Spectral graph theory efficiently characterises ventilation heterogeneity in lung airway networks.

Carl A. Whitfield[1,2†], Peter Latimer[3], Alex Horsley[2], Jim M. Wild[4], Guilhem J. Collier[4], and Oliver E. Jensen[1]

[1]Department of Mathematics, University of Manchester, Manchester, UK
[2]Division of Inflammation, Immunity and Respiratory Medicine, University of Manchester, Manchester, UK
[3] Department of Physics and Astronomy, University of Manchester, Manchester, UK
[4]POLARIS, Imaging Sciences, Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK
[†]e-mail: carl.whitfield@manchester.ac.uk

April 15, 2020

Abstract

This paper introduces a linear operator for the purposes of quantifying the spectral properties of transport within resistive trees, such as airflow in lung airway networks. The operator, which we call the Maury matrix, acts only on the terminal nodes of the tree and is equivalent to the adjacency matrix of a complete graph summarising the relationships between all pairs of terminal nodes. We show that the eigenmodes of the Maury operator have a direct physical interpretation as the relaxation, or resistive, modes of the network. We apply these findings to both idealised and image-based models of ventilation in lung airway trees and show that the spectral properties of the Maury matrix characterise the flow asymmetry in these networks more concisely than the Laplacian modes, and that eigenvector centrality in the Maury spectrum is closely related to the phenomenon of ventilation heterogeneity caused by airway narrowing or obstruction. This method has applications in dimensionality reduction in simulations of lung mechanics, as well as for characterisation of models of the airway tree derived from medical images.

1 Background

In healthy human lungs, the airways form a bifurcating tree where, on average, around the first 16 generations of airways are purely conductive and serve to transport gas from the mouth to the alveolar region where the majority of gas exchange takes place. The conducting airways terminate in ∼30,000 respiratory units (or acini) [1], where diffusion becomes the dominant transport mechanism. Resistive flow in the conducting airways, as well as tissue compliance, determines the ventilation mechanics and hence requires a high-dimensional mathematical model to be represented accurately [2, 3, 4]. Efficient dimensionality reduction algorithms are needed to improve the computation time and tractability of using such models in inverse problems in the future.

Furthermore, understanding the relationship between structure and function in the lungs is crucial to identifying lung disease physiology and its progression. One particular marker of progression in obstructive

lung conditions is "ventilation heterogeneity" (VH), or the uneven distribution of fresh gas in the lungs, which can result from obstructions in or narrowing of the airways [5, 6, 7]. The methods presented here provide a new approach to characterise the resistance structure and its effect on VH. This will be useful for visualisation and characterisation of complex airway models based on patient computed tomography (CT) images, lung casts or micro-CT scans of excised lungs.

Graph theory perspectives have proven useful in characterising, classifying and understanding transport in various examples of biological networks [8, 9, 10]. The approach we take in this paper uses concepts from spectral graph theory, where a linear operator describing a particular physical property or characteristic of the network is decomposed into its eigenspectum. This has proved a useful for dimensionality reduction and visualisation of large data sets in various areas of science [11] and has a number of real-world applications [12] including diffusive transport networks [13], but has not previously been applied to the airways. Some standard operators, such as the adjacency and Laplacian operators, are generic to all networks and can be adapted to characterise different processes or uncover important motifs in the network [14]. However, for lung airway networks and other resistive trees, these operators are not necessarily the optimal representation of the physical processes of interest on the network, as we show here.

In this paper we demonstrate that the resistive properties of realistic tree networks can be characterised by a tree-specific operator that we call the Maury matrix, first introduced in ref. [15]. This matrix provides a complete description of the linear resistance relations on the network. Furthermore, this operator provides a new method for dimensionality reduction in modelling these resistance trees, which we use to calculate gas flow and VH in the lung airway network. In section 2, we formulate the transport problem on an airway network in terms of the conductance Laplacian and the Maury matrix, and show how each can be approximated via spectral reduction. We show how VH is evaluated for a set of realistic airway network models (four of which are based on CT imaging). In section 3, we compare spectral approximations of the Laplacian and Maury operators, and show in particular how the reduced Maury operator efficiently captures spatial patterns of VH.

2 Methods

Unlike in several other examples of biological transport networks, mammalian airways develop robustly into tree networks [1], such that they contain no cycles and airways can be defined hierarchically by parent airways branching into children airways. Therefore, we model a tree network $\mathcal{N} = \{\mathcal{V}, \mathcal{E}\}$ consisting of a set of nodes (or vertices), $\mathcal{V} = \{v_i\}$, and a set of branches (or edges) representing airways, $\mathcal{E} = \{e_j = (v_i, v_{i'})\}$. The intrinsic hierarchy in tree networks means that we can consistently define each branch's orientation to point from its more proximal node v_i to the more distal node $v_{i'}$ based on its graph distance from the root node v_1.

The set of terminal nodes is denoted $\mathcal{T} \subset \mathcal{V}$ and we order the indices i so that the terminal nodes are at the end of the list of nodes $\mathcal{T} = \{v_{|\mathcal{V}|-|\mathcal{T}|+1}, \ldots, v_{|\mathcal{V}|}\}$, where $|.|$ indicates the number of objects in the set. The tree topology means that every branch e_j has a unique distal node $v_{i'}$ so we index the branches such that $j = i' - 1$. In the conducting airways of the lung, the root node represents the entrance from the upper airway, which is connected to the trachea that we label as branch e_1 (and so the carina is v_2). See figure 1(a) for an overview of this notation. The distribution of flow through the airway network in the lung is primarily dependent on the interplay between resistance and compliance of the airways and tissue. In this paper we focus on new ways to model and analyse the resistance of the airway network, and its implications

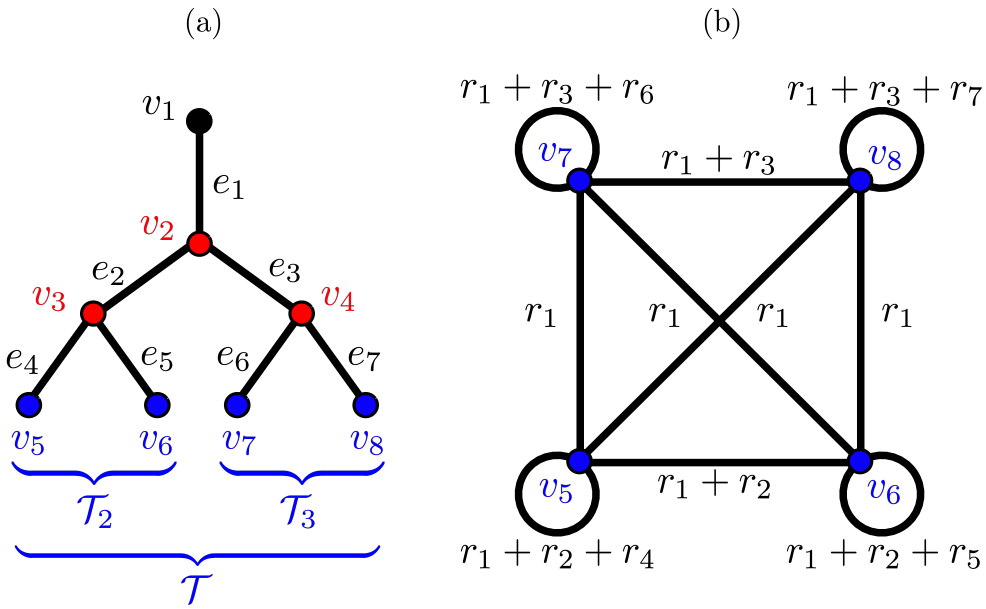

Figure 1: (a) Sketch of an airway tree network with nodes $\mathcal{V} = \{v_1, \ldots, v_8\}$ and branches $\mathcal{E} = \{e_1, \ldots, e_7\}$. The set of internal nodes $\mathcal{V}_{\text{int}} = \{v_2, v_3, v_4\}$ is coloured red and the set of terminal nodes $\mathcal{T} = \{v_5, \ldots, v_8\}$ is coloured blue. The sets $\mathcal{T}_2$ and $\mathcal{T}_3$ denote the subsets of $\mathcal{T}$ descended from branchs e_2 and e_3 respectively. (b) The network shown here has weighted adjacency matrix $\mathbf{R}$ where the nodes are the terminal nodes $\mathcal{T}$ from figure (a). The branch weights are labelled where r_j corresponds to the resistance of branch e_j in figure (a).

for the distribution of gas flow.

2.1 The Conductance Laplacian

The network incidence matrix $\mathbf{N} \in \mathbb{Z}^{|\mathcal{V}| \times |\mathcal{E}|}$ maps the nodes of the network to their associated branches. Its entries are

$$N_{i,j} = \begin{cases} -1 & \text{if} \quad e_j = \{v_{i'}, v_i\} \\ 1 & \text{if} \quad e_j = \{v_i, v_{i'}\} \\ 0 & \text{if} \quad v_i \notin e_j \,. \end{cases} \tag{1}$$

We model flow through the conducting airways of the lung by a linear resistance equation so that the branch flux is given by

$$\mathbf{q} = \text{diag}(\mathbf{r})^{-1} \mathbf{N}^T \mathbf{P} \tag{2}$$

where $\mathbf{q} \in \mathbb{R}^{|\mathcal{E}|}$ is the vector of branch fluxes, $\mathbf{r} \in \mathbb{R}^{|\mathcal{E}|}$ is the vector of branch resistances and $\mathbf{P} \in \mathbb{R}^{|\mathcal{V}|}$ is the vector of node pressures. The incidence matrix $\mathbf{N}$ is a discrete boundary operator, mapping 1-chains (branch quantities) to 0-chains (node quantities), and is the discrete analogue of the divergence operator in vector calculus [16]. Similarly, $\mathbf{N}^T$ is analogous to the gradient operator and performs the reverse mapping.

Applying $\mathbf{N}$ to (2) gives

$$\mathbf{L}\mathbf{P} = \mathbf{Q}, \tag{3}$$

where $\mathbf{L} = \mathbf{N}\mathrm{diag}(\mathbf{r})^{-1}\mathbf{N}^T$ is the network conductance Laplacian. The node flux $\mathbf{Q} = \mathbf{N}\mathbf{q}$ for each node v_i is equal to the sum of fluxes into node i from the branches connected to it. Assuming incompressibility, these are zero except on any nodes that are sources or sinks. In the lung, gas can only enter or exit the conducting airways via the upper airway or the transitional bronchioles (which feed the acini), so henceforth, we assume that the root v_1 and terminal nodes $v_i \in \mathcal{T}$ are the only sources or sinks. Therefore, the vectors of node and branch quantities can be written as

$$\mathbf{P} = \begin{pmatrix} P_1 \\ \mathbf{P}_{\mathrm{int}} \\ \mathbf{P}_{\mathrm{term}} \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} Q_1 \\ 0 \\ \mathbf{Q}_{\mathrm{term}} \end{pmatrix}, \quad \mathbf{q} = \begin{pmatrix} q_1 \\ \mathbf{q}_{\mathrm{int}} \\ \mathbf{q}_{\mathrm{term}} \end{pmatrix}, \quad \mathbf{r} = \begin{pmatrix} r_1 \\ \mathbf{r}_{\mathrm{int}} \\ \mathbf{r}_{\mathrm{term}} \end{pmatrix}. \tag{4}$$

Here, the separate blocks of $\mathbf{P}$ and $\mathbf{Q}$ distinguish the quantities defined on the root node v_1, internal nodes $\mathcal{V}_{\mathrm{int}} = \{v_2, \ldots, v_{|\mathcal{V}|-|\mathcal{T}|}\}$, and terminal nodes $\mathcal{T} = \{v_{|\mathcal{V}|-|\mathcal{T}|+1}, \ldots, v_{|\mathcal{V}|}\}$. Likewise, $\mathbf{q}$ and $\mathbf{r}$ are separated into quantities defined on the root branch e_1, the internal branches $\mathcal{E}_{\mathrm{int}} = \{e_2, \ldots, e_{|\mathcal{V}|-|\mathcal{T}|-1}\}$ and terminal branches $\mathcal{E}_{\mathrm{term}} = \{e_{|\mathcal{V}|-|\mathcal{T}|}, \ldots, e_{|\mathcal{V}|-1}\}$. In block form, the incidence and Laplacian operators are

$$\mathbf{N} = \begin{pmatrix} 1 & 0 & 0 \\ -\mathbf{a} & \mathbf{N}_{\mathrm{int}} & \mathbf{B} \\ 0 & 0 & -\mathbf{I} \end{pmatrix}, \quad \mathbf{L} = \begin{pmatrix} r_1^{-1} & -r_1^{-1}\mathbf{a}^T & 0 \\ -r_1^{-1}\mathbf{a} & \mathbf{L}_{\mathrm{int}} & -\mathbf{B}\boldsymbol{\Gamma}_{\mathrm{term}} \\ 0 & -\boldsymbol{\Gamma}_{\mathrm{term}}\mathbf{B}^T & \boldsymbol{\Gamma}_{\mathrm{term}} \end{pmatrix}. \tag{5}$$

Here $\mathbf{a} = (1, 0, \ldots, 0)^T$ and $\mathbf{I} \in \mathbb{R}^{|\mathcal{T}|\times|\mathcal{E}_{\mathrm{term}}|}$ is the identity matrix. The rows of the matrices $\mathbf{N}_{\mathrm{int}} \in \mathbb{R}^{|\mathcal{V}_{\mathrm{int}}|\times|\mathcal{E}_{\mathrm{int}}|}$ and $\mathbf{B} \in \mathbb{R}^{|\mathcal{V}_{\mathrm{int}}|\times|\mathcal{E}_{\mathrm{term}}|}$ indicate the connections of the internal nodes $v_i \in \mathcal{V}_{\mathrm{int}}$ to the internal branches $e_j \in \mathcal{E}_{\mathrm{int}}$ and the terminal branches $e_j \in \mathcal{E}_{\mathrm{term}}$ respectively. For brevity, we have denoted $\mathrm{diag}(\mathbf{r}_{\mathrm{term}})^{-1} \equiv \boldsymbol{\Gamma}_{\mathrm{term}}$. The internal Laplacian operator is given by $\mathbf{L}_{\mathrm{int}} = r_1^{-1}\mathbf{a}\mathbf{a}^T + \mathbf{N}_{\mathrm{int}}\mathrm{diag}(\mathbf{r}_{\mathrm{int}})^{-1}\mathbf{N}_{\mathrm{int}}^T + \mathbf{B}\boldsymbol{\Gamma}_{\mathrm{term}}\mathbf{B}^T$.

The relation $\sum_i N_{ij} = 0$ holds for all j, and so $\mathbf{N}^T\mathbf{e}_{|\mathcal{V}|} = 0$ where $\mathbf{e}_M \equiv (1, \ldots, 1)^T \in \mathbb{R}^M$ denotes the all-one vector of size M, and so $\mathbf{e}_{|\mathcal{V}|}$ is the vector identifying all nodes. Thus, $\mathbf{e}_{|\mathcal{V}|}$ is a zero-eigenvalue mode of $\mathbf{L}$, and multiplying (3) by $\mathbf{e}_{|\mathcal{V}|}^T$ gives the condition for global mass conservation $Q_1 = -\sum_{v_i\in\mathcal{T}} Q_i$. In fully connected networks, as considered here, this is the only zero-eigenvalue mode and so the conductance Laplacian $\mathbf{L}$ is rank $|\mathcal{V}| - 1$. Therefore, if $\mathbf{P}$ is a valid solution to (3) then so is $\mathbf{P} + c\,\mathbf{e}_{|\mathcal{V}|}$ for any $c \in \mathbb{R}$. It follows that $|\mathcal{T}| + 1$ boundary conditions are required to define the $|\mathcal{V}| + |\mathcal{T}|$ unknowns in (3) ($|\mathcal{V}|$ node pressures and $|\mathcal{T}|$ terminal fluxes).

In the context of lungs we are generally interested in pressure boundary conditions on P_1 and $\mathbf{P}_{\mathrm{term}}$. To remove the arbitrary pressure constant, we first reformulate in terms of the pressure drop relative to the entry node v_1 such that $\Delta\mathbf{P} = P_1\mathbf{e}_{|\mathcal{V}|} - \mathbf{P}$, and so $\Delta\mathbf{P} = (0, \Delta\mathbf{P}_{\mathrm{int}}, \Delta\mathbf{P}_{\mathrm{term}})^T$. Thus, given $\Delta\mathbf{P}_{\mathrm{term}}$ and inserting (4) into (3), the system of equations to be solved is

$$\mathbf{L}_{\mathrm{int}}\Delta\mathbf{P}_{\mathrm{int}} = \mathbf{B}\boldsymbol{\Gamma}_{\mathrm{term}}\Delta\mathbf{P}_{\mathrm{term}}, \tag{6}$$

$$\mathbf{Q}_{\mathrm{term}} = \boldsymbol{\Gamma}_{\mathrm{term}}\left(\mathbf{B}^T\Delta\mathbf{P}_{\mathrm{int}} - \Delta\mathbf{P}_{\mathrm{term}}\right). \tag{7}$$

Equation (6) can be solved efficiently for $\Delta\mathbf{P}_{\mathrm{int}}$ by numerical linear algebra methods due to the sparsity of the operator $\mathbf{L}_{\mathrm{int}}$. In general, for connected networks, $\mathbf{L}_{\mathrm{int}}$ is invertible and so substituting (6) into (7) gives

the terminal fluxes in terms of the pressure boundary conditions

$$\mathbf{Q}_{\text{term}} = \mathbf{\Gamma}_{\text{term}} \left[\mathbf{B}^T \mathbf{L}_{\text{int}}^{-1} \mathbf{B} \mathbf{\Gamma}_{\text{term}} - \mathbf{I} \right] \Delta \mathbf{P}_{\text{term}}. \tag{8}$$

In practice, we solve this system of equations by implementing the "SparseLU" method from the C++ library Eigen [29], based on LU factorisation, to solve (6) for $\Delta \mathbf{P}_{\text{int}}$ given boundary conditions on $\Delta \mathbf{P}_{\text{term}}$ and then substitute the solution for $\Delta \mathbf{P}_{\text{int}}$ into (7).

2.1.1 Associated spectra of the Conductance Laplacian

The solution to (6) can also be expressed in terms of the eigenbasis of the operator $\mathbf{L}_{\text{int}}$. In general the internal Laplacian $\mathbf{L}_{\text{int}}$ has $|\mathcal{V}| - |\mathcal{T}| - 1$ non-zero eigenvalues that we label $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_{|\mathcal{V}|-|\mathcal{T}|-1}$ with corresponding eigenvectors $\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_{|\mathcal{V}|-|\mathcal{T}|-1}$, where the hat indicates a normalised vector such that $\hat{\mathbf{u}}_k^T \hat{\mathbf{u}}_{k'} = \delta_{kk'}$ and δ is the Kronecker delta. Substituting the eigen-decomposition of $\mathbf{L}_{\text{int}}$ into (6) and solving for $\Delta \mathbf{P}_{\text{int}}$ gives

$$\Delta \mathbf{P}_{\text{int}} = \sum_{k=1}^{|\mathcal{V}|-|\mathcal{T}|-1} \left(\frac{\hat{\mathbf{u}}_k^T \mathbf{B} \mathbf{\Gamma}_{\text{term}} \Delta \mathbf{P}_{\text{term}}}{\lambda_k} \right) \hat{\mathbf{u}}_k. \tag{9}$$

Substituting this solution into (7) gives the corresponding solution for $\mathbf{Q}_{\text{term}}$.

Dimensionality reduction can in principle be achieved using spectral methods by reconstructing the system with a subset of the eigenmodes. We can measure the relative dominance of each mode in the decomposition in (9) by the size of the coefficient $|p_k| = \left| \hat{\mathbf{u}}_k^T \mathbf{B} \mathbf{\Gamma}_{\text{term}} \Delta \mathbf{P}_{\text{term}} / \lambda_k \right|$. However, the relative weight of these modes is dependent on the boundary condition $\Delta \mathbf{P}_{\text{term}}$. A more general measure of how well a subset of the eigenmodes captures the behaviour of the whole operator, independent of boundary conditions, is to consider how closely these modes approximate the inverse of the operator, $\mathbf{L}_{\text{int}}$, which we measure by the normalised distance δ_L [11] given by

$$\delta_L(M) = \frac{||\mathbf{L}_{\text{int}}^{-1} - \sum_{k=1}^{M} \lambda_k^{-1} \hat{\mathbf{u}}_k \hat{\mathbf{u}}_k^T||_F}{||\mathbf{L}_{\text{int}}^{-1}||_F} = \sqrt{1 - \left(\sum_{k=1}^{M} \lambda_k^{-2} \right) \Big/ \left(\sum_{k=1}^{|\mathcal{V}|-|\mathcal{T}|-1} \lambda_k^{-2} \right)}, \tag{10}$$

where $||.||_F$ indicates the Frobenius norm and $M \in \{1, \ldots, |\mathcal{V}| - |\mathcal{T}| - 1\}$ is the number of modes used in the approximation. Evidently, the modes with the smallest eigenvalues contribute the most to this reconstruction, however depending on the boundary conditions used they may not be the most dominant in the reconstruction of the solutions $\Delta \mathbf{P}_{\text{int}}$ and $\mathbf{Q}_{\text{term}}$.

2.2 The Maury matrix

An alternative representation of the resistive airway tree network is presented in [15] for the case of a symmetric dyadic tree. Here, we generalise this operator to all connected trees and name it the Maury matrix in reference to [15]. To derive this operator, we begin by noting that connected sub-trees of the airway network can be defined due to the hierarchical structure of trees. We define the matrix $\mathbf{S} = (\mathbf{s}_1, \ldots, s_{|\mathcal{E}|})$ where each column $\mathbf{s}_j$ maps the branch e_j to a subset of nodes $\mathcal{V}_j \subset \{v_2, \ldots, v_{|\mathcal{V}|}\}$ descended from it such

that

$$S_{ij} = \begin{cases} 1 & \text{if } v_{i+1} \in \mathcal{V}_j \\ 0 & \text{if } v_{i+1} \notin \mathcal{V}_j \end{cases}. \tag{11}$$

The first column of $\mathbf{S}$ is $\mathbf{s}_1 = \mathbf{e}_{|\mathcal{V}|-1}$ as all nodes except v_1 are descended from the root branch. The column vectors $\mathbf{s}_j$ satisfy the relation $\mathbf{s}_j = \mathbf{i}_j + \sum_{n'=1}^{n} \mathbf{s}_{j_{n'}}$ where $\{e_{j_1}, \ldots, e_{j_n}\}$ are the n child branches of e_j, and $\mathbf{i}_j$ is the jth column of the identity matrix $\mathbf{I} \in \mathbb{R}^{(|\mathcal{V}|-1)\times(|\mathcal{V}|-1)}$. This relation follows from noting that each sub-tree $\mathcal{E}_j$ is a union of its child subtrees and its connecting node, such that $\mathcal{V}_j = \{v_{j+1}, \mathcal{V}_{j_1}, \ldots, \mathcal{V}_{j_n}\}$. It follows that $\mathbf{S}\mathbf{N}^T = (\mathbf{e}_{|\mathcal{V}|-1}, -\mathbf{I})$ and therefore

$$\mathbf{N} = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ & -(\mathbf{S}^T)^{-1} & & \end{pmatrix}, \tag{12}$$

Substituting (12) into $\mathbf{N}\mathbf{q} = \mathbf{Q}$ gives $\mathbf{q} = -\mathbf{S}^T(\mathbf{0}, \mathbf{Q}_{\text{term}})^T \equiv \mathbf{T}^T\mathbf{Q}_{\text{term}}$, where the rows of $\mathbf{T} \in \mathbb{R}^{|\mathcal{T}|\times|\mathcal{E}|}$ are comprised of the final $|\mathcal{T}|$ rows of the matrix $\mathbf{S}$. Substituting this into (2) and multiplying both sides by $-\mathbf{S}\,\text{diag}(\mathbf{r})$ gives

$$-\begin{pmatrix} \Delta\mathbf{P}_{\text{int}} \\ \Delta\mathbf{P}_{\text{term}} \end{pmatrix} = \mathbf{S}\,\text{diag}(\mathbf{r})\mathbf{S}^T \begin{pmatrix} 0 \\ \mathbf{Q}_{\text{term}} \end{pmatrix}. \tag{13}$$

Comparing (13) to (3) we see that $\mathbf{S}^T\text{diag}(\mathbf{r})\mathbf{S} \equiv \mathbf{L}_0^{-1}$ where $\mathbf{L}_0^{-1}$ is the conductance Laplacian $\mathbf{L}$ with the first row and column removed (or the reduced Laplacian, as features in the matrix-tree theorem [17]). Finally, the bottom $|\mathcal{T}|$ rows of (13) directly relate the terminal pressure drops to the fluxes

$$\mathbf{R}\mathbf{Q}_{\text{term}} = -\Delta\mathbf{P}_{\text{term}}, \tag{14}$$

where $\mathbf{R} = \mathbf{T}\text{diag}(\mathbf{r})\mathbf{T}^T$ is the aforementioned Maury operator. Comparing with (8) gives

$$\mathbf{R} = \left[\mathbf{\Gamma}_{\text{term}}\left(\mathbf{B}^T\mathbf{L}_{\text{int}}^{-1}\mathbf{B}\mathbf{\Gamma}_{\text{term}} - \mathbf{I}\right)\right]^{-1}. \tag{15}$$

The entries of the terminal node map T_{ij} are 1 if the terminal node $v_{1+|\mathcal{V}|_{\text{int}}+i}$ is descended from branch e_j and zero otherwise (see (11)). We denote the set of subset of terminal nodes descended from e_j as $\mathcal{T}_j \subset \mathcal{T}$ (see figure 1(a)). The Maury matrix is symmetric and in index notation its components are

$$R_{i,i'} = \sum_{e_j \in \mathcal{P}_{1\to i} \cap \mathcal{P}_{1\to i'}} r_j, \tag{16}$$

where $\mathcal{P}_{1\to i} = \{e_{j_1}, e_{j_2}, \ldots, e_{j_N}\}$ is the unique direct path on the tree from v_1 to a terminal node $v_i \in \mathcal{T}$. Each term $R_{i,i'}$ is equivalent to the resistance distance [18] between the node v_1 and node $v_{i_{LCA}}$ such that $v_{i_{LCA}}$ is the lowest common ancestor (LCA) node of the terminal nodes v_i and $v_{i'}$.

A direct solve of (14) is an inefficient method for solving the system of equations (compared to (6) and (7)) because the matrix $\mathbf{R}$ has, in general, $|\mathcal{T}|^2$ non-zero entries and so cannot benefit from any algorithms optimised for the factorisation of sparse matrices. Nonetheless, this does not inhibit computation of the spectra of $\mathbf{R}$ because we can decompose $\mathbf{R}$ into sparse, non-square matrices. Details of the numerical methods used are given in section 1.3 of the Supplementary Text.

2.2.1 Spectral properties of the Maury matrix

Diagonalising the Maury matrix $\mathbf{R}$ into its normalised eigenbasis and substituting into (14) gives

$$\left(\mathbf{Q}_{\text{term}}^T \hat{\mathbf{v}}_k\right) \mu_k = -\Delta \mathbf{P}_{\text{term}}^T \hat{\mathbf{v}}_k, \tag{17}$$

for each eigenvalue μ_k and associated eigenvector $\hat{\mathbf{v}}_k$. The matrix $\mathbf{R}$ has the form of an adjacency matrix for a complete network of resistors (with self-loops) relating the terminal nodes (as shown figure 1(b)). Therefore, as shown in (17), the modes act as resistors in parallel with resistance μ_k, each subject to the pressure drop $\Delta \mathbf{P}_{\text{term}} \cdot \hat{\mathbf{v}}_k$. As it is a weighted adjacency matrix, the matrix $\mathbf{R}$ is non-negative, regular and symmetric and so by the Perron–Frobenius theorem its largest eigenvalue mode μ_1 is unique and is the only eigenvector $\hat{\mathbf{v}}_1$ in the spectrum of $\mathbf{R}$ that can be expressed as an all-positive vector. Additionally, this implies that the eigenvalues of the Maury matrix are all positive, and so we order the indices by $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{|\mathcal{T}|}$ (the opposite ordering to modes of the Laplacian).

From (17) the solution for the terminal node fluxes is

$$\mathbf{Q}_{\text{term}} = -\sum_{k=1}^{|\mathcal{T}|} \left(\frac{\Delta \mathbf{P}_{\text{term}}^T \hat{\mathbf{v}}_k}{\mu_k} \right) \hat{\mathbf{v}}_k \tag{18}$$

and so an approximation to $\mathbf{Q}_{\text{term}}^{(\text{approx})}$ is given by limiting the sum in (18) to some subset of the mode spectrum. The magnitude of the coefficient $|q_k| = \left|\Delta \mathbf{P}_{\text{term}}^T \hat{\mathbf{v}}_k / \mu_k\right|$ quantifies the relative importance of each term in the expansion of (18). As for the internal Laplacian operator, we measure the accuracy of such an approximation for arbitrary boundary conditions by the convergence of the inverse of $\mathbf{R}$ such that

$$\delta_R(M) \equiv \sqrt{1 - \left(\sum_{k=|\mathcal{T}|-M+1}^{|\mathcal{T}|} \mu_k^{-2} \right) \Big/ \left(\sum_{k=|\mathcal{T}|-M+1}^{|\mathcal{T}|} \mu_k^{-2} \right)}, \tag{19}$$

where M is the number of (smallest eigenvalue) modes used in the expansion. We will compare this with $\delta_L(M)$ for specific networks to compare the general efficiency of dimensionality reduction for these two operators.

2.3 Dimensionality reduction in simulations of ventilation in the lungs

In models of ventilation in the lungs, the effects of lung tissue compliance generally dominate the mechanics compared to airway resistance, except in cases where there is airway narrowing or blockage [19]. Consider the simplest model of ventilation where each terminal node $v_i \in \mathcal{T}$ is connected to an elastic bag of volume $V_i(t)$ and elastance κ with all bags subject to the same pleural pressure $P_{\text{pl}}(t)$ driving the breathing motion at time t. The terminal fluxes are related to the bag volumes via $\mathbf{Q}_{\text{term}}(t) = -\dot{\mathbf{V}}(t)$ and the terminal pressures are $\mathbf{P}_{\text{term}} = P_{\text{pl}}(t)\mathbf{e}_{|\mathcal{T}|} + \kappa \mathbf{V}(t)$. Then (14) becomes

$$\mathbf{R}\dot{\mathbf{V}}(t) + \kappa \mathbf{V}(t) = -P_{\text{pl}}(t)\mathbf{e}_{|\mathcal{T}|}. \tag{20}$$

We are interested in characterising how the distribution of ventilation is affected by the airway resistance, and so have assumed that the lung unit compliances are all equal to κ. Diagonalising $\mathbf{R}$ into its eigenbasis

in (20) gives

$$\mu_k \dot{a}_k(t) + \kappa a_k(t) = -P_{\mathrm{pl}}(t)\, \mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k, \tag{21}$$

where $\mathbf{V}(t) = \sum_k a_k(t)\hat{\mathbf{v}}_k$. Thus each mode of $\mathbf{R}$ is characterised by an independent ODE with a resistance term μ_k and elastance κ driven by a pressure term proportional to $\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k$. Solving (21) for $a_k(t)$ gives

$$\mathbf{V}(t) = -\frac{P_{\mathrm{pl}}(t)}{\kappa}\mathbf{e}_{|\mathcal{T}|} + \sum_k \left[\frac{\left(\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k\right)}{\kappa} e^{-\frac{\kappa}{\mu_k}t} \int_0^t \dot{P}_{\mathrm{pl}}(t') e^{\frac{\kappa}{\mu_k}t'} \mathrm{d}t' \right] \hat{\mathbf{v}}_k. \tag{22}$$

The first term in the solution is the elastic contribution from all units, while the second term is the resistive contribution. The sum in the resistive term can be approximated by the largest eigenvalue modes with a cutoff on the order $\mu_k/(\kappa\tau) \sim O(1)$ (where τ is the typical breath timescale), since modes satisfying $\mu_k/(\kappa\tau) \ll 1$, will be dominated by lung unit compliance (the first term in (22)).

In the simulations used in this paper we take the pleural pressure to be sinusoidal such that $P_{\mathrm{pl}} = P_{\mathrm{pl0}} + P_{\mathrm{s}} \sin(2\pi t/\tau)$. Substituting into (22) we see that the periodic solution (when $t \gg \mu_1/\kappa$) is

$$\begin{aligned} \mathbf{V}(t) = &-\frac{1}{\kappa}\left(P_{\mathrm{pl0}} + P_s \sin\frac{2\pi t}{\tau}\right)\mathbf{e}_{|\mathcal{T}|} \\ &+ P_s \sum_k \frac{\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k}{\kappa}\left[\frac{2\pi\mu_k/(\kappa\tau)}{1+(2\pi\mu_k/(\kappa\tau))^2}\right]\left(\cos\frac{2\pi t}{\tau} + \frac{2\pi\mu_k}{\kappa\tau}\sin\frac{2\pi t}{\tau}\right)\hat{\mathbf{v}}_k. \end{aligned} \tag{23}$$

We define the ventilation of each terminal unit as $\Delta V_i = \max(V_i(t)) - \min(V_i(t))$, where the maximum and minimum are taken over a single breath cycle $0 \le t < \tau$, which gives

$$\Delta V_i = 2P_s \left[\left(\frac{1}{\kappa_i} - \sum_k \frac{\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k}{\kappa}\left[\frac{(2\pi\mu_k/(\kappa\tau))^2}{1+(2\pi\mu_k/(\kappa\tau))^2}\right]\hat{v}_{k,i}\right)^2 + \left(\sum_k \frac{\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k}{\kappa}\left[\frac{2\pi\mu_k/(\kappa\tau)}{1+(2\pi\mu_k/(\kappa\tau))^2}\right]\hat{v}_{k,i}\right)^2 \right]^{1/2}. \tag{24}$$

Inefficient ventilation of the lungs can occur if there are blockages or narrowing of the airway lumen, which can be caused by numerous pathophysiological factors, resulting in increased resistance of those airways and inhomogeneous delivery of gas to the lung units, commonly known as ventilation heterogeneity (VH). To characterise this in simulations we measure ventilation heterogeneity via the coefficient of variation of the ventilation σ_V across all terminal units,

$$\sigma_V = \frac{\sqrt{\sum_i (\Delta V_i - \overline{\Delta V})^2}}{\overline{\Delta V}}, \quad \text{where} \quad \overline{\Delta V} = \frac{1}{|\mathcal{T}|}\sum_i \Delta V_i\,. \tag{25}$$

Therefore the above decomposition gives a method to approximate the ventilation efficiency of an airway network given κ and τ by truncating the sums in (24) to $M < |\mathcal{T}|$ modes of the Maury matrix. While this method is specifically derived for sinusoidal breathing, it provides a useful qualitative measure of relative efficiency that can be used for comparing different networks and is easily computed using a subset of the full spectrum of the Maury operator.

2.4 Airway Models

In this paper we use idealised models of airway geometry and models derived from CT images, which we now briefly describe.

First, we adopt the Weibel-type model used in [20, 21] to systematically test the effect of asymmetry in resistance on the spectral properties of the Maury and Laplace operators. This idealised model describes a dyadic tree where each airway branches into major and minor daughters such that

$$\frac{a_{\text{maj}}}{a_{\text{par}}} = \frac{l_{\text{maj}}}{l_{\text{par}}} = \left(\frac{1+A}{2}\right)^{\frac{1}{3}}, \qquad \frac{a_{\text{min}}}{a_{\text{par}}} = \frac{l_{\text{min}}}{l_{\text{par}}} = \left(\frac{1-A}{2}\right)^{\frac{1}{3}}, \tag{26}$$

where a and l are the radii and lengths respectively of the parent and (major and minor) daughter branches as labelled by the subscripts. The parameter $0 \leq A < 1$ quantifies the asymmetry in branching ($A = (1-2r)$ in terms of the notation in [20]). The cube root results in the two daughter airways having the same combined effective Poiseuille resistance and volume as the symmetric ($A = 0$) case. In this model we assume a length-to-diameter ratio of 3 for all airways and a fixed number of divisions N between the root branch and the terminal branches (resulting in $N + 1$ generations of airways).

In real lungs, branches terminate in a wide range of generations. To characterise this effect we use the Horsfield model 1 [22] derived from cast measurements of an adult male, with airways isotropically scaled to 60% of their original volume to a lung with $\sim$ 3L functional residual capacity (approximately that of the average adult). The model is terminated at Horsfield order 4 (as defined in [22]) because this corresponds to the transitional bronchioles as defined by Weibel [1] and gives $29,240$ terminal airways, approximately the number of acini in the lung [23]. In this model, parent airways branch into daughters that are in different 'Horsfield orders' (generations counting up from the terminal nodes) and thus larger branches (with smaller resistance) give rise to more terminal nodes. The radii and lengths are Gaussian random variables with mean taken from the data in [22] and a standard deviation of either 10 or 20% of the mean to approximate the variation in healthy lungs, which we label H10 and H20 respectively.

Finally, we also use four image-based (IB) models derived from computed tomography (CT) images. These use volume-filling branching algorithms [24, 25] to populate the distal generations of airways in each lobe of the lung, which cannot be resolved from CT. The radius of these distal airways is assumed from an exponential fit to lung cast data, as in [25], such that

$$\log D(z) = (z - N)\log(R_d H) + \log D_N, \tag{27}$$

where $D(z)$ is the airway diameter, z is the Horsfield order of the airway, N is the Horsfield order of the most distal ancestor airway taken from CT and its diameter D_N. The parameter $R_d H$ is taken to be 1.15 as in [25]. The four IB models used here (labelled IB1-4) are from CT scans of children with cystic fibrosis and normal lung function (measured by spirometry) from a previously published study [26] (IB2 is shown in figures 4 and 5 below). The number of airways successfully acquired from CT depends on the image resolution, and was approximately between 3-5 generations of bifurcations from the trachea, and was acquired using the open-source software Pulmonary Toolkit [27]. The algorithm to generate these models used the C++ library `stl_reader` [28] and is detailed in section 1.2 of the Supplementary Text.

As a simple first approximation, the branch resistance is assumed to follow Poiseuille's law and so $r = 8\pi\nu l/a^4$ where ν is the viscosity of air. To compute the spectra of the operators we use the Eigen [29] and Spectra [30] libraries in C++. The source codes of the algorithms used are available in online repositories

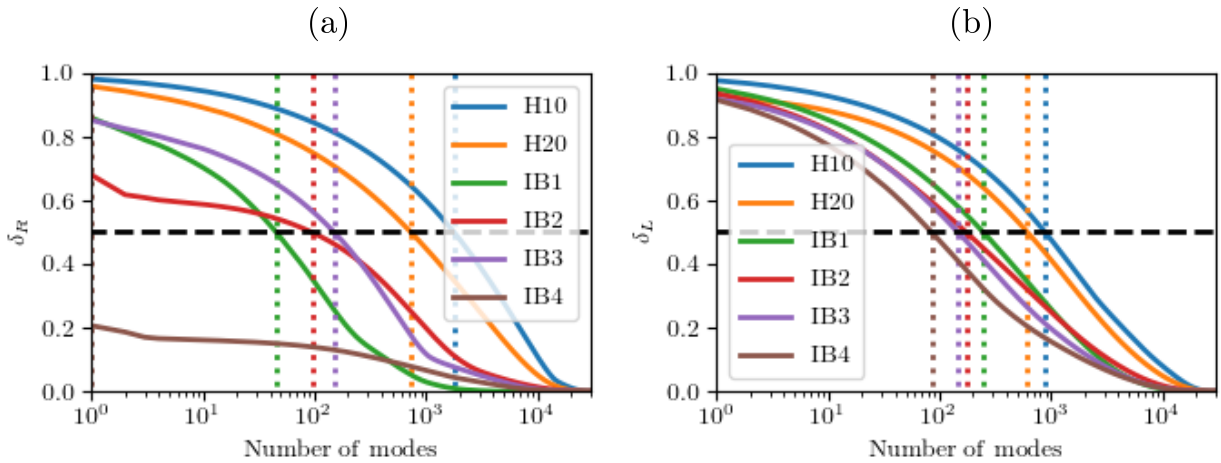

Figure 2: Normalised discrepancy in the cumulative estimate of the inverse of the (a) Maury operator and (b) internal Laplacian operator using the number of eigenmodes indicated on the x-axis where modes are ordered from smallest to largest eigenvalues. These are computed from the sums in (19) and (10) respectively. The networks H10 and H20 are the Horsfield model 1 with independent Gaussian random variables with standard deviations of 10% and 20% respectively added to the original airway diameter and length. Models IB1 through to IB4 are image-based trees derived from CT data as outlined in section 2.4. The vertical dashed lines indicate the point where each curve passes 50% accuracy.

[31, 32].

3 Results

In the following section we first study the spectra of the Maury operator for a number of idealised and image-based airway models and compare with the spectra of the internal Laplacian operator. Then, in the final results section we demonstrate how (24) can be used to produce a low-dimensional approximation of simulations and characterise realistic lung networks.

3.1 Comparison of Maury and internal Laplacian operator mode structures

The Maury and internal Laplacian operators encode similar information regarding the resistance of the network, as shown by the relation in (15). However, they represent the resistance relations differently, and thus the eigenvectors of the two operators relate to different properties of the tree (see (9)) and (18)). Therefore, we wish to know which spectrum is more efficient for realistic lung airway networks.

Figure 2 shows the relative accuracy of the inverse approximation for each operator (from (19) and (10)) plotted against the number of eigenmodes used in the approximation. The Maury decomposition is more variable between networks, and in particular performs better (i.e. a smaller number of modes are required to capture more of the behaviour) for the image-based airway networks. In particular, the network labelled IB4 only requires a single mode to approximate 80% of the Maury inverse. This mode has a small eigenvalue and hence represents a "short-circuit" that, given appropriate boundary conditions, may account for the majority of the flow on the network. In contrast, the Frobenius norm of the inverse Laplace operator is largely comprised of high-resistance modes, and so the corresponding eigenvectors are centralised on high-resistance motifs in the network. The mode reconstruction is more accurate in the image-based networks

than the Horsfield networks for either operator, which is due to the increased variance of resistance in these networks, however the effect is stronger for the Maury operator.

3.1.1 Characterisation of network resistance with fixed pressure boundary conditions

We now consider fixed boundary conditions for the pressure-drop $\Delta P_{\text{term}} = \mathbf{e}_{|\mathcal{T}|}$, such that flow is directed from the mouth (represented by the root node v_1) to the alveolar region (represented by the terminal nodes $v_i \in \mathcal{T}$). Therefore, any flow heterogeneity that results is due to the distribution of resistance in the networks. To comprehensively cover a range of possible networks, we apply these conditions to the Weibel-type model described by (26) with $N = 9$ divisions (512 terminal nodes), representative of approximately 1/64th of the conducting airways [1].

The mode weights $|q_k|$ and $|p_k|$ (see (9) and (18)) of the spectra of the Maury and internal Laplacian operators respectively rank the modes in terms of their relative importance to the solution flow and pressure solutions respectively. As the network asymmetry parameter A is changed from $A = 0$ (symmetric) to $A = 0.98$ (all-but-one path having very large resistance), the dominant vectors of the Maury matrix transition from highest resistance modes (smallest index) to the lowest resistance modes (largest index), see figure 3(a). This characterises the transition from a flow solution that is homogeneously spread over all the terminal nodes and one that is concentrated on a small collection of terminal nodes that terminate the low-resistance paths in the tree. The $A = 0$ case is considered in more detail in section 1.1 and 1.2 of the Supplementary Text.

For all values of A there is a small subset of modes in the Maury spectrum that dominate the flow solution (figure 3(a)). In contrast, the Laplace spectrum does not demonstrate any strong ordering of mode dominance for the majority of A values and clear separation of modes occurs only at small values of A (figure 3(b)). For larger A values the Maury spectrum outperforms the Laplace spectrum in terms of the number of modes required to reconstruct the solutions (figure 3(c)). However, the difference is small in human lungs, where it is estimated that $A \approx 0.35$ [20]. The Maury matrix performs best at extreme values of A, but for all A values a 75% accuracy can be achieved using at most 10 eigenmodes ($\sim$2% of the spectrum).

3.2 Practical application to simulations of lung ventilation mechanics

As shown in section 2.3, a simple mechanical model of regular tidal breathing can be generated by including a dynamic pressure boundary condition at the terminal nodes that takes into account the elastic recoil of the acini. Then, it is the largest eigenvalue modes of the Maury matrix that most significantly impact on the ventilation profile, and the distribution of ventilation is well approximated by a subset of these modes using the expansion in (24). The largest eigenvalue modes correspond to high resistance paths in the tree, as demonstrated systematically in section 2.1 of the Supplementary Text, and so are also useful for characterising complex models of airway structure.

3.2.1 Dimensionality reduction in simulations and characterisation of realistic airway networks

First, we consider the Horsfield model 1 networks H10 and H20, which provide a realistic model of the asymmetric branching airway network. Choosing parameters for elastance κ and time-scale τ that are representative of an average adult (listed in figure 4), all eigenvalues of the Maury matrix satisfied $\mu_k/(\kappa\tau) < 0.1$. Therefore, the ventilation is approximately homogeneous in these networks and the airway resistance has little effect on ventilation.

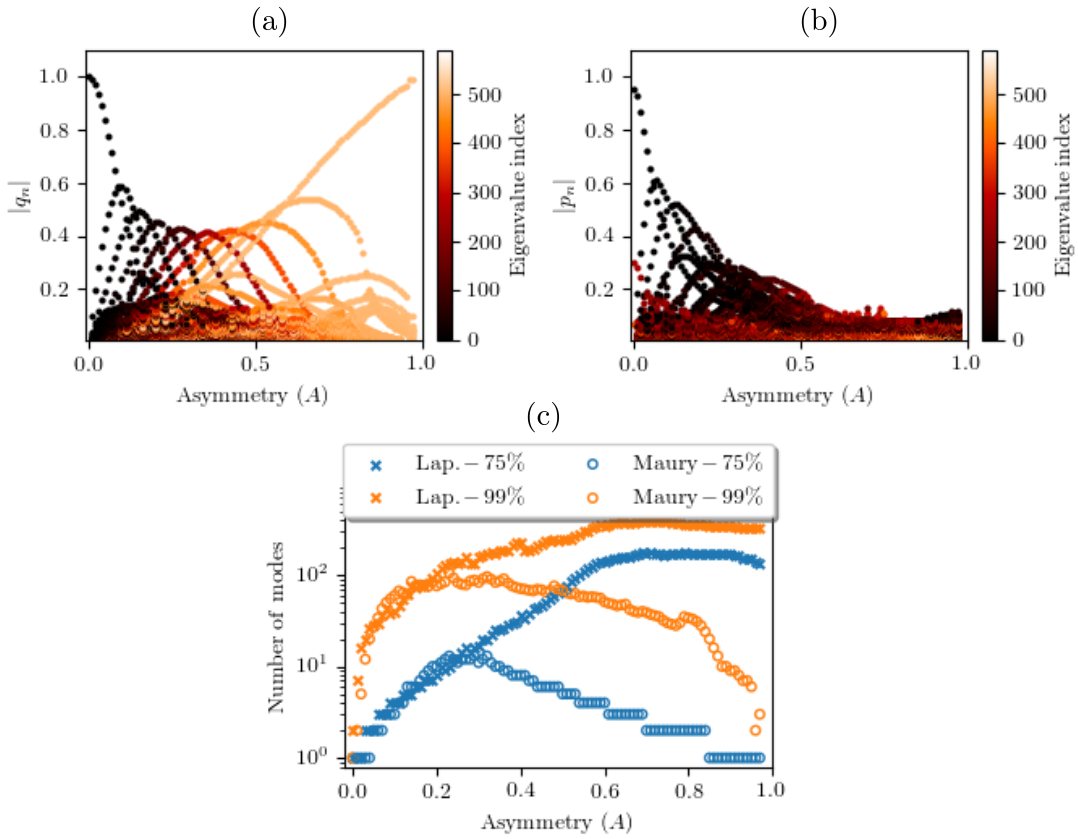

Figure 3: Relative magnitude of contributions of each eigenmode to the sums in equations (a) (18) and (b) (9) of a 10-generation Weibel network versus the A value of that network. Each dot represents the relative contribution of one of the 512 modes for the tree corresponding to that particular value of A, coloured by mode index; (a) Maury matrix (eigenvalues ordered largest to smallest) and (b) internal Laplacian (eigenvalues ordered smallest to largest). (c) The minimum number of modes required to achieve particular approximations the sums in (18) and (9) for the Maury and internal Laplace operators respectively.

To generate networks with high-resistance paths, we systematically applied constrictions to the Horsfield network by selecting a fraction f of the airways (at random) in a given Horsfield order g and reducing their radius by a factor between 50% and 95% (drawn from a uniform random distribution between these limits). The networks generated are idealised examples of the effects of obstructive lung disease, with different networks for disease centred on proximal ($g = 15$), central ($g = 10$), or distal ($g = 5$) airways. In the constricted networks, we found that as the number of random constrictions is increased (either by increasing f or reducing g), the number of independent large-eigenvalue modes of the Maury matrix rises (figure 4(a)). This agrees with our interpretation of these modes being centred on high-resistance paths in the tree.

The VH parameter σ_V from (25) does not depend strongly on g in this airway model, and so its variation in figure 4(b) is mainly due to the range of f values used (larger f results in more constricted airways and so greater VH). Regardless of f and g values, we see that using only large-eigenvalue modes in (25) gives a very good approximation of the simulated VH for all values of f and g. These reconstructions use at most 1% of the full Maury spectrum, and so represent a significant reduction in dimensionality.

Figure 4(c) shows the effect on VH of changing κ in image-based models of ventilation and compares it to the prediction made using spectral decomposition. For each network we chose a range of realistic values

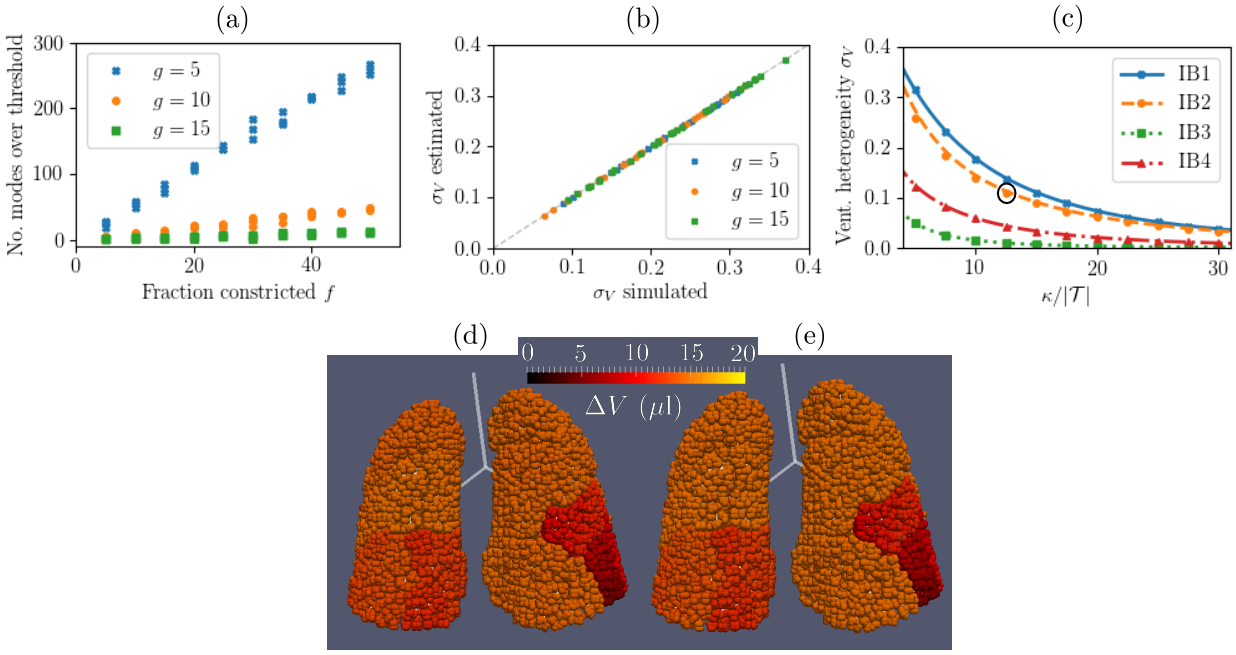

Figure 4: (a) Number of large eigenvalues (defined as $\mu_k > 0.1\kappa\tau$) of the Maury matrix plotted against the fraction f of a given Horsfield order g constricted for the H10 airway tree network with added constrictions. The legend indicates which Horsfield order g constrictions were applied to for each tree. The total number of modes is 29,240. (b) Comparison of VH parameter σ_V for all networks in (a) computed through full simulation (x-axis) and estimated using just the largest-eigenvalue modes of the Maury matrix (y-axis) as identified in figure (a). All of the calculations in figures (a) and (b) used $\kappa/|\mathcal{T}| = 5$cm H_20 L^{-1} and $\tau = 4$s. (c) Simulated σ_V versus $\kappa/|\mathcal{T}|$ for the networks IB1-4 as labelled. Markers indicate results from numerical integration of (6) and (7) while lines indicate approximations using modes corresponding to the largest 1% of the eigenvalues in the Maury spectrum. (d) Visualisation of $\Delta\mathbf{V}$ in the simulation of IB2 with $\kappa/|\mathcal{T}| = 12.5\text{cmH}_2\text{0L}^{-1}$ (circled in (c)). (e) The prediction of ΔV for the same simulation using the 1% of the modes (corresponding to the largest 299 eigenvalues of the Maury matrix). See Supplementary Video 1 for a full 360° view of (d) and (e).

of the total lung elastance $\kappa/|\mathcal{T}|$ for children from 5-30cm H_20 L^{-1} [33, 34, 35]. As κ is reduced, airway resistance has a more significant contribution to the dynamics in (20), and so VH is increased due to the heterogeneity of network resistance. The variation between the networks is characteristic of the different network structures, with IB1 and IB2 showing significantly more VH than IB3 and IB4.

Based on our findings in figure 4(a) and (b), we calculated only the largest 1% of eigenvalues for each network and used these to estimate σ_V in figure 4(c), making comparison with full simulations. The approximations perform marginally worse for smaller values of κ, however overall the agreement is very strong. Since the networks remain the same for different κ values, once the modes had been calculated (which in this case takes ∼20s on a single processor) parameter sweeps over κ or τ were performed analytically and so did not require repeated simulation (each of which also takes ∼20s on the same single processor when using a sinusoidal pleural pressure, but on the order of minutes for non-linear and non-analytic pleural-pressure). Therefore, we were able to predict the relationship between VH and lung elastance accurately and more efficiently using this method.

An explicit example of this reduced dimensionality approximation is given in figures 4(d) and 4(e) for the network IB2, corresponding to the highlighted simulation in 4(c), which shows distributions of ventilation

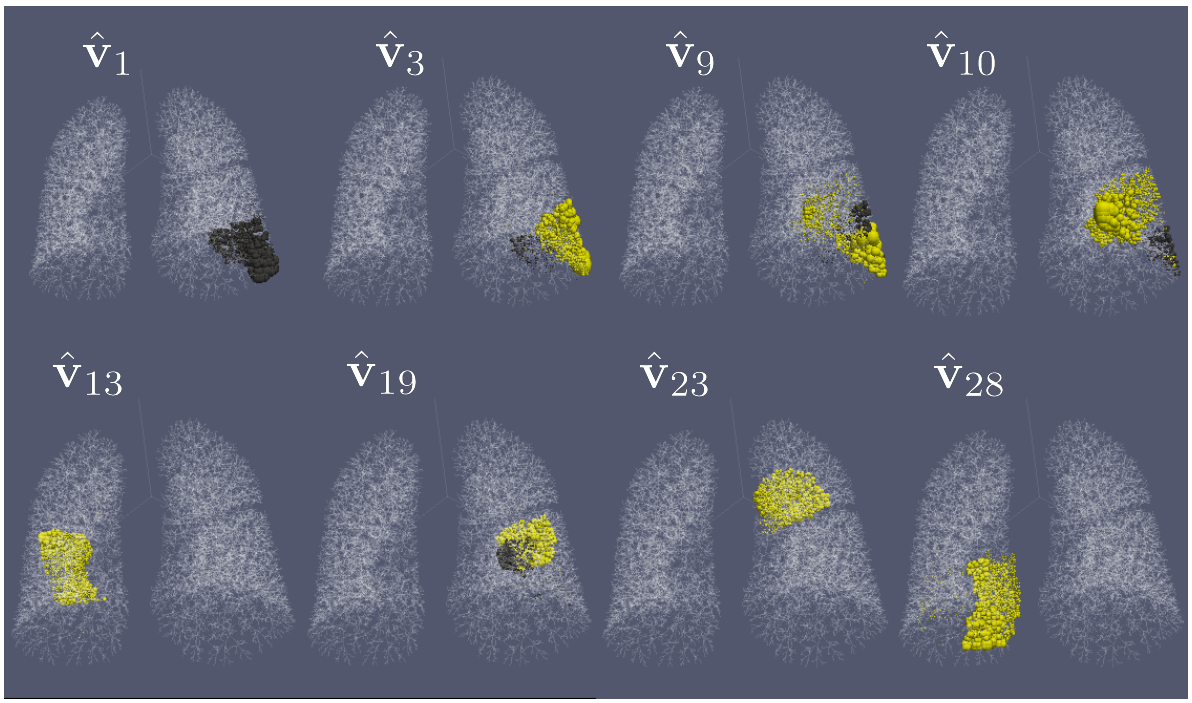

Figure 5: Visualisation of a selection of the eigenvectors corresponding to the largest eigenvalues of the Maury matrix for the IB2 network. The eigenvectors plotted are those with the largest values of $\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k$ from the subset of modes that satisfy $\mu_k > 0.1\kappa\tau$ for the simulation values used in figure 4(d). In each plot, sphere size is proportional to the magnitude of eigenvector entries on all terminal nodes for the eigenvector labelled. The locations of the eigenvector entries highlight regions of impaired ventilation (according to the model) with the larger eigenvalue modes showing the least well-ventilated regions. The colours represent eigenvector entries with opposite sign. See Supplementary Video 2 for a full 360° view of this figure.

ΔV_i (see (24)) over terminal units computed by full simulation (figure 4(d)) and the reduced model (figure 4(e)). The heterogeneity in ventilation is a result of the airway geometry and topology generated by the space-filling branching algorithm, as well as the branch radii which scales with its parent airway from the CT image, and is captured very efficiently by the reduced model.

3.2.2 Characterisation of realistic airway networks

Large-eigenvalue (high-resistance) eigenvectors of the Maury matrix are centralised on those terminal nodes most affected. Therefore, the eigen-decomposition of the Maury matrix is a powerful tool for visualising and ranking the least ventilated regions, regardless of the chosen value of lung elastance κ. Figure 5 visualises a subset of the largest eigenvalue eigenvectors of the IB2 network by the size of spheres superimposed onto its terminal nodes. Comparing this to figure 4(d) we see that these are localised on the least ventilated terminal nodes. Furthermore, the largest eigenvalue modes highlight the regions of poorest ventilation, as the eigenvalue is a measure of the resistance of the mode. The modes in figure 5 are selected by their value of $\mathbf{e}_{|\mathcal{T}|}^T \hat{\mathbf{v}}_k$, which is the coefficient in the mode sum in (24), so that they are all important in this reconstruction (for small enough $\kappa\tau$).

4 Discussion

In section 3.1 we showed that for a range of lung airway networks the spectrum of the Maury operator is a more efficient tool than the Laplacian operator for dimensionality reduction in the calculation of flow through the network. Under fixed-pressure boundary conditions, the dominant flow eigenmodes can occur throughout the spectrum, and depend on the resistance structure of the network. Therefore, it is not possible *a priori* to calculate only the most dominant modes without computing the whole spectrum. However, when we extend the Maury matrix construction to simulations of lung ventilation (including elastic recoil of the acini), it is only the largest eigenvalue modes that significantly contribute to the distribution of gas delivery and hence the extent of VH. These modes are generally the most efficient to compute using methods based on power iteration, as used here. Therefore, this provides a method for efficient prediction and characterisation of poorly ventilated regions of the lung based on the geometry and topology of the airway tree. The method has significant advantages over previous approaches of dimensionality reduction in this area, which generally involve replacing parts of the tree with symmetric models [36, 37] based on "trumpet" models [38] of the airway geometry. These methods remove the asymmetry of the smaller airways, and sacrifice some of the complexity of the system to reduce dimensionality, which is not the case for the spectral graph theory methods used here.

The main limitation of the work presented here is the restriction to linear resistance relations, as well as the rigidity of airway branches. First, the Reynolds number in the large airways exceeds the threshold for turbulence, and even flow in the smaller airways is affected by inertia and better described by a non-linear pressure-drop relation [39]. Additionally, airways themselves are compliant, albeit much less so than the acinar units [40], which affects the distribution of flow. Nonetheless, the work presented here is an important first step to going beyond simple "reduction by symmetry" methods of dimensionality reduction in lung airway models. In the limit of tidal or shallow breathing, it is likely that assumptions of linear pressure-drop relations (potentially different from the Poiseuille assumption used here) are a good approximation, in which case the methods presented here can tell us a great deal about the resistance properties of the airway network independent of choices of other specific simulation parameters (e.g. tissue elastance, breath volume, etc.).

The results shown here are also limited to artificial lung airway geometries, and use a combination of idealised cast-based models and algorithm-generated image-based networks. Therefore, the simulations and results in this paper do not predict how lungs will behave in specific states of health or disease. The models used here do not contain enough information about the small airways or tissue compliance to make patient-specific predictions of VH, which was measured in the original study by ^{3}He MRI [26]. However, the methods introduced in this paper can be adapted to patient-specific modelling of airway mechanics. In particular, the dimensionality reduction demonstrated here will enable faster computation that can be used to make inverse problems requiring multiple model realisations computationally tractable. In future work we plan to combine CT and ^{3}He MRI data to develop patient-specific models of lung structure and function, and this research will make it much more feasible by reducing the computational time required to estimate the ventilation distribution for a particular candidate airway network.

A further application of the findings of this paper is in characterisation of complex airway tree models. The largest eigenvalue modes of the Maury matrix are localised on regions that have higher resistance, and hence poorer ventilation. Since these eigenmodes are also often the most efficient to calculate (due to the nature of power iteration methods) this provides an efficient method to visualise "problem" lung regions. The relative importance of a given mode depends both on its eigenvalue (larger eigenvalues indicating larger resistance per terminal node) and the specific boundary conditions considered.

The methods in this paper are generic to resistance trees, and hence we expect this to have applications outside of the application of lung mechanics. First, other resistance trees in biology, such as the pulmonary vasculature, can be analysed using the same principles. Second, in discrete mathematics, the steady diffusion problem is identical to the linear resistance problem, therefore diffusion on tree networks can also be studied within this framework. For example, transport of gas in the respiratory zone of the lungs is dominated by diffusion on the tree network of acinar ducts, and so in future work we aim to use this formulation to produce low dimensional representations of models of these networks.

In summary, this paper has demonstrated how spectral graph theory provides a powerful tool for dimensionality reduction in the analysis of lung ventilation. This is achieved by reformulating the physical model for gas transport using a linear operator (the Maury matrix) that captures patterns of ventilation heterogeneity more efficiently than the traditional Laplacian operator, by exploiting the tree-like structure of the network. This approach shows potential for the description of a wider class of transport processes on tree-like networks.

Acknowledgements

CW was supported by the UK Medical Research Council (MR/ R024944/1) and PL was funded by a University of Manchester internship. Imaging data was collected by the POLARIS group at University of Sheffield and was funded by the UK Medical Research Council (MR/M008894/1).

References

[1] Weibel ER. 1963 Morphometry of the Human Lung. Springer-Verlag Berlin Heidelberg.

[2] Swan AJ, Clark AR, Tawhai MH. 2012 A computational model of the topographic distribution of ventilation in healthy human lungs. J. Theor. Biol. 300, 222–23. (doi:10.1016/j.jtbi.2012.01.042)

[3] Kim M, Bordas R, Vos W, Hartley RA, Brightling CE, Kay D, Grau V, Burrowes KS. 2015 Dynamic flow characteristics in normal and asthmatic lungs. Int. J. Numer. Meth. Bio. 31, e02731. (doi:10.1002/cnm.2730)

[4] Roth CJ, Ismail M, Yoshihara L, Wall WA. 2016 A comprehensive computational human lung model incorporating inter-acinar dependencies: Application to spontaneous breathing and mechanical ventilation. Int. J. Numer. Meth. Bio. 33, e02787. (doi:10.1002/cnm.2787)

[5] Downie SR, Salome CM, Verbanck S, Thompson B, Berend N, King GG. 2007 Ventilation heterogeneity is a major determinant of airway hyperresponsiveness in asthma, independent of airway inflammation. Thorax 62, 684–689. (doi:10.1136/thx.2006.069682)

[6] Horsley A, Wild JM. 2015 Ventilation heterogeneity and the benefits and challenges of multiple breath washout testing in patients with cystic fibrosis. Paediatr. Respir. Rev. 16, 15–18. (doi:10.1016/j.prrv.2015.07.010)

[7] Farrow CE, Salome CM, Harris BE, Bailey DL, Berend N, King GG. 2017 Peripheral ventilation heterogeneity determines the extent of bronchoconstriction in asthma. J. Appl. Physiol. 123, 1188–1194. (doi:10.1152/japplphysiol.00640.2016)

[8] Katifori E, Magnasco MO. 2012 Quantifying Loopy Network Architectures. PLoS ONE 7, e37994. (doi:10.1371/journal.pone.0037994)

[9] Lee SH, Fricker MD, Porter MA. 2016 Mesoscale analyses of fungal networks as an approach for quantifying phenotypic traits. J. Complex Netw. 5, 145–159. (doi:10.1093/comnet/cnv034)

[10] Meigel FJ, Alim K. 2018 Flow rate of transport network controls uniform metabolite supply to tissue. J. R. Soc. Interface 15, 20180075. (doi:10.1098/rsif.2018.0075)

[11] Spielman DA. 2007 Spectral Graph Theory and its Applications. In 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS07), IEEE. (doi:10.1109/focs.2007.56)

[12] Forrow A, Woodhouse FG, Dunkel J. 2018 Functional Control of Network Dynamics Using Designed Laplacian Spectra. Physical Review X 8, 041043. (doi:10.1103/physrevx.8.041043)

[13] Johnson OK, Lund JM, Critchfield TR. 2018 Spectral graph theory for characterization and homogenization of grain boundary networks. Acta Mater. 146, 42–54. (doi:10.1016/j.actamat.2017.11.054)

[14] Brouwer AE, Haemers WH. 2011 Spectra of Graphs Springer-Verlag Berlin Heidelberg.

[15] Maury B. 2013 The Respiratory System in Equations. Springer Milan, Milano.

[16] Grady LJ, Polimeni JR. 2010 Discrete Calculus. Springer London. (doi:10.1007/978-1-84996-290-2)

[17] Tutte WT. 2001 *Graph Theory* Cambridge : Cambridge University Press.

[18] Klein DJ, Randič M. 1993 Resistance distance. J. Math. Chem. 12, 81–95. (doi:10.1007/bf01164627)

[19] Kang W, Clark AR, Tawhai MH. 2018 Gravity outweighs the contribution of structure to passive ventilation-perfusion matching in the supine adult human lung. J. Appl. Physiol. 124, 23–33. (doi:10.1152/japplphysiol.00791.2016)

[20] Majumdar A, Alencar AM, Buldyrev SV, Hantos Z, Lutchen KR, Stanley HE, Suki B. 2005 Relating Airway Diameter Distributions to Regular Branching Asymmetry in the Lung. Phys. Rev. Lett. 95, 168101. (doi:10.1103/physrevlett.95.168101)

[21] Henry FS, Llapur CJ, Tsuda A, Tepper RS. 2012 Numerical Modelling and Analysis of Peripheral Airway Asymmetry and Ventilation in the Human Adult Lung. J. Biomech. Eng. 134, 061001. (doi:10.1115/1.4006809)

[22] Horsfield K, Dart G, Olson DE, Filley GF, Cumming G. 1971 Models of the human bronchial tree. J. Appl. Physiol. 31, 207–217. (doi:10.1152/jappl.1971.31.2.207)

[23] Haefeli-Bleuer B, Weibel ER. 1988 Morphometry of the human pulmonary acinus. Anat. Record 220, 401–414. (doi:10.1002/ar.1092200410)

[24] Tawhai MH, Pullan AJ, Hunter PJ. 2000 Generation of an Anatomically Based Three-Dimensional Model of the Conducting Airways. Ann. Biomed. Eng. 28, 793–802. (doi:10.1114/1.1289457)

[25] Bordas R, Lefevre C, Veeckmans B, Pitt-Francis J, Fetita C, Brightling CE, Kay D, Siddiqui S, Burrowes KS. 2015 Development and Analysis of Patient-Based Complete Conducting Airways Models. PLOS ONE 10, e0144105. (doi:10.1371/journal.pone.0144105)

[26] Marshall H et al. 2017 Detection of early subclinical lung disease in children with cystic fibrosis by lung ventilation imaging with hyperpolarised gas MRI. Thorax 72, 760–762. (doi:10.1136/thoraxjnl-2016-208948)

[27] Doel T. 2018. Pulmonary Toolkit v1.0. Link: https://github.com/tomdoel/pulmonarytoolkit .

[28] Reiter S. 2018. stl_reader. Link: https://github.com/sreiter/stl_reader.

[29] Guennebaud G, Jacob B, Others. 2010 Eigen v3. Link: http://eigen.tuxfamily.org.

[30] Qiu Y. 2015 Spectra. Link: https://spectralib.org/.

[31] Whitfield CA. 2020 Spectral properties of airway networks v1.0.2, 2020. Link: https://github.com/CarlWhitfield/Spectral_decomposition (doi:10.5281/zenodo.3708114).

[32] Whitfield CA. 2020 Space filling airway branching algorithm v1.0.1, 2020. Link: https://github.com/CarlWhitfield/Space_filling_tree (doi:10.5281/zenodo.3709073).

[33] Cook CD, Helliesen PJ, Agathon S. 1958 Relation Between Mechanics of Respiration, Lung Size and Body Size From Birth to Young Adulthood. J. Appl. Physiol. 13, 349–352. (doi:10.1152/jappl.1958.13.3.349)

[34] Sharp JT, Druz WS, Balagot RC, Bandelin VR, Danon J. 1970 Total respiratory compliance in infants and children. J. Appl. Physiol. 29, 775–779. (doi:10.1152/jappl.1970.29.6.775)

[35] Zapletal A, Paul T, Samanek M. 1976 Pulmonary elasticity in children and adolescents. J. Appl. Physiol. 40, 953–961. (doi:10.1152/jappl.1976.40.6.953)

[36] Whitfield CA, Horsley A, Jensen OE. 2018 Modelling structural determinants of ventilation heterogeneity: A perturbative approach. PLOS ONE 13, e0208049. (doi:10.1371/journal.pone.0208049)

[37] Hasler D, Anagnostopoulou P, Nyilas S, Latzin P, Schittny J, Obrist D. 2019 A multi-scale model of gas transport in the lung to study heterogeneous lung ventilation during the multiple-breath washout test. PLOS Computational Biology 15, e1007079. (doi:10.1371/journal.pcbi.1007079)

[38] Paiva M. 1973 Gas transport in the human lung. J. Appl. Physiol. 35, 401–410. (doi:10.1152/jappl.1973.35.3.401)

[39] Pedley TJ, Schroter RC, Sudlow MF. 1970 Energy losses and pressure drop in models of human airways. Resp. Physiol. 9, 371–386. (doi:10.1016/0034-5687(70)90093-9)

[40] Lambert RK, Wilson TA, Hyatt RE, Rodarte JR. 1982 A computational model for expiratory flow. J. Appl. Physiol. 52, 44–56. (doi:10.1152/jappl.1982.52.1.44)