From dynamical lung ventilation data to plugs distribution in asthma – A numerical diagnosis tool

Nicolas Pozin, Spyridon Montesantos, Ira Katz, Marine Pichelin, Irene Vignon-Clementel, Céline Grandmont

From dynamical lung ventilation data to plugs distribution in asthma – A numerical diagnosis tool

N. Pozin[1,2,3], S. Montesantos[3], I. Katz[3,4], M. Pichelin[3], I. Vignon-Clementel[1,2], C. Grandmont[1,2]

[1]INRIA Paris, 2 Rue Simone Iff, 75012 Paris, France
[2]Sorbonne Universités, UPMC Univ. Paris 6, Laboratoire Jacques-Louis Lions,75252 Paris, France
[3]Medical R&D, WBL Healthcare, Air Liquide Santé International, 1 Chemin de la Porte des Loges, 78350 Les Loges-en-Josas, France
[4]Department of Mechanical Engineering, Lafayette College, Easton, PA, 18042, USA

Abstract

In asthma and COPD, some airways of the tracheo-bronchial tree can be constricted, from moderate narrowing up to closure. Those pathological patterns affect the lung ventilation distribution. While some imaging techniques enable visualization and quantification of constrictions in proximal generations, no non-invasive technique provides precise insights on what happens in more distal areas. In this work, we propose a process that exploits lung ventilation measures to access positions of airways closures in the tree. This identification approach combines the lung ventilation model in which a $0D$ tree is strongly coupled to a $3D$ parenchyma description along with a machine learning approach. Based on synthetic data generated with typical temporal and spatial resolutions as well as reconstruction errors, we obtain very encouraging results with a detection rate higher than 90%.

1. Introduction

Lung provides the organism with fresh air. It is made of a porous elastic media, the parenchyma, irrigated with gas through the tracheo-bronchial tree, which is a non symmetric dyadic structure. Some pathologies such as asthma or COPD can induce airway constriction [**1**], [**2**] and generate ventilation defects [**3**], [**4**]. Response to Heliox treatment may depend on airway closure severity and position within the lung (generational and regional position) [**5**]. Thus knowing their locations could give predictive insights on the treatment efficiency. Another application would be to predict aerosol deposition, which can be greatly affected by bronchoconstrictions [**6**].

Some dynamic imaging techniques such as $4D$ computed tomography (4D-CT) [**7**], [**8**] and $4D$ magnetic resonance imaging (4D-MRI) [**9**], [**10**] can give insights on lung ventilation distribution along the breathing cycle. High resolution computed tomography (HRCT) shows the presence of constrictions in asthmatic patients [**11**], [**12**] but only proximal airways are easily accessible with current resolutions. However, histological analysis demonstrates that more distal branches can be constricted [**13**], [**14**].

In this work, we propose an approach that exploits $4D$ dynamical ventilation data to identify which regions of the tracheo-bronchial tree are severely bronchoconstricted. The process combines a lung ventilation model presented in [**15**], in which a $0D$ model of the tracheo-bronchial tree is coupled to a $3D$ continuous parenchyma, along with a plug identification step that includes a supervised machine learning technique. Approaches combining machine learning and modeling have been proposed to speed up calculations [**16**] or to improve diagnosis quality [**17**]. To the best of our knowledge, this study is the first attempt to use lung ventilation data in order to get insights on the tree structure.

In a first part, the global methodology is presented. Then, the ventilation model and the theoretical background of the plug identification approach are described. Results on synthetic data are finally shown.

2. Direct ventilation model and identification problem approaches

2.1. Global methodology

The tracheo-bronchial tree is a dyadic structure starting from large upper airways, referred to as proximal, down to smaller branches, referred to as distal. Some pathologies affect the tree structure; in asthma for instance, airways can be constricted. This goes from frequent moderate constrictions all over the tree with airway radius reduction of 10 to 25% in average [**18**], [**2**] up to closures [**14**], [**12**] in a much lower proportion. These latter severe constrictions will be designed as plugs. As observed experimentally through dynamic $4D$ images [**7**], [**9**], [**10**], stenoses affect the lung ventilation distribution. It takes more energy for the gas to flow through a constricted path, and downstream parenchyma regions are less irrigated.
Simulation results presented in subsection 4.1 indicate that ventilation defects caused by constrictions are mainly due to plugs rather than moderate narrowings, although the former are less frequent. Thus, ventilation information provided by $4D$ scans may give insights on plugs distribution. While the ventilation model described in subsection 2.2 enables to compute the flow distribution knowing the stenoses repartition within the tree, we aim at using the ventilation data (as those provided by medical images for instance) in order to determine which airways are plugged. This is an identification problem for which unknowns are the tree branches radii (see subsection 2.3). There is inter-variability in airway dimensions [**19**] and typical normal radii values are within a range. Here the healthy variability is neglected (see *Remark 4*) and each airway in a healthy state is assumed to have a known representative value. The ratio between the pathological radius and the one in healthy configuration then reveals the constriction.
Radii are accessed through the determination of airway resistances. As explained in subsection 2.2, flows at the tree exits and pressure drops from the tree entrance to the exits are linked through airway resistances by:

$$AQ = P$$

where $Q = (q_i)_i$ with q_i the flow at the i^{th} exit. The i^{th} element of P is the pressure drop from the tree entrance to the i^{th} exit. Matrix A is defined by (7) and includes information on the tree resistance distribution. Knowing Q and P, one can get insights on airway resistances and hence on the constrictions. From $4D$ scans, registration techniques can provide the parenchyma displacement field and finally Q through equation (6) (see subsection 2.2). In this study, we propose to determine P through a minimization problem (see subsection 2.3) based on the ventilation model described in [**15**] and subsection 2.2. In subsection 2.3, it is shown that for the identification problem to be well-posed the maximum number of unknowns is, at least, two times smaller than the number of airways within the tree. Thus only a sub-set of the branches can have an unknown radius. To select them, an *a priori* prediction process is proposed. Based on ventilation features extracted from $4D$ scans (see subsection 2.4.1), a machine learning technique determines which airways are more likely to be constricted (see subsection 2.4.2). If the prediction process is good, non-selected airways are little constricted at most. Since their radius cannot be set among unknowns, they are assumed to have known dimensions that are set in the normal physiological range. The *a priori* prediction step leads to a system that contains as many unknown resistances as equations. Though, this system may not be invertible. An extra step, so called *resistance removal step*, is proposed so as to get an invertible extraction. For reasons detailed in *Remark 2*, the resistance removal step is performed before the minimization problem. The identification problem can then be solved, outputting a constriction prediction for the unknown set (see subsection 2.3). This is the *inversion step*. From there, plugged airways i.e. these with severe constrictions, are identified. See Figure 1 for a synthetic scheme describing the whole

process. Note that in this study, $4D$ ventilation data are simulated with the tree-parenchyma coupled model presented in [**15**] (see subsection 2.2): these are synthetic data. Some noise is added to these data so as to reproduce realistic measurements.
In the next section, the ventilation model is described.

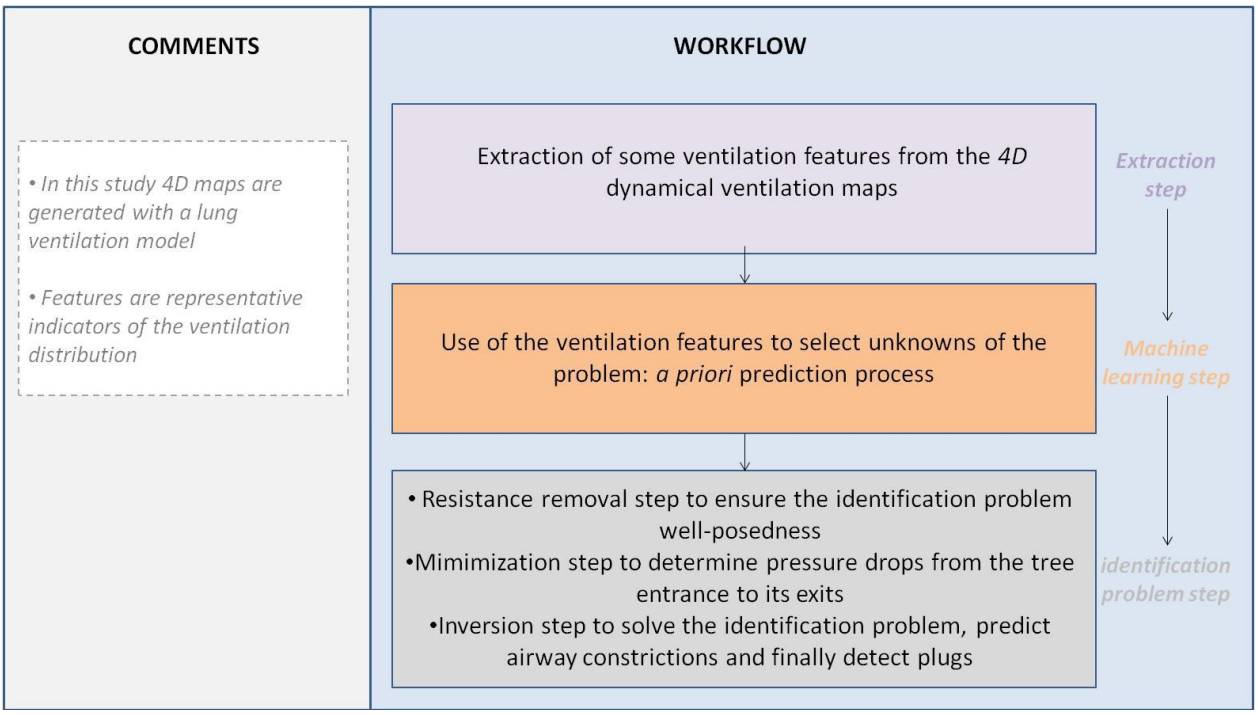

Figure 1: global process to determine the plug distribution in the tree based on dynamic ventilation data.

2.2. Direct ventilation problem

Models for both the parenchyma and tree used in this study as well as the coupling between them are precisely described in [**15**].
Under tidal breathing, Mach numbers in the tree are low [**20**] and the flow is assumed to be incompressible. In these conditions, it is reasonable to consider airways, the tree branches, as rigid cylindrical pipes [**15**]. Each branch is defined by its length L and radius r. Viscous and inertial effects induce energy dissipations. The pressure drop Δp and the flow q in an airway are linked through a non linear coefficient $R(q)$, called resistance, by $\Delta p = R(q)q$. This coefficient is defined as in [**21**] by the Pedley model [**22**]:

$$R(q) = \begin{cases} \gamma\left(Re\,\frac{2r}{L}\right)^{\frac{1}{2}} R_{pois}, & \gamma\left(Re\,\frac{2r}{L}\right)^{\frac{1}{2}} > 1, \\ R_{pois}, & \gamma\left(Re\,\frac{2r}{L}\right)^{\frac{1}{2}} < 1, \end{cases} \tag{1}$$

where $\gamma = 0.327$, Re is the Reynolds number defined by $Re = \frac{2\rho|q|}{\mu_f \pi r}$ with ρ the fluid density, μ_f its dynamic viscosity and the laminar Poiseuille resistance is $R_{pois} = \frac{8\mu_f L}{\pi r^4}$. The term $\gamma\left(Re\,\frac{2r}{L}\right)^{\frac{1}{2}}$ is associated with inertial losses.

Remark 1: note that (1) has been designed for inspiration and for one bifurcation angle only. Following [**23**] we use it independently of the angle and we choose to also use it during expiration.

The parenchyma is a porous medium. Following [**24**], we treat it as a homogenized isotropic media occupying a domain Ω, which is the configuration at rest. In tidal breathing conditions, i.e breathing at rest with normal lung expansion, displacements are slow and with limited amplitude. We thus assume it follows a linear elastic Hook law. The related stress tensor is given by:

$$\sigma_{mat}(\boldsymbol{u}) = 2\mu\varepsilon(\boldsymbol{u}) + \lambda tr\big(\varepsilon(\boldsymbol{u})\big)I,$$

where λ and μ are the effective Lamé parameters, $\boldsymbol{u}$ is the parenchyma displacement defined in the reference state domain Ω, and ε the linear strain tensor defined by $\varepsilon(\boldsymbol{u}) = \frac{1}{2}(\nabla\boldsymbol{u} + {}^{T}\nabla\boldsymbol{u})$.
If an airway collapses, its radius is decreased inducing a resistance steep rise (1). Whatever the elastic properties of the downstream region, it is more energy consuming and hence more difficult for the patient to push or pull flow through this branch. To take this phenomenon into account, the tree and parenchyma have to be coupled. In the following, to model the closure of an airway, we homogeneously reduce its radius.
In [**15**] we strongly couple the tree and parenchyma. Assuming $\Omega = \bigcup_{i=1}^{N}\Omega_i$ and $\forall i,j\ \Omega_i \bigcap \Omega_j = \emptyset$ where Ω_i is the sub-region of Ω irrigated by the i^{th} tree exit ex_i (see Figure 2), and neglecting gravity, we obtain the following strong formulation:

$$div(\sigma_{mat}) - \boldsymbol{\nabla}\, p_{tree} = \rho_{par}\frac{\partial^2\boldsymbol{u}}{\partial t^2}, \quad in\ \Omega. \tag{2}$$

Here ρ_{par} is the parenchyma density and p_{tree} is the piecewise constant function defined by:

$$p_{tree} = \sum_{i=1}^{N}(p_{trachea} - p_i)\mathbb{1}_{\Omega_i} \tag{3}$$

where $\mathbb{1}_{\Omega_i}$ is the characteristic function on Ω_i, $p_{trachea}$ and p_i are the pressure at the trachea entrance and at the i^{th} tree exit, respectively. The tree induces on each terminal region Ω_i an apparent pressure $p_{tree_{/\,\Omega_i}}$ which opposes volume variations (see Figure 2). The more the resistance on the path going from the trachea down to exit i denoted ex_i, the more the pressure drop along this path and the more the apparent pressure. In case of airway closure, the fed related region expands or contracts at the price of an increased energy spending [**5**]. In the following, without loss of generality, we assume the pressure at the trachea is equal to zero. Dirichlet boundary conditions are prescribed on the parenchyma surface $\partial\Omega$, the surface displacement being given by $\boldsymbol{u} = \boldsymbol{u_D}$ on $\partial\Omega$. The variationnal formulation associated with (2) is: for any $\boldsymbol{w}$ sufficiently smooth such that $\boldsymbol{w} = 0$ on $\partial\Omega$,

$$\rho_{par}\ \int_{\Omega}\ddot{\boldsymbol{u}}\cdot\boldsymbol{w} + \int_{\Omega}\sigma_{mat}(\boldsymbol{u}){:}\,\varepsilon(\boldsymbol{w}) - \int_{\Omega}p_{tree}\ I{:}\,\varepsilon(\boldsymbol{w}) = 0, \tag{4}$$

where

$$\begin{pmatrix} p_{tree_{/\,\Omega_1}} \\ \vdots \\ p_{tree_{/\,\Omega_N}} \end{pmatrix} = P = AQ. \tag{5}$$

Vector Q is defined by $Q = (q_i)_{1\le i\le N}$ with q_i the flow at the i^{th} exit. Note that under the assumption of small displacement around the reference state, we have

$$q_i = q_i(\dot{\boldsymbol{u}}) = \int_{\partial\Omega_i} \dot{\boldsymbol{u}} \cdot \boldsymbol{n}\,. \tag{6}$$

Matrix A includes information on airways resistances and thus on their dimensions. We denote T_i the set of indices of airways going from the trachea down to the i^{th} exit $\forall i \in [\![1; N]\!]$. Set T_i is defined by:

$$T_i = \{(n,k) \text{ / branch } (n,k) \in \text{path from the trachea down to exit } i\}$$

and thus T_i defines a unique path. Index n refers to the airway generation starting at zero for the trachea. Index k orders the airways of a given generation. A tree is said to be complete if each path contains the same number of generations. On a complete tree, at a given generation n, the index k runs through the set $[\![0; 2^n - 1]\!]$. From a complete tree one can extract a dyadic subtree as illustrated on Figure 2. Denoting T_{ij} the intersection set between T_i and T_j we have $A = \left[A_{ij}\right]_{1\le i,j\le N}$ with

$$A_{ij} = \sum_{(n,k)\in T_{ij}} R_{n,k}, \tag{7}$$

with $R_{n,k}$ the resistance of branch (n,k) (see Figure 2). These resistances are time dependent (see equation (1)).

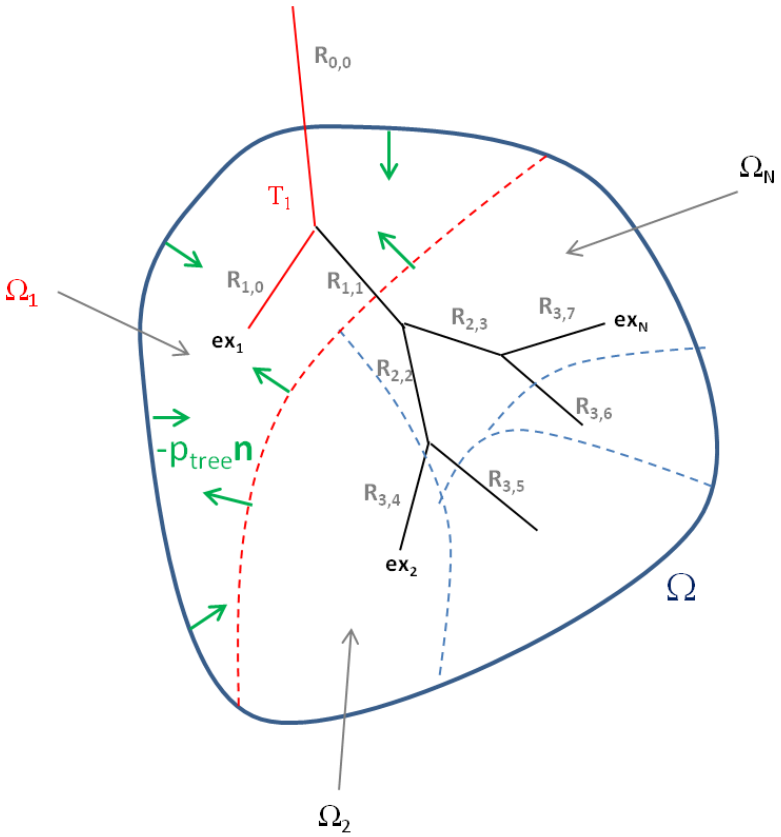

Figure 2: Schematic of parenchyma subdivision in non intersecting regions Ω_i, each one being irrigated by a tree path T_i. Representation on a non-complete tree. Each airway (n,k) has a resistance $R_{n,k}$. Green arrows illustrate the apparent pressure p_{tree} exerted by the tree through the coupling on region Ω_1 when it expands.

By considering the time and finite element (FE) discretizations introduced in [**15**], we obtain the following linear system:

$$E\,Y^{m+1} + B^T P^{m+1} = F^{m+1} \tag{8}$$

with P^{m+1} given by

$$P^{m+1} = A^m B Y^{m+1}.$$

Vector Y is the discrete FE representation of the velocity field $\dot{\boldsymbol{u}}$, E stands for the FE mass and elastic matrices and depends on the parenchyma mechanical properties. The right hand side F contains terms relative to surface displacement and velocity, thus depending on $\boldsymbol{u_D}$ and $\dot{\boldsymbol{u}}_{\boldsymbol{D}}$. Details can be found in [**15**]. Matrix B is defined by $B = \left[\int_{\Omega_i} div(\boldsymbol{e_j})\right]_{1\leq i\leq N, 1\leq j\leq n_{dof}}$ where $\boldsymbol{e_j}$ is the j^{th} element of the finite element basis and n_{dof} the number of finite element degrees of freedom of the problem; it links the discrete velocity field and the flow distribution Q through

$$BY^{m+1} = Q^{m+1}. \tag{9}$$

As mentioned in subsection 2.1, some airways can be constricted in asthmatic lungs, from numerous moderate closures to a few severe constrictions. As in [**5**], based on a literature review, we propose to simulate asthma remodeling by including these two types of constrictions with stochastic laws. To simulate a broncho-constriction, the airway radius is homogeneously divided by a factor so-called *constriction ratio*. To model asthma, moderate constriction ratios are generated according to a log-normal law with average $log(1.2)$, standard deviation 0.3 and within bounds $[1; 3]$. In addition, $n_{closure}$ severe constrictions are imposed, $n_{closure}$ being a random integer. These plugs are applied at randomly chosen positions of the tree and have constriction ratios randomly selected with a uniform law in the interval $[4; 10]$.
Some lung regions irrigated by constricted paths may be poorly ventilated, as seen with 4D-CT [**7**] and 4D-MRI [**3**] images. Knowing the tree structure with constrictions severity and position, which is an information included in matrix A, the parenchyma mechanical properties ρ_{par}, λ, μ, and the parenchyma surface displacement field over time $\boldsymbol{u_D}$, one can compute the displacement field $\boldsymbol{u}$ within the lung and deduce the flow distribution within the tree. Though, the positions and severity of closures in the tree as well as mechanical properties are not *a priori* known. Some high-resolution imaging techniques enable to see constrictions in proximal airways typically only until generation 5 or 6 [**11**]. In the frame of this study, we assume that the lung parenchyma mechanical properties are known. This is not the case in practice but some techniques such as magnetic resonance elastography [**25**] may give relevant estimations. As described in subsection 2.1, we propose to use dynamic ventilation data to get insights on the constriction distribution and identify plugged airways. This approach will be referred to as *identification problem*.

2.3. Identification problem

Function p_{tree} is linked to the flow distribution Q through the resistance matrix A (see equation (5)). If both p_{tree} and Q were known, one could get insights on matrix A, so on airway resistances (see (7)) and finally on airway radii through equation (1). This identification is done following a three-step approach. Image analysis on $4D$ data gives access to the parenchyma displacement field $\boldsymbol{u}$ (see subsection 2.4.1) from which Q can be computed given a parenchyma partition (6). Knowing the displacement field $\boldsymbol{u}$ along the breathing cycle, and assuming that both the parenchyma mechanical properties and the tree topology are known, one can use the ventilation model described in the previous subsection to define a function p^*_{tree} satisfying a minimization problem. This is the minimization step. Pressure drops from the tree entrance down to the tree exits are linked to the flows within the tree through airway resistances, and the related system contains as many resistances as there are airways within the tree. However, from the knowledge of Q and p^*_{tree}, there is not enough information to recover every airway resistance. Based on the process described in this subsection, the number of unknowns is reduced so as to get an invertible system. This is the resistance removal step. In a last step, this system is inverted and airway radii are determined. This is the inversion step. Plugged airways are thus identified.

Minimization step:
From the knowledge of the displacement field $\boldsymbol{u}$ and assuming that the mechanical properties of the parenchyma are known, we are going to define a tree pressure p^*_{tree} that is piecewise constant on each terminal region Ω_i. The idea, based on the direct problem (2), is to minimize the distance between $\boldsymbol{\nabla} p^*{}_{tree}$ and $\mathfrak{D} = div(\sigma_{mat}) - \rho_{par}\frac{\partial^2 \boldsymbol{u}}{\partial t^2}$ which is known by assumption. After time discretization, this minimization problem reads formally, at each time step:

$$\begin{cases} \min\limits_{p^*_{tree}\in \mathrm{K_h}} \left|\left|\boldsymbol{\nabla} p^*_{tree} - \mathfrak{D}_d\right|\right|, \\ K_h = \left\{ p^*_{tree} = \sum_i p_{tree\,/\Omega_i} \mathbb{1}_{\Omega_i} \;/\; p^*_{tree\,/\Omega_h} = p_h \right\} \end{cases}$$

where h and p_h are given region index and the corresponding pressure drop, $\mathfrak{D}_d$ is a time discretization of $\mathfrak{D}$. Note that we have prescribed the value of p^*_{tree} on region Ω_h to ensure the uniqueness of p^*_{tree} which is otherwise defined up to an additive constant. Note also that $\mathfrak{D}$ depends on the measurement of $\boldsymbol{u}$ as well as on the mechanical model that is chosen to describe the parenchyma behavior. The measures and the chosen model inaccuracies are two possible sources of errors. After discretization, we get the discretized related minimization problem $(\mathcal{P}_d)$ defined at each time step m by:

$$(\mathcal{P}_d): \begin{cases} \min\limits_{P^{*m}\in \mathbb{K}_\mathrm{h}} \left|\left|B^T P^{*m} - S^m\right|\right|^2 \\ \mathbb{K}_\mathrm{h} = \{\, V \in \mathbb{R}^N / \, V_h = p_h\} \end{cases} \tag{10}$$

where P^{*m}_h is the h^{th} component of P^{*m} and S^m is defined following (8) by

$$S^m = F^{m+1} - E\, Y^{m+1}\,.$$

Note that B^T is the FE operator corresponding to the gradient operator acting on piecewise constant functions on the partition $(\Omega_i)_i$. Since $dim(ker(B^T)) = 1$ (see Appendix A.1), P^{*m} is defined up to a constant vector. This is the reason why a constraint on P^{*m}_h is added. Let κ be the Lagrange multiplier associated to this pressure constraint and Π_h the linear form that maps a vector of $\mathbb{R}^N$ into its h^{th} component. For the sake of clarity, temporal indices are removed in following equations. The Lagrange function associated with (10) is:

$$\mathcal{L}(P^*, \kappa) = \left|\left|B^T P^* - S\right|\right|^2 + \kappa(\Pi_h(P^*) - p_h).$$

Solution P^* of $(\mathcal{P}_d)$ is such that:

$$\nabla_{P,\kappa}\mathcal{L}(P^*, \kappa) = 0 \Leftrightarrow \begin{cases} BB^T P^* - BS = 0, \\ \Pi_h(P^*) - p_h = 0. \end{cases} \tag{11}$$

The pressure constraint is enforced as described in subsection 3.2 by penalizing equation $BB^T P^* = BS$.
To solve $(\mathcal{P}_d)$, the pressure drop p_h along one path T_h has to be known. Which path to use is determined based on the *a priori* prediction process described in subsection 2.4. Along this path we suppose airway dimensions are known (see *Remark 2*). To get the pressure drops through each airway, flows shall also be known (see equation (1)). Since the gas is incompressible and because airways are assumed to be rigid, the flow in branch (n, k) equals the sum of the flows in its daughters:

$$q_{n,k} = q_{n+1,2k} + q_{n+1,2k+1}. \tag{12}$$

So the determination of exit flows (see subsection 2.4.1) enables by recursion to assess flows in every airway. Knowing airway dimensions and flows along T_h, we can finally compute p_h as $p_h = -\sum_{(n,k)\in T_h} R_{n,k}\, q_{n,k}$.
Note that the minimization problem $(\mathcal{P}_d)$ is solved at each time step. Knowing flows and pressures, we now want to get insights on the resistance distribution.

Resistance removal step:
In the following, we look for a resistance distribution such that vectors $A\,Q$ and P are close (see equation (5)) where both P and Q are known. This leads to a linear system for which resistances $R_{n,k}$ are the unknowns. Once they are determined, since flows and airway lengths are supposed to be known, equation (1) gives the radii $r_{n,k}$ from which we can deduce the radii reduction ratios defined by

$$red_{n,k} = \frac{r_{n,k}}{r_{healthy}{}_{n,k}}, \tag{13}$$

and finally identify plugged airways. The smaller the reduction ratio, the more closed the airway.
The pressure drop along any path T_i is given by:

$$\sum_{(n,k)\in T_i} R_{n,k} q_{n,k} = -p_i. \tag{14}$$

The tree is dyadic: if it has N exits, there are $2N-1$ airways. Thus the system (14) has size N and contains $2N-1$ unknowns. It is not invertible, thus the number of unknowns has to be reduced. The *a priori* prediction process described in subsection 2.4 provides for each airway a prediction of the radius reduction ratio (13) seen as a "constriction likeliness", based on which airways can be sorted. The lower the likeliness, the closer the airway is to its healthy configuration. From this ranking, the N unknown resistances associated to the airways that are most likely to be plugged are kept, the $N-1$ airways associated to the remaining unknowns are assumed to have healthy radius. At this stage, system (5) contains as many unknowns as equations. However, there is no guarantee for these equations to be independent. In practice, the system is invertible if the pressure at the nodes of unknown resistances are known, i.e $R_{n,k}$ can be determined if $p_{n,k}$ and $p_{n-1,\left[\frac{k}{2}\right]}$ are known. Notation $[k]$ refers to the integer part of k.
Let U_N be the set of indices of airways which resistances are unknown. For each path T_i from the tree mother branch to its i^{th} exit, let us denote $T_{i_N}^u = T_i \cap U_N$. Then, $\forall i \in [\![1; N]\!]$, (14) writes:

$$\sum_{(n,k)\in T_{i_N}^u} R_{n,k} q_{n,k} = -p_i - \sum_{(n,k)\in T_i \setminus T_{i_N}^u} R_{n,k} q_{n,k}\,.$$

Let us set a new indexation for airways with unknown radius. We denote ζ_N a bijective function that maps elements of U_N to the integer interval $[\![1; N]\!]$ based on the ranking provided by the *a priori* prediction process described in subsection 2.4: index 1 is associated to the airway that is predicted with highest radius reduction ratio i.e. the least likely to be constricted, and so on till index N associated to the airway predicted with lowest radius reduction ratio i.e. the most likely to be constricted. By extension, $\forall (n,k) \in U_N$ we note $R_{n,k} = R_i$ with $i = \zeta_N(n,k)$ and $q_{n,k} = q_i$ (see Figure 4).
Let Z_N be the matrix of $\mathcal{M}_N(\mathbb{R})$ defined by:

$$Z_{N\,ij} = \begin{cases} 1 & if\ \ {\zeta_N}^{-1}(j) \in T_{i_N}^u, \\ 0 & else. \end{cases} \tag{15}$$

A given line i of Z_N contains 1 at position j if the two following conditions are satisfied: R_j belongs to the set of unknown resistances, and the airway associated to R_j belongs to path T_i; otherwise the line contains 0 at position j.

We also define the vectors $G_N = (R_i q_i)_{i=1,\dots,N}$ and $H_N = \left(-p_i - \sum_{(n,k)\in T_i \setminus T_{i_N}^u} R_{n,k} q_{n,k}\right)_{i=1,\dots,N}$. With these notations system (14) is equivalent to:

$$Z_N G_N = H_N. \tag{16}$$

Information on airway resistances and hence on their radius is contained in G_N. One has to solve system (16) to recover the resistances knowing the right hand side H_N. However, matrix Z_N may not be invertible. The strategy is thus to extract from Z_N an invertible matrix, which will be of maximal rank. To decide which resistance shall be removed from the unknowns, we define an iterative process taking advantage of the *a priori* prediction made on the likeliness constriction of a branch. The higher the predicted reduction ratio for an airway, the less likely this airway is constricted. At iteration 1 the resistance R_1 of the airway with highest radius reduction ratio is removed with highest priority and the radius of the associated airway is set as "healthy". This corresponds to removing the first column of Z_N. The resulting matrix is not squared and contains at least two collinear rows, otherwise Z_N would have been invertible (see demonstration in Appendix A.2); to keep a square system, one of these collinear rows is removed, defining matrix Z_{N-1}. Vectors G_N and H_N are modified accordingly, defining G_{N-1} and H_{N-1}. If $rank(Z_{N-1}) = N-1$, Z_{N-1} is invertible and unknown resistances can be determined. If $rank(Z_{N-1}) = rank(Z_N) < N-1$, Z_{N-1} is not invertible and we proceed to iteration 2. Otherwise, in case $rank(Z_{N-1}) < rank(Z_N)$, R_1 is set back among the unknowns and the second airway resistance that is less likely to be plugged, R_2, is removed; the process goes on till $rank(Z_{N-1}) = rank(Z_N)$. Then the algorithm moves to iteration 2 and so on. Doing so, airways predicted as severely constricted by the *a priori* prediction process are less likely to be removed from the unknowns. Note that for reasons explained in *Remark 2*, we enforce the resistance of airway $(0,0)$ among the unknown set and prevent it from being removed. The algorithm is described in Figure 3.

In the following, for a vector V of $\mathbb{R}^p$ and an integer $c \leq p$, we denote by $V\{c\}$ the vector of $\mathbb{R}^{p-1}$ obtained by removing the c^{th} element of V. For a matrix M of $\mathcal{M}_p(\mathbb{R})$, we denote by $M\{c_1, c_2\}$ the matrix of $\mathcal{M}_{p-1}(\mathbb{R})$ obtained by removing the c_1^{th} line and c_2^{th} column of M. For a matrix M we denote $M(c,:)$ the c^{th} row of M. Vector and matrix indexations start at one.

Initialization:

- Index exits in $\{1, \dots, N\}$
- $l = N$
- $Z_l = Z_N$
- $k = 0$

While $rank(Z_l) \neq l$

- **step 1** $k \leftarrow k + 1$
  If $\zeta_l^{-1}(k) = (0,0)$
  - $k \leftarrow k + 1$

  $R_{0,0}$ is kept as unknown
- **step 2** $i = 0, j = 0, s = 0$
  While $s = 0$
  - $i = i + 1$
  - ***For*** $j' = i + 1$ *to* l
    - ***If*** $Z_l(j', :)$ collinear to $Z_l(i, :)$
      - $s = 1$
      - $j = j'$
      - $Z_{l-1}^j = Z_l\{j, k\}$
      - ***Exit for***

  Sub-system extraction

  End while

  If $rank\left(Z_{l-1}^j\right) = rank(Z_l)$
  - ***If*** $rank\left(Z_{l-1}^j\right) = l - 1$
    - Set R_k as healthy
    - $Z_{l-1} = Z_{l-1}^j$
    - $G_{l-1} = G_l\{k\}$
    - $H_{l-1} = H_l\{j\}$
    - $l \leftarrow l - 1$

    The sub-system is invertible. End
  - ***Else***
    - $Z_{l-1} = Z_{l-1}^j$
    - $U_{l-1} = U_l\{\zeta_l^{-1}(k)\}$
    - Set ζ_{l-1} a function from U_{l-1} to $[\![1; l-1]\!]$ based on the *a priori* prediction process*
    - Reindex exits $\{1, \dots, j-1\} \cup \{j+1, \dots, l\}$ in the interval $\{1, \dots, l-1\}$
    - $l \leftarrow l - 1$

    The sub-system is not invertible. Continue extraction: ***step*** **1**

  Else if $rank\left(Z_{l-1}^j\right) < rank(Z_l)$
  - Proceed to step 1

  Wrong extraction go back to ***step*** **1**

*** ζ_{l-1} is chosen such that unknown R_1 is the resistance associated with the airway that has highest predicted reduction ratio by the *a priori* prediction process, R_2 with the airway with second highest predicted reduction ratio,… and R_{l-1} with the airway with lowest predicted reduction ratio.**

Figure 3: Resistance removal algorithm for unknowns selection. Resistances are removed one by one according to the predicted radius reduction ratio in order to make the system (16) invertible.

This algorithm converges: indeed if $l = 1$ then $Z = Z_1 = [1]$ which is invertible. See Figure 4 for a schematic illustration of the process.

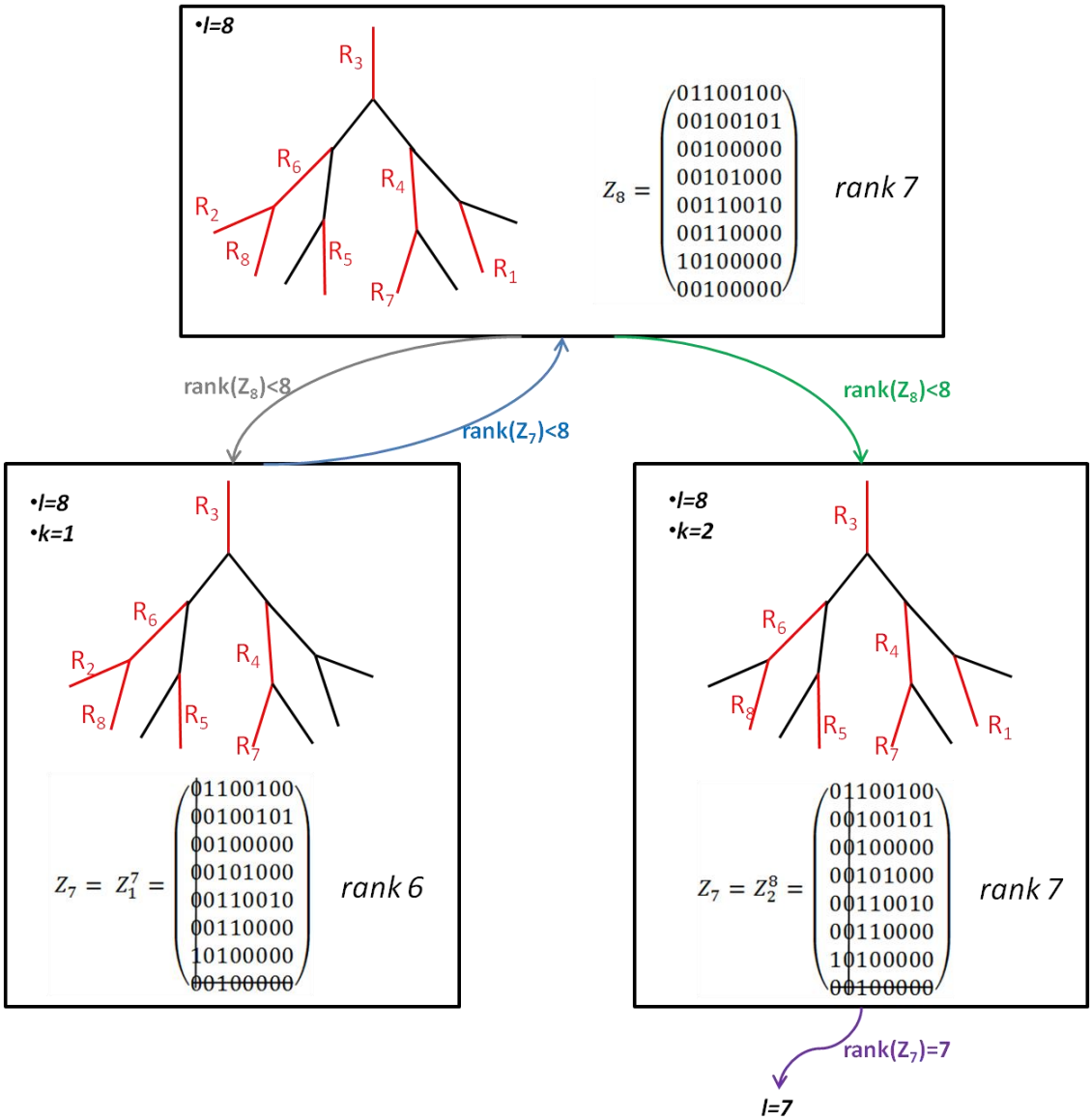

Figure 4: illustration of the resistance removal algorithm on an eight-exit tree. Branches in red are the ones with unknown resistance. Matrix Z_8 of the initial system is not invertible (top square). Following the process described in Figure 3, R_1 is first removed from the set of unknowns, a line and a column are removed from Z_8 leading to Z_7 (bottom left square); rank(Z_7)< rank(Z_8) so R_1 is set back among unknowns (top square) and R_2 is removed instead (bottom right square). Resulting reduced matrix Z_7 has now the same rank as Z_8 so the removal is confirmed. At the end of the process the tree contains several paths with $R_{0,0}$ ($= R_3$) as only unknown, so a pressure along a path can be computed so as to solve the minimization problem (10), see *Remark 2*.

Inversion step:

If Z_l is invertible we have, from (16), $G_l = Z_l^{-1} H_l$ and since the flows in all airways are known we can then determine the unknown resistances. We have $\forall i \in [\![1; l]\!], G_{l_i} = R_i q_i$. Then the radii can be determined from the Pedley formula (1).

$$r_i = \begin{cases} \left[\frac{q_i}{G_{b_i}} \frac{8\mu_f L}{\pi}\right]^{\frac{1}{4}}, & if \quad \gamma \left(\frac{4\rho|q_i|}{\mu_f \pi L}\right)^{\frac{1}{2}} < 1, \\ \left[\frac{q_i}{G_{b_i}} \gamma 16 \left(\frac{\rho \mu_f |q_i| L_i}{\pi^3}\right)^{\frac{1}{2}}\right]^{\frac{1}{4}}, & else. \end{cases}$$

Recall that for each airway with resistance in the unknown set, a radius is computed at each time step, and the final prediction is taken as the time average of these values. Note that the resistance removal step is performed one time only: the resulting invertible matrix Z_l does not depend on the time step.
The identification problem step will be all the more accurate that resistances removed from the process are associated to airways with healthy dimensions. In practice, most of them are moderately constricted. Assuming they are healthy will induce some errors in the right hand side H_l and thus some prediction errors on unknowns, since they will have to account for a higher pressure drop on their paths.

Remark 2: To solve (10), the pressure drop along a path has to be imposed. To compute the pressure drop p_h along a path T_h, radii of airways belonging to T_h need to be known (recall that lengths and flows are supposed to be known). After the *a priori* detection process, to make sure that at least one path contains airways with known dimensions only, the resistance of airway $(0,0)$ is enforced among the unknowns of system (14) and prevented to be removed during the resistance removal step (described in Figure 3). Doing so, one path T_h of the tree will contain no other unknown than airway $(0,0)$. Indeed, if $R_{0,0}$ is enforced among unknowns, one has to access $p_{1,0}$ to ensure invertibility. To that end, there needs to be a path with no unknown from node $(1,0)$ to an exit. In practice, airway $(0,0)$ would be the trachea. Its dimensions can be measured with bronchoscopy for example and enforcing $R_{0,0}$ among unknowns is an artifice designed to enable p_h computation. So dimensions of airways along T_h are all known and the pressure drop p_h along this path can be computed. Note that the minimization step is thus performed after the resistance removal step.

Remark 3: Note that by solving an extracted system from (16) we do not ensure (14) is satisfied. To verify (14) some compatibility conditions on the right hand side would be needed. The last invertible extracted system $Z_l G_l = H_l$ could be enriched by adding unknown resistances while keeping the system invertible. But our first aim here is not to get a resistance distribution that enables to recover the displacement field measured on the 4D data. The goal of this study is rather to identify plugged airways. Note however that if the resistance of an airway that is actually plugged is removed from the set of unknowns, it will not be identified as a plug.

To summarize, the identification problem aims at determining which airways are severely constricted. It is made of three steps and performed after the *a priori detection* process. Pressure drops are linked to the flows within the tree through airway resistances by system (14). This system contains more unknown resistances than independent equations. The unknown set has to be reduced: this is the resistance removal step, which is performed one time and is valid for every time step. The removal step ensures the existence of a path through which the pressure drop along the tree can be computed. The minimization problem $(\mathcal{P}_d)$ can then be solved to determine the pressure drops from the tree entrance to all its exits. Now that exit pressures are known, the system extracted from (14) is inverted at each time step.
As mentioned before, an unknown selection process is needed to determine which resistance to remove from the unknown set. This process shall keep, among others, the airways that are actually plugged and remove some healthy or moderately constricted airways, i.e. some that have negligible effect on ventilation distribution. Indeed, as mentioned in subsection 2.1, asthma affects the lung structure from frequent moderate radius constrictions to a few airway closures. Using the ventilation distribution information provided by dynamic images, we propose a way to make an *a priori* prediction on airways constriction likeliness (see subsection 2.4).

Remark 4: We aim to keep in the unknown set resistances associated to the most constricted airways. If the *a priori* prediction process provides an appropriate ranking, airways whose resistances do not belong to that set are moderately constricted at most. Their radius is supposed to have a typical "healthy" value known from morphometric studies (see subsection 3.1.3), although normal airway dimensions are submitted to an inter- and intra- variability [**19**]. However, possible radius variations within this "healthy" or "normal" range are much lower than these induced by severe bronchoconstrictions. The errors caused by

considering an arbitrary “healthy” radius are thus expected to have little impact on the pressure drop distribution compared to plugs.

2.4. *A priori* prediction process

Following the process described in subsection 2.3, to get insights on plugs distribution, only a sub-set of airway resistances can be considered as unknowns. The tracheo-bronchial tree has a dyadic structure: if it contains N exits it has $2N-1$ airways. At most N have their resistance initially set as unknown (see subsection 2.3). One could try all the combinations of N airways such that system (16) is invertible, invert the system for each combination, and keep as final result the one that minimizes the distance between vectors AQ and P. However, it can be shown that there are more than 2^{N-1} of these combinations: this makes this approach costly when the tree contains a few hundred exits as done in this study. Thus, before the identification problem step, an *a priori* prediction process is required to determine which branches are likely to be constricted and which ones rather have close to healthy dimensions. These latter are removed with highest priority. The prediction process described in this section is based on dynamical ventilation data analysis. It uses a machine learning technique to map the observed ventilation distribution with constrictions positions and strengths within the tree.

2.4.1. Ventilation information extraction

Ventilation distribution, i.e. flows within the tree, is related to the tree structure. In the following, some features characterizing the ventilation are set as inputs of a machine learning algorithm so as to determine which airways are likely to be plugged. The ventilation distribution can be determined from the parenchyma velocity field. Indeed, from (6) one can compute tree exit flows. Under the incompressibility assumption, the flow distribution within the tree is then obtained (12).

2.4.2. Machine learning process

The machine learning approach proceeds in two steps (see Figure 5). First a correspondence is established between ventilation features extracted from $4D$ ventilation data and known airways radius reduction ratios (13) that are associated to these features. This mapping is built on a large synthetic database, so-called *learning base*, so as to capture the diversity of ventilation distribution patterns. This is the *learning step*. In a second step, for a dynamical ventilation map associated to an unknown constrictions distribution, ventilation features are extracted, and airway reduction ratios are determined. This is the target *prediction step*.

Ventilation features:
Which ventilation features to consider is the object of this subsection. Bronchoconstrictions influence the ventilation distribution. To quantify a disease severity and get insights on its geometrical distribution, ventilation pathological patterns can be compared to the ones that would be obtained if there were no constrictions. In this study we assume that for a given asthmatic patient, lung mechanical properties are known. We also consider we know the tree topology, i.e. nodes positions, and typical dimensions airways would have if they were not constricted. So a “healthy” tree configuration is available. We finally assume we can access the lung geometry and the parenchyma surface displacement along breathing. These are the necessary input data to use the tree-parenchyma coupled model described in subsection 2.2. A reference “healthy” ventilation distribution can be simulated. In the following, variables with subscript “asthma” refer to the ventilation in an asthmatic case while “healthy” refers to the one simulated when the tree has no bronchoconstriction.

The machine learning step aims at determining, given a ventilation distribution, which airways are likely to be constricted and with which severity. To do so, we need to define at the airway level some features that correlate with the ventilation distribution. As noted above (also see subsection 4.3), we know the flow rate evolution and hence the volume of gas transiting in every airway over a breathing cycle. For the airway (n,k) we note respectively $V_{n,k}^{max}$ and $V_{n,k}^{min}$ the maximum and minimum of $V_{n,k}(t) = \int_0^t q_{n,k}$.
We define the following feature:

$$c_{n,k} = \frac{\left(V_{n,k}^{max} - V_{n,k}^{min}\right)_{asthma} - \left(V_{n,k}^{max} - V_{n,k}^{min}\right)_{healthy}}{\left(V_{n,k}^{max} - V_{n,k}^{min}\right)_{healthy}}.$$

If $c_{n,k} < 0$, less flow goes through airway (n,k) than in the healthy reference case. This may indicate there are constrictions in the paths going through (n,k). They could be upstream, downstream or both. The value of $c_{n,k}$ correlates with the constriction severity. To get more insight on where constrictions could occur, we define:

$$\alpha_{n,k} = \begin{cases} 1, & if\ c_{n-1,\left[\frac{k}{2}\right]} > 0, \\ 0, & else. \end{cases}$$

Note that airway $\left(n-1,\left[\frac{k}{2}\right]\right)$ is the mother of (n,k). A value of 1 indicates that there is more flow going through the mother than in the healthy case. Thus this branch is less likely to be affected and constriction may rather be downstream. On the contrary, when the value is 0, the mother or ascending airways may rather be affected. In the same spirit we define:

$$\begin{cases} \beta_{n,2k} = \begin{cases} 1, & if\ c_{n,2k+1} > 0, \\ 0, & else, \end{cases} \\ \beta_{n,2k+1} = \begin{cases} 1, & if\ c_{n,2k} > 0, \\ 0, & else. \end{cases} \end{cases}$$

If $c_{n,k} < 0$ and $\beta_{n,k} = 0$, both airway (n,k) and its sister have reduced flows compared to the healthy case, so some constrictions may exist upstream. To add information on the downstream environment we define:

$$\gamma_{n,k} = \begin{cases} 2, & if\ c_{n+1,2k} > 0\ and\ c_{n+1,2k+1} > 0, \\ 0, & if\ c_{n+1,2k} < 0\ and\ c_{n+1,2k+1} < 0, \\ 1, & else. \end{cases}$$

Airways $(n+1,2k)$ and $(n+1,2k+1)$ are the daughters of airway (n,k). For a value of 0, both daughters have reduced flows compared to the healthy case so airway (n,k) or ascending branches may be constricted. If the value is 1 some constrictions would rather occur downstream.

We also define the distribution ratio:

$$\begin{cases} \delta_{n,2k} = \dfrac{\left(V_{n,2k}^{max} - V_{n,2k}^{min}\right)_{asthma}}{\left(V_{n,2k}^{max} - V_{n,2k}^{min}\right)_{asthma} + \left(V_{n,2k+1}^{max} - V_{n,2k+1}^{min}\right)_{asthma}}, \\ \qquad\qquad \dfrac{\left(V_{n,2k}^{max} - V_{n,2k}^{min}\right)_{healthy}}{\left(V_{n,2k}^{max} - V_{n,2k}^{min}\right)_{healthy} + \left(V_{n,2k+1}^{max} - V_{n,2k+1}^{min}\right)_{healthy}} \\ \delta_{n,2k+1} = \dfrac{\left(V_{n,2k+1}^{max} - V_{n,2k+1}^{min}\right)_{asthma}}{\left(V_{n,2k+1}^{max} - V_{n,2k+1}^{min}\right)_{asthma} + \left(V_{n,2k}^{max} - V_{n,2k}^{min}\right)_{asthma}}. \\ \qquad\qquad \dfrac{\left(V_{n,2k+1}^{max} - V_{n,2k+1}^{min}\right)_{healthy}}{\left(V_{n,2k+1}^{max} - V_{n,2k+1}^{min}\right)_{healthy} + \left(V_{n,2k}^{max} - V_{n,2k}^{min}\right)_{healthy}} \end{cases}$$

It compares the volume distribution between an airway and its sister. If $\delta_{n,k}$ is significantly smaller than 1, airway (n,k) or some downstream branches may be constricted. These features may behave differently depending on the airway generation as the effect of constrictions depends on the flow which itself decreases with the generation level [**5**]. So we add the generation level n as a feature.
We use a random forest regressor algorithm to map for each airway (n,k) its feature vector $[c_{n,k}, \alpha_{n,k}, \beta_{n,k}, \gamma_{n,k}, \delta_{n,k}, n]$ and its radius reduction ratio $\frac{r_{n,k}}{r_{healthy_{n,k}}}$. Quality of the prediction results is shown in subsection 4.2. The random forest method was chosen because it is fast and insensitive to noise.

Remark 5: Note that ventilation distribution is also influenced by factors that do not appear among the selected features. In particular, it is a function of the position within the lung: for instance, in tidal regime, displacements are more important around the diaphragm than near the lung apex. To make the prediction more accurate, one could add as features the lobar position of the airway, or even its exact position within the lung. Though doing so would dramatically increase the problem dimensionality. Rather than providing a precise prediction of the constriction ratio, we aim at predicting a good ranking of plugged airways (see results subsection 4.2). For that reason, the output of the machine learning algorithm may be referred to as "*constriction likeliness*" rather than radius reduction ratio. Airways could have been sorted based on another criterion. For instance, one could have enforced proximal airways in the unknown set (see subsection 2.3) by modulating the prediction with the airway generation.

Data base construction and constriction likeliness prediction:
To build a learning base the following approach is adopted: first the parenchyma displacement field is extracted from the $4D$ dynamical map, so-called *target map*, associated to a tree whose plug distribution has to be determined. The extracted surface displacement is used as Dirichlet boundary conditions (namely $\boldsymbol{u_D}$). Knowing the tree structure, the lung mechanical properties, and assuming that the parenchyma behavior can be accurately described by a Hooke law, a healthy ventilation simulation is performed based on the tree-parenchyma coupled model described in subsection 2.2. Imposing stochastic constrictions distributions as described in subsection 2.2, synthetic pathological ventilations are simulated with the same given Dirichlet boundary conditions. So all synthetic patients breathe with the same surface displacement. From these synthetic data, airway ventilation features described in subsection 2.4.2 are computed. Note that the associated radius reduction ratios are known since the trees have been generated. The learning phase is performed on this synthetic database by mapping the features and the related known airway constrictions. If the tree contains $2N-1$ airways and if m_p pathological cases are simulated, there are $(2N-1) * m_p$ couples [airway features; radius reduction ratio] in the learning base.

This data base is then the basis to predict airway radius reduction ratios associated to the target $4D$ dynamical map. As mentioned in subsection 2.1, in this study the target map is synthetically generated with the tree-parenchyma coupled model.

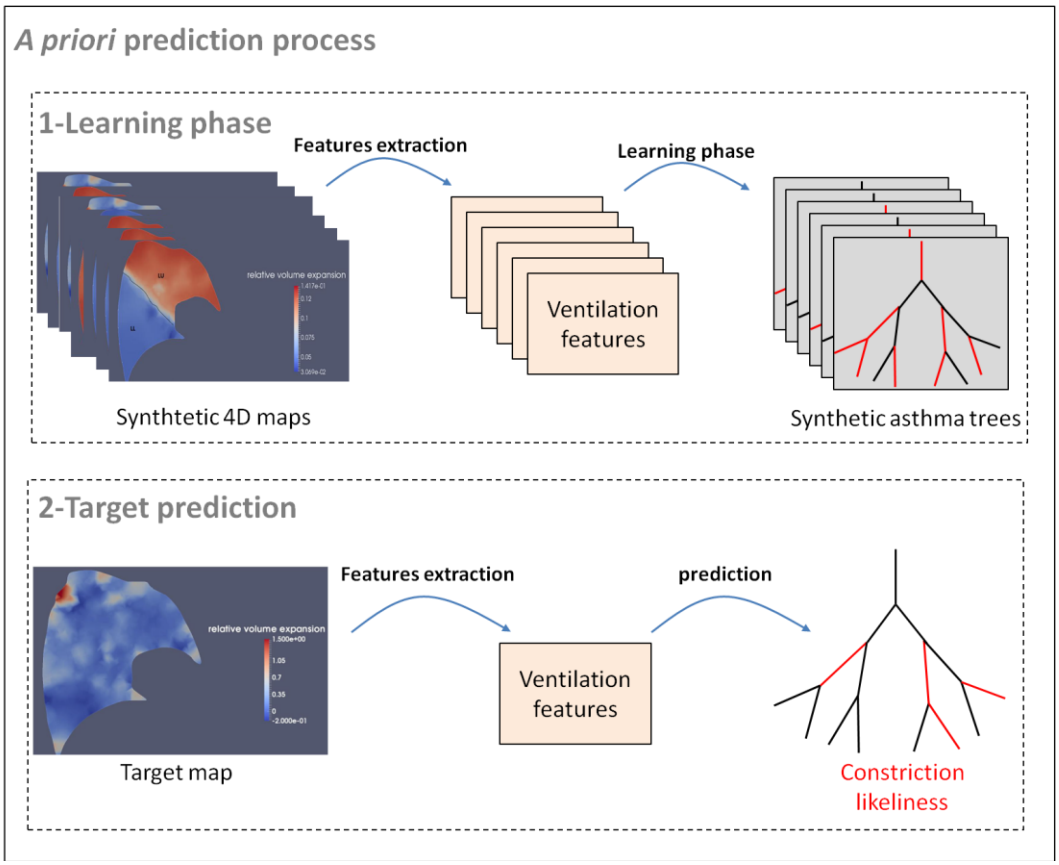

Figure 5: Schematic of the *a priori* prediction process. In the learning phase, a mapping is built between ventilation features extracted from synthetic $4D$ ventilation data and airways constriction ratio in related asthmatic tress. Note that the Dirichlet boundary conditions and the tree geometry used to generate the synthetic data are those of the target ventilation map. Then, airways constriction likeliness is predicted for the tree associated to the target ventilation map. If the process is good, plugged airways (in red) have a high constriction likeliness.

2.5. Global approach summary

In this section we summarize the overall process to go from a $4D$ dynamical map and tree topology to the tree plugs distribution, see Figure 6. For the identification problem (*step 4,* see subsection 2.3) to be well-posed, not every airway resistances can be considered as unknowns. An *a priori* prediction process (*step 3*) is used to select a subset of airways that are more likely constricted and shall hence be put in the unknown set. Based on a machine learning technique (see subsection 2.4.2) ventilation features extracted from the target ventilation map (*step 1,* see subsection 2.4.1) are used to predict constriction likeliness for each airway. To that end, a learning phase is necessary. To generate synthetic data for the learning phase, the tree-parenchyma ventilation model presented in [**15**] and subsection 2.2 is used. The extracted parenchyma surface displacement is used as Dirichlet boundary conditions. Ventilation simulations are performed with a healthy tree model and various asthmatic configurations (*step 2,* see subsection 2.4.2). The healthy tree model is built following subsection 3.1.3: segmented branches are used along with the parenchyma envelope to generate a complete tree geometry. Asthmatic configurations are obtained by imposing stochastic constrictions as described in subsection 3.1.3. Using the *a priori* constriction likeliness ranking, one can finally proceed to the identification problem step (*step 4*) which is done in three stages. The system linking flows and tree pressures through airway resistances is *a priori* not invertible. An

invertible extraction is first obtained thanks to the resistance removal step. Then a minimization problem is solved to determine the exit tree pressures. The resulting system is inverted and plugs are finally identified.

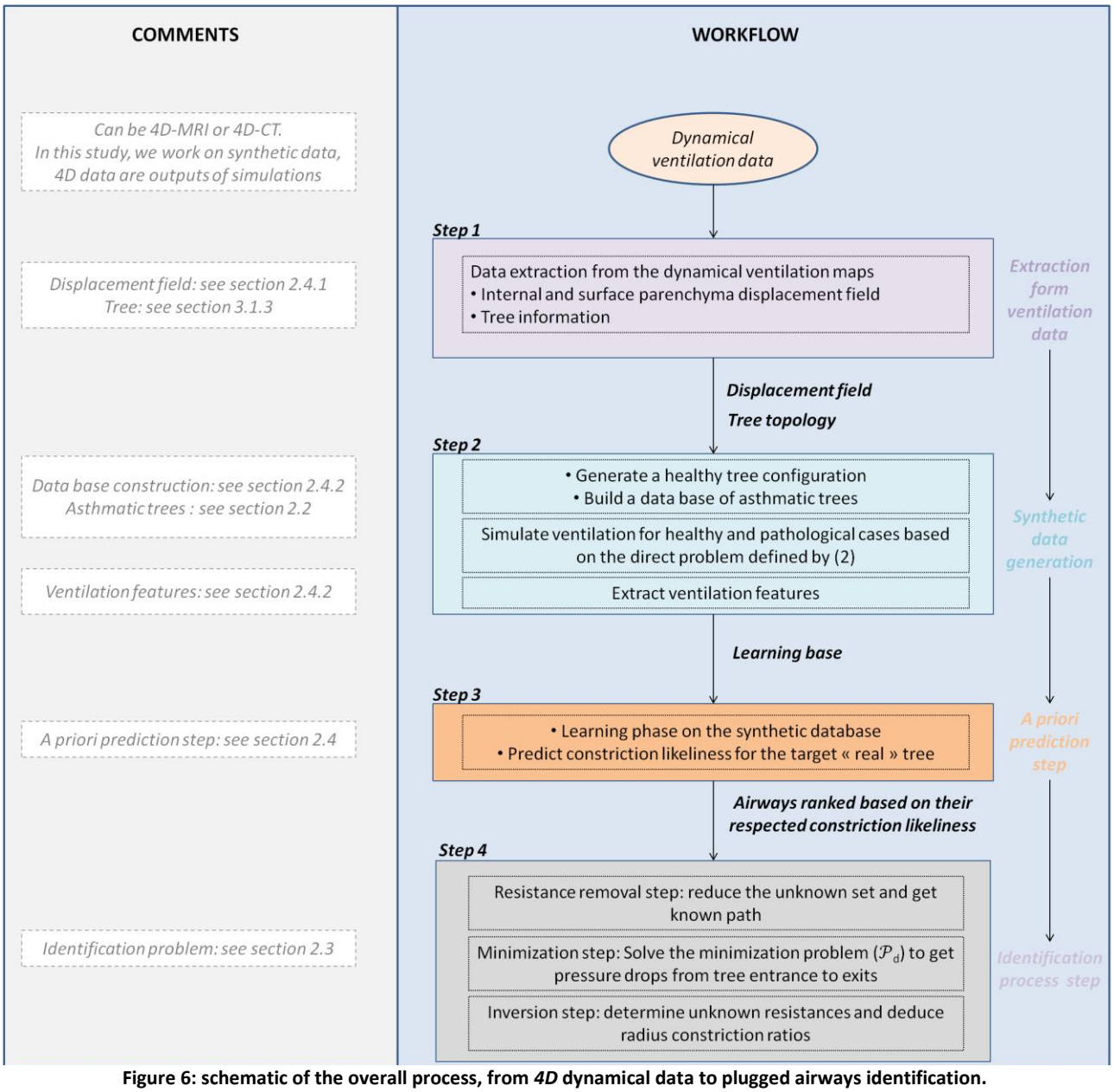

Figure 6: schematic of the overall process, from *4D* dynamical data to plugged airways identification. Detailed version of Figure 1.

3. Numerical methods and patient-specific data exploitation

In this section we give insights on patient-specific data construction, numerical strategies and systems resolution.

3.1. Input data construction

For the process described in Figure 6, the lung geometry, its displacement over the breathing cycle and the tree structure are needed.

3.1.1. Lung geometry

From $4D$ scans the lung envelope can be segmented. We work on a human left lung geometry. Segmented HRCT data acquired in [**26**] are used to build a patient-specific tree and a parenchyma geometry.
A surface mesh of the parenchyma is generated with Meshlab [**27**]. From there, a tetrahedral volume mesh is built with GMSH [**28**]. It is subdivided based on the tree structure in as many regions as there are terminal branches (see Figure 7) according to an algorithm described in [**15**]. The mesh contains 250 594 tetrahedrons, which corresponds to achievable 4D-CT spatial resolution (see subsection 3.1.2). There are 140 091 degrees of freedom.

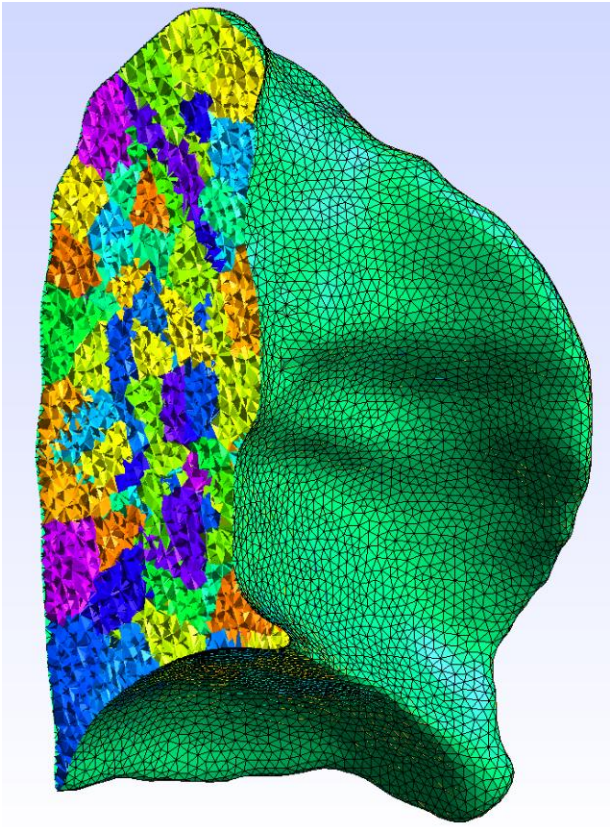

Figure 7: left lung mesh subdivision based on a ten-generation tree structure. The domain is divided into 477 terminal regions. Each color corresponds to a different region. The average volume of each region at the reference state is 2.2 ml.

3.1.2. Temporal and spatial resolutions

Synthetic $4D$ ventilation data generated from the tree-parenchyma coupled model described in [**15**] and subsection 2.2 are used in this study. Spatial and temporal resolutions are chosen coherently with what available machines can provide. Simulations performed to generate these data are done with a small time step of 0.04 s to ensure precision but only a few time points are then kept to cope with experimental temporal resolutions.
The parenchyma surface Dirichlet boundary condition is built based on two static inflation states segmented from [**29**]: mean lung volume (MLV), here considered as the reference state, and total lung capacity (TLC). A non-linear registration is performed with deformetrica [**30**] to map MLV and TLC parenchyma surfaces. In this study we work in tidal breathing conditions, i.e for respiration at rest and with normal volume expansion. The MLV to TLC registered displacement field is bounded and sinusoidal dynamics with time period 4s are prescribed so as to impose a typical tidal volume evolution. This surface displacement is imposed as Dirichlet boundary conditions to the tree-parenchyma coupled model. Here we

work on synthetic data: the displacement field $\boldsymbol{u}(t)$ is the output of a simulation based on the mechanical model described in subsection 2.2. Note that, as mentioned in subsection 2.3, one could work on real $4D$ data from which a displacement field can be extracted through registration. Depending on the scanner and reconstruction techniques, 4D-CT and 4D-MRI lead to various spatial and temporal resolutions [**31**], [**32**], [**33**], [**34**]. We assume a spatial resolution of 4 mm^3 isotropic and we use a mesh that is refined accordingly (see subsection 3.1.1). Temporal resolution is 0.4 s (i.e. ten images per breathing cycle with the considered dynamic). Synthetic ventilation data are thus generated every 0.4 s to perform the machine learning step (see results in subsection 4.2). Information from the ten resulting time points is aggregated to proceed with the identification problem step (see results in subsection 4.3).
Whatever the technique, registration is not exempt from errors. In the following we conduct the identification process both on noise-free and noisy data. Based on [**31**], [**32**], [**33**], noisy data are generated by multiplying at each time step m each component i of Y^m, FE approximation of the velocity field at time $m\Delta t$**,** by a factor $(1+\theta_i^m)$ where θ_i^m is generated according to a uniform law in the interval $[-0.3; 0.3]$.

3.1.3. Tree structure

From $4D$ scans, upper airways can be segmented and used along with the parenchyma surface to propagate a conductive tree model with a patient-specific space filling approach [**35**], see Figure 8.

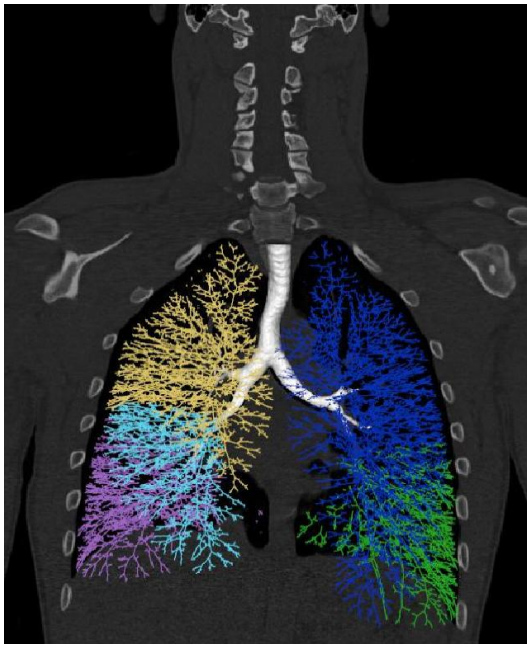

Figure 8: Space filling tree obtained based on the work presented in [35**]. Segmented upper airways and lung envelope are used along to propagate a patient-specific $0D$ model of the conductive tree structure. Each color corresponds to a different lobe. Figure taken from [**15**].**

The resulting structure contains a faithful representation for segmented upper branches, and a tree representation model for the remaining airways. This model is built based on both general physiological knowledge and patient data; it generates nodes positions, airway lengths and radii. While stenoses may be seen on segmented upper airways [**11**], the propagated part of the tree does not account for pathological cases. Indeed there is no *a priori* knowledge on the distribution of constricted branches, which is specific to each patient. Nodes positions and airways lengths of the propagated part are thus from now on known. Airways radii $r_{healthy_{n,k}}$ are in a healthy configuration range (see *Remark 4*). Note that since upper airways geometry is known from segmentation, their resistance would not need to be included among the unknowns of the identification problem described in subsection 2.3. Though, in this study, we work on synthetic pathological trees with stochastic constrictions generated along the process described in subsection 2.2, and we consider all branches radii including proximal ones are *a priori* unknowns.

In the following, we work on ten-generation space filling trees. Downstream sub-trees are treated as symmetric units and characterized by equivalent resistances. Each unit is irrigated by one airway referred to as *tree exit*. Constrictions in such subtree would be accounted for by a resistance increase of the related exit. The resulting tree on the studied left lung contains 954 airways among which 477 are terminal branches. When simulating asthma, we generate $n_{closure}$ plugs with $n_{closure}$ randomly chosen in the interval $[\![1; 40]\!]$ (see subsection 2.1), this interval being set based on non published data from [**35**].

3.1.4. Mechanical properties

In this study, the parenchyma mechanical properties are assumed to be homogeneous and are set within the ranges provided in [**36**]: Young's modulus $E = 1256\,Pa$ and Poisson ratio $\nu = 0.4$ where E and ν are linked to the Lamé parameters by $\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2(1+\nu)}$. Parenchyma density is $\rho_{par} = 100\ kg.m^{-3}$. Air density and viscosity at 37°C are respectively 1.125 kg.m^{-3} and 1.895 10^{-5} Pa.s.

3.2. Numerical methods

Direct problem:
System (8) is solved through a conjugate gradient descent with Jacobi preconditioning, as described in [**15**].

Machine learning:
The machine learning step is performed with a random forest regressor. The number of elements in the forest is chosen to be 50 and other parameters (as proposed by scikit-learn library [**37**]) are set to their default value. In order to predict the airway radius reduction ratios of a patient, synthetic data are generated as described in subsection 2.4.2 and used for the learning phase.
In this study, we work on one healthy tree and 49 asthmatic ones, the mechanical parameters and constitutive relation being the same in every cases. We adopt a leave-one-out strategy: iteratively, one of these pathological trees is considered as unknown in the sense that the aim is to predict the radius reduction ratios of its airways. The 48 other ones are used to build the learning base, which thus contains more than 45 000 airways. Prediction results are shown in subsection 4.2.
Airways are ranked based on their predicted ratios, and this ranking is used for the unknown selection (see subsection 2.3). As mentioned in subsection 3.1.2, the identification problem is performed on both noise-free and noisy synthetic ventilation data.

Identification problem:
To solve the minimization problem $(\mathcal{P}_d)$ (see (10)) with constraint $P_h = p_h$, system (11) is penalized as follows:

$$BB^T_{penalized} = BB^T + \frac{1}{\varepsilon} I_h$$

where

$$[I_h]_{ij} = \begin{cases} 1, & if\ i = j = h, \\ 0, & else \end{cases}$$

with $\varepsilon > 0$ and small. The penalized matrix $BB^T_{penalized}$ is invertible. For the right handside of (11):

$$BS_{penalized} = BS + \frac{1}{\varepsilon} V_h,$$

$$[V_h]_i = \begin{cases} \frac{1}{\varepsilon} p_h, & if\ i = h, \\ 0, & else. \end{cases}$$

We finally solve $BB^T_{penalized} P = BS_{penalized}$ with the conjugate gradient method. Then, as described in subsection 2.3, an extracted system of (16) is solved ten times along the breathing cycle (see temporal resolution in subsection 3.1.2). At each resolution, radius reduction ratios of airways belonging to the unknown set are computed. If the average radius reduction ratio defined by (13) over the ten time points is below 0.3, the airway is classified as a plug. This bound is set based on the radius reduction ratio distributions (see subsection 2.2). Some airways that are actually severely constricted are indeed predicted as plugged (they are true positive), some are missed (thus are true negative), and some that are only moderately constricted are classified as plugged (thus are false positive).

4. Results

In this section we present some results relative to the direct problem (see subsection 2.2), the machine learning step (see subsection 4.2) and the identification problem (see subsection 4.3) on a human left lung, which thus contains two lobes, the left upper lobe (LU) and the left lower lobe (LL).

4.1. Results for the direct problem

Results of tree-parenchyma coupled model presented in subsection 2.2 have been extensively described in [**15**]. Here we present the results of two typical cases: a healthy case and a configuration in which a plug with radius reduction ratio 1/7 is applied on the branch that feeds lobe LU. Results shown here are plotted after a few cycles so that the periodic regime has been reached. In Figure 9 the volume expansion from the reference state is plotted for both cases. Recall that a Dirichlet boundary condition with sinusoidal dynamics is imposed (see subsection 3.1.2). This prescribes the total lung volume variation, which is thus also the same in the pathological case. Though, air distribution between both lobes is affected. In the stenosis case, lobe LU is less ventilated in favor of lobe LL. Note also that, along the breathing cycle, LU does not expel all the gas it contains: there is air trapping which is typical of asthma and COPD [**38**]. Finally, due to the resistance increase, ventilation of the diseased region is shifted in time. This has been reported in [**39**]. In Figure 10 we plot the scalar field $div(\boldsymbol{u})$ which represents the local relative volume expansion for the same both cases.

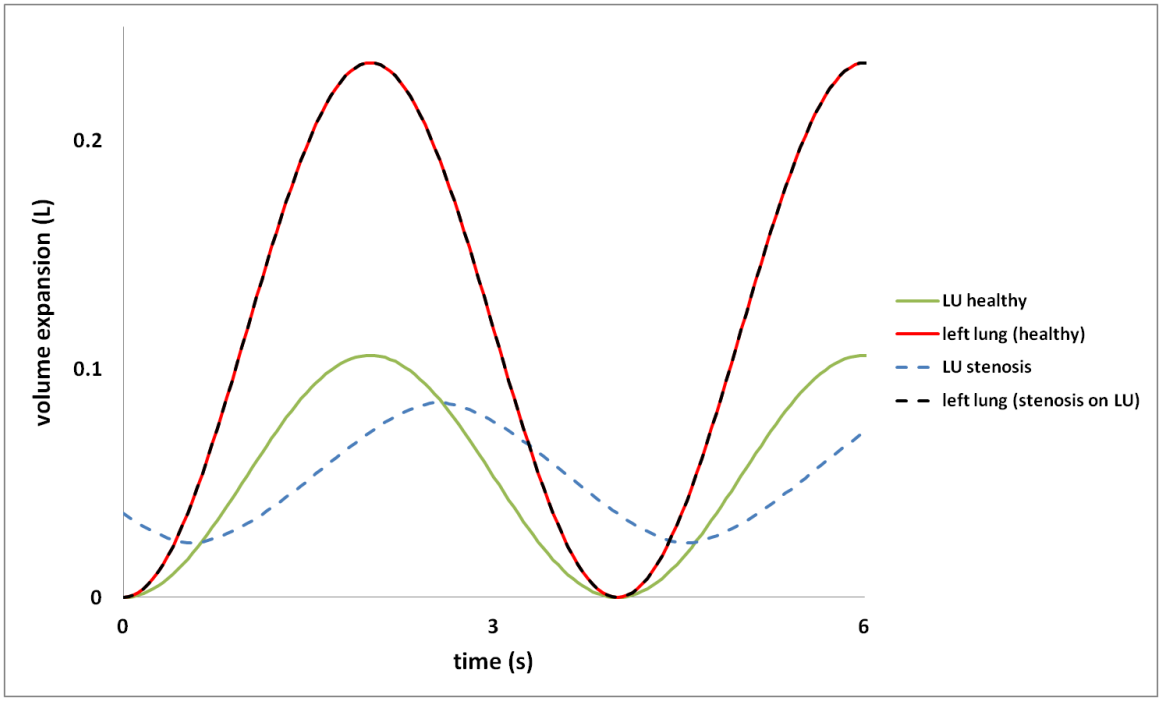

Figure 9: left lung and LU lobe volume expansions from the reference state in a healthy configuration and when a bronchoconstriction with constriction ratio 7 is applied to the branch that feeds lobe LU. Dirichlet boundary conditions with sinusoidal dynamics are applied to the system.

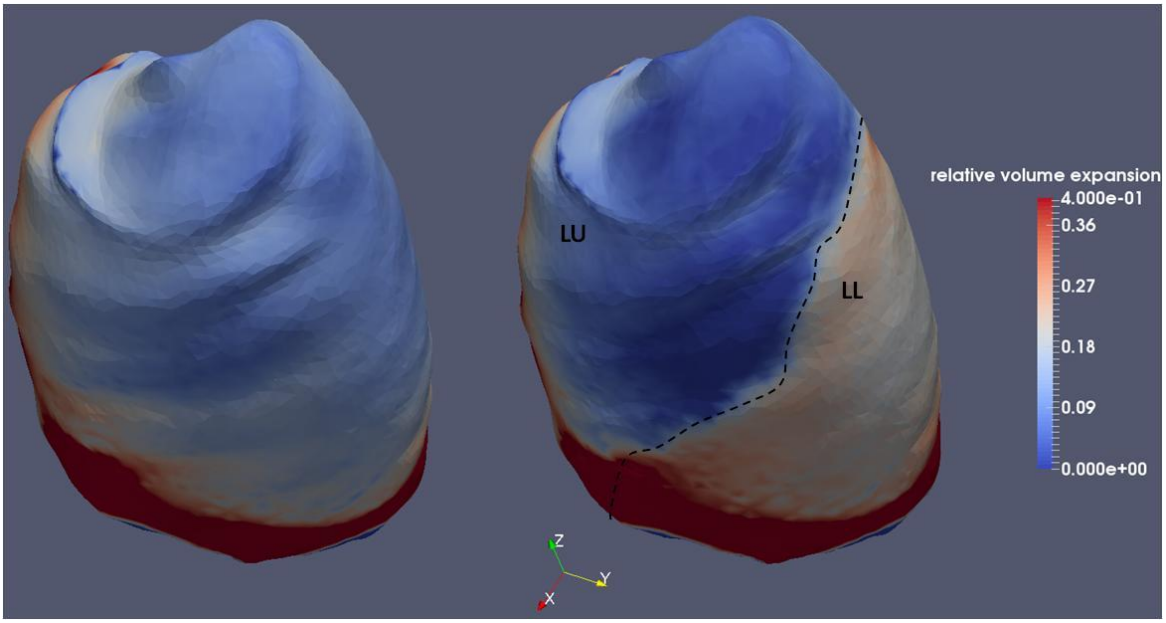

Figure 10: local relative volume expansion 3D map on a left lung geometry. Dirichlet boundary conditions with sinusoidal evolution are imposed. On the left side a healthy configuration is simulated. On the right side a bronchoconstriction with constriction ratio 7 is applied on the branch that feeds lobe LU. Plot at time=1s. Figure taken from [15].

Coherently with previous observations, lobe LU expands less in the pathological case. Note that diaphragmatic motion induces higher expansion in lower regions of the lung.

As explained in subsection 2.1, an asthmatic tree contains both few severe constrictions and numerous moderate constrictions. These closures modify the ventilation distribution. Here we want to compare the influence of moderate and severe bronchoconstrictions on ventilation distribution. To do so, we first build 49 stochastic asthmatic trees following the process described in subsection 2.2. Then we generate so-called "plugs only" trees built from the asthmatic trees by removing moderate constrictions, and "plugs free" trees built from asthmatic trees by removing the plugs. We compare first the ventilation distribution

obtained in the asthma and the "plugs free" cases, and then the ventilation distribution obtained in the asthma and the "plugs only" cases. In Figure 11, we plot for each of the 49 trees the ratio

$$diff_{plugsOnly} = \frac{100}{N}\sum_{i=0}^{N}\left|\frac{\left|V_i^{max} - V_i^{min}\right|_{onlyPlugs} - \left|V_i^{max} - V_i^{min}\right|_{asthma}}{\left|V_i^{max} - V_i^{min}\right|_{asthma}}\right|$$

where index i refers to tree exit i. The sum is over the N tree exit airways, here 477, and notations V_{max} and V_{min} have been defined in subsection 2.4.2. Similarly we define $diff_{plugsFree}$.

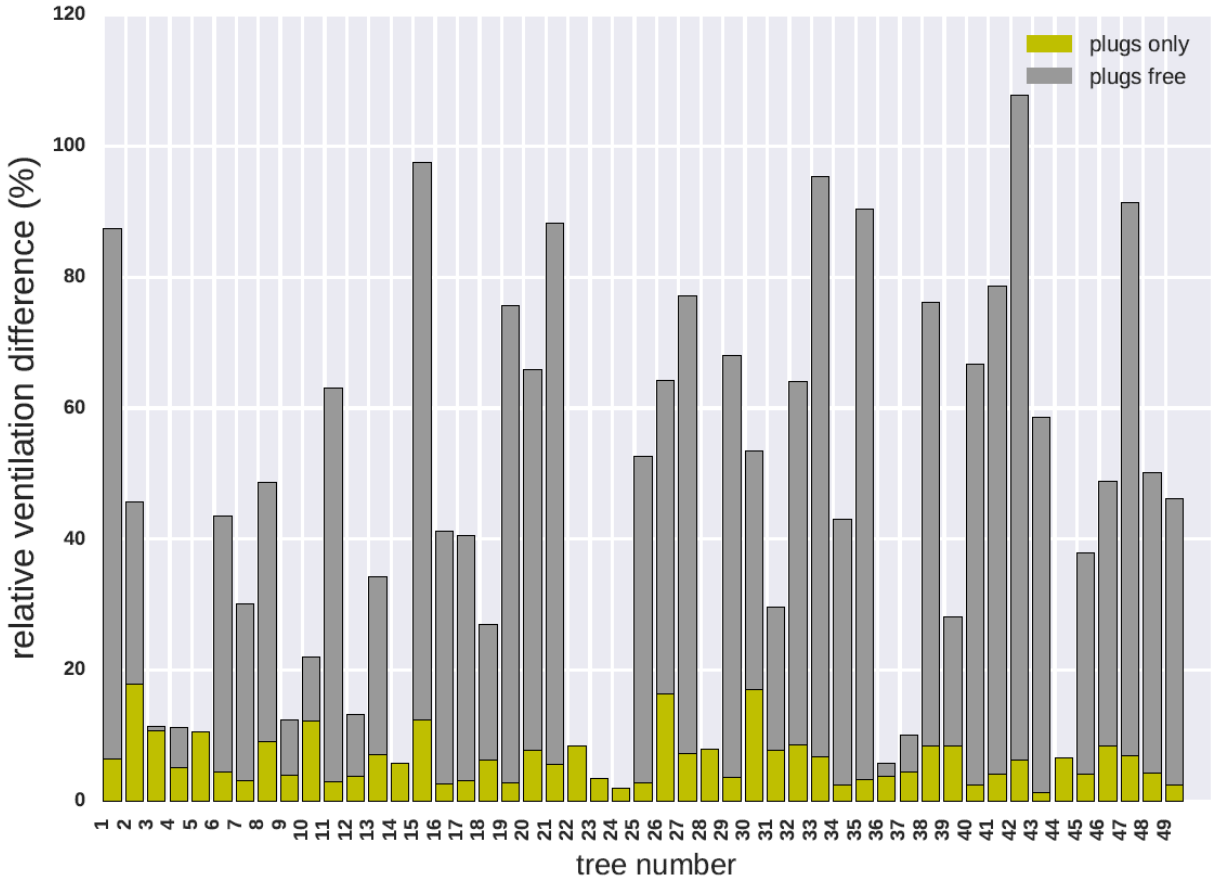

Figure 11: Plot of $diff_{plugsOnly}$ (yellow) and $diff_{plugsFree}$ (grey) for each of the 49 asthma trees.

We observe that in all cases the ventilation distribution obtained with plugs only is close to the one obtained with both plugs and moderate constrictions (6% difference on average). In most cases there is a high discrepancy between the asthma and the "no plugs" ventilation (45% difference on average). As a conclusion, plugs are much more influential than moderate constrictions on ventilation distribution in asthma. In some cases (trees 14 or 22 for example), plugs have little influence on the ventilation distribution. These are configurations in which the number of plugs is in fact low (see Figure 12). This gives ground to get insights on their position and severity based on $4D$ ventilation data.

4.2. Results for the machine learning step

In this section, we assess the quality of the machine learning step described in subsection 2.4.2. As explained in subsection 2.3, not all the tree resistances can be considered as unknowns of the process. First, among the 954 airways, N=477 have to be selected. A radius reduction ratio (which is in range $[0; 1]$, see equation (13)) is predicted for each airway and branches are sorted in increasing order. The smaller the ratio, the more likely the airway is constricted. In Figure 12, we plot, for each of the 49 synthetic asthma

trees, the total number of plugs and among them the number for which the predicted radius reduction ratio is sorted among the first 477.

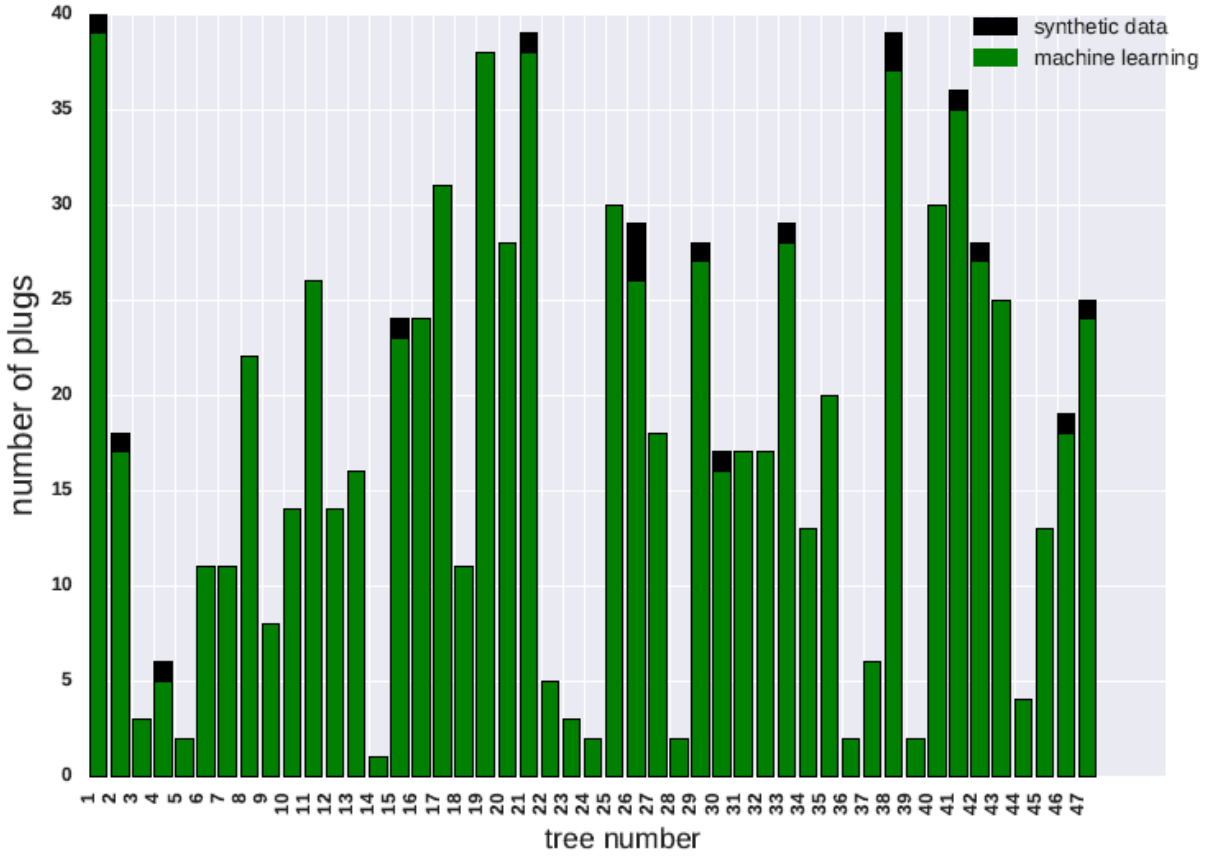

Figure 12: Machine learning classification success. The machine learning step predicts for each airway a radius reduction ratio. A low ratio corresponds to a severe constriction. Airways are sorted in increasing order with respect to their predicted ratio. Thus, each airway has a rank; the smaller the rank the smaller the predicted radius reduction ratio. For each asthmatic tree are plotted the total number of plugs (ground truth as they come from synthetic data), and among them the number of plugs that are classified with radius constriction ratio in the first half by the machine learning step. Trees contain 954 airways.

At this stage, for all trees, a large majority of the plugged airways if not all have their resistances kept in the unknown set, 98.5% in average (see Table 1). Among these N unknowns, some will be removed to get an invertible extraction of system (16) (see subsection 2.3) based on the constriction likeliness ranking provided by the machine learning step. For airway with rank k, we call $\frac{k}{N}$ its normalized rank. The smaller the normalized rank, the less likely the airway is removed from the unknown set. The *a priori* prediction process is judged to be good when plugged airways have a low normalized rank and are thus likely to be kept in the final unknown set. Note that in average on the 49 trees, after the resistance removal process (see Figure 3), there are 250 kept unknowns over the 954 initial tree airways.

In Figure 13 we plot, for each tree, the average normalized rank of plugged airways obtained when ventilation features are extracted from noise-free and noisy $4D$ ventilation data. Note that the average normalized rank depends on the number of plugs within the tree which goes from 1 to 40 (see subsection 3.1.3). If a tree contains one plug only, as for tree number 24, the best achievable average rank is obtained when the plug ranks first and its value is 1/954≈0.1%. If a tree contains 40 plugs, as for tree number 1, the best achievable average rank is obtained when the plugs rank among the 40 first and its value is 20/954≈2%.

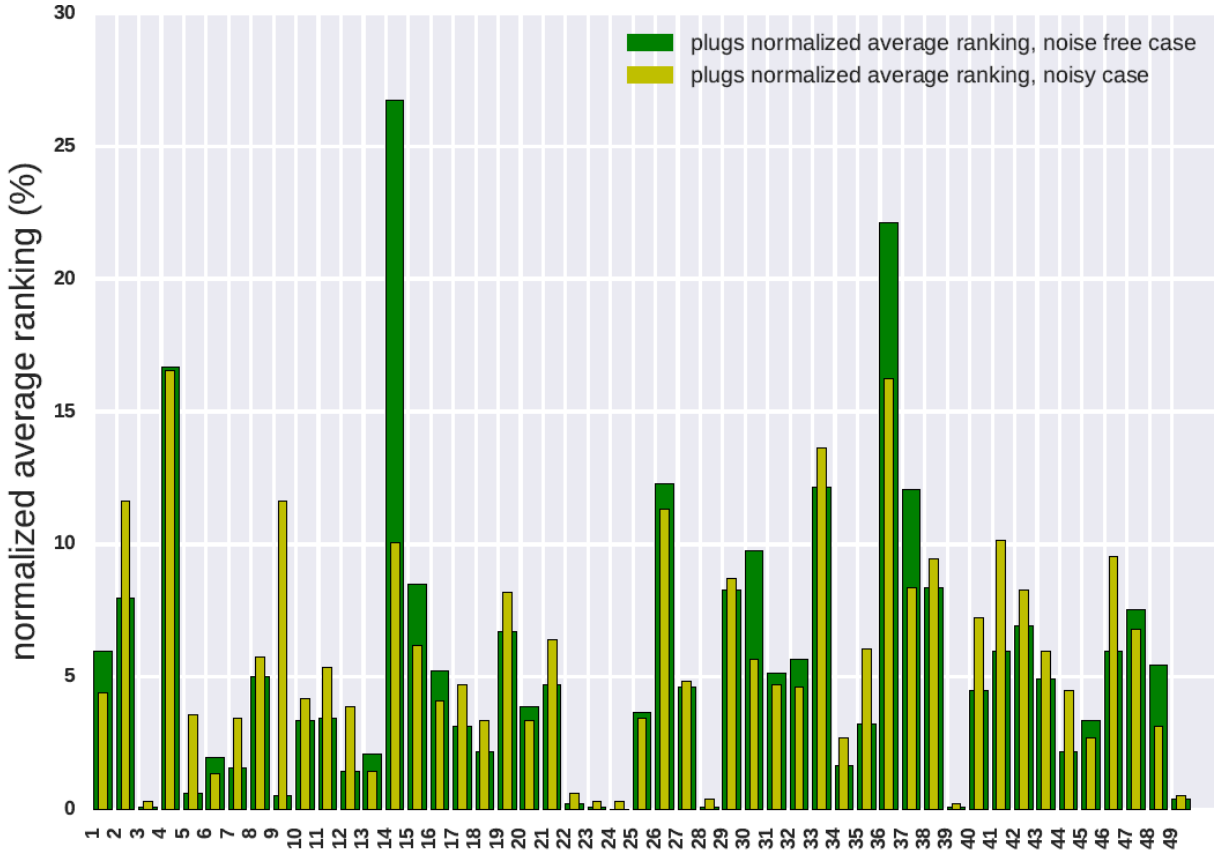

Figure 13: Normalized average ranking of plugged airways by the machine learning step. The machine learning step predicts for each airway a radius reduction ratio. A low ratio corresponds to a severe constriction. Airways are sorted in increasing order according to their predicted ratio. Thus, each airway has a rank; the smaller the rank, the smaller the predicted constriction ratio. The average predicted rank of plugged airways divided by 954, the total number of airways, is plotted for each asthmatic tree. The lower this normalized average rank, the more likely resistances of plugged airways will be kept in the unknown set during the removal process (see subsection *2.3*). In green (resp. yellow) results obtained when ventilation features are extracted from a noise-free (resp. noisy) $4D$ ventilation data.

The machine learning step is consistent with noise. In about 90% of the cases plugged airways are ranked in the first decile, for both the noise-free and noisy case. The average rank is good, thus airways that are actually plugged are not likely to be removed during the selection process described in Figure 4. Although the machine learning step provides good ranking, it does not accurately predict airways constriction. The identification problem is needed to determine which branches are actually plugged.

4.3. Results for the identification problem step

In this section are presented the results of the identification problem step (see subsection 4.3) without and with noise added to the velocity field coming from 4D data. As mentioned in section 2.1 the inputs come from synthetic data generated by the tree-parenchyma coupled model (see subsection 2.2). Simulations are performed over one breathing cycle. The identification problem classification process is performed as described in subsection 2.3.

In Figure 14, we plot for each asthmatic tree the number of plugged airways that are properly identified as plugged by the identification problem step divided by the total number of plugged airways, and the number of false positive divided by the total number of plugged airways. As explained in subsection 3.2, an airway is classified as plugged by the identification problem if its average computed radius reduction ratio is below 0.3. Most of the plugs are detected. In some cases, airways radius reduction ratios are too much overestimated by the identification problem step and they are not classified as plugged although they should be. For a few airways, the ratio is too much under-estimated and they are classified as plugged

although they should not be. Thus, they appear as false positive. The identification problem detection rate quantifies the proportion of plugged airways that are properly classified after the identification problem step. The false positive detection rate quantifies the proportion of airways classified as plugs while they are actually moderately constricted at most.

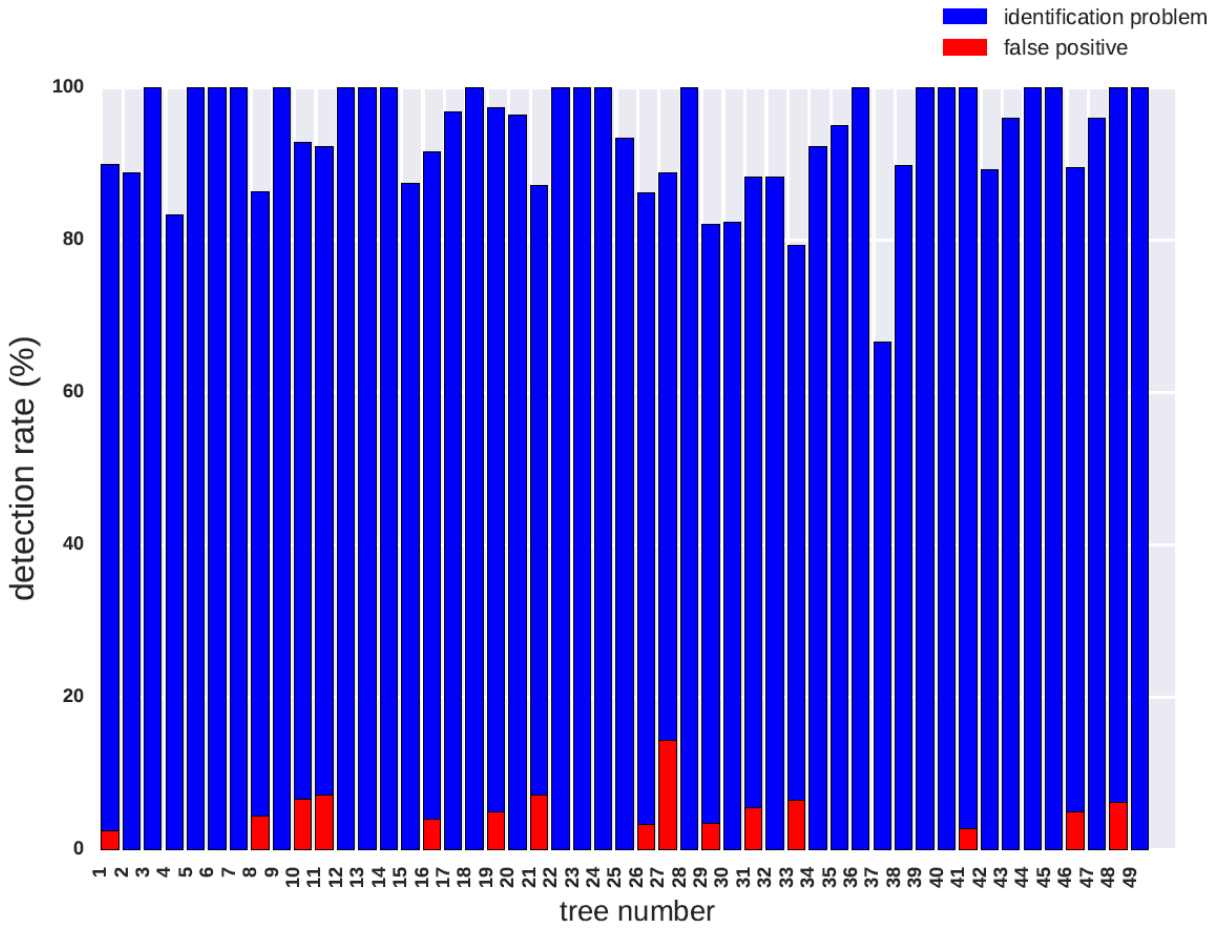

Figure 14: plugs detection for the noise-free case. For each of the asthmatic tree are plotted: in blue the proportion of plugged airways that are indeed predicted as plugged by the identification problem step (true positive), and in red the proportion that are predicted as plugged by the identification problem step whereas in practice they are moderately constricted at most (false positive).

Average detection rates are summed up in Table 1. A more detailed analysis shows that some plugs have little influence on ventilation distribution: they will logically not be detected by the process. This can be because they are in regions where some other plugs have more impact or because they affect downstream generations where low flows result in low pressure drops even in case resistances are increased (see Figure 18). When analyzing the detection rate, one shall consider that if the tree contains a few plugs only, one missed airway notably depreciates the classification rate. In tree 37 for example, 4 plugs over 6 are actually properly classified.

Note that the space-filling tree is a model. Specifying which precise airway is plugged would require a perfect knowledge of the tree architecture even in downstream generations, which is not achievable. More reasonably, we can deduce from the analysis insights on which lung regions and tree generations are affected by plugs. This is useful information when predicting particle deposition [**6**] or response to Heliox treatments [**5**]. In Figure 15, we plot the true positive detection rate as a function of the false positive detection rate in two cases: assuming a false positive is actually properly classified if its mother or daughters airways are plugged, and the strict case from Figure 14. As summarized in Table 1 and Figure 17 the detection rate is slightly improved by 0.5% while the false positive rate is reduced by 29%.

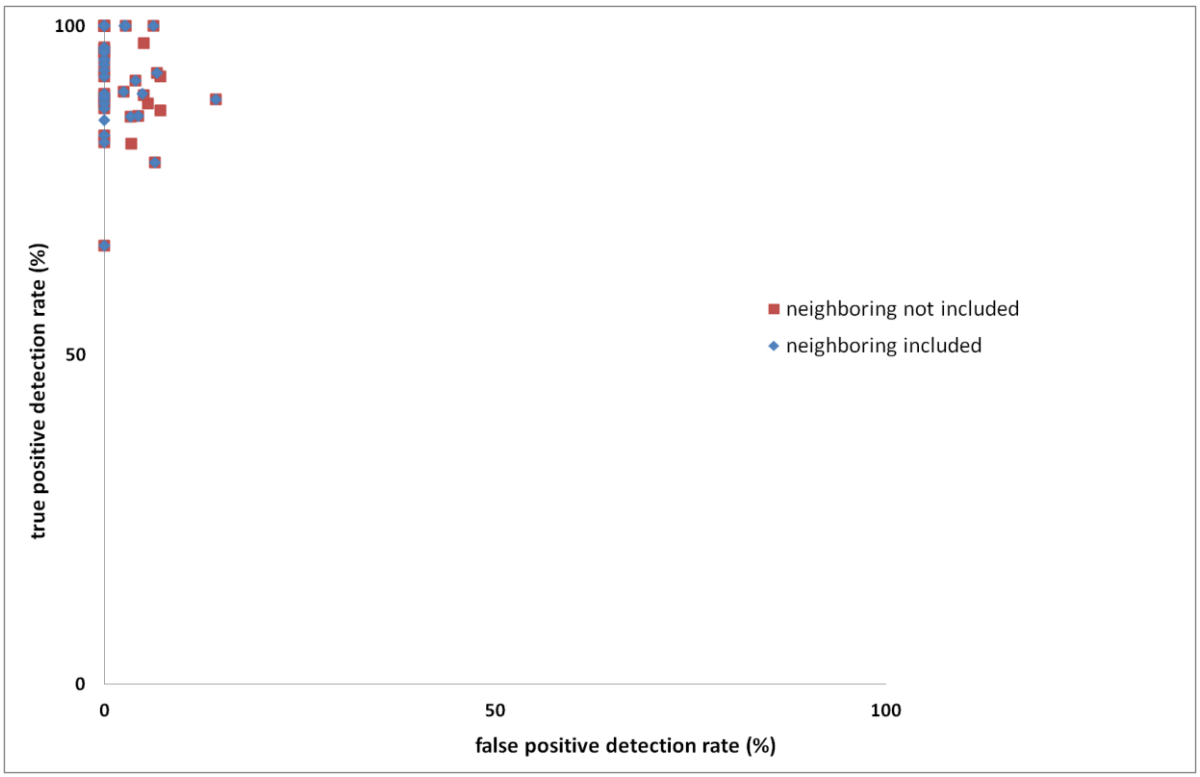

Figure 15: true positive detection rate as a function of false positive detection rate plotted for each of the asthmatic tree. In blue, results obtained when including airway neighboring on the analysis: it is considered that if a non severely constricted airway is predicted as plugged by the identification problem step and has a mother or daughter that is actually plugged, this airway is a true positive rather than a false positive. In blue results obtained when the neighboring is not included in the analysis. In both cases, the identification problem is performed on noise-free data.

In the following, we perform a similar analysis in case the velocity field used as input of the identification problem step is noisy (see subsection 3.1.2), shown in Figure 16. As expected, detection rates are reduced compared to the case without noise. However a large majority (86.5%) of the plugs are detected and the amount of false positive remains low (see Table 1).

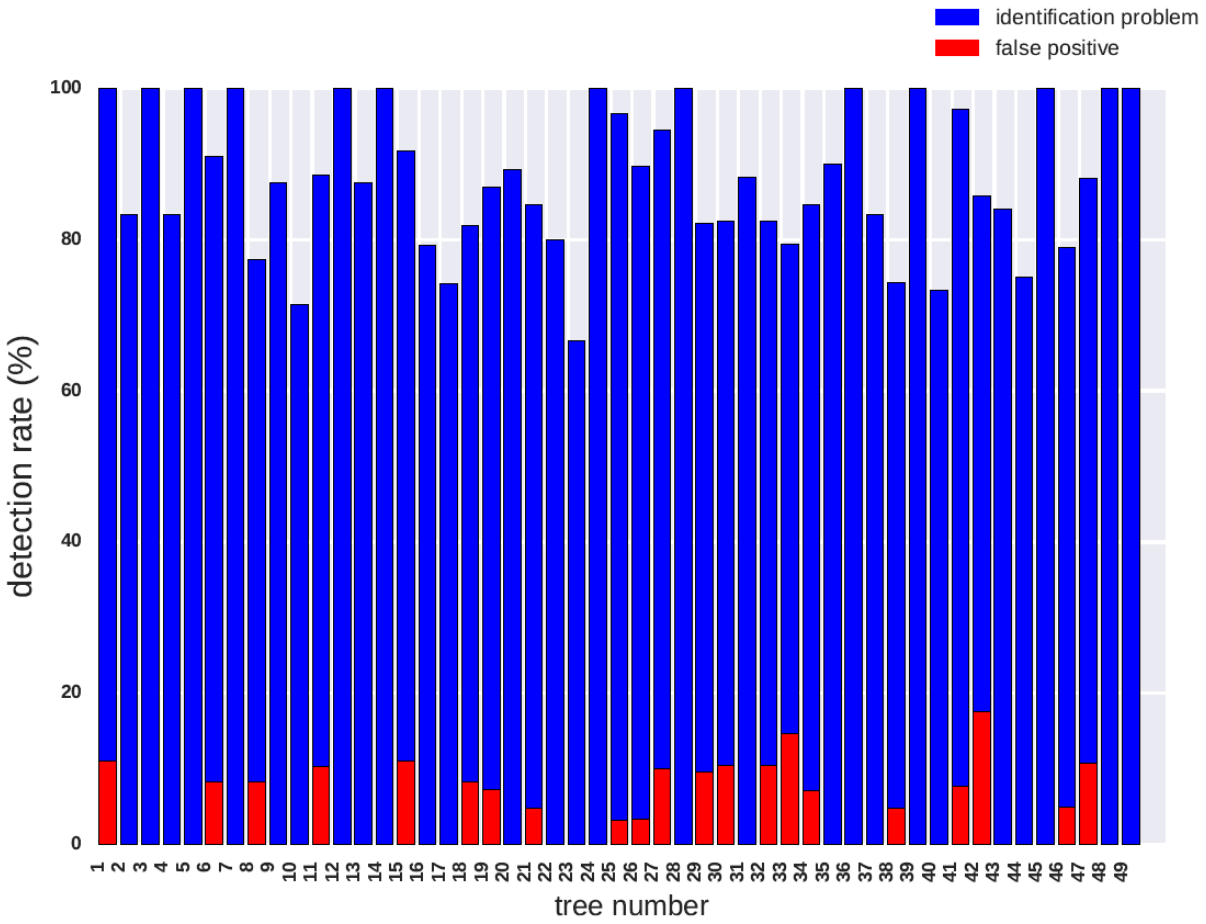

Figure 16: plugs detection ratio for noisy displacement data. For each of the asthmatic tree are plotted in blue the proportion of plugged airways that are indeed predicted as plugged after the identification problem step, in red the proportion that are predicted as plugged by the identification problem step whereas in practice they are moderately constricted at most; these are false positive. If a moderately constricted airway is predicted as plugged and it has a mother or daughter that is actually plugged, classification is judged to be good and this airway is not considered as false positive.

In Figure 17, detection rates obtained in the noise-free and noisy cases are compared.

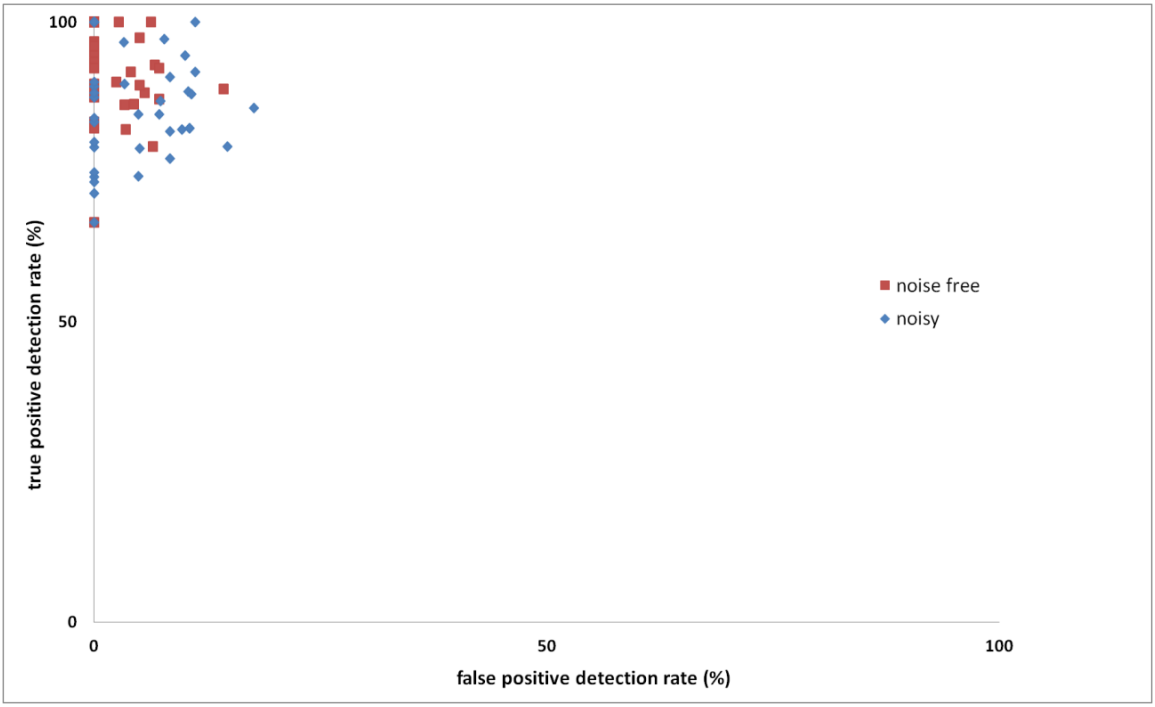

Figure 17: true positive detection rate as a function of false positive detection rate plotted for each of the asthmatic tree. In red, results obtained when the inverse problem step is performed on noise-free data, in blue results obtained when it is performed on noisy data.

	Noise-free			***With noise***		
	ML	IP	FP	ML	IP	FP
No neighbors	98.5	93.8	1.7	98.5	85.2	6.0
Neighbors	98.5	94.3	1.2	98.5	88.0	3.8

Table 1: Summary of plug classification and detection results, in %, by the machine learning step (ML), by the identification problem step (IP), and proportion of false positive (FP). Note that at the ML step, a plug is considered to be properly classified if its predicted radius reduction ratio ranks in the first half (see subsection 3.2) whereas a plug is considered to be properly classified by the IP step if its average computed radius reduction ratio over the ten time points is under 0.3 (see subsection 3.2). Results given for both noise-free and noisy displacement data.

In the following, we analyze the detection rate by generation. Plugs in upper airways have a high influence on ventilation distribution [**5**]. This is not the case when constrictions occur in distal generations. It is thus expected that dynamical ventilation data enable to access constriction information in proximal tree regions, which are critical, rather than in downstream areas which are less impacting. In Figure 18 are plotted the identification problem detection rates per generation. Note that we work on a half lung, fed by an airway belonging to generation one. In simulated cases, there is no constriction on this airway since there needs to be an unknown-free path for the system to be invertible (see subsection 2.3). In practice, bronchoscopy enables to trace plugs in the most proximal airways [**40**], which can thus be removed from the unknown set.

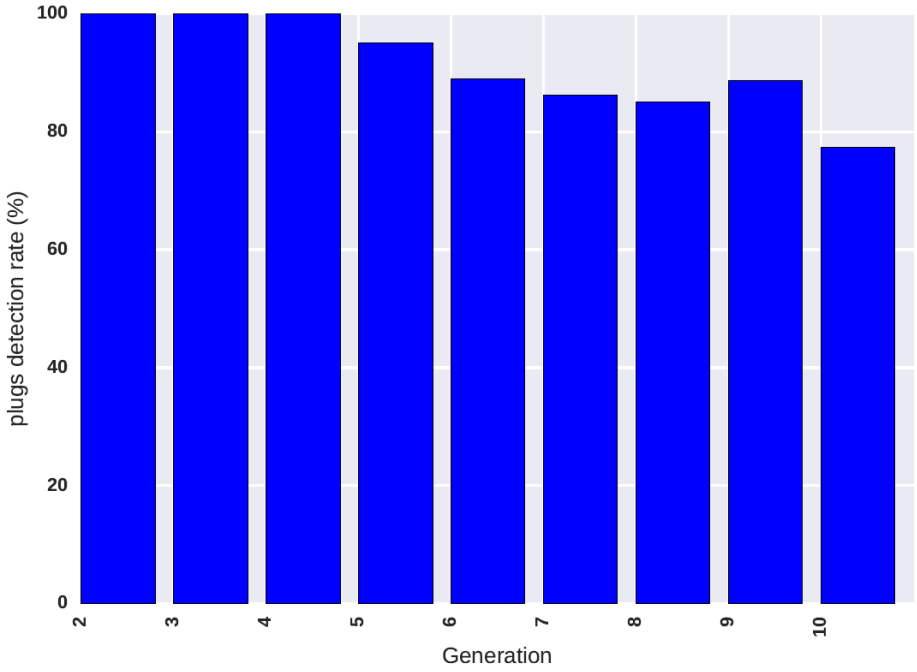

Figure 18: plugs detection rate per generation, for noisy displacement data. For each generation starting at two (one being the trachea) are plotted the identification problem plug detection rates in average on the 49 asthmatic trees.

Plugs in proximal generations are all detected. The detection rate drops under 80% at generation 10. The tree has a dyadic structure, and thus there are about two times more airways at generation $n+1$ that at generation n. So a bad detection rate at a distal generation has a large influence on the overall detection rate result although plugs in these airways have less impact on ventilation distribution. In Table 1 detection rates are summarized. When not considering small airways (commonly defined as these with internal diameter inferior to 2mm [**13**], starting around generation 8 [**19**]) more than 90% of the plugs are detected by the process.

5. Limits and conclusions

Limitations of the ventilation model used in this study are pointed out in [**15**]. We assume a linear elastic behavior of the lung parenchyma which is likely to be valid for low tidal breathing frequency and displacement regimes. Although there is no consensus on which constitutive relation shall be considered [**41**], [**42**], [**43**], the actual parenchyma law is much more complex. The identification problem approach proposed in this chapter gives insights on the tree structure assuming mechanical properties are known, which is not the case in practice, here representative values from the literature being assumed.
As mentioned in subsection 3.1.3, we assume the generated space-filling tree is a proper approximation of the real one in terms of topology. Although it has good statistical properties [**44**], this remains a model. Validation of this topology and of the predicted constriction ratios on imaging data of high resolution is warranted.
Another limitation of our work is, as stated in subsection 2.2, the use of the Pedley resistance model at expiration. In addition airways are assumed to be rigid which is not the case in practice. Their dimensions may evolve along the breathing cycle, though this has little impact in tidal regime (see[**15**]) and a quasi-static evolution would be compatible with our framework.

In conclusion, we proposed a plug detection approach that couples machine learning and ventilation modeling. Considering common 4D-CT/4D-MRI temporal and spatial resolutions we obtain very encouraging results. More than 85% of the plugs are detected and when not considering severe

constrictions in small airways, which have a lower impact on ventilation distribution, the detection rate reaches more than 90%.
Some in-vivo techniques aim at determining the parenchyma mechanical properties based on image analysis [**45**], though in cases where the ventilation is strongly affected by pathological patterns within the tree, the tree-parenchyma coupling shall be taken into account. Indeed, some regions with normal mechanical behavior may not inflate because they are irrigated by constricted paths. To that end identification problems for parenchyma mechanical properties and plugs distribution may be coupled. Finally, in this study, we tested the process with synthetic data, but future works shall include application to real imaging and validation based on experimental data.

Appendix

A.1 On the dimension of $ker(B^T)$

Let us compute the kernel of B^T (see subsection 2.2 for the definition of matrix B). For a vector $P = \begin{pmatrix} p_1 \\ \vdots \\ p_N \end{pmatrix}$ we have

$$(B^T P)_j = \left(\sum_i p_i \int_{\Omega_i} div(\boldsymbol{e_j}) \right)_j$$

where $(\boldsymbol{e_j})_j$ are the finite element basis functions. With the imposed Dirichlet boundary conditions, they are such that $\int_\Omega div(\boldsymbol{e_j}) = 0$. Let P be a non null vector such that $B^T P = 0$. Let us consider the k^{th} element of $B^T P$:

$$(B^T P)_k = \sum_i p_i \int_{\Omega_i} div(\boldsymbol{e_k}).$$

For e_k with support included in Ω_l, we have for $i \neq l$ that $\int_{\Omega_i} div(\boldsymbol{e_k}) = \int_\Omega div(\boldsymbol{e_k}) = \int_{\partial\Omega} \boldsymbol{e_k}.\boldsymbol{n} = 0$ since $\boldsymbol{e_k} = 0$ on $\partial\Omega$. And $\int_{\Omega_l} div(\boldsymbol{e_k}) = \int_\Omega div(\boldsymbol{e_k}) = 0$. If $\boldsymbol{e_k}$ has a support intersecting the border between regions Ω_l and Ω_j then $\int_{\Omega_j \cup \Omega_l} div(\boldsymbol{e_k}) = 0$ and

$$(B^T P)_k = p_l \int_{\Omega_l} div(\boldsymbol{e_k}) + p_j \int_{\Omega_j} div(\boldsymbol{e_k}) = (p_l - p_j) \int_{\Omega_l} div(\boldsymbol{e_k})$$

So $(B^T P)_k = 0 \Rightarrow p_l = p_j$. As this is true for any frontier between neighboring regions, we conclude

$$B^T P = 0 \Rightarrow \exists\, p_0 \in \mathbb{R} \,/\, P = p_0 \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

The reverse is trivial.

So $\ker(B^T) = span\left\{ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right\}$ and $dim(ker(B^T)) = 1$.

A.2 On the reason why if Z_b is not invertible, it has some collinear lines.

Let Z_b the matrix defined by (15) (see subsection 2.3). A given line i of Z_b contains 1 at position j if the two following conditions are satisfied: resistance R_j belongs to the set of unknown resistances, and the airway associated to R_j belongs to path T_i; otherwise the line contains 0 at position j.
Lemma: assume $rank(Z_b) < b$, then any line of Z_b cannot be linear combination of several of the others and thus, necessarily, Z_b has some collinear lines.

Proof of the lemma: we denote $\mathcal{U}$ the set of airways $(a_i)_{i=1\ldots b-1}$ excluding the mother branch and which resistances are unknowns. A path T_l goes from the mother branch down to exit l. Along this path, only a subset $\mathcal{A}_l$ of airways belongs to $\mathcal{U}$. Let us denote g_l the l^{th} line of Z_b.
Let us note, $m_l = \sum_k [Z_b]_{lk} = \sum_k g_{l_k} = cardinal(\mathcal{A}_l)$ and $\mathcal{G} = \{g_l \,/\, \forall j \neq l, m_l \geq m_j\}$ the set of lines containing most ones. We randomly select a line g_s in $\mathcal{G}$. We will show that g_s cannot be a linear combination of several vectors $g_{t\neq s}$. Let a_d be the most distal airway of $\mathcal{A}_s$. Path T_s is the only one that contains a_d. If another path Γ_q contained a_d, because a_d is the most distal branch of T_s, the tree structure would impose $\mathcal{A}_s \subset \mathcal{A}_q$. If $\mathcal{A}_s \neq \mathcal{A}_q$ then $m_q > m_s$ which is contradiction with the fact $g_s \in \mathcal{G}$. Thus $\mathcal{A}_s = \mathcal{A}_q$. Finally we have that $g_{s_d} = 1$ and $\forall j \neq s, g_{j_d} = 0$, so g_s cannot be a linear combination of several $g_{t,\forall t\neq s}$. Repeating the same process $b-1$ times with vectors $g_{t\neq s}$ we conclude that any line of Z_b cannot be linear combination of several the others.

References

[1] C Kotaru et al., "Morphometric changes after thermal and methacholine bronchoprovocations," *J Appl Physiol*, no. 98(3):1028-36, 1985.

[2] M Montaudon et al., "Bronchial measurements in patients with asthma: comparison of quantitative thin-section CT findings with those in healthy subjects and correlation with pathologic findings," *Radiology*, vol. 253(3):844-53, 2009.

[3] D Leary et al., "Hyperpolarized He3 magnetic resonance imaging ventilation defects in asthma: relationship to airway mechanics," no. 2016 Apr;4(7), 2016.

[4] J, Parameswaran, H Lui, M Albert, and K Lutchen, "Linking ventilation heterogeneity quatified via hyperpolarized He3 MRI to dynamic lung mechanics and airway hyperresponsiveness," *PLos One*, no. 16;10(11):e0142738, 2015.

[5] N Pozin et al., "Calculated ventilation and effort distriubtion as a measure of respiratory disease and Heliox effectiveness (submitted)," *JBiomech*, 2017.

[6] S Taheriann, H Rahai, BZ Gomez, T Waddington, and JR Bonifacio, "Tracheal stenosis: a CFD approach for evaluation of drug delivery.," in *ASME International Mechanical Engineering Congress and Exposition*, 2015.

[7] N Jahani et al., "Comparison of regional ventilation and lug deformation for asthmatic vs. healthy human lungs using 4d-CT image data," *J Appl Physiol*, no. 15;119(10):1064-74, 2016.

[8] N Jahani et al., "A four-dimensional computed tomography comparison of healthy and asthmatic human lung," *J Biomech*, no. 3;56-102-110, 2017.

[9] C Kolb et al., "Regional lung ventilation analysis using temporally resolved magnetic resonance imaging," *J Comput Assist Tomogr*, no. 40(6):899-906, 2016.

[10] J Biederer et al., "MRI of the lung (2/3). Why. when. how?," *Insights Imaging*, no. 3(4):355-71, 2012.

[11] P Grenier, C Fetita, and PY Brillet, "Quantitative computed tomography imaging of airway remodelling in severe asthma," *Quant Imaging Med Surg*, no. 6(1):76-83, 2016.

[12] S Montesantos et al., "Airway morphology from high resolution computed tomography in healthy subjects and patients with moderate persistent asthma," *The anatomical record*, no. 296:852-866, 2013.

[13] PR Burgel, J de Blic, and P Chanez, "Update on the roles of distal airways in ashtma," *Eur Respir Rev*, no. 18:80-95, 2009.

[14] DH Carr, S Hibon, M Rubens, and KF Chung, "Peripheral airways obstruction on high-resolution computed tomography in chronic severe asthma," *Respir Med*, no. 92(3):448-53, 1998.

[15] N Pozin et al., "A tree-parenchyma coupled model for lung ventilation simulation," no. 10.1002/cnm.2873, 2017.

[16] S Sankaran, L Gady, and CA Taylor, "Fast computation of hemodynamic sensitivity to lumen segmentation uncertainty," *IEEE trans med imaging*, no. 34(12):2562-71, 2015.

[17] A Lungu et al., "Diagnosis of pulmonary hypertension from magnetic resonance imaging-based computational models and decision tree analysis," *Plum circ*, no. 6(2):181-90, 2016.

[18] T Aikawa, S Shimura, H Sasaki, M Ebina, and T Takishima, "Marked goblet cell hyperplasia with mucus accumulation in the airways of patients who died of severe acute asthma attack," *Chest*, no. 101(4):916-21, 1992.

[19] TT Soong, P Nicolaides, CP Yu, and SC Soong, "A statistical description of the human tracheobronchial tree geometry," *Respir physiol*, no. 37(2):161-72, 1979.

[20] B Maury, *The Respiratory System in Equations*.: Springer, 2013, vol. 7.

[21] M Ismail, A Comerford, and WA Wall, "Coupled and reduced dimensional modeling of respiratoy mechanics during spontaneous breathing," *Int. J. Numer. Meth. Biomed. Engng*, no. 29(2013):1285-1305, 2013.

[22] TJ Pedley, RC Schroter, and MF Sudlow, "Energy losses and pressure drops in models of human airways," *Respir Physiol*, no. 9(3):371-86, 1970.

[23] MY Kang, J Hwang, and JW Lee, "Effect of geometric variations on pressure loss for a model bifurcation of the human lung airway," *J.Biomech*, no. 44(6):1196-9, 2011.

[24] P Cazeaux and C Grandmont, "Homogenization of a multiscale viscoelastic model with nonlocal damping, application to the human lungs," *Mathematical Models and Methods in Applied Sciences*, vol. 25(06):1125–1177, 2012.

[25] YK Mariappan, KJ Glaser, and RL Ehman, "Helium-3 magnetic resonance elastography of the lung," *Clin anat*, no. 23(5):497-511, 2016.

[26] EE Greenblatt et al., "Analysis of three-dimensional aerosol deposition in pharmacologically relevant terms: beyound black or white ROIs," *J Aerosol Med Pulm Drig Deliv*, no. 28(2):116-29, 2014.

[27] Meshlab - developped with the support of 3D-CoForm project - meshlab.sourceforge.net.

[28] C Geuzaine and JF Remacle, "Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing," *Int. J. Numer. Meth. Engng*, vol. 79(11):1309-1331, 2009.

[29] EE Greenblatt et al., "Analysis of three-dimensional aerosol deposition in pharmacologically relevant terms: beyound black or white ROIs," *J Aerosol Med Pulm Drig Deliv*, no. 28(2):116-29, 2014.

[30] S Durrleman et al., "Morphometry of anatomical shape complexes with dense deformations and sparse parameters," *NeuroImage*, vol. 101:35-49, 2014.

[31] V Boldea, GC Sharp, SB Jiang, and D Sarrut, "4D-CT lung motion estimation with deformable registration: quantification of motion nonlinearity and hysterisis," *Med Phys*, no. 35(3):1008-18, 2008.

[32] A Hilsmann et al., "Deformable 4DCT lung registration with vessel bifurcations," *International journal of computer assisted radiology and surgery*, no. 39015418, 2007.

[33] J Vandemeulebroucke et al., "Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT," *Med Phys*, no. 39(2):1006-15, 2012.

[34] J Cai, Z Chang, and Z Wang, "Four-dimensional magnetic resonance imaging (4D-MRI) using image-based respiratory surrogate: A feasibility study," *Med Phys*, no. 38(12):6384-94, 2011.

[35] S Montesantos, I Katz, M Pichelin, and G Caillibote, "The creation and statistical evaluation of a deterministic model of the human bronchial tree from HRCT images," *PLOS one*, 2016.

[36] MA Epstein and JR Ligas, *Respiratory biomechanics*, Springer-Verlag, Ed., 1990.

[37] F Pedregosa et al., "Scikit-learn: machine learning in python," *JMLR*, no. 12(2011):2825-2830, 2011.

[38] F Laurent, V Latrabe, C Raherison, R Marthan, and JM Tunon-de-Lara, "Functional significance of air trapping detected in moderate asthma," *Eur Radiol*, no. 10(9):1404-10, 2000.

[39] EE Greenblatt, JP Butler, JG Venegas, and T Winkler, "Pendelluft in the bronchial tree," *J Appl Physiol*, no. 117(9):979-88, 2014.

[40] H Hoppe et al., "Grading airway stenosis down to the segmental level using virtual bronchoscopy," *Chest*, no. 125(2):704-11, 2004.

[41] L Berger et al., "A poroelastic model coupled to a fluid network with applications in lung modelling," *Int J Numer Method Biomed Eng*, vol. 32(1):e02731, 2015.

[42] L Yoshihara, CJ Roth, and WA Wall, "Fluid-structure interaction including volumetric coupling with homogenised subdomains for modeling respiratory mechanics," *Int J Numer Method Biomed Eng*, no. 10.1002/cnm.2812, 2016.

[43] B Suki, D Stamenovic, and R Hubmayr, "Lung Parenchymal Mechanics," *Compr Physiol*, vol. 1(3):1317-1351, 2011.

[44] MH Tawhai et al., "CT-based geometry analysis and finite element models of the human and ovine bronchial tree," *J Appl Physiol*, no. 97(6):2310-21, 2004.

[45] AD Freed and DR Einstein, "An implicit elastic theory for lung parenchyma," *Int J Eg Sci*, no. 2013 Hen;62:31-47, 2013.