knowledge, the first of its kind. Several previous imaging studies [92, 38, 62] have reported that segmentation leads to false loops and extra branches, but have not been successfully addressed how to remove these from the generated graphs. A previous study by Miyawaki et al. [77] addressed how to handle trifurcations in bronchial networks, but did not address how to remove loops and false branches, an issue that likely would arise if more generations were included in the network.

One of the most significant benefits of our method is its ability to determine the inlets and outlets of the network automatically. While several pulmonary modeling studies [96, 92, 42] have had success in using open source centerline algorithms, such as the Vascular Modeling ToolKit [43], the studies required manual identification of the terminal vessels of the tree. This quickly becomes infeasible for large networks, such as the pulmonary tree, as there exist hundreds to thousands of terminating arteries. Our methodology identifies all terminal vessels in the tree automatically, making network reconstruction more efficient. This technique can be applied to human-based pulmonary studies, and could also be applied to other biological networks that have a rapid branching pattern, such as the liver [95] or the brain [52].

Another use of the results reported here is to provide a validation of geometry that can be used in studies generating artificial networks, e.g., as was done by Clark et al. [86], who used a CT image to generate large vessels combined with a space-filling algorithm to generate small vessels.

6.5.2 Structured trees

The results presented here provide a first attempt to validate the structured tree model, first introduced by Olufsen [29]. While several previous studies have implemented the structured tree in the pulmonary circulation [55, 46], the algorithm has never been validated for subject-specific networks. Moreover, previous studies utilizing the structured tree model [55, 46, 65] were formulated using three fractal parameters: the radius relation (also known as Murray’s exponent) $\xi$, the asymmetry ratio $\gamma$, and the area ratio $\eta$. These values are used in concert to construct estimates of $\alpha$ and $\beta$ using literature values aggregated from multiple patients. Our methods extract $\alpha$ and $\beta$ directly from the data, minimizing the possible variability that is introduced by coupling network data and literature-based fractal properties. Some key findings in this study are that the parent-daughter radius scaling factors ($\alpha$ and $\beta$) remain relatively constant throughout the lung, and that the subject-specific values agree with values reported by Olufsen [31] ($\alpha = 0.9, \beta = 0.6$).

In this study, we assume that $\alpha$ and $\beta$ are constant, but it can be argued that the lung has two zones separated by $r \approx 150\mu$m. This result agrees with suggestions by Clipp et al. [46] analyzing
pulmonary arterial casts in a lamb. This study predicted $\alpha$ and $\beta$, but did not estimate the $l_{rr}$. Instead, they used optimization to best match measured pressure waveforms with model predictions from a 1D fluids model.

A clear benefit of our method is that we do not require assumptions of fractal parameters based on literature values, making it easy to extract subject-specific, region-specific, and disease-specific values. Results reported here did not identify differences between lobes or with disease. This may be a result of genetic similarity among laboratory animals, or it could be proof that the morphometry of the vasculature favors specific optimality principles.

Another accomplishment of our algorithm is detection of the principal pathway. In the systemic arteries, the major arteries (e.g., the major aortic branches) are named, making it easy to determine what vessels to include in a modeling study. The same is not valid for the pulmonary network. The MPA artery branches into two vessels that transport blood to the left and right lung. At this stage, it is essential to perfuse every lobe of the lung, but as shown in Figure 6.7, significant variation is observed between individuals, making it challenging to identify specific principal pathway vessels. In this study, we used a scaling law, including all connected vessels with radius at least 40% of the root vessel radius in the principal pathway. This method is advantageous over previous methods, e.g., by Molthen et al. [35], who only identified one pathway in each lung. Moreover, this method, coupled with the calculation of the radius using the interquartile mean, is robust to image segmentation uncertainty, where vessel radii might be larger depending on image segmentation parameters [92].

6.5.3 Fluid dynamics

The results presented here highlight the advantages of a multiscale modeling approach. Using the image generated principal pathway network, we get a subject-specific geometry in which we can solve nonlinear fluid dynamics equations. By combining this geometry with a subject-specific structured tree model, it is possible to retain physiological features of the large arteries while still modeling microcirculation structure and function. This approach generates a sophisticated model at a relatively low computational cost, advancing previous studies (e.g., [86]) that used large networks but did not account for inertial effects in the large vessels, or studies by Colebank et al. [93, 92] that used lumped parameter boundary conditions and therefore was unable to predict dynamics in the microcirculation. The latter is of importance when using computational methods to analyze disease or to study model-based approaches to treatment when disease progression affects the microcirculation. Moreover, our predictions of microvascular hemodynamics account for nonlinear rheologic effects via an empirical hematocrit-radius dependent viscosity. This allows for a more accurate description of microvascular resistance, which has been known to depend on blood hematocrit and increase with a reduction in blood vessel area [58, 21, 24].

In contrast to previous pulmonary studies utilizing the structured tree model [55, 46], our results are the first to couple a large pulmonary tree segmented from imaging data with a structured tree model of the microcirculation that is also conditioned on data. Our unique approach of including the largest vessels in the principal pathway ensures that the model transports blood to all major
lobes of the lung without having to solve computationally expensive, nonlinear equations in the entire pulmonary tree. The identification of the principal pathway is critical, in particular in studies aiming at predicting pulmonary perfusion. This fact has only been addressed in a few studies [35, 33], but has been known for years. For example, an early study by West et al. [9] showed that the hydrostatic pressure differences in the lung are the driving factor in heterogeneous lung perfusion. While we do not account for lobe-specific pressure differences, the model presented is able to predict pressure and flow within different lobes of the lungs, and can be validated with experimental results in future studies. Our predictions in the MPA show that the hypertensive predictions are, on average, closer to the data than the control predictions. By increasing the large and small artery stiffness, the computational model can predict accurate systolic pressure values in hypertension induced by hypoxia. Several previous studies [36, 51, 58] have documented that increased small artery stiffness is a biomarker of pulmonary hypertension, and it is suspected to play a significant role in the increase in right ventricular afterload. Hypertensive results in Figure 6.7 are created in part by increasing small artery stiffness by several orders of magnitude relative to the control stiffness (Table 6.3). This increase in small artery stiffness agrees with current physiological knowledge and previous modeling studies of the pulmonary circulation [55, 65]. The flexibility of adjusting both large and small artery stiffness in the model allows systolic, diastolic, and pulse pressure values to be altered based on downstream tissue characteristics, making the parameters easier to interpret than standard Windkessel boundary conditions [96, 92]. Pressure predictions in both control and hypertensive networks, shown in Figure 6.9, decrease rapidly from the MPA down to the terminal arteries, and show a slower decay in the structured tree predictions. While this is uncharacteristic of the systemic circulation (e.g., [31]), the pulmonary circulation shows a uniform pressure drop across the rapidly branching arterial tree [76]. The principal pathway flow predictions in Figure 6.8 show a similar behavior, dropping in magnitude rapidly from the MPA to the terminal arteries. In contrast to the pressure, the flow distribution in the structured tree tends to be consistent in both the $\alpha$ and $\beta$ branches, whereas the mean pressure drop is more predominant in the $\alpha$ branch than the $\beta$. These results are similar to the findings by Olufsen et al. [55], showing a more pronounced drop in the $\alpha$ side of the tree as this side contains more vessels. The Reynolds number values computed are in line with previous mouse studies [45], reporting a measured Reynolds number of 61 in the abdominal aorta of mice. The hypertensive mice had a smaller value of Re than the control, which is due to the fact that induced hypertension leads to decreased cardiac output ($\overline{q}$) and larger area ($\overline{A}$). These values suggest that the viscous forces are dominant at the terminal end of the structured tree and could be modeled using Stokes equation. Pressure and flow predictions can be computed along any pathway within the structured tree. For simplicity, we only show predictions from the $\alpha$ and $\beta$ branches (see Figure 5b), including vessels scaling the root vessel by $\alpha, \alpha^2, ..., \alpha^n$ and $\beta, \beta^2, ..., \beta^m$, where $n$ and $m$ are the number of branches satisfying $r_{\text{root}}\alpha^n < r_{\text{min}}$ and $r_{\text{root}}\beta^m < r_{\text{min}}$. Since these pathways represent vessels with the largest (the $\alpha$-branch) and smallest (the $\beta$-branch) radii, pressures along any other pathway will fall between the ones predicted by these pathways. This was demonstrated in our previous study [96].
6.5.4 Limitations

There are several limitations of the study and the results presented here. The computational model uses constant subject-specific values for $\alpha$ and $\beta$, but the data shown in Figure 6.6 could be fitted by a sigmoidal or bi-modal function. Finding appropriate functional forms for these two fractal parameters will be pursued in future studies to ensure that the network becomes more symmetric for smaller radius values (i.e., closer to the capillaries). Also, the structured tree model only include bifurcations, but some trifurcations are present in the data (type SC2-SC4 in Figure 6.4d). However, only $1.67\% \pm 0.61$ of the junctions in the control and $1.86\% \pm 0.59$ in the hypertensive mice were non-bifurcations (mostly trifurcations, with one quadfurcation in one of the hypertensive networks). Therefore, we skipped non-bifurcating junctions when calculating $\alpha$ and $\beta$ in the networks.

We observed more vessels in the hypertensive networks. This is likely a result of the imaging/segmentation process. In these animals, as expected, vessels expand, and therefore more vessels are visible in the images (above the threshold for imaging), one way to prevent this is to limit the segmentation to a specific number of vessels or generations. The use of contrast agent for perfusing the lung tissue assists in vessel detection, but imaging done on human PH patients or in animals in-vivo is typically done with little to no contrast. Non-contrast imaging lacks clarity and will be susceptible to higher uncertainty due to natural fluctuations in pulmonary pressure, for example, during respiration [46]. Moreover, the same constant perfusion pressures were used in both control and hypertensive animals, even though hypertensive animals likely have stiffer arteries. The constant inflation pressure used for image extraction is not characteristic of the dynamic, oscillating pressure seen in-vivo, which should be considered when conducting human studies. Moreover, we did not quantify the branching angles necessary to project fluid predictions in 3D, nor did we employ formal parameter estimation or sensitivity analysis techniques, mostly since imaging and hemodynamic data are not from the same animals. However, the techniques presented here can easily be extended to account for these factors, as discussed in some of our previous studies [92, 93, 96]. Our results show that the length-to-radius ratio $L/r$ has little effect, while the parameters $\alpha$ and $\beta$ have a more significant effect. Changing these parameters leads to an average discrepancy of 2.3 and 8.5 mmHg in the control and hypertensive mean MPA, respectively, when parameters are varied within the bounds reported in Table 6.2. In this study we kept $\alpha$ and $\beta$ constant, but as indicated on Figure 6.6, better results may be obtained if $\alpha$ and $\beta$ are varied with radius. It should be noted that the variation given in Table 6.2 comes from comparing values over the three animals, so rather than reporting variation in data, the results shown in Figure 6.9 illustrate the importance of using boundary conditions informed by data. More work is needed to further study this variation, preferably using imaging and hemodynamic data from the same animal or person.

Finally, to match experimental conditions, this study used a zero-impedance at the terminal branches of the structured trees. In-vivo, the arterioles continue branching into the capillaries, which do not have zero pressure. Therefore, if the model presented here is translated to analyze in-vivo data, this condition should be modified. As discussed by Clipp et al.[46], the capillary pressure in the lung is not only non-zero, but varies over the respiratory cycle and, if the model represents a
subject in the up-right position, also varies with gravity. This implies that the terminal impedance is likely heterogeneous oscillating with the respiratory cycle and varying throughout the lung.

### 6.6 Conclusion

In this study, we developed a semi-automated method for extracting directed graphs from micro-CT images and used this data to extract a principal pathway and fractal scaling parameters from control and hypertensive mice. These parameters were used to inform a 1D fluids model predicting pressure in the large and small vessels. Results show that the fractal scaling parameters $\alpha$ and $\beta$ do not vary significantly between animals or with disease, but that the length-to-radius ratio $\ell_{rr}$ was lower in the hypertensive animals. Pressure predictions in the principal pathways were within the range of measured values, and results tuning the vessel stiffness reveal, as expected, that the hypertensive animals have stiffer vessels than the control animals. Also, in both groups the small vessels are stiffer than the large vessels. Our study into network extraction and correction gives us a more thorough understanding of the arterial network structure, and our graph extraction method opens the door for future analysis of in-vivo human or animal arterial networks from the lung or any other organ system.

### 6.7 Acknowledgements

The study was supported in part by the National Science Foundation via awards NSF-DMS 1246991 and 1615820, and the American Heart Association Predoctoral fellowship 19PRE34380459. In addition, we would like to thank Naomi Chesler, University of Wisconsin-Madison for making micro-CT images and hemodynamic waveforms available for this study.
This chapter includes a study conducted in conjunction with a summer REU program for which I served as a graduate student mentor. This study was supervised by Dr. Mette Olufsen and Dr. Radmila Sazdanovic. Students in this program included Mariam Kharbat (North Carolina State University), Natalie Johnston (Vassar College), Ian Livengood (North Carolina Agricultural and Technical State University), Miya Spinella (University of Massachusetts Dartmouth), and Robert Sternquist (Enloe High School). In this study, I was responsible for supervising the implementation of the height filtration, troubleshooting the pruning algorithms, and writing up the data collection methods, persistent homology background, and results. This chapter is written as a journal manuscript which we intend to submit for publication, and therefore repeats some of the image analysis and persistent homology background discussed in Chapters 4 and 5.

7.1 Abstract

Pulmonary hypertension (PH), defined by a mean pulmonary arterial blood pressure above 20 mmHg, is a cardiovascular disease impacting the pulmonary vasculature. PH is accompanied by
vascular remodeling, wherein vessels become stiffer, large vessels dilate, and smaller vessels constrict. Some types of PH, including hypoxia-induced PH (HPH), lead to microvascular rarefaction. The goal of this study is to analyze the change in pulmonary arterial network morphometry in the presence of HPH. To do so, we use novel methods from topological data analysis (TDA), employing persistent homology to quantify arterial network morphometry for control and hypertensive mice. These methods are used to characterize arterial trees extracted from micro-computed tomography (micro-CT) images. To compare results between control and hypertensive animals, we normalize generated networks using three pruning algorithms. This proof-of-concept study shows that the pruning methods effects the spatial tree statistics and complexities of the trees. Results show that HPH trees have higher depth and that the directional complexities correlate with branch number, except for trees pruned by vessel radius, where the left and anterior complexity are lower compared to control trees. While more data is required to make a conclusion about the overall effect of HPH on network topology, this study provides a framework for analyzing the topology of biological networks and is a step towards the extraction of relevant information for diagnosing and detecting HPH

Keywords: pulmonary hypertension, vascular remodeling, image segmentation, tree pruning, Strahler order, persistent homology

7.2 Introduction

Cardiovascular diseases (CVD) are the leading cause of death in the world. Data from the World Health Organization shows that in 2019, CVD claimed an estimated 17.9 million lives, accounting for 32% of deaths worldwide. While CVD encompasses many disease types, a common trait is that these diseases are associated with remodeling of the vasculature and heart. This causes structural and functional problems, impacting the blood vessels and the heart. Advanced medical imaging has emerged as a valuable tool for studying CVD, but fails to provide a complete understanding of how disease progression impacts the morphometry of the vascular network.

One CVD is pulmonary hypertension (PH), defined by a high blood pressure ($\geq 20$ mmHg) in the main pulmonary artery [67]. 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. PH encompasses five subtypes, and while early diagnosis and targeted treatment can improve quality of life by delaying severe complications, all types of PH but one have no cure [67]. This study aims at quantifying morphological changes using persistent homology, a topological data analysis (TDA) method, to provide numerical descriptors that could aid in the diagnostic process.

PH is associated with vascular remodeling, changes in the structure of pulmonary arteries [35, 51]. In particular, hypoxia-related PH (HPH) starts in the arterioles, stiffening and constricting the vessels, then migrating to the larger arteries, which stiffen and dilate. These structural changes vessels have been studied and characterized before [11, 34], but less is known about how the network morphometry changes. In this study, we apply topological data analysis to graphs of pulmonary
arterial networks and see how their properties change in the presence of HPH.

**Pulmonary arterial morphometry** characterizes the network properties in space. Since early contributions by Murray in 1926 [1], many researchers (e.g., [15, 10, 13, 29, 31, 51, 35, 73, 102]) have examined the pulmonary arterial network. This network has received significant attention, as it exhibits a branching structure to fill a well-defined volume. The early studies by Murray [1] and Zamir [15] devise optimality principles, describing the pulmonary arterial network as branching to minimize pumping power and lumen volume. Like many other studies, including ours, these assume that the arterial network bifurcates and that the dimensions of the two daughter vessels can be determined as functions of the parent vessel. Essential results from these studies include a power law defining how vessel radii change across a bifurcation, asymmetry ratio relating the radii of the two daughter vessels, and an area ratio relating the combined cross-sectional area of the daughter vessels to that of the parent vessel. Other studies by Singhal et al. [10], Horsfield [13], and Olufsen et al. [29, 31] incorporate data from lung casts to devise relations between parent and daughter vessels. The casts are generated by injecting liquid resin into the arterial networks. 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 measuring the vessel dimensions. More recent studies by Molthen et al. [35], Vanderpool et al. [51], Davidoiu et al. [73] and Chambers et al. [102] use medical imaging to accurately and efficiently extract geometric information from pulmonary arterial networks.

**Persistent homology** refers to a TDA technique wherein combinatorial structures are successively built from a data set, and their homology is used to define descriptors of the shape of the data. These novel techniques have been used to analyze the arterial networks in the brain [70] and bronchial networks [79], but to our knowledge, our study is the first to use persistent homology to characterize pulmonary arterial networks. Bendich et al. [70] use persistent homology to analyze human brain arterial trees generated from a tube-tracking segmentation algorithm on magnetic resonance images (MRI). Using these trees, they compute degree-0 persistent homology and find to be strongly correlated with age, while degree-1 persistence is strongly correlated with sex. In [79], persistent homology is used to compare lung CT images from healthy and chronic obstructive pulmonary disease (COPD) patients. Degree-0 persistent homology is able to distinguish the patient groups, while degree-2 persistent homology is able to distinguish inspiratory from expiratory images.

Motivated by these studies, we compute persistent homology using the labeled spatial trees we extract from pulmonary arterial scans of mice. We compute the degree-0 persistent homology of these trees using a height filtration in all 3 dimensions in forward and reverse directions (resulting in 6 applications of filtration per spatial tree). We compute the total persistence in each direction, known as the **directional complexity**, and compare these values between control and HPH networks.
7.3 Methods

This study follows the protocol we described in [102] to extract labeled spatial trees representative of pulmonary arterial networks of control and HPH mice [51]. Our overall objective is to identify markers that can characterize differences between control and HPH arterial networks extracted from micro-CT images. This is done in three steps. First, we generate a 3D rendered network from the image, constructing a labeled spatial tree with branches, representing vessels, connected at vertices, representing junctions. Each edge is labeled by its length, spatial orientation, and radius, and each vertex by its coordinate in $\mathbb{R}^3$. Second, we compare the spatial trees obtained from scans of control animals to those with HPH. To obtain biologically significant information from these spatial trees, it is essential to compare only corresponding parts of control and HPH trees. In HPH, the diameter of the large vessels increases [102], making more vessels visible in the 3D rendered network. To compare HPH networks with controls, we use three pruning algorithms to normalize the trees, ensuring that we compare corresponding subtrees of vessels. Third, we use persistent homology to characterize the normalized networks. Inspired by Belchi et al. [79], we employ a height filtration in 3D (6 directions), compute degree-0 persistent homology, and compare the directional complexities of trees from control and HPH mice.

7.3.1 Imaging protocol

The images used in this study were made available to us by the Chesler Lab, University of California-Irvine [51]. Our study uses data from six C57BL6/J mice, each aged 10-12 weeks and weighing 25.8 ± 2.3 grams. Of the six mice, 3 were controls and 3 were induced with HPH by placing them in an hypoxic environment (FiO2 reduced by half) for 10 days. The mice were mediated with (52 mg/kg body weight) pentobarbital sodium and euthanized by exsanguination before the lungs were extracted. After extraction, the lungs were imaged following the protocol described in [51]. First, a cannula (PE-90 tubing, 1.27 mm outer and 0.86 mm inner diameter) was positioned in the main pulmonary artery (MPA) well above the first arterial bifurcation. After, 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). PFOB has been shown to be an excellent alternative to other artificial blood substitutes and inhibits vasoactivity. The lungs were further prepared by periodically adjusting the intravascular pressure from 0 to 25 mmHg multiple times. After preparation, the arterial pressure would remain constant as the lungs are rotated around the X-ray beam at 1° increments to obtain 360 planar images. The lungs were imaged at arterial pressures of 6.3, 7.4, 13.0, and 17.4 mmHg (Figure 7.2). The planar images for each pressure were then reconstructed using the Feldkamp cone-beam algorithm, converted into a three-dimensional volumetric dataset, and formatted as a Digital Imaging and Communications in Medicine (DICOM) 3.0 image (Figure 7.1(a)).
7.3.2 Segmentation and skeletonization

Once the micro-CT scan images are obtained, labeled spatial trees are extracted following the protocol described in [102] and illustrated in Figure 7.1. Each image is a voxel complex, a finite set of voxels. Voxels are cubes of image data analogous to the “pixels” in a 2D image. Each image has the dimensions $(497 \times 497 \times 497)$ voxels, where the spatial resolutions for each image varied between $30-40 \mu m$. Each voxel $v$ in a voxel complex has spatial coordinates $(x_v, y_v, z_v)$ and intensity value $I(v) \in [0, 255]$ with 0 and 255 equalling the intensity of a black voxel and a white voxel, respectively.
<table>
<thead>
<tr>
<th>Pressure (mmHg)</th>
<th>6.3</th>
<th>7.4</th>
<th>13.0</th>
<th>17.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertensive</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 7.2 The labeled spatial trees from a control mouse (top row) and a HPH mouse (bottom row) extracted from images taken at the 4 different arterial pressures. As pressure increases, the number of branches increases, since a higher arterial pressure perfuses the imaging contrast further down the arterial tree.

**Segmentation:** This is the process whereby the voxels in a voxel complex are partitioned into the set of *foreground* voxels, which represent the arteries, and *background* voxels. Details about the segmentation process are provided in [102]. We segment the images using the open-source program 3D Slicer [53, 63, 104], employing a combination of segmentation techniques, including global thresholding, median smoothing, and manual editing. Global thresholding includes all voxels with intensities \( I(v) \in [\tau_{\text{min}}, \tau_{\text{max}}] \) in the foreground. The pulmonary arteries are the only anatomical structures captured in the image, so \( \tau_{\text{max}} = 255 \) for all images and \( \tau_{\text{min}} \) is adjusted using an *ad hoc* method 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 hypobaric cylinder can distort the appearance of the image, minimal manual editing is required to remove voxels from the foreground which do not represent the true arterial structure. The result of segmentation is a foreground which can be visualized as a 3D rendering of the pulmonary arteries. The constructed foreground is now referred to as the *segmented network*, and each *segmented artery* is a portion of the segmented network that lies between two junctions (Figure 7.1(b)).

**Distance map:** To obtain the dimensions for the segmented arteries, we generate a *distance map*, an associated voxel complex with the same spatial dimensions as the original image that encodes a distance measurement for every voxel. For each voxel \( v \) in the image, we define the distance function

\[
d(v) = \min_{u \in \text{background}} \| u - v \|_2, \tag{7.1}
\]
where \( \| u - v \|_2 \) is the Euclidean distance from \( v \) to each background voxel, \( u \). In other words, \( d(v) \) is the distance from the voxel \( v \) to the nearest background voxel. For each \( v \) in the image, there exists a corresponding voxel \( v_D \) in the distance map with the same spatial coordinates \((x_v, y_v, z_v)\) and intensity \( I(v_D) = d(v) \). This distance map, as well as the skeleton and labeled spatial tree, are obtained using the Spatial Graph Extractor Github repository (SGEXT) in the Digital Geometry Tools and Algorithms Library (DGtal) [80, 83, 84].

**Skeletonization:** The skeleton is a voxel complex that is a thinned representation of the segmented arteries. In the skeleton, the branching structure of the segmented arteries is preserved where each branch is one voxel in width and centered in the corresponding artery. The skeleton is obtained by iteratively removing voxels from the segmented arteries via Couprie and Bertrand’s “Asymmetric Thinning” algorithm, effectively “shaving” down the segmented arteries into a thinned structure [102, 72]. Centering is accomplished by, when faced with a choice of which voxel to keep in the skeleton, always choosing the voxel \( v \) for which \( d(v) \) is maximum, which is the most centered voxel. This results in a centered skeleton in the segmented arteries (Figure 7.1(b)).

### 7.3.3 Labeled spatial tree

Each skeleton is used to generate a *spatial graph*, a collection of *vertices* in \( \mathbb{R}^3 \) connected by a collection of *edges*. Every voxel in the skeleton corresponds to a vertex in the spatial graph and is connected by edges to all adjacent vertices. Each vertex is denoted with a numerical ID \( A \) and has 3D coordinates \((x_A, y_A, z_A)\). Each edge connects 2 vertices \( A \) and \( B \), and is denoted \( e_{AB} \). If \( \deg(A) = 1 \), \( A \) is either the vertex at the inlet of the MPA (called \( A_{\text{root}} \)) or a *leaf*, which is 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 *branching point* and represents the point where one segmented *parent* artery branches into multiple *daughter* arteries. Typically, branching points have degree 3, as 98-99% of junctions are bifurcations [102]. All edges are oriented in the direction of blood flow (away from \( A_{\text{root}} \)). With this orientation in place, we can refer to \( e_{AB} \) as an edge from start vertex \( A \) to end vertex \( B \).

Constructing the spatial graph is subject to errors, but can be corrected according to the protocol introduced in [102]. Errors include edges that do not accurately represent the segmented arteries, small cycles that arise when 3 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. However, as described in [102], small cycles and duplicate edges/points can be removed automatically, with cycles broken by removing the longest edge in the cycle. After implementing the corrections, the graph becomes a *spatial tree* \( T \) (Figure 7.1(c)).

Each vertex in \( T \) is labeled with the radius of the corresponding segmented artery at that point, which is determined by the distance map introduced in Section 7.3.2. Recall that the skeleton is centered in the segmented arteries. Therefore, for voxel \( v \) in the skeleton, the measurement \( d(v) \) calculated using equation (7.1) gives an estimate for the radius of the corresponding artery centered at the point \((x_v, y_v, z_v)\) in voxels. Hence, every vertex \( A \) in \( T \) can be labeled with radius \( r_A = d(v) \) for the corresponding voxel \( v \) in the skeleton such that \((x_v, y_v, z_v) = (x_A, y_A, z_A)\).
Every segmented artery is represented by a collection of edges connecting consecutive degree 2 vertices between either 2 branching points or a branching point and a leaf. We refer to each such collection of edges as a branch of $T$ such that each branch corresponds to a singular artery in the segmented arteries. Branches are labeled with an overall radius and length measurement in microns ($\mu m$) from the corresponding artery. To convert the voxel measurements into $\mu m$, all dimensions, $x$, $y$, and $z$, are multiplied by the scaling factor $\lambda_T$, obtained by dividing the cannula’s radius, 430 $\mu m$, by the average of the first 5 radius measurements of the branch representing the MPA, which are the radii of vertices $A_{root}, A_1, \ldots, A_4$.

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

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

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

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

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

### 7.3.4 Spatial tree statistics

To quantify the spatial trees, we computed the number of branches, number of leaves, tree depth, and Strahler order. We define tree depth to be the number of branches in the longest direct path from $A_{root}$ to any leaf. The Strahler order (SO) of a branch is an indicator of its level within the tree [22]. Terminal branches all have $SO = 1$. If two joining daughter vessels have equal $SO$ (i.e. $SO_{d_1} = SO_{d_2}$), then the $SO$ parent vessel is equal to $SO_p = SO_{d_1} + 1$. Otherwise, the $SO$ of the parent vessel is $SO_p = \max \{SO_{d_1}, SO_{d_2}\}$ (Figure 7.3). The Strahler order of a tree refers to the maximum Strahler order of any branch within that tree, which will be determined by the first branch. Since the first branch of the trees in this study represents the MPA, we denote this $SO_{MPA}$.

Results, reported in Table 7.1, show that HPH trees have significantly more branches and leaves than the control trees. While the pulmonary vascular remodeling observed in HPH has been characterized by increased vessel stiffness and dilation of large arteries [102], one effect that is not typically observed is angiogenesis, the development of new blood vessels. In fact, studies show that HPH animals have fewer arterioles (arteries with radius $\leq 50 \mu m$) than control [51]. Since the contrast used in the micro-CT images travels through dilated vessels more easily, more vessels are captured in the segmented arteries of HPH mice compared to the control. Consequently, the increased number of branches in the HPH trees is an artifact of the imaging process and does not translate to the HPH mice actually having more pulmonary arteries. Additionally, the minimum radius of the vessels
captured during segmentation is approximately 39 \( \mu m \), with the vast majority of vessels having a radius greater than 50 \( \mu m \) (97.3% of vessels). Since arterioles typically have radii of 25 – 50 \( \mu m \), this illustrates that hardly any arterioles are captured in the images, explaining why we do not observe the same microvascular rarefaction described in [51].

7.3.5 Spatial tree pruning

The goal of this research is to capture the properties of the arterial trees that distinguish HPH mice from the control. Due to the imaging effects discussed above, tree pruning techniques are implemented to allow us to compare only the corresponding branches in both groups. If the pulmonary trees branched symmetrically, trees could be pruned in a straightforward way until they have the same number of generations. However, pulmonary arterial networks are asymmetrical due to the left lung being significantly smaller than the right. Consequently, a different pruning approach is necessary to normalize these trees. In this study, we examine 3 pruning methods wherein terminal branches are systematically removed to allow better comparison between control and HPH trees. Two of the pruning techniques rely on computing the Strahler order of the branches, and the third method is based on vessel radius (Figure 7.4).

**Maximum Strahler order pruning:** We observed that for all pressures except 17.2 mmHg, the HPH trees had a higher Strahler order \((SO_{MPA} = 7)\) than in control trees \((SO_{MPA} = 6)\). In these cases, the method of maximum Strahler order pruning is applied to the HPH trees to obtain their largest subtrees with \(SO_{MPA} = 6\). All pairs of terminating sister branches (i.e. \(SO_{d1} = SO_{d2} = 1\)) are removed from a HPH tree, and the Strahler order of the tree is recomputed. This process is repeated until the

![Figure 7.3 The Strahler ordering system illustrated on a small tree. All terminal branches have \(SO = 1\). Whenever 2 branches of the same Strahler order meet, their parent has Strahler order which is 1 higher than that of its daughters. If 2 branches of different Strahler order meet, their parent is labeled with the maximum Strahler order of its daughters.](image-url)
Figure 7.4 Examples of pruned trees for each of our 3 pruning algorithms. Pruned trees are shown in light blue, overlaid on their original trees in dark blue. Any dark blue branches that are visible represent branches that are removed from the original tree during pruning. The pruned trees are obtained via the 3 pruning algorithms defined in Section 7.3.5: (a) minimum Strahler order pruning, (b) maximum Strahler order pruning, & (c) radius pruning.

<table>
<thead>
<tr>
<th></th>
<th>(a) Minimum Strahler order pruning</th>
<th>(b) Maximum Strahler order pruning</th>
<th>(c) Radius pruning</th>
</tr>
</thead>
<tbody>
<tr>
<td># branches: original</td>
<td>2350</td>
<td>2350</td>
<td>2350</td>
</tr>
<tr>
<td># branches: pruned</td>
<td>704</td>
<td>1606</td>
<td>1599</td>
</tr>
<tr>
<td># branches: removed</td>
<td>1646</td>
<td>744</td>
<td>751</td>
</tr>
</tbody>
</table>

HPH tree has $SO_M_{PA} = 6$, the same as the control trees. Note that for the pressure of 17.2 mmHg, all control and HPH trees had $SO_M_{PA} = 7$ already, so maximum Strahler order pruning is not applied at this pressure. We stop the algorithm as soon as the HPH tree has $SO_M_{PA} = 6$.

**Minimum Strahler order pruning**: Trees with any given Strahler order can vary greatly in their properties, including number of branches and tree depth. Therefore, we also perform *minimum Strahler order pruning* of both the control and HPH trees at each pressure to examine the minimal subtrees with $SO_M_{PA} = 6$. All pairs of terminating sister branches (i.e. $SO_{d1} = SO_{d2} = 1$) are removed from a tree, Strahler orders are recomputed, and the process continues until the next iteration of branch removal would cause $SO_M_{PA} = 5$. The result is a set of control and HPH trees which are the smallest subtrees of their originals with $SO_M_{PA} = 6$.

**Radius pruning (RP)**: This pruning technique (RP) involves removing pairs of terminal daughter branches from HPH trees that have a radius less than or equal to a certain radius threshold $\tau_r$. This threshold is specific to each HPH tree, determined by observing which $\tau_r$ value removed enough branches such that the number of branches for the tree is close to the average number of branches for all the control trees, which differed between the four contrast pressures. After radius pruning, all bifurcations with terminal daughters have at least one daughter with $r \geq \tau_r$.

### 7.3.6 Directional complexity

This study uses topological data analysis (TDA) to examine the spatial features of our labeled spatial trees. TDA provides extensive toolbox of data analysis methods for learning about the shape of data.
In this study, we use a tool called *persistent homology* to analyze our control and HPH spatial trees [16, 23, 30]. To be precise, we compute the degree-0 persistent homology of a height filtration, and use this to create a barcode and compute a topological marker known as directional complexity. These techniques are robust to noise [100, 64], making them useful for biological applications.

Relevant topological definitions and background are given in Chapter 5. In this section, we provide a brief overview of some of those key concepts and describe the particular topological methods carried out in this study. In summary, computing persistent homology involves building a sequence of topological spaces upon a data set, computing the corresponding homology, and tracking a topological summary that describes the shape of the data.

**Simplicial homology:** A spatial tree is an example of a *simplicial complex*, the formal sum of building blocks called *simplices*. A *k*-simplex is a convex hull of *k* + 1 vertices [32] (in particular, 0-simplex is a point, a 1-simplex is edge, a 2-simplex is a triangular face, and a 3-simplex is a filled tetrahedron). Each labeled spatial tree in this study amounts to a set of points (0-simplices) attached by a set of edges (1-simplices). Recall that these edges are directed away from the root of the tree to represent the direction of blood flow. Likewise, when computing simplicial homology, the *k*-simplices of a simplicial complex are imbued with orientation.

For a given simplicial complex *S*, the *chain vector space*, *C*<sub>*k*</sub>(*S*) is the vector space whose basis is the set of oriented *k*-simplices in *S* [32]. The elements *c* ∈ *C*<sub>*k*</sub> are called a *k*-chains, *c* = ∑<sub>*i*</sub> *a*<sub>*i*</sub>*σ*<sub>*i*</sub>, where each *a*<sub>*i*</sub> is an integer and each *σ*<sub>*i*</sub> is an oriented *k*-simplex from *S*. The boundary map, ∂<sub>*k*</sub> : *C*<sub>*k*</sub> → *C*<sub>*k*−1</sub>, sends *k*-dimensional chains to their (*k* − 1)-dimensional boundaries.

Let *L*<sub>*k*</sub> be the kernel of ∂<sub>*k*</sub> and *B*<sub>*k*</sub> be the image of ∂<sub>*k*+1</sub>. The *k*<sup>*th*</sup> *simplicial homology group*, *H*<sub>*k*</sub>, is the quotient group

\[ H_k = L_k / B_k = \ker(\partial_k) / \text{im}(\partial_{k+1}). \] (7.5)

The *k*<sup>*th*</sup> *Betti number*, *β*<sub>*k*</sub>, is the dimension of the homology *H*<sub>*k*</sub>. It counts the number of *k*-dimensional “holes” in a simplicial complex [64]. We define a *k*-feature to be a property of a simplicial complex which is counted by *β*<sub>*k*</sub>. The Betti numbers *β*<sub>0</sub>, *β*<sub>1</sub>, and *β*<sub>2</sub> count the number of disjoint connected components, loops, and spheres *S*<sup>2</sup>, respectively.

**Persistent homology:** The idea behind persistent homology is to iteratively build nested simplicial complexes which connect data points based on proximity. Such a set of nested simplicial complexes is called a *filtration* and is governed by a distance function *ε*. For an increasing sequence of *ε* values *ε*<sub>1</sub>,..., *ε*<sub>*m*</sub>, the filtration *S*<sub>*ε*<sub>1</sub></sub> ⊂ *S*<sub>*ε*<sub>2</sub></sub> ⊂ ... ⊂ *S*<sub>*ε*<sub>*m*</sub></sub> is built upon the data. *Degree-*<sub>*k*</sub> *persistent homology* is found by computing *H*<sub>*k*</sub>(*S*<sub>*ε*<sub>1</sub></sub>)..., *H*<sub>*k*</sub>(*S*<sub>*ε*<sub>*m*</sub></sub>) and examining the evolution of their *k*<sup>*th*</sup> Betti numbers. For any *i* = 1,..., *m*, the Betti number *β*<sub>*k*</sub>(*S*<sub>*ε*</sub>) counts the number of *k*-features in *S*<sub>*ε*</sub>. If a feature appears for the first time in a certain *S*<sub>*ε*</sub>, but is no longer present in *S*<sub>*ε*</sub> for some *j* > *i*, then the *birth* of that feature is *ε*<sub>*B*</sub> = *ε*</sub>*i*</sub> and the *death* is *ε*<sub>*D*</sub> = *ε*</sub>*j*</sub>. The *persistence* of a feature is *P* = *ε*<sub>*D*</sub> − *ε*<sub>*B*</sub>. More “persistent” features are ones present for a large range of *ε* values. The last complex created in a filtration will have the topology of a point.
**Figure 7.5** The 6 height filtration directions, 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.

**Barcodes:** The evolution of the homology $H_k(S_{\epsilon})$ can be visualized by creating a degree-$k$ *barcode*. Such a barcode, $B_\epsilon$, is a diagram which contains 1 bar for each new $k$-feature detected in the filtration. Each bar spans the length from $\epsilon_B$ to $\epsilon_D$ for that feature (Figure 7.6). Longer bars represent features with greater persistence in the chosen filtration, while shorter bars are often interpreted to represent noise, though this depends on the application.

**Height filtration:** In this study, we compute the degree-0 persistent homology of spatial trees with respect to the *height filtration*. Given a spatial tree $T$ with vertices $V$ and edges $E$, the height filtration dictates that each $S_{\epsilon_i}$ includes all vertices $v \in V$ and edges $e \in E$ for which the value of our governing distance function $\epsilon(\cdot)$ is at most $\epsilon_i$. The exact definition of $\epsilon(\cdot)$ depends on which direction we want to examine in our filtration. There are 6 filtration directions used in this study: the positive/negative $x$-directions ($+x/-x$), the positive/negative $y$-directions ($+y/-y$), the positive/negative $z$-directions ($+z/-z$) (Figure 7.5). For a given filtration direction $\pm \xi$, $\epsilon$ of a vertex $v \in V$ with spatial coordinates $(x_v, y_v, z_v)$ is defined as

$$
\epsilon(v) = \begin{cases} 
\xi_v, & \text{if direction } = +\xi \\
\xi_T^{\max} - \xi_v, & \text{if direction } = -\xi
\end{cases}
$$

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

$$
\epsilon(e) = \max \{\epsilon(v_1), \epsilon(v_2)\}.
$$
For example, a height filtration in the $+z$ direction indicates that the filtration moves along the $z$-axis in the positive direction, and each $S_{\epsilon_i}$ includes all vertices with $z_v \leq \epsilon_i$ and all edges with $\max\{z_{v_1}, z_{v_2}\} \leq \epsilon_i$. Conversely, a direction of $-z$ indicates that the filtration moves along the $z$-axis in the negative direction, and each $S_{\epsilon_i}$ includes all vertices with $z_T^{max} - z_v \leq \epsilon_i$ and all edges with $\max\{z_T^{max} - z_{v_1}, z_T^{max} - z_{v_2}\} \leq \epsilon_i$. An illustration of the $-z$ height filtration is shown in Figure 7.6.

**Figure 7.6** Directional filtration in the $z$-direction illustrated on a labeled spatial tree. (a) shows the tree with several examples of the subtrees $S_t$ for different values of $t$, which is the distance in the $z$-direction from the top of the tree. These subtrees are labeled with their 0 degree Betti numbers $\beta_0$. At $t = 50$, $S_{50}$ is a portion of the root branch of the tree, which is 1 connected component and thus, $\beta_0 = 1$. $S_{80}$ includes the root branch and a portion of the first bifurcation. This is still 1 connected component and thus, $\beta_0 = 1$. $S_{100}$ includes 3 connected components, labeled in the bottom box. Therefore, $\beta_0 = 3$ for $S_{100}$.

**Directional complexity**: For each tree $T$ and filtration direction $\pm \xi$, the degree-0 barcode is computed by tracking the number of disjoint connected components in the nested set of subgraphs at each $\epsilon_i$ (Figure 7.6). The *directional complexity* (DC) of $T$ in each direction is the total persistence of all 0-features in the filtration, computed by summing the lengths of the bars in the barcode,

$$DC = \sum_{k \text{ bars in barcode}} (d_k - b_k).$$  \hspace{1cm} (7.6)

Directional complexities are named relative to murine anatomy and captures the length and occurrence of branches in the named direction (Figure 7.5). For example, the $-z$ filtration yields *anterior complexity* which refers to branches pointing toward the head of the mouse (Figure 7.6).
7.4 Results

Table 7.1 contains spatial tree statistics of the original and pruned trees in this study. These include number of branches, number of leaves, tree depth, and Strahler order. Table 7.2 shows the trees’ directional complexities generated from the degree-0 barcode from the height filtrations in 6 directions. Arteries were imaged at the 4 different contrast pressures, 6.3, 7.4, 13.0, and 17.4 mmHg, with results for trees from all pressures shown. All values in these tables are reported as mean ± standard deviation unless otherwise stated.

7.4.1 Spatial tree statistics

The number of branches, number of leaves, and tree depth were all lower in the original control trees than in the original HPH trees. All HPH trees had Strahler order $SO_{MPA} = 7$, while control trees had $SO_{MPA} = 6$ at all pressures except 17.2 mmHg, for which they had $SO_{MPA} = 7$. To generate normalized trees, we applied maximum Strahler order pruning to the HPH trees at all pressures except 17.2 mmHg, since the control and HPH trees already had the same Strahler order at that pressure. In these pruned HPH trees, the values in 7.1 were closer to but still greater than the controls’ values at pressures 6.3 and 7.4 mmHg. At 13.0 mmHg, the control trees had greater number of branches and leaves than the maximum Strahler order pruned HPH trees, but still had a lower tree depth.

We applied minimum Strahler order pruning to both control and HPH trees down to the minimal sized subtrees that had $SO_{MPA} = 6$. At pressures 6.3, 7.4, and 13.0 mmHg, the pruned control trees had a greater number of branches and leaves than the pruned HPH trees. At 17.2 mmHg, these values were higher in the pruned HPH trees, but by a very small margin (7 more branches and 2 more leaves on average in pruned HPH than pruned control). The tree depth is higher in the minimum Strahler order pruned HPH trees than control for all pressures except 13.0 mmHg, when the pruned control average tree depth is higher by 0.4.

Radius pruning is applied to HPH trees only, bringing the number of branches and leaves rather close to those of control trees. On average, the radius-pruned HPH trees had more branches than the control trees at 6.3 and 17.2 mmHg, fewer branches at 13.0 mmHg, and the same number of branches at 7.4 mmHg. The control and radius-pruned HPH trees differed by between 0-19 branches on average. The radius-pruned HPH trees had a greater number of leaves than the control trees at all pressures, though they only differed by between 1-10 leaves on average. The depth of radius-pruned HPH trees is higher than that of control trees at all pressures. The Strahler order of all radius-pruned HPH trees is 1 higher than controls at 17.2 mmHg. At 13.0 and 7.4 mmHg, one of the radius pruned HPH trees had Strahler order 6 (matching controls), while 2 maintained Strahler order 7. At 6.3 mmHg, the Strahler order for all radius-pruned HPH trees is the same as controls.
Table 7.1 Spatial tree statistics from micro-CT images at 4 different contrast pressures. 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 $SO\_\text{MPA} = 7$, and the other two had $SO\_\text{MPA} = 6$. All other values are reported as mean ± standard deviation. Data for HPH\^M at pressure 17.2 mmHg is missing because the control and original HPH trees at this pressure already have the same Strahler order.

<table>
<thead>
<tr>
<th>Type</th>
<th># Branches</th>
<th># Leaves</th>
<th>Tree depth</th>
<th>Strahler order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pressure 6.3 mmHg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1018 ± 151</td>
<td>513 ± 77</td>
<td>25.3 ± 1.5</td>
<td>6</td>
</tr>
<tr>
<td>HPH</td>
<td>2196 ± 778</td>
<td>1107 ± 395</td>
<td>32.3 ± 4.0</td>
<td>7</td>
</tr>
<tr>
<td>HPH^R</td>
<td>1022 ± 53</td>
<td>516 ± 29</td>
<td>30.7 ± 3.5</td>
<td>6</td>
</tr>
<tr>
<td>HPH^M</td>
<td>1519 ± 530</td>
<td>766 ± 270</td>
<td>31.3 ± 4.0</td>
<td>6</td>
</tr>
<tr>
<td>C^m</td>
<td>829 ± 262</td>
<td>418 ± 132</td>
<td>24.7 ± 1.5</td>
<td>6</td>
</tr>
<tr>
<td>HPH^m</td>
<td>729 ± 178</td>
<td>368 ± 91</td>
<td>28.7 ± 3.5</td>
<td>6</td>
</tr>
<tr>
<td><strong>Pressure 7.4 mmHg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1122 ± 193</td>
<td>565 ± 98</td>
<td>28.0 ± 5.2</td>
<td>6</td>
</tr>
<tr>
<td>HPH</td>
<td>2362 ± 769</td>
<td>1193 ± 390</td>
<td>31.7 ± 2.9</td>
<td>7</td>
</tr>
<tr>
<td>HPH^R</td>
<td>1122 ± 82</td>
<td>567 ± 42</td>
<td>30.0 ± 1.7</td>
<td>6.3 ± 0.6</td>
</tr>
<tr>
<td>HPH^M</td>
<td>1337 ± 346</td>
<td>676 ± 175</td>
<td>30.0 ± 2.6</td>
<td>6</td>
</tr>
<tr>
<td>C^m</td>
<td>837 ± 289</td>
<td>422 ± 146</td>
<td>27.0 ± 5.5</td>
<td>6</td>
</tr>
<tr>
<td>HPH^m</td>
<td>728 ± 106</td>
<td>367 ± 53</td>
<td>27.7 ± 2.1</td>
<td>6</td>
</tr>
<tr>
<td><strong>Pressure 13.0 mmHg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>1693 ± 372</td>
<td>854 ± 191</td>
<td>30.3 ± 1.5</td>
<td>6</td>
</tr>
<tr>
<td>HPH</td>
<td>2730 ± 972</td>
<td>1379 ± 495</td>
<td>33.0 ± 3.6</td>
<td>7</td>
</tr>
<tr>
<td>HPH^R</td>
<td>1689 ± 206</td>
<td>855 ± 108</td>
<td>32.0 ± 2.6</td>
<td>6.3 ± 0.6</td>
</tr>
<tr>
<td>HPH^M</td>
<td>1394 ± 306</td>
<td>706 ± 156</td>
<td>31.0 ± 3.0</td>
<td>6</td>
</tr>
<tr>
<td>C^m</td>
<td>961 ± 171</td>
<td>485 ± 84</td>
<td>28.7 ± 2.5</td>
<td>6</td>
</tr>
<tr>
<td>HPH^m</td>
<td>725 ± 84</td>
<td>367 ± 43</td>
<td>28.3 ± 2.5</td>
<td>6</td>
</tr>
<tr>
<td><strong>Pressure 17.2 mmHg</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>2573 ± 517</td>
<td>1301 ± 263</td>
<td>32.0 ± 4.4</td>
<td>7</td>
</tr>
<tr>
<td>HPH</td>
<td>3239 ± 1103</td>
<td>1639 ± 564</td>
<td>35.0 ± 4.0</td>
<td>7</td>
</tr>
<tr>
<td>HPH^R</td>
<td>2592 ± 563</td>
<td>1312 ± 289</td>
<td>34.7 ± 3.5</td>
<td>7</td>
</tr>
<tr>
<td>HPH^M</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>C^m</td>
<td>728 ± 170</td>
<td>369 ± 84</td>
<td>27.7 ± 4.6</td>
<td>6</td>
</tr>
<tr>
<td>HPH^m</td>
<td>735 ± 163</td>
<td>371 ± 83</td>
<td>29.7 ± 3.5</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 7.2 Directional complexities of trees from micro-CT images at 4 different contrast pressures in 6 different directions. Abbreviations used are “C” for control trees, “HPH” for unpruned HPH trees, “HPHR” for radius-pruned HPH trees, “HPHM” for maximal Strahler-order pruned HPH trees, and “Cm” and “HPHm” for the minimal Strahler-order pruned control and HPH trees, respectively. All values are reported as mean ± standard deviation. Data for HPHM at pressure 17.2 mmHg is missing because the control and original HPH trees at this pressure already have the same Strahler order.

<table>
<thead>
<tr>
<th>Type</th>
<th>Right</th>
<th>Left</th>
<th>Ventral</th>
<th>Dorsal</th>
<th>Posterior</th>
<th>Anterior</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pressure 6.3 mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>3145 ± 235</td>
<td>2679 ± 530</td>
<td>2546 ± 456</td>
<td>3329 ± 494</td>
<td>4502 ± 551</td>
<td>1648 ± 423</td>
</tr>
<tr>
<td>HPH</td>
<td>5028 ± 1641</td>
<td>5222 ± 1600</td>
<td>4460 ± 1341</td>
<td>5843 ± 1495</td>
<td>6716 ± 1339</td>
<td>3659 ± 1369</td>
</tr>
<tr>
<td>HPHR</td>
<td>2582 ± 338</td>
<td>2717 ± 153</td>
<td>2160 ± 85</td>
<td>3198 ± 25</td>
<td>3788 ± 155</td>
<td>1678 ± 54</td>
</tr>
<tr>
<td>HPHM</td>
<td>3745 ± 1249</td>
<td>3840 ± 1160</td>
<td>3189 ± 925</td>
<td>4383 ± 1126</td>
<td>5043 ± 980</td>
<td>2601 ± 950</td>
</tr>
<tr>
<td>Cm</td>
<td>2627 ± 645</td>
<td>2249 ± 820</td>
<td>2120 ± 553</td>
<td>2801 ± 949</td>
<td>3836 ± 830</td>
<td>1363 ± 645</td>
</tr>
<tr>
<td>HPHm</td>
<td>1941 ± 595</td>
<td>1968 ± 394</td>
<td>1530 ± 246</td>
<td>2431 ± 402</td>
<td>2930 ± 274</td>
<td>1195 ± 293</td>
</tr>
<tr>
<td>Pressure 7.4 mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>3357 ± 299</td>
<td>2887 ± 421</td>
<td>2784 ± 559</td>
<td>3534 ± 264</td>
<td>4718 ± 638</td>
<td>1810 ± 288</td>
</tr>
<tr>
<td>HPH</td>
<td>5311 ± 1634</td>
<td>5549 ± 1418</td>
<td>4765 ± 1143</td>
<td>6135 ± 1209</td>
<td>7070 ± 1262</td>
<td>3904 ± 1221</td>
</tr>
<tr>
<td>HPHR</td>
<td>2814 ± 386</td>
<td>2906 ± 61</td>
<td>2339 ± 118</td>
<td>3435 ± 158</td>
<td>4074 ± 159</td>
<td>1852 ± 110</td>
</tr>
<tr>
<td>HPHM</td>
<td>3313 ± 731</td>
<td>3427 ± 815</td>
<td>2804 ± 636</td>
<td>3984 ± 700</td>
<td>4593 ± 767</td>
<td>2302 ± 613</td>
</tr>
<tr>
<td>Cm</td>
<td>2527 ± 807</td>
<td>2244 ± 843</td>
<td>2084 ± 753</td>
<td>2724 ± 975</td>
<td>3742 ± 1021</td>
<td>1353 ± 654</td>
</tr>
<tr>
<td>HPHm</td>
<td>1917 ± 256</td>
<td>1991 ± 414</td>
<td>1544 ± 313</td>
<td>2432 ± 204</td>
<td>2980 ± 318</td>
<td>1197 ± 207</td>
</tr>
<tr>
<td>Pressure 13.0 mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>4254 ± 706</td>
<td>3962 ± 872</td>
<td>3771 ± 500</td>
<td>4563 ± 1075</td>
<td>5821 ± 628</td>
<td>2728 ± 906</td>
</tr>
<tr>
<td>HPH</td>
<td>5912 ± 1859</td>
<td>6203 ± 2241</td>
<td>5368 ± 1397</td>
<td>6797 ± 1481</td>
<td>7700 ± 1600</td>
<td>4445 ± 1426</td>
</tr>
<tr>
<td>HPHR</td>
<td>4130 ± 314</td>
<td>4035 ± 530</td>
<td>3454 ± 218</td>
<td>3771 ± 218</td>
<td>4752 ± 156</td>
<td>2828 ± 233</td>
</tr>
<tr>
<td>HPHM</td>
<td>3397 ± 441</td>
<td>3547 ± 840</td>
<td>2926 ± 733</td>
<td>4059 ± 711</td>
<td>4643 ± 793</td>
<td>2361 ± 529</td>
</tr>
<tr>
<td>Cm</td>
<td>2702 ± 450</td>
<td>2367 ± 349</td>
<td>2270 ± 508</td>
<td>2879 ± 364</td>
<td>3812 ± 619</td>
<td>1493 ± 223</td>
</tr>
<tr>
<td>HPHm</td>
<td>1894 ± 256</td>
<td>1973 ± 250</td>
<td>1489 ± 190</td>
<td>2400 ± 50</td>
<td>2886 ± 173</td>
<td>1183 ± 75</td>
</tr>
<tr>
<td>Pressure 17.2 mmHg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>5677 ± 796</td>
<td>5462 ± 737</td>
<td>5197 ± 531</td>
<td>6110 ± 1135</td>
<td>7400 ± 487</td>
<td>4040 ± 955</td>
</tr>
<tr>
<td>HPH</td>
<td>6619 ± 2148</td>
<td>7002 ± 2064</td>
<td>6134 ± 1655</td>
<td>7647 ± 1854</td>
<td>8455 ± 1891</td>
<td>5191 ± 1659</td>
</tr>
<tr>
<td>HPHR</td>
<td>5623 ± 1284</td>
<td>5797 ± 1091</td>
<td>5039 ± 761</td>
<td>6476 ± 944</td>
<td>4208 ± 811</td>
<td>7256 ± 922</td>
</tr>
<tr>
<td>HPHM</td>
<td>1950 ± 356</td>
<td>1681 ± 250</td>
<td>1644 ± 345</td>
<td>2132 ± 358</td>
<td>2801 ± 434</td>
<td>1068 ± 308</td>
</tr>
<tr>
<td>Cm</td>
<td>1810 ± 470</td>
<td>1925 ± 413</td>
<td>1465 ± 311</td>
<td>2365 ± 265</td>
<td>2767 ± 310</td>
<td>1144 ± 195</td>
</tr>
</tbody>
</table>

87
7.4.2 Directional complexity

All directional complexities are higher in the original HPH trees than in the original control trees. In the maximum Strahler order pruned HPH trees, all directional complexities are higher than controls at 6.3 mmHg but lower than controls at 13.0 mmHg. At 7.4 mmHg, right and posterior complexities are higher in control trees than in the radius-pruned HPH trees; all other directional complexities are higher in the radius-pruned HPH trees at this pressure.

For the minimum Strahler order pruned trees, all directional complexities are higher in pruned control trees than in pruned HPH trees for pressures 6.3, 7.4, and 13.0 mmHg. At 17.2 mmHg, right, posterior, and ventral complexities are higher in control than in HPH, while left, dorsal, and anterior complexities higher in the radius-pruned HPH trees at this pressure.

The radius-pruned HPH trees had higher right, posterior, and ventral complexities than controls at all pressures and higher dorsal complexity than controls at all pressures except 17.2 mmHg. The left and anterior complexities are higher in control trees than in the radius-pruned HPH trees.

7.5 Discussion

This study characterizes control and HPH spatial trees, quantifying the number of branches and leaves, tree depth, Strahler order, and directional complexities. We found that the tree depth is higher in HPH trees compared to control, a phenomenon which is preserved after pruning the trees. This difference can be explained by vascular remodeling in HPH, causing the large arteries to stiffen and dilate [51, 102]. However, at higher contrast pressures (13.0 and 17.2 mmHg) the control and HPH trees become more similar. This can be explained by the fact that in control animals, the vessels are more compliant and expand more when perfused at higher pressures. The latter is also observed with the minimal Strahler order pruning at the second highest pressure (13.0 mmHg), where the control trees’ depth exceeds that of the HPH trees.

The radius pruned HPH trees have higher left and anterior complexity compared to the controls. This means that in HPH, more vessels point toward the head and the left. This may indicate that the network is adapting to effectively transport blood to these regions. Overall, the directional complexities correlate with number of branches. However, the HPH spatial trees have more branches due to increased vessel radius, which allows the contrast to perfuse more arteries. By design, the radius-pruning algorithm makes the number of branches in HPH trees close to that of controls. Therefore, the directional complexity results from these trees are more significant, as they will be less influenced by branch count. Moreover, the control trees have higher right, ventral, dorsal, and posterior complexities than the radius-pruned HPH trees, with the exception of dorsal complexity at the highest pressure (17.2 mmHg). This exception reflects our previous finding that higher pressure trees are more difficult to distinguish.

Our use of degree-0 persistent homology via the height filtration is inspired by Belchi et al. [79], who found that humans with mild to moderate COPD had lower upwards complexity of the bronchial network compared to healthy subjects. In this context, “upwards” complexity refers to
bronchi pointing toward the head. Since HPH is the type of PH most commonly associated with COPD [67] and the pulmonary arteries branch in a manner similar to the bronchi, we implemented the same filtration as [79] on the murine spatial trees. Belchi et al. [79] found that upwards complexity of the bronchi is lower in disease; this is contrary to our findings that anterior complexity is higher in HPH. However, mice are quadrupeds and their lungs are rotated compared to humans. Thus, dorsal complexity might be a more apt comparison to human upwards complexity because it is the direction which counters gravity. We found that the dorsal complexity is lower in the radius pruned HPH trees compared to control trees at the three lowest pressures.

Finally, Belchi et al. [79] found that the choice of filtration direction do not influence results, reporting upwards complexity only. In our study, we did detect a difference in some of the directional complexities in the pruned trees, which could be reflective of the differences between human and mouse lung structures. Moreover, Belchi et al. [79] examined the airways, which could explain the differences in our results.

### 7.5.1 Limitations and future work

The major limitation in this study is a lack of data. We analyzed data from three control and HPH subjects. TDA methods are designed for analysis of large datasets, yet for physiological studies it can be difficult to conduct experiments with large number of animals. One way to circumvent this is by using machine learning to generate surrogate networks representative of the actual data, an approach we plan to pursue in future studies. This approach was used in the persistent homology study by Bendich et al. [70] analyzing brain arterial networks. Using machine learning they generated 98 trees from repeated iterations of a tube-tracking algorithm. We have had some success in generating simulated networks by attaching self-similar structured trees to principal branches of our labeled spatial trees [102], though these trees did not capture the three-dimensional nature of arterial branching. Using angles extracted from our spatial trees, we plan to generate networks with more similar branching properties to the trees found in this study and implement our methods to see if stronger conclusions can be drawn.

It is also critical to test our methods on trees extracted from human lung images. Such images can be difficult to obtain due to privacy concerns of patients. Moreover, segmenting these images is challenging because clinical CT images have lower resolution and in vivo images also contain airways and veins, making it difficult to isolate the arteries. The mouse images analyzed here were ideal. The CT images were gathered at high resolution and the pulmonary arteries were excised, i.e., the arterial network was the anatomical objects in the images.

### 7.6 Conclusion

In this study, we extracted spatial trees from micro-CT images representing the pulmonary arterial networks in control and HPH mice. There is a large discrepancy between branch counts between the two groups, which we hypothesize is an imaging consequence rather than evidence of angiogenesis,
as microvascular rarefaction has been observed in mice with HPH [51]. We devised pruning algorithms based on vessel radius and Strahler order, each with their own advantages and disadvantages. Guided by a study conducted on similar networks [79], we computed the directional complexities using a degree-0 persistent homology of a height filtration. Overall, we found that tree depth is larger in the HPH trees, directional complexity correlates with the branch count, and that the left and dorsal complexities are lower in HPH. Finally, we noticed that results at low perfusion pressures reveal more differences, likely because at higher pressure increased elasticity in healthy animals cause vessels to dilate more making them appear similar to HPH animals. This study serves as a proof-of-concept for the use of TDA to detect differences in diseased vascular networks.

7.7 Acknowledgements

This study was initiated during a summer research experience for undergraduates at North Carolina State University, funded by NSA Contract: H98230-20-1-0259. The authors also thank Naomi Chesler, University of California-Irvine, for making micro-CT images available for this study.
This chapter includes preliminary results for a work-in-progress study on the use of approximate Bayesian computation (ABC) to generate 3D structured trees which align with data. Using novel algorithms, we assign direction vectors to vessels at a junction using optimally placed change points. Arterial branching angles in 3D are computed using these vectors, and ABC analysis is conducted to determine acceptable angle values for generating realistic 3D structured trees. We intend to submit this work for publication this summer.

8.1 Introduction

The pulmonary arterial network is a rapidly branching structure that transports blood throughout the lungs, ensuring that each lobe is reached. As previously stated (Section 4.6), blood flow dynamics in the network can be studied with a multiscale model, where large arteries’ dimensions are taken from data and small vessels are represented by self-similar structured trees. We defined such structured trees by their radius scaling factors and length-to-radius ratios in [102], but branching angles were not taken into account due to the 1D nature of the fluid dynamics model.

Characterizing the branching angles of the pulmonary arteries is crucial for understanding its space-filling properties and generating realistic structured trees in 3D. Before angles can be
extracted, vectors must be assigned to indicate the directions of the parent and daughter vessels at a junction. We define such vectors using a novel approach involving change points which captures the branching angles at a junction even in vessels which change direction along their length. For each daughter vessel at a bifurcation, we use these vectors to calculate the in-plane angle $\psi$ between the daughter and parent vectors and the rotational out-of-plane angle $\theta$. With these angles, we can generate entire 3D trees along the framework in [85]. To do so using the exact angle measures at each junction, simply recreating the exact junctions in the data, is computationally expensive and does not result in a self-similar structured tree. Therefore, we use approximate Bayesian computation (ABC) analysis to determine which angle values should be used. This analysis compares simulated trees to the data, calculates their difference in summary statistics with a distance function, and generates posterior distributions of acceptable angle values that produce trees within a certain distance threshold of the original data [48, 78].

8.2 Vessel vectors

As previously mentioned (Chapters 4, 6, & 7), each micro-CT image studied for this dissertation depicts the pulmonary arteries of either a control or a pulmonary hypertensive (PH) mouse. Through the protocol in [102], we obtained labeled spatial trees which represent these arterial networks. In these trees, each vertex has spatial coordinates $(x, y, z)$ and is either a node, which collectively refers to "leaves" (deg = 1) and "branching points" (deg ≥ 3), or an edge point (deg = 2). A branch in the spatial tree is a path of edge points between two nodes that represents a single arterial section between junctions (or between a junction and a terminal end). Approximately 98-99% of all junctions in the spatial trees were bifurcations, where a single parent branch $D_0$ splits into 2 daughter branches $D_1$ (daughter with larger radius) and $D_2$. These branches connect at a common branching point called the hub node, $H$, representing the point at which an artery splits in two. Each $D_i$ also contains another node, $N_{D_i}$, on the end opposite the hub node (Figure 8.1). To define branching angles at a bifurcation, we first must define vectors $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2$ for the direction of the parent and daughter vessels from the hub node. For each branch $D_i = (H, N_{D_i} : e_1, e_2, ..., e_k)$, we use linear regression to find the vector $\mathbf{v}_i$ which minimizes the relative error $J$ between $\mathbf{v}_i$ and the points $\{H, e_1, e_2, ..., e_k, N_{D_i}\}$. This is accomplished using an application of the singular-value decomposition.

**SVD:** Any $n \times p$ real-valued matrix $X$ can be decomposed using three matrices $U$, $S$, and $V$ which each have certain properties. This factorization, called the singular-value decomposition (SVD), is

$$X = U S V^T$$

where $U$, $S$, and $V$ are of sizes $n \times n$, $n \times p$, and $p \times p$, respectively. If $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$ are the
nonzero eigenvalues of $X^T X$ (also of $XX^T$), then the entries $s_{ij}$ of $S$ are given by

$$s_{ij} = \begin{cases} 
0 & i \neq j \\
\sqrt{\lambda_i} & i = j.
\end{cases}$$

The columns of $U$ are the eigenvectors $U_1, ..., U_n$ of $XX^T$, while the columns of $V$ are the eigenvectors $V_1, ..., V_p$ of $X^TX$. Additionally, the matrices $U$ and $V$ are orthogonal [108].

**Linear regression:** Consider a set of points represented by the column vectors $\{x_1, x_2, ..., x_p\} \in \mathbb{R}^3$. Suppose $X$ is a $3 \times p$ matrix given by

$$X = \begin{bmatrix} 
x_1 - x' & x_2 - x' & ... & x_p - x' 
\end{bmatrix},$$

for some anchor point $x' \in \mathbb{R}^3$, with SVD given by

$$X = USV^T,$$

then the line of best fit for the points $\{x_1, x_2, ..., x_p\}$ that passes through $x'$ is in the direction of the vector $U_1$, the first column of $U$. The sum of squared errors between the points and the line of best fit is given by $SSE = s_2^2 + s_3^2$, the sum of the squares of the two lowest singular values in $S$ [25].

To fit a vector to a branch $D$, we use this SVD result to our advantage. Algorithm 1 describes this process, taking the branch’s points $\{H, e_1, e_2, ..., e_k, N_D\}$ as inputs and providing the vessel vector $v$ and relative error $J = SSE/(k + 2)$ as outputs.
**Change points:** Some branches do not follow a straight trajectory along their entire length, bending considerably due to their morphology or errors in imaging. For example, $D_1$ in Figure 8.1 points slightly right and steeply upward during its first 7 edge points, then has a noticeable bend and points more toward the right for the rest of the branch. In these instances, defining a single vector to encompass the branch’s direction may not be the best strategy. The purpose of defining vessel vectors is to compute branching angles at a junction. If edge points further from the junction are influencing this vector’s direction, it may be useful to insert one or more change points to divide the branch into sections and fit a different vector in each section.

Consider a branch $D = (H, N_D : e_1, e_2, ..., e_k)$. A change point can be inserted at any of the edge points $e_i$. Inserting a change point yields two vectors, $v_1$ and $v_2$. The vector $v_1$ starts at the hub node $H$ and minimizes the relative error $J_1$ between $v_1$ and the points $\{H, e_1, ..., e_i\}$, and $v_2$ starts at $e_i$ and minimizes the relative error $J_2$ between $v_2$ and the points $\{e_i, ..., e_k, N_D\}$. Algorithm 2 details how 1 change point can be placed in the optimal location along a branch, minimizing $J = J_1 + J_2$.

It is clear that if we make every $e_i$ a change point, we will have an error $J = 0$. However, this is computationally expensive and defeats the purpose of using linear regression at all. In this study, we allow for at most 2 change points (Algorithm 3), and choose their location via Algorithm 4. The optimization is based on the **Bayesian information criterion (BIC)**

$$\text{BIC} = P \ln(k + 2) + 2 \ln(J),$$

where $k = \#$ of edge points, $J = \text{relative error}$, and $P = \#$ of parameters calculated by the model. Placing $n$ change points yields $P = 2n + 1$ parameter calculations: the optimal locations of the $n$ change points, and the $n + 1$ best fit vectors. Thus, the use of change points increases the $P \ln(k + 2)$ term in (8.1), penalizing our BIC. This is to prevent us from unnecessarily complicating the model; additional change points will not be added unless they lower the BIC.

With Algorithm 4, we can obtain the vessel vectors $v_0, v_1, v_2$ for the parent and 2 daughter branches any bifurcation. Regardless of how many change points were placed, $v_0, v_1, v_2$ are always taken to be the vectors which intersect the hub node. Figure 8.2 shows an example junction’s vessel vectors when up to 0, 1, and 2 change points are allowed to be placed on its branches.

![Figure 8.2](image-url) Vessel vectors $v_0, v_1, v_2$ generated for an example bifurcation when at most A) 0, B) 1, and C) 2 change points are allowed.
Note that in Figure 8.2C, even though 2 change points were allowed, it was only optimal to add 2 change points to the daughter 1 branch. This agrees with the BIC criterion penalizing increased model complexity. We found that choosing 2 change points is only optimal in branches with 2 edge points, allowing for a perfect fit. This happens rarely, as shown in Table 8.1.

Table 8.1 Breakdown of vessel vectors with \( n = 0, 1, 2 \) change points for 3 control and 3 PH mice (imaged at contrast pressure 6.3 mmHg). Note that the number of total vessel vectors does not equal the total number of branches because a branch may be a daughter of one junction and parent of another, in which case its vessel vector is computed twice for different hub nodes.

<table>
<thead>
<tr>
<th>Mouse</th>
<th># of branches</th>
<th># (%) of vessel vectors with ( n ) optimal change points</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( n = 0 )</td>
</tr>
<tr>
<td>Control 1</td>
<td>1095</td>
<td>584 (36.6%)</td>
</tr>
<tr>
<td>Control 2</td>
<td>844</td>
<td>417 (33.6%)</td>
</tr>
<tr>
<td>Control 3</td>
<td>1114</td>
<td>552 (33.9%)</td>
</tr>
<tr>
<td>PH 1</td>
<td>2106</td>
<td>1259 (40.7%)</td>
</tr>
<tr>
<td>PH 2</td>
<td>3015</td>
<td>1808 (41.2%)</td>
</tr>
<tr>
<td>PH 3</td>
<td>1467</td>
<td>798 (36.7%)</td>
</tr>
</tbody>
</table>

8.3 Branching angles

Four branching angles govern the direction of the daughter pulmonary arteries as they split from their parent. Tawhai et al. [37] defined these angles, with \( \psi_{1,2} \) being the branching angle between each daughter and the parent branch within their common plane, and \( \theta_{1,2} \) being the out-of-plane rotational angles between each daughter and a specified reference plane. Using the vessel vectors \( v_0, v_1 \) and \( v_2 \), these angles are computed at each bifurcation in our spatial trees. Then, kernel density estimation is used to approximate their distributions [82]. From these distributions, we sample angle values to generate structured trees using the algorithm in Lee et al [85].

Angle Definitions: Figure 8.3 shows the four branching angles \( \psi_{1,2} \) and \( \theta_{1,2} \) at a bifurcation with hub node \( H \) and vessel vectors \( v_0, v_1, \) and \( v_2 \). The vector \( A \) specifies vessel vector of the other daughter branch from the previous bifurcation, i.e., the parent branch’s sibling, referred to as the “aunt” vector. This vector is used to define a reference plane for the \( \theta_{1,2} \) rotations, as described in [85].

The vessel vectors \( v_0, v_1 \) and \( v_2 \) are in \( \mathbb{R}^3 \), i.e., they are each defined in terms of their direction along the standard \( x, y, \) and \( z \) axes. To compute branching angles, it is advantageous to reorient vectors of a junction in an adjusted coordinate frame, where the reversed unit parent vector specifies the adjusted \( \tilde{z} \)-axis. The adjusted axes are defined as

\[
\tilde{z} = - \frac{v_0}{||v_0||}, \quad \tilde{x} = \frac{A - (A \cdot \tilde{z}) \tilde{z}}{||A - (A \cdot \tilde{z}) \tilde{z}||}, \quad \tilde{y} = \tilde{z} \times \tilde{x}.
\]
The $\tilde{x}$-axis is perpendicular to the $\tilde{z}$-axis and lies in the parent-aunt plane, labeled the $\tilde{x}\tilde{z}$-plane in Figure 8.3. The in-plane angles $\psi_{1,2}$ are the angles between the daughter vectors and $\tilde{z}$, defined by

$$
\psi_i = \cos^{-1}\left( \frac{\tilde{z} \cdot \mathbf{v}_i}{||\tilde{z}|| \cdot ||\mathbf{v}_i||} \right) \quad \text{for } i = 1, 2. \tag{8.3}
$$

The out-of-plane angles $\theta_{1,2}$ are the clockwise angles of rotation between each daughter vector and the $\tilde{x}\tilde{z}$-plane about the $\tilde{z}$-axis. These are defined as

$$
\theta_i = \pi - \tan^{-1}\left( \frac{\mathbf{v}_i \cdot \tilde{y}}{\mathbf{v}_i \cdot \tilde{x}} \right) \quad \text{for } i = 1, 2. \tag{8.4}
$$

The in-plane angles $\psi_{1,2}$ lie in the range $\psi_j \in [0, \pi]$, and we adjust the out-of-plane angles to lie in the range $\theta_i \in (-\pi, \pi]$. A positive $\theta_i$ value indicates that the vector lies above the $\tilde{x}\tilde{z}$-plane in the adjusted frame, while a negative $\theta_i$ value indicated that the vector lies below the $\tilde{x}\tilde{z}$-plane.

**Figure 8.3** Example bifurcation at hub node $H$ with parent vessel vectors $\mathbf{v}_0$ and daughters $\mathbf{v}_1$ and $\mathbf{v}_2$. A coordinate frame $\tilde{x}, \tilde{y}, \tilde{z}$ is shown in blue, where the origin is the hub node $H$. The $\tilde{z}$ axis is aligned in the reverse direction of the parent vector $\mathbf{v}_0$. The $\tilde{x}$-axis is defined as the vector perpendicular to $\tilde{z}$ in the plane defined by $\tilde{z}$ and $A$ (the "aunt" vector, the other daughter vector from the previous hub node). The $\tilde{y}$ is the cross product of $\tilde{z}$ and $\tilde{x}$. The branching angles $\psi_j$ are computed as the arccosine of the normalized dot product of $\tilde{z}$ and $\mathbf{v}_j$, as defined in equation (8.3). The angles $\theta_i$ are computed using the arctangent of the ratio $(\mathbf{v}_t \cdot \tilde{y})/(\mathbf{v}_t \cdot \tilde{x})$, as defined in equation (8.4).
Figure 8.4 shows histograms and distributions of the angles extracted from our three control and three PH spatial trees. The distributions are calculated using the kernel density estimator (KDE), which approximates the probability density function. For angle data $\Theta = \{\Theta_1, ..., \Theta_N\}$ (being values of $\psi_1, \psi_2, \theta_1,$ or $\theta_2$), KDE is given by

$$\text{KDE}(x) = \frac{1}{nh} \sum_{i=1}^{N} K \left( \frac{x - \Theta_i}{h} \right)$$

(8.5)

where $K$ is the normal kernel and $h$ is the optimal bandwidth [82]. Note that this distribution is calculated over the interval of $[0, \pi]$, but distributions for angle data should be periodic in nature. More work is needed to generalize this to fit periodic data.

We did not find a significant difference between the angle measures between the control and PH mice. Additionally, we found that the angles $\theta_{1,2}$ had bimodal distributions, exhibiting fairly symmetric behavior from $(-\pi, 0]$ and $[0, \pi]$. This bimodal distribution can be explained by the fact that in 88% of junctions, one of $\theta_i$ was positive and the other was negative, with it being equally likely that $\theta_1$ or $\theta_2$ would be the negative one. To account for this phenomenon when generating structured trees, we pulled the $\theta_i$ from a distribution of $|\theta_i|$ angle values and handled the sign of $\theta_i$ separately. Note that the periodic extension of the KDE distribution of the absolute values $|\theta_{1,2}|$ (shown in Figure 8.4C) is not differentiable.

**Figure 8.4** Histograms (blue: control, orange: HPH) and KDE distributions (green lines) of angle measures for control and PH mice. A) $\psi$ angles, B) $\theta$ angles, C) absolute value of $|\theta|$ angles.
Generating 3D Trees: To rotate vector \( \mathbf{v} \) clockwise about a vector \( \mathbf{u} \) by an angle \( \Theta \), multiply \( \mathbf{v} \) by the rotation matrix \( R(\mathbf{u}, \Theta) \),

\[
R(\mathbf{u}, \Theta) = \begin{bmatrix}
\cos \Theta + u_x^2 (1 - \cos \Theta) & u_x u_y (1 - \cos \Theta) - u_z \sin \Theta & u_x u_z (1 - \cos \Theta) + u_y \sin \Theta \\
 u_x u_y (1 - \cos \Theta) + u_z \sin \Theta & \cos \Theta + u_y^2 (1 - \cos \Theta) & u_y u_z (1 - \cos \Theta) - u_x \sin \Theta \\
u_x u_z (1 - \cos \Theta) - u_y \sin \Theta & u_y u_z (1 - \cos \Theta) + u_x \sin \Theta & \cos \Theta + u_z^2 (1 - \cos \Theta)
\end{bmatrix}.
\]

At a junction with parent and aunt vectors \( \mathbf{v}_0 \) and \( \mathbf{A} \), the branching angles \( \psi_{1,2} \) and \( \theta_{1,2} \) can be used to generate daughters \( \mathbf{v}_{1,2} \) by defining the adjusted coordinate frame (equation (8.2)) and performing two subsequent rotations,

\[
\tilde{\mathbf{v}}_i = R(\tilde{y}, \psi_i) \tilde{z} \\
\mathbf{v}_i = R(\tilde{z}, \theta_i) \tilde{v}_i.
\]

Using branching angles pulled from the KDE distributions (Figure 8.4), we generated structured trees off of the principal pathway for the three control and three PH mice. The principal pathway was defined as the largest connected subtree where all vessels’ radii were at least 40\% of the root radius, as in [102]. Angles \( \psi_1, \psi_2, \theta_1 \) were positive, and \( \theta_2 \) was assigned to be negative in 88\% of junctions\(^1\). There is some variability inherent in choosing angles from the distribution; Figure 8.5 shows results from two sample tree generations for each mouse.

### 8.4 ABC analysis

The trees in Figure 8.5 were generated using branching angles pulled from the KDE distributions (Figure 8.4). They look reasonable at first glance, but the high number of branches and the 3D nature of the trees make it difficult to visually judge how well they fit the data. We need a more robust way of comparing simulated trees to the original data to determine which parameter values are needed to generate the 3D structured trees that most closely resemble the actual data. This can be accomplished via approximate Bayesian computation (ABC), a family of parameter inference methods which can be used when data simulation is more straightforward than computing parameter likelihoods [48, 78]. By generating simulated trees, we compare them to an observed tree in terms of a distance function on a set of summary statistics from both trees. Angle measures for which this distance falls below a certain threshold are accepted and become part of the new posterior distribution; otherwise, they are rejected. The lower this acceptable distance threshold, the closer our simulated trees are to the prior observed tree. However, a low threshold can result in a low acceptance rate, providing us with very little data to form the posterior distribution.

**ABC Rejection Sampling:** Techniques of ABC generally follow the same framework. Using observed data, a prior distribution for parameters is defined. Simulated data is generated by sampling param-

---

\(^1\)Although we found that \( \theta_1 \) and \( \theta_2 \) were equally likely to be negative, we also found that choosing one over the other did not heavily impact the appearance of our trees.
Figure 8.5 Trees generated for control (C) and hypertensive (H) mice. The red branches belong to the principal pathway, and in gray are the original branches of the spatial trees. The generated trees are in blue and were created by pulling $\psi_1, \psi_2, \theta_1,$ and $\theta_2$ from the KDE distributions in Figure 8.4, as well as using the structured tree parameters from [102]: $\alpha = 0.86, \beta = 0.67$, control $\ell_{rr} = 13.4e^{-0.00771r}$, and hypertensive $\ell_{rr} = 10.9e^{-0.00797r}$.

The simplest ABC method is the ABC Rejection Sampling technique, which we have included in Algorithm 5. In this algorithm, parameters that are not accepted are simply rejected. In some more sophisticated ABC algorithms, unaccepted values are used to initialize the parameters for the next iteration [48, 78].

Preliminary Results: We have implemented ABC rejection sampling on a small example tree shown in Figure 8.6A. For simplicity, we begin with a 2D tree, but we plan to extend these results to trees in 3D in the future.

Defining appropriate summary statistics for ABC rejection sampling is important for judging the fitness of the simulated data. Choosing the wrong summary statistics can lead to outcomes that, while they fall below the distance threshold, do not accurately represent the observed data. In our analysis, we tried using the leaf coordinates as our summary statistics and minimizing their change in position in the simulated trees from the example tree. While this produced closely aligned trees, it was not feasible to extend this to larger subsets of the actual spatial trees without implementing an additional graph-matching algorithm.

Rather than attempting to align the leaves, we then focused on ensuring that the simulated trees would perfuse a similar area of the lung as the original tree. For our 2D tree, $T$ (with vertices $V$), we
determined that the summary statistics \( S(T) \) should be

\[
S(T) = \{A, XL, YL\}, \text{ where}
\]

\[
A = \text{area of the convex polygon created by the root and leaves}
\]

\[
XL = \text{maximum distance across the tree in the x direction}
\]

\[
= \max_{v_1, v_2 \in V} (x_{v_1} - x_{v_2})
\]

\[
YL = \text{maximum distance across the tree in the y direction}
\]

\[
= \max_{v_1, v_2 \in V} (y_{v_1} - y_{v_2}).
\]

To compare summary statistics of the observed tree, \( T_0 \), to a generated tree, \( T^* \), we defined the distance function \( d_2(T_0, T^*) \) to be

\[
d_2(S_2(T_0), S_2(T^*)) = \frac{(A_0 - A^*)^2}{A_0^2} + \frac{(XL_0 - XL^*)^2}{XL_0^2} + \frac{(YL_0 - YL^*)^2}{YL_0^2}.
\]

Minimizing this distance function ensures that the accepted parameters produce trees which perfuse a similar area and are not significantly taller or wider than the original tree.

Our example tree (Figure 8.6A) was generated using \( \alpha = 0.86, \beta = 0.67, \) and \( \ell_{rr} = 3.59. \) At each bifurcation, \( \psi_1 = \pi/3, \psi_2 = -\pi/3, \) and \( \theta_{1,2} = 0 \) because the tree was in 2D. We performed ABC Rejection Sampling, generating simulated trees with the same \( \alpha, \beta, \ell_{rr} \) and \( \theta_{1,2} \), sampling \( \psi_{1,2} \sim \text{unif} (0, \pi) \)). As \( \epsilon \) decreased, the posterior distributions for \( \psi_{1,2} \) became increasingly narrow around \( \pi/3 \) and \( -\pi/3 \), as we would expect (Figure 8.6B). Lowering \( \epsilon \) produced simulated trees which more closely aligned with the example tree (Figure 8.6C), but very few parameters were accepted into the posterior distribution.

### 8.5 Conclusion

Analyzing branching angles provides a better understanding of the morphometry in the pulmonary arterial tree. To obtain these angles from data, it is necessary to define vectors representing the parent and daughter directions at each bifurcation. While many studies define such a vector in terms of a branch's start and end vertex, we used a novel approach based on the strategic implementation of change points to ensure that the vectors captured the local branching behavior and were not swayed by bends in the vessels further from the hub node. We extracted 4 branching angles at each junction and used these angles to generate 3D structured trees off of the principal pathway of pulmonary arterial networks in control and PH mice.

Moving forward, we plan to apply ABC analysis in these 3D spatial trees to obtain more fitting structured trees for fluid dynamics simulations. We have had some success implementing ABC Rejection Sampling in 2D, and have obtained posterior distributions which aligned with what we expected. In our extension to 3D, we plan to implement one of the more sophisticated ABC
Figure 8.6 A) 2D tree example, generated from parameters $\alpha = 0.86$, $\beta = 0.67$, $\ell_{rr} = 3.59$, $\psi_1 = \pi/3$, $\psi_2 = -\pi/3$, $\theta_{1,2} = 0$. B) Prior and posterior distributions for $\psi_{1,2}$ for different acceptance thresholds $\epsilon$. The acceptance rate, $ar$, is the % of trees generated whose parameters were accepted by the ABC rejection sampler, which decreases with $\epsilon$. C) Two representative trees for each $\epsilon$. The sample tree is shown in blue. The simulated trees share the same root branch with the example, and their subsequent branches are shown in red.

algorithms which can help with the low acceptance rate we experienced. The ability to generate realistic simulated data is useful for complex fluid dynamics simulations as well as studies where a lack of data is a major limitation.

8.6 Appendix: Algorithms

Algorithm 1: $[v, J] = \text{LinReg}(\{H, e_1, e_2, ..., e_k, N_D\})$

**Data:** Branch vertices $\{H, e_1, e_2, ..., e_k, N_D\} \in \mathbb{R}^3$

**Result:** Vessel vector $v \in \mathbb{R}^3$, relative error $J \in \mathbb{R}$

\[
X := [H - H \mid e_1 - H \mid e_2 - H \mid ... \mid e_k - H \mid N_D - H]
\]

$[U, S, V] := \text{svd}(X)\dagger$

$v := [U_{11}, U_{21}, U_{31}]^T$

$SSE := S_{22}^2 + S_{33}^2$

$J := SSE/(k + 2)$

\dagger Matlab's singular value decomposition function. $\text{svd}(X) = [U, S, V]$ such that $X = USV^T$. 

101
Algorithm 2: \([v_1, v_2, CP, J] = \text{OneChangePoint}\{H, e_1, e_2, ..., e_k, ND\}\)

**Data:** Branch vertices \([H, e_1, e_2, ..., e_k, ND] \in \mathbb{R}^3\)

**Result:** Vessel vectors \(v_1, v_2 \in \mathbb{R}^3\), change point \(CP \in \mathbb{R}^3\), relative error \(J \in \mathbb{R}\)

Errors := {}

for \(i = 1\) to \(k\) do

\(A := \{H, e_1, ..., e_i\}\)
\(B := \{e_i, ..., e_k, ND\}\)
\([v_1, J_1] := \text{LinReg}(A)\)
\([v_2, J_2] := \text{LinReg}(B)\)

IterationSummary := \([v_1, v_2, J_1 + J_2]\)
Errors ← IterationSummary

\([J, index] := \min(\text{Errors}[: , 3])\)
\(CP := e_{\text{index}}\)
\(v_1 := \text{Errors}[\text{index}, 1], \quad v_2 := \text{Errors}[\text{index}, 2]\)

Algorithm 3: \([v_1, v_2, v_3, CP_1, CP_2, J] = \text{TwoChangePoints}\{H, e_1, e_2, ..., e_k, ND\}\)

**Data:** Branch vertices \([H, e_1, e_2, ..., e_k, ND] \in \mathbb{R}^3\)

**Result:** Vessel vectors \(v_1, v_2, v_3 \in \mathbb{R}^3\), change points \(CP_1, CP_2 \in \mathbb{R}^3\), relative error \(J \in \mathbb{R}\)

Errors := []

for \(i = 1\) to \(k - 1\) do

\(A := \{H, e_1, ..., e_i\}\)
\(CP_1 := e_i\)
\([v_1, J_1] := \text{LinReg}(A)\)

\(B := \{e_i, ..., e_k, ND\}\)
\([v_2, v_3, CP_2, J] := \text{OneChangePoint}(B)\)

IterationSummary := \([v_1, v_2, v_3, CP_1, CP_2, J_1 + J]\)
Errors ← IterationSummary

\([J, index] := \min(\text{Errors}[:, 6])\)
\(v_1 := \text{Errors}[\text{index}, 1], \quad v_2 := \text{Errors}[\text{index}, 2], \quad v_3 := \text{Errors}[\text{index}, 3]\)
\(CP_1 := \text{Errors}[\text{index}, 4], \quad CP_2 := \text{Errors}[\text{index}, 5]\)
**Algorithm 4:** \([v, n_{\text{CP}}] = \text{ChooseVesselVector}([H, e_1, e_2, \ldots, e_k, N_D])\)

**Data:** Branch vertices \([H, e_1, e_2, \ldots, e_k, N_D] \in \mathbb{R}^3\)

**Result:** Vessel vector \(v \in \mathbb{R}^3\), # of change points \(n_{\text{CP}} \in \mathbb{R}\)

\[
[v_{\text{LR}}, J_{\text{LR}}] = \text{LinReg}([H, e_1, e_2, \ldots, e_k, N_D])
\]

\[
[v_{1\text{CP}}, v_{2\text{CP}}, J_{1\text{CP}}] = \text{OneChangePoint}([H, e_1, e_2, \ldots, e_k, N_D])
\]

\[
[v_{2\text{CP}}, v_{3\text{CP}}, J_{2\text{CP}}] = \text{TwoChangePoints}([H, e_1, e_2, \ldots, e_k, N_D])
\]

\[
\text{BIC}_{LR} := \ln(k+2) + 2\ln(J_{LR})
\]

\[
\text{BIC}_{1\text{CP}} := 3\ln(k+2) + 2\ln(J_{1\text{CP}})
\]

\[
\text{BIC}_{2\text{CP}} := 5\ln(k+2) + 2\ln(J_{2\text{CP}})
\]

\[
\text{BIC}_{\text{min}} := \min(\text{BIC}_{LR}, \text{BIC}_{1\text{CP}}, \text{BIC}_{2\text{CP}})
\]

if \(\text{BIC}_{\text{min}} = \text{BIC}_{LR}\) then

\[
v := v_{\text{LR}}, \quad n_{\text{CP}} := 0
\]

else if \(\text{BIC}_{\text{min}} = \text{BIC}_{1\text{CP}}\) then

\[
v := v_{1\text{CP}}, \quad n_{\text{CP}} := 1
\]

else

\[
v := v_{2\text{CP}}, \quad n_{\text{CP}} := 2
\]

**Algorithm 5:** \([\pi^*(\Theta)] = \text{ABCRejectionSampler}(x_0, \pi(\Theta), \epsilon, I)\)

**Data:** Observed data \(x_0\), prior parameter distribution \(\pi(\Theta)\), acceptance threshold \(\epsilon > 0\), number of iterations \(I\)

**Result:** Posterior distribution \(\pi^*(\Theta)\)

\[
S_0 := S(x_0)
\]

Accepted := []

for \(i = 1\) to \(I\) do

Sample \(\Theta^* \leftarrow \pi(\Theta)\)

Generate \(x^* \leftarrow \Theta^*\)

\(S^* := S(x^*)\)

\(D := d(S_0, S^*)\)

if \(D \leq \epsilon\) then

\[
\text{Accepted} \leftarrow \Theta^*
\]

†Other important a priori components include defining the summary statistics \(S(x)\) to be examined and the distance function \(d(S_0, S^*)\).
CONCLUSION

The pulmonary arteries form a complex network structure, the mathematical properties of which have been studied for decades. In the presence of group III pulmonary hypertension (PH), these arteries become stiffer, large vessels dilate, and small vessels constrict, a phenomenon called vascular remodeling [51, 102]. Through studying the morphology of the pulmonary arterial network, we can quantify these changes and gain a better understanding of the effects of PH.

In this dissertation, I aimed to characterize the morphological properties of pulmonary arterial networks so as to optimize the formation of the realistic structured trees which can be used to conduct 1D fluid dynamics simulations. I analyzed micro-computed tomography scans (micro-CT) of lungs extracted from control and group III PH mice using a multi-step semi-automated algorithm developed in collaboration with Kitware, Inc. Through a combination of image segmentation, skeletonization, and graph extraction, I obtained labeled spatial trees representing pulmonary arterial networks and extracted parameters necessary to generate structured trees, including radius scaling factors $\alpha$ and $\beta$, length-to-radius ratio $\ell_{rr}$, and 3D branching angles $\psi_{1,2}$ and $\theta_{1,2}$. Of these parameters, I found that only $\ell_{rr}$ differed significantly between control and PH mice, indicating that increased vessel radius in the large arteries was the dominant morphological effect. Increased vessel stiffness was also observed in the PH mice [102].

Additional analyses were performed using the topological data analysis technique of persistent homology, which detects the shape of data by building a filtration of simplicial complexes and tracking their topological properties. Inspired by Belchi et al.'s work on bronchial networks [79], my collaborators and I computed degree-0 persistent homology of a height filtration, analyzing the complexity of the trees in different directions. The PH trees had significantly more branches than the control trees due to their increased vessel radii, so various pruning algorithms were implemented...
to normalize the trees and make a more apt comparison. Tree depth was higher in the PH trees, and PH trees pruned by vessel radius exhibited lower left and anterior complexities than control trees. Additionally, the properties of control trees obtained from higher pressure images appeared to become more similar to the PH trees due to their vessels being more inflated.

The lack of available data was a key limitation of the persistent homology study. To improve this study, approximate Bayesian computation (ABC) can be used to generate a large set of simulated trees representing a real pulmonary arterial network. Characterizing arterial branching angles is crucial to generating trees which accurately exhibit the network's space-filling properties. For the control and PH spatial trees, I extracted 3D branching angles at each junction to obtain distributions for the angles exhibited in these trees. Linear regression with change points was implemented to define vectors at each junction that captured the local trajectories of the branches; angles were calculated based on these vectors. While preliminary simulated trees in 3D look reasonable, ABC analysis can provide parameters which allow for the generation of trees which are within a measurable threshold of the real data. Such analysis was implemented on a 2D tree with promising results, and the next step would be to implement this analysis on trees in 3D.

All methods described in this dissertation are widely applicable and not dependent on the type of data. They can be used to study networks from other organ systems, species, and imaging formats. For example, we have used the process described in Chapter 4 to segment and graph the airways and aorta from human chest CT scans, though working with these images is not without its challenges. In the short-term, we plan to continue the work of this dissertation by publishing our ABC analysis and persistent homology study this year. However, the long-term goal of this research would be to define quantitative properties which, when observed in medical images, could alert clinicians about early disease development, aiding in early detection and better patient outcomes.
REFERENCES


