Mathematical Medicine and Biology: A Journal of the IMA (2023) **40**, 291–307
https://doi.org/10.1093/imammb/dqad006
Advance Access publication on 29 September 2023

Which airways should we treat? Structure–function relationships and estimation of the singular input modes from the forward model alone

Graham M. Donovan*
Department of Mathematics, The University of Auckland, Private Bag 92019, 1142, Auckland, New Zealand
*Corresponding author. Email: g.donovan@auckland.ac.nz

[Received on 21 February 2023; revised on 10 September 2023; accepted on 26 September 2023]

Structure–function relationships occur throughout the sciences. Motivated by optimization of such systems, we develop a framework for estimating the input modes from the singular value decomposition from the action of the forward operator alone. These can then be used to determine the input (structure) changes, which induce the largest output (function) changes. The accuracy of the estimate is determined by reference to the method of snapshots. The proposed method is demonstrated on several example problems, and finally used to approximate the optimal airway treatment set for a problem in respiratory physiology.

Keywords: singular value decomposition; randomized linear algebra.

1. Introduction

Structure–function relationships are a key concept in biology and physiology. The central idea is that there is 'a relationship between the structure of a biological entity, and the function of that entity' (Michael, 2021). This concept occurs throughout the sciences, and one area in which it is particularly prominent is in respiratory physiology: e.g. how does the structure of the lung and airways affect lung function, and in particular, how do changes of that structure in disease lead to dysfunction (Hoffman & McLennan, 1997; Skloot, 2017)?

In this paper, we outline a framework in which we think of the structure–function relationship as a mapping between the structure of an entity on the one hand, and the resulting function on the other. Doing so enables us to bring to bear some mathematical tools, such as the singular value decomposition (SVD), which can then be used to answer important questions about the biological system at hand.

We motivate this approach with an example from respiratory physiology. Asthma is a widespread disease in which airway narrowing is driven by contraction of the band of *airway smooth muscle* (ASM) surrounding the airways. Although a substantial majority of asthma cases can be controlled with the available suite of pharmacological therapies, the subset of cases that do not respond to these therapies suffers from significant mortality and morbidity. One relatively new treatment, *bronchial thermoplasty* (BT), targets the ASM layer by directly heating the airways to a level intended to permanently ablate the ASM, without causing undue damage to other airway structures. This is done directly by bronchoscope and radiofrequency catheter, which means that the treatment is limited to a relatively small number of relatively large airways.[1]

[1] Large airways because of the size of the catheter (Hackmann *et al.*, 2022), a small number because of limited treatment time.

More recently, efforts have been made to personalize BT by selecting an optimal set of airways for treatment for each patient, e.g. based on pre-treatment imaging of the patient (Donovan *et al.*, 2019; Hall *et al.*, 2020) rather than using a fixed set. This has the potential to improve patient outcomes by both increasing efficacy of the treatment, and by reducing side effects. The question, then, is this: given information about the pre-treatment airway structure of an individual patient, how can we determine the optimal set of treatment airways? This question motivates the proposed approach to estimating the singular modes from the forward model.

The paper is organized as follows. In the next section, we outline the mathematical framework and proposed algorithms. These are then demonstrated on a series of example problems to demonstrate the strengths and limitations of the method; finally, we apply the proposed method directly to the BT airway selection problem.

2. Framework

The framework is as follows. Suppose that we have a model,[2] which relates the structure and function of the system at hand

$$\vec{\text{function}} = \mathbf{M}\text{odel}\left(\vec{\text{structure}}\right).$$

For concreteness, we will think of both structure and function as being finite-dimensional, if only as a discrete approximation, and write

$$\vec{y} = M(\vec{x})$$

where

$$\vec{x} \in \mathbb{R}^n, \vec{y} \in \mathbb{R}^m$$

represent the structure and function, respectively, and $m, n > 1$.

We make several important assumptions:

1. We assume that we have only the action of M

$$x \mapsto M(x)$$

   and that we do not have access to the action of the adjoint operator $y \mapsto M^{\dagger}(y)$. That is, the model is essentially a black box; we can only alter the input (structure) and see how that changes the output (function).

2. We also assume that the computational cost of the forward model evaluation is the dominant effect. That is, we don't care about the linear algebra costs of manipulating the output in comparison. Moreover, we assume that we cannot afford, in terms of computational cost, $\mathcal{O}(n)$, evaluations of the forward model (the number of input dimensions).

[2] One might naturally think of a mathematical model to begin with, but there is nothing to preclude use of a *physical* model so long as there is adequate control of the inputs.

Within this context, consider the linearization of the model operator about a reference configuration x_0

$$M(x + x_0) \approx Ax + M(x_0), \tag{2.1}$$

where A is an m-by-n matrix. Bear in mind that because we only assume the action of M, we cannot form A explicitly at this stage. (Note also that the linear approximation is a significant assumption, which we will discuss on more depth later on.)

Now we seek the SVD, or more specifically the *truncated* SVD of A given by

$$A = U_r \Sigma_r V_r^\top. \tag{2.2}$$

Briefly, the 'full' SVD ($A = U\Sigma V^\top$) consists of the input modes in the matrix V, the singular values along the diagonal of the square matrix Σ and the output modes in U. The SVD itself makes no approximation, but a reduced-rank approximation is possible by including only the largest singular values (and the associated input and output modes). Thus, the truncated SVD is truncated at rank $r \ll m, n$; the accuracy of the approximation clearly depends upon the spectrum of the singular values, and how much of the 'energy' of the singular values is captured within the truncation (Brunton *et al.*, 2022).

In particular, motivated by our question about which airways to treat, we want the input modes (V_r) and the associated singular values. These, in essence, tell us which changes to the input (structure) induce the largest changes in the output (function), and by how much (the associated singular value).

However, this is not straightforward to achieve, because we do not actually have the matrix A itself; recall that we assume only access to the forward action[3] $x \mapsto M(x)$.

Of course, one could create a naïve reconstruction of A with sufficient evaluations of the forward action by sampling the input space, e.g.

$$M(x_0 + \delta\{I_n\}) \approx \delta\{A\} + M(x_0) \tag{2.3}$$

and so

$$\{A\} \approx \frac{M(x_0 + \delta\{I_n\}) - M(x_0)}{\delta}, \tag{2.4}$$

where the notation $\{I_n\}$ indicates that the action is taken column-by-column. It may be more intuitive to write this explicitly as

$$A \approx \frac{1}{\delta}\left[M(x_0 + \delta e_1), M(x_0 + \delta e_2), \ldots, M(x_0 + \delta e_n)\right] - M(x_0)/\delta, \tag{2.5}$$

where e_k are the usual orthogonal basis functions (or columns of I_n). This requires $n + 1$ evaluations of the forward action $x \mapsto M(x)$. Here we have used the standard basis but transformations are possible; more anon.

[3] There are methods that can calculate, or approximate, the SVD from the action of the forward model $x \mapsto M(x)$ *and* the action of the adjoint operator $y \mapsto M^\dagger(y)$, such as Krylov subspace methods (Stoll, 2012) or the randomized SVD (rSVD) (Halko *et al.*, 2011). However, because we do not assume access to the action of the adjoint, we cannot make use of these methods.

What if we can't afford n evaluations? For our lung application, if one thinks of the input (structure) space as being the ASM content of the conducting airways, then $n \sim 60,000$, but computational costs mean that the practical limit is much lower. Let us suppose then we have a computational budget of $N \ll n$ evaluations of the forward action. In general, without the full matrix A, or access to the adjoint operator, we cannot expect to construct the SVD input modes, but nonetheless we seek a way to make an informed partial sampling. Thus, we truncate to N input dimensions as

$$A_N \approx \frac{M\left(x_0 + \delta T_N(I_n)\right) - M(x_0)}{\delta}, \tag{2.6}$$

where T_N allows a basis transformation informed by the problem at hand. We will revisit the choice of T_N later; for now, one can imagine a fully naïve version where $T_N(I_n) = I_N$. Again, this can be written out column-by-column as

$$A \approx \frac{1}{\delta}\left[M(x_0 + \delta T_N(e_1)), M(x_0 + \delta T_N(e_2)), \ldots, M(x_0 + \delta T_N(e_N))\right] - M(x_0)/\delta. \tag{2.7}$$

However, this leaves us with no information about the accuracy of the truncation. Is N big enough? Is $T_N()$ well-chosen? How can we know the answer to these questions?

Our proposed approach is to use the *method of snapshots* (Sirovich, 1987) to sample the range of M and thus estimate the *output* modes and spectrum.[4] Then comparison between the output modes and spectrum from the naïve method, and from the method of snapshots, gives us an estimator for the quality of our input mode approximation.

Here we briefly review the method of snapshots, originally due to Sirovich (1987) in the context of fluid dynamics (but see also Brunton *et al.*, 2022, for a more general presentation). We begin with random vectors

$$\omega_1, \omega_2, \ldots, \omega_N, \tag{2.8}$$

where $\omega_j \in \mathbb{R}^n$ (i.e. in this context the input/structure space). These are usually drawn from a suitable Gaussian distribution. We then compute the action of these random samples $M(\omega_i + x_0)$ to assemble the matrix

$$Z = \left[M(\omega_1 + x_0) - M(x_0), M(\omega_2 + x_0) - M(x_0), \ldots, M(\omega_N - x_0) - M(x_0)\right]. \tag{2.9}$$

Finally, we compute the SVD of $ZZ^\top$ as

$$ZZ^\top = U_{ss}\Sigma_{ss}V_{ss}^\top \tag{2.10}$$

where here the subscript 'ss' is used to indicate that these are the (s)nap(s)hot singular values and vectors. The input modes (V_{ss}) are not usable (see above), but the output modes and spectrum (U_{ss} and Σ_{ss}) approximate the true output modes and spectrum, and thus can be used as an estimator of the accuracy of the direct approximation. To do so, we define an error measure $\mathcal{E}$ between the direct output modes and

[4] The method of snapshots cannot, itself, estimate the input modes; essentially, this is because we lack information about the adjoint and as such are only estimating the range (Brunton *et al.*, 2022; Sirovich, 1987).

Algorithm 1 Basic sampling algorithm

set error threshold ϵ

$j \leftarrow 1$

while $\mathcal{E}\left(\left\{U_j^{ss}, \Sigma_j^{ss}\right\}, \left\{U_j^d, \Sigma_j^d\right\}\right) > \epsilon$ **do**

draw snapshot random sample ω_j

compute sample $M(x_0 + \omega_j)$:

$\{Z\}_j \leftarrow M(x_0 + \omega_j)$

compute SVD of $ZZ^\top$, obtain U_j^{ss} and Σ_j^{ss}

compute sample $M\left(x_0 + \left\{T(I_n)\right\}_j\right)$:

$\{A\}_j \leftarrow M\left(x_0 + \left\{T(I_n)\right\}_j\right)$

compute SVD of A, obtain U_j^d, Σ_j^d, and V_j^d

$j \leftarrow j + 1$

end while

return Σ_j^d, and V_j^d

spectrum, and the method of snapshots output modes and spectrum, i.e. $\mathcal{E}\left(\left\{U_j^{ss}, \Sigma_j^{ss}\right\}, \left\{U_j^d, \Sigma_j^d\right\}\right)$. This allows for generalization (which we will need later), but for now one could imagine a simple projection of the leading modes as

$$\mathcal{E}\left(\left\{U_j^{ss}, \Sigma_j^{ss}\right\}, \left\{U_j^d, \Sigma_j^d\right\}\right) = 1 - \left|\left[U_j^{ss}\right]_1 \cdot \left[U_j^d\right]_1\right|. \tag{2.11}$$

We will return to discussing different choices of convergence criterion after demonstrating the basic approach.

The basic algorithm, then, is to sequentially increase the sampling space of the forward action both in the naïve and snapshot directions. The naïve approach provides an estimate of the input modes, but without any intrinsic ability to estimate the accuracy of the approximation; the method of snapshots more efficiently samples the output space, and so comparison of the output directions allows estimation of the error. The algorithm is outlined in Algorithm 1.

Note that the computational cost is two evaluations of the forward action $x \mapsto M(x)$ per iteration. One obvious refinement to improve this efficiency is to stop drawing new snapshots once the method of snapshots estimate of the output modes/spectrum $\left(\left\{U_j^{ss}, \Sigma_j^{ss}\right\}\right)$ is stable. One such version is given in Algorithm 2. For clarity, the two estimates are carried out separately, but of course they can also be performed in parallel.

Algorithm 2 Sampling algorithm with limited snapshots

set error thresholds ϵ, ϵ_{ss}

$j \leftarrow 1$

while $\mathcal{E}\left(\left\{U_j^{ss}, \Sigma_j^{ss}\right\}, \left\{U_{j-1}^{ss}, \Sigma_{j-1}^{ss}\right\}\right) > \epsilon_{ss}$ **do**

draw snapshot random sample ω_j

compute sample $M(x_0 + \omega_j)$:

$\{Z\}_j \leftarrow M(x_0 + \omega_j)$

compute SVD of $ZZ^\top$, obtain U_j^{ss} and Σ_j^{ss}

end while

$J \leftarrow j$

$j \leftarrow 1$

while $\mathcal{E}\left(\left\{U_J^{ss}, \Sigma_J^{ss}\right\}, \left\{U_j^{d}, \Sigma_j^{d}\right\}\right) > \epsilon$ **bold**

compute sample $M\left(x_0 + \left\{T(I_n)\right\}_j\right)$:

$\{A\}_j \leftarrow M\left(x_0 + \left\{T(I_n)\right\}_j\right)$

compute SVD of A, obtain U_j^d, Σ_j^d, and V_j^d

$j \leftarrow j + 1$

end while

return Σ_j^d, and V_j^d

2.1 *Parallelization*

This approach is straightforward to parallelize because the computational cost is dominated by the sampling (by assumption), and the samples are independent. Thus, a block-parallel version of Algorithm 1 is given as follows in Algorithm 3. Clearly, it is also possible to use both block-parallel and limited snapshot sampling, though we do not explicitly describe this as it is a straightforward combination of Algorithms 2 and 3.

3. Applications

We now demonstrate the proposed approach on a series of examples. These begin with toy models, which are constructed to illustrate some of the strengths and limitations of the method but without any particular physical significance, before proceeding through two more realistic examples drawn from the literature, and finally to the problem of airway selection for BT with which we motived the approach in the introduction.

Algorithm 3 Block parallel sampling algorithm

set error threshold ϵ,

N parallel tasks

$j \leftarrow 1$

while $\mathcal{E}\left(\left\{U_j^{ss}, \Sigma_j^{ss}\right\}, \left\{U_j^d, \Sigma_j^d\right\}\right) > \epsilon$ **do**

for $k = 1 \ldots N$ parallel tasks **do**

draw snapshot random sample ω_{j+k}

compute sample $M(x_0 + \omega_{j+k})$:

$\{Z\}_{j+k} \leftarrow M(x_0 + \omega_{j+k})$

compute sample $M\left(x_0 + \left\{T(I_n)\right\}_{j+k}\right)$:

$\{A\}_{j+k} \leftarrow M\left(x_0 + \left\{T(I_n)\right\}_{j+k}\right)$

end for

$j \leftarrow j + N$

compute SVD of $ZZ^\top$, obtain U_j^{ss} and Σ_j^{ss}

compute SVD of A, obtain U_j^d, Σ_j^d, and V_j^d

end while

return Σ_j^d, and V_j^d

3.1 *Toy models*

The following two examples are constructed to illustrate properties of the proposed method, and have no particular relevance as models of any real processes.

3.1.1 *Oscillating modes.*

For our first example, we take

$$M(x_i) = \sum_{j=1}^{N} \cos(j x_i) \frac{j}{N} \sum_{k=1}^{N} \sin\left(\pi z_k\right) x_k + \epsilon \mathcal{N}(0,1), \tag{3.1}$$

where $z_i = x_i = i/N$ and $\mathcal{N}(0,1)$ is the standard normal random variable. This is designed to have oscillating output and input modes. Application of the method to this example is shown in Fig. 1. The bottom panel shows the evolution of the error $\mathcal{E}$ with each iteration; the upper panels show the leading modes and spectrum at various points, as indicated by the annotation arrows. Although the leading output mode is accurate with only a relatively small number of iterations, the input mode is not well-captured until the majority of the input directions has been sampled, at least using a naïve sampling. However, this

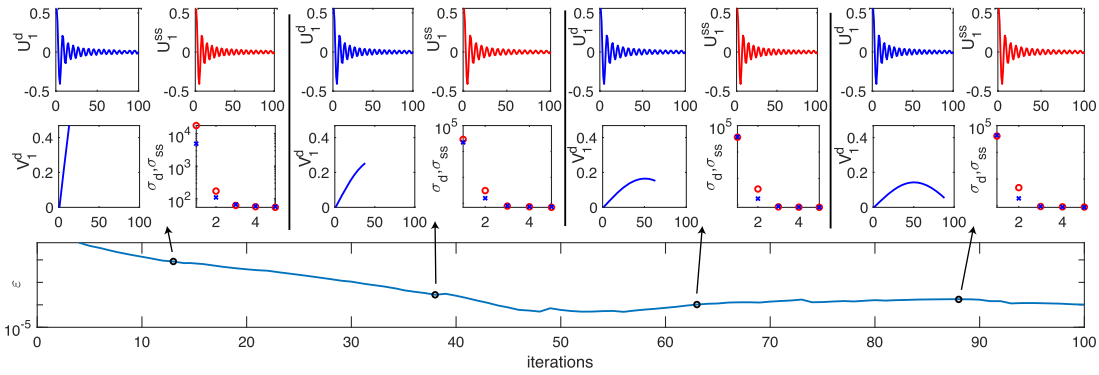


FIG. 1. Example 1: oscillating modes. The sampled modes and spectrum are shown at four evenly spaced points (upper panels), as indicated by the black circles in the lower panel. At each selected point, the leading direct output mode U_1^d, leading snapshot output mode U_1^{ss}, leading direct input mode V_1^d and SVD spectra σ_d (blue) and σ_{ss} (red) are shown. Here $m = n = 100$, in the upper panels the x axes indicate x_j for the modes and the ordering of the singular values for σ. Output from the direct approximation are shown in blue, and from the method of snapshots in red.

is clearly indicated by examination of $\mathcal{E}$ with increasing iterations. In this case, complete convergence is obtained only after sampling the entire input space, but nonetheless examining the error measure gives and estimate of the convergence for incomplete sampling. Here, and in the following examples, we have used

$$\mathcal{E} = 1 - \left[U_j^{ss}\right]_1 \cdot \left[U_j^d\right]_1 / \left|\left|\left[U_j^{ss}\right]_1\right|\right|_2 + \left|1 - \frac{\sigma_{ss}^{1,j} / \sum \sigma_j^{ss}}{\sigma_d^{1,j} / \sum \sigma_j^d}\right| \tag{3.2}$$

where $\sigma = \mathrm{tr}\,(\Sigma)$ are the singular values; this accounts for both the leading mode and the energy in the leading singular value.

Of course, a simple transformation of the input basis (i.e. choice of $T_N(\cdot)$) alters the performance markedly. By taking $T_j = \sin(j\pi x)$, which maps the leading input mode into the first direction, we can obtain the expected near-instantaneous convergence, shown in Fig. 2 (the layout is the same as Fig. 1 but with annotations omitted). Although the input modes are correct after the first iteration, the estimates of the output mode and spectrum by the method of snapshots (and therefore the error relative to these) require a few more iterations to stabilize.

3.1.2 *Localized modes.* For our second example, we take

$$M(x(s)) = \int_0^1 x(s)\exp\left(-c(s-1/3)^2\right)\mathrm{d}s\Bigg[\exp\left(-c(s-1/3)^2\right)\left\{\int_0^1 x(s)\exp\left(-c(s-2/3)^2\right)\mathrm{d}s\right\}^{\alpha}$$
$$+\exp\left(-c(s-2/3)^2\right)\int_0^1 x(s)\exp\left(-c(s-1/3)^2\right)\mathrm{d}s\Bigg],$$

which, by design, has Gaussian localized modes. The parameter α controls the behaviour in an important way. When $\alpha = 1$, mixing of the leading modes occurs, as shown in Fig. 3. The two largest singular

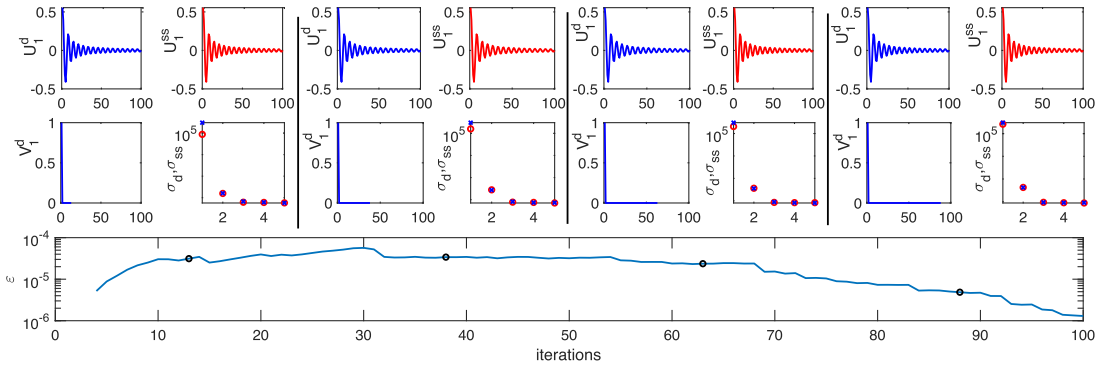


FIG. 2. Example 1b: Oscillating modes, subject to change of basis. Layout as in Fig. 1, for details see Fig. 1 caption.

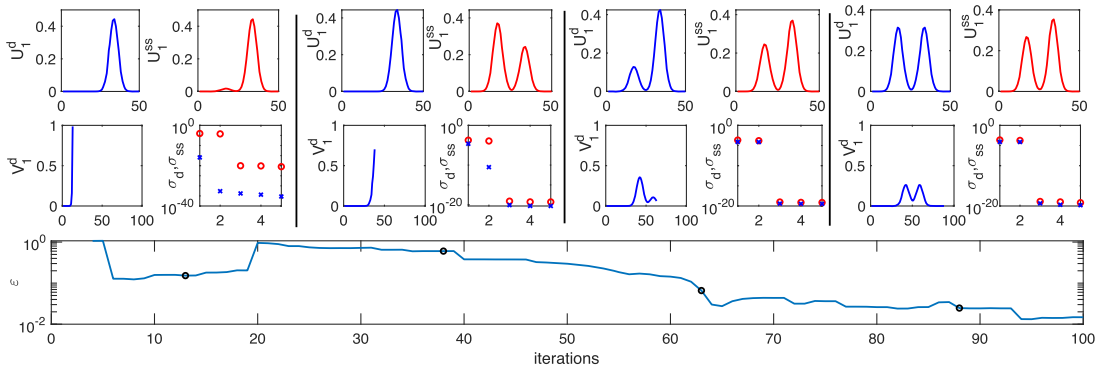


FIG. 3. Example 2: mixed localized modes ($\alpha = 1$). Discretized with 100 input components and 50 output components. Layout as in Figs 1 and 2, for details see Fig. 1 caption.

values are equal, and the associated modes then appear in arbitrary linear combinations. This obviously compromises a convergence metric based on the leading mode and singular value alone, which can clearly be seen in the poor performance of $\mathcal{E}$. Here the continuous operator has been discretized with 100 input components (uniformly on $[0, 1]$) and 50 output components.

Of course, this is precisely the role of the parameter α; for $\alpha \neq 1$, the leading mode is isolated, see Fig. 4, and $\mathcal{E}$ clearly indicates when the localized input mode is captured.

An alternative approach would be to make a different choice of the error measure $\mathcal{E}$ to account for the mode mixing.

3.2 *Laser system*

The next example is not biological but drawn from the literature of fibre optic communications systems (Donovan & Kath, 2011). Briefly, pulse propagation through optical fibre is governed by the nonlinear

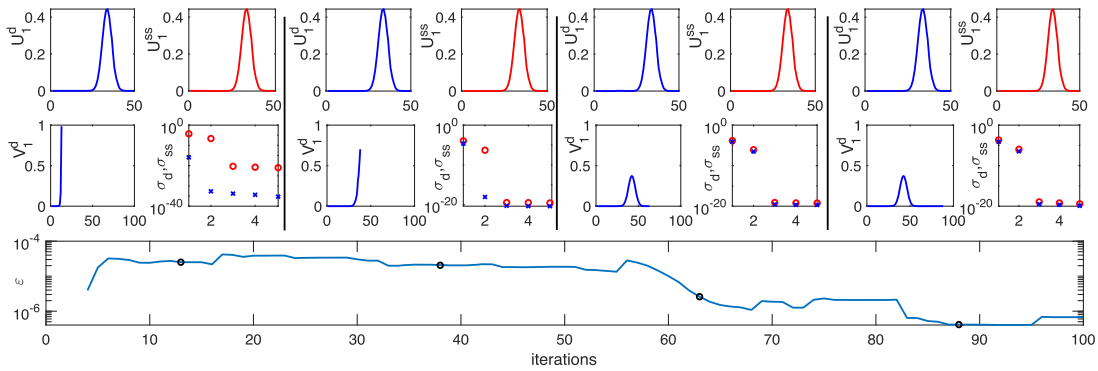


FIG. 4. Example 2B: localized modes, unmixed ($\alpha = 2$). Layout as in Figs 1,2 and 3, for details see Fig. 1 caption.

Schrödinger equation

$$\frac{\partial u}{\partial z} - \frac{i}{2}\beta(z)\frac{\partial^2 u}{\partial t^2} - i\gamma(z)|u|^2 u = 0, \tag{3.3}$$

where $u(t, z)$ is the slowly varying envelope, $\beta(z)$ is the dispersion coefficient of the fibre and $\gamma(z)$ is the nonlinear coefficient of the fibre. We use a soliton initial condition[5]

$$u(t, 0) = A\text{sech}(A(t - T))e^{i\Omega t + i\Phi} \tag{3.4}$$

with soliton parameters $A = 4, T = \Omega = \Phi = 0, N = 256$ discrete modes and propagation distance $z_{max} = 6$. Eq. 3.3 is solved numerically using the split-step Fourier method (Sinkin *et al.*, 2003).

We take the input mapping in terms of perturbations to the fibre properties as

$$\beta(z) = 1 - \delta x_j, \text{ where} j = \lfloor \frac{z}{z_{max}} N \rfloor$$

$$\gamma(z) = 1.$$

Finally, phase detection is used for the output mapping as

$$M = |u(t, z_{max})| \,\text{atan}\left(\frac{Re(u(t, z_{max}))}{Im(u(t, z_{max}))}\right). \tag{3.5}$$

Thus, the input–output mapping takes perturbations to the fibre structure and returns alterations to the final pulse shape (at $z = z_{max}$). This particular example is selected because the modes can be found directly via the action of the adjoint operator for comparison (Donovan & Kath, 2011). Application of the current methodology is shown in Fig. 5.

[5] Note that in an optics context, t is retarded time and z propagation distance, thus the usual intuitive role of spatial and temporal variables is reversed: the initial condition is a function of t at $z = 0$, and this propagates forward in z.

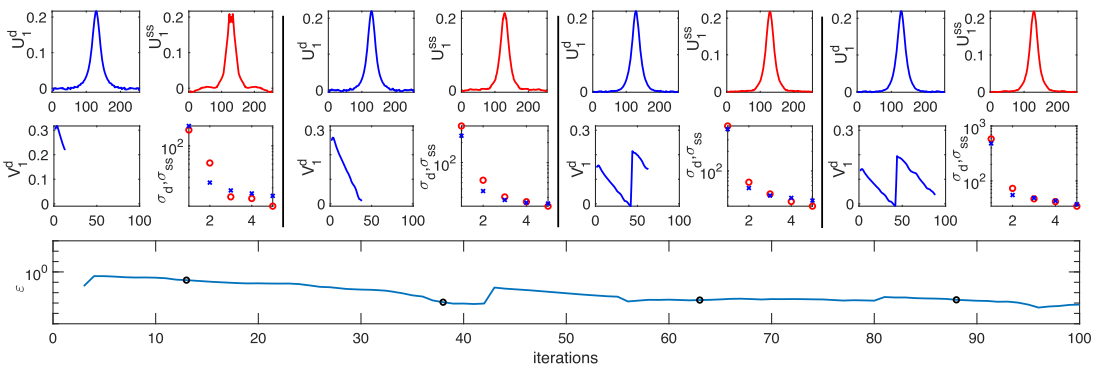


FIG. 5. Example 3: application to an optical communication system with evolution governed by the nonlinear Schrödinger equation with $\delta = 1$. Layout as in Figs 1–4, for details see Fig. 1 caption.

The output modes converge quickly, with the limiting step being the number of iterations required for the method of snapshots. As with examples 1 and 2B, the limiting factor for the input modes is localized coverage of the input basis.

3.3 *Network infection model*

The penultimate example is a network infection model in which we seek the optimal treatment locations on the network (Postlethwaite, 2022). The network itself is a Watts–Strogatz network (Watts & Strogatz, 1998) with mean degree 4 and parameter $\beta = 0.04$, which controls the regularity of the structure.[6] For a fixed, discrete timestep the infection and recovery probabilities at the i^{th} node are

$$P_I^i = \lambda(1 - 0.9\nu^i)n_I^i \quad \text{for uninfected nodes, and}$$
$$P_R = \mu \quad \text{for infected nodes,}$$

where n_I^i is the number of infected neighbours of the i^{th} node and ν^i is the vaccination status $[0, 1]$ of each node. The input (structure) space is vaccination of each node, and the output space is the final steady-state status of those who have been infected.

The results are shown in Fig. 6, in which it can be seen that approximately 80 iterations are required to accurately capture the leading modes, which is clearly shown by the stability of the input modes over time, the similarity of the output and snapshot modes, and perhaps most importantly by $\mathcal{E}$.

In this particular context, one might reasonably imagine that the leading input mode corresponds to the optimal nodes based on a graph-theoretic measure such as eigenvalue centrality of the adjacency matrix (e.g. Shakarian *et al.*, 2015). Interestingly, the leading mode does not directly equate to any of the common centrality measures; however, because we wish to focus on a broader context than network-based models, we will not dwell overlong on the the precise relationship here.

[6] $\beta = 0$ is a regular lattice.

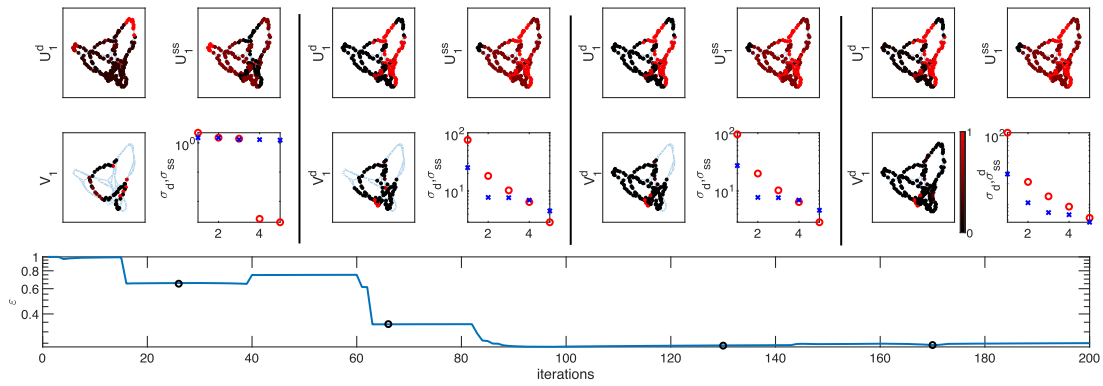


FIG. 6. Example 4: network infection model. Here the parameter values are $\lambda = 0.8$, $\mu = 0.2$ on a graph with $N = 200$. As previously the upper panels are a uniformly spaced sampling of the iteration space. Layout as in Figs 1–4, for details see Fig. 1 caption. Note that the colourbar for the network plots is shown only once on the right-hand side for compactness. Additionally, the graph layout is chosen for visual clarity, and the greyed out nodes in the V^d plots are those not yet sampled. The input modes are in terms of treatment (v^i) and the output modes in terms of infection.

3.4 *Optimal airway selection for bronchial thermoplasty*

Finally, we apply the method to the original motivating problem: optimal airway selection in BT. The dynamical model of airway constriction, flow conservation and pressure balance in the airway tree is based on Donovan (2017) and ideas from Anafi & Wilson (2001); Venegas *et al.* (2005) and similar to Stewart & Jensen (2015). A brief summary is given here; for full details, the reader is referred to Donovan (2017). We write $\boldsymbol{r}$ for the vector of airway lumen radii, and $\boldsymbol{p}$ and $\boldsymbol{q}$ for pressure and flow vectors, respectively, and for the evolution in time of the i^{th} airway we then have

$$\dot{r}_i = \rho\left(\boldsymbol{\phi}(r_i;\boldsymbol{r},\boldsymbol{p},\boldsymbol{q}) - r_i\right), \tag{3.6}$$

where ρ controls the airway relaxation time scale. The function ϕ describes static airway behaviour by composition as $\phi = R(P_{tm}(r))$ where $R(P)$ is the so-called Lambert model describing the pressure-radius relationship using matched rectangular hyperbolae (Lambert *et al.*, 1982). The function $P(r)$ gives airway transmural pressure as a function of the radius as

$$P_{tm}(r_i) = p_{mid_i} - \frac{\kappa R_{ref}}{r_i} + 2\mu_i\left(\left(\frac{R_{ref} - r_i}{R_{ref}}\right) + 1.5\left(\frac{R_{ref} - r_i}{R_{ref}}\right)^2\right), \tag{3.7}$$

where the second and third terms are the ASM and *parenchymal tethering* pressures. In particular, the parameter κ incorporates both smooth muscle mass and activation, and so is altered by BT treatment of each airway. The parameter p_{mid_i} is the mid-airway pressure, and R_{ref} is the reference radius (Donovan, 2016).

In the conducting airways, the flow conservation constraints are

$$q_m = q_{d_1} + q_{d_2} \tag{3.8}$$

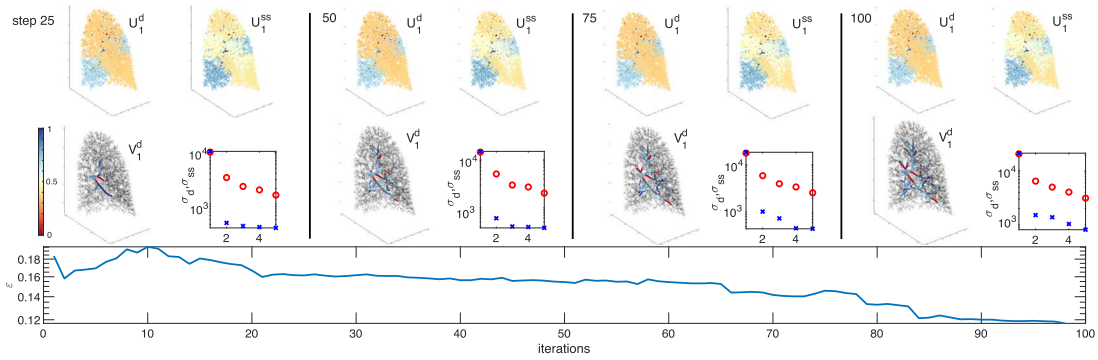


FIG. 7. Estimating input modes and the optimal airway treatment set for bronchial thermoplasty. Layout as in previous figures. In this example, 21 airways area treated in both the naïve and optimal treatment sets. As previously the upper panels are a uniformlyspaced sampling of the iteration space. Note that the colourbar for the network plots is shown only once on the left-hand side for compactness. The greyed out airways in the V^d plots are those not yet sampled by the direct approach. The input modes are in terms of treatment (i.e. κ) and the output modes in terms of flow (q).

where the notation indicates the mother and two daughter branches at each junction. In each airway, we assume Poiseuille flow

$$\Delta p_i = \alpha_i r_i^{-4} q_i, \tag{3.9}$$

where the constants α_i absorb all dependencies aside from radius, pressure and flow (Donovan, 2016, 2017). Following Donovan *et al.* (2018), we assume a 75% reduction in ASM in the BT treated airways and a concomitant reduction in total wall area, although the precise figure is not without controversy (Chernyavsky *et al.*, 2018).

In terms of our input–output (structure–function) mapping, the input space is the airways to be treated directly, and the output is the flow at steady state (periodically averaged for breathing). During optimization, the input space is treated as continuous on $[0, 1]$ but discretized for determination of the final treatment set by ranking the optimal mode. The computational geometry is patient-specific and acquired from a combination of CT imaging and biobank data as previously described (Donovan *et al.*, 2018) and similar to Leary *et al.* (2014).

Although the input space (all conducting airways) is in theory as large as 30,000, this is severely constrained by both computational constraints and (in reality) the size of the bronchoscope relative to the airways. However, within the smaller set of treatable airways, we still wish to find the optimal set. To do so, we apply Algorithm 3 with sequential treatment basis starting from the largest airways, with results given in Fig. 7. The comparative 'naïve' treatment set here is a simple ordering of the largest treatable airways.

Several things are apparent. The first is that, even at the maximum number of iterations computed, the direct output modes do not fully capture all aspects of the snapshot output modes (at least in this particular subject). However, there is a degree of stability in both sets of output modes as the number of iterations increases, suggesting that this behaviour lies outside of the control of the treatable central airways. It is also important to recall here the assumption of linearity of the operator, which may prevent convergence of the two set of output modes both in general and in this particular application: airways do not act only in isolation but exhibit network and cascade effects (Alencar *et al.*, 2002). Thus, the isolated

sampling of the direct approach will not capture some of these effects. It is possible that this could be addressed by a transformed sampling basis, but it is not clear *a priori* what the optimal sampling basis should be.

The input modes and spectrum also display a degree of stability over time (iterations), which suggest an optimal treatment set. We then tested this optimal treatment set, compared with a naïve treatment set consisting of the same number of treated airways. The comparison between the naïve and targeted treatments are shown in Fig. 8, both in terms of the dose-response curve (panel (a)) showing how airway resistance (William Thorpe & Bates, 1997) increases with contractile agonist (which stimulates the ASM to contract) as well as in terms of the flow patterns across the lung at maximal ASM activation (panels (c)–(e)). By design, the method should have captured the largest change in the flow pattern (the output/function space); however, it can also be seen that this corresponds to improved lung function across the agonist dose range. Observe that the naïve treatment response is paradoxical, which is not uncommon (Donovan, 2018), but recall also that resistance is only an indirect outcome and the flow distribution is the direct target. Moreover, it is worth noting that the targeted treatment set is not something so simple as selecting the airways with the greatest ASM content (panel (b)) but rather it is the result of the complex interplay between interrelated airways. Similarly, no obvious predictor is apparent in terms of total wall area (data not shown). Here ASM is normalized by the square of the basement membrane perimeter as the most appropriate index of airway size (Donovan *et al.*, 2023).

4. Discussion

In this manuscript, we have proposed an iterative approach to estimating the SVD input modes from the action of the forward operator alone. This was motivated by a particular problem in respiratory physiology (e.g. the optimal airway treatment set in BT) and also applied to several test problems to demonstrate its strengths and weaknesses. It may be applicable to similar problems in different application areas.

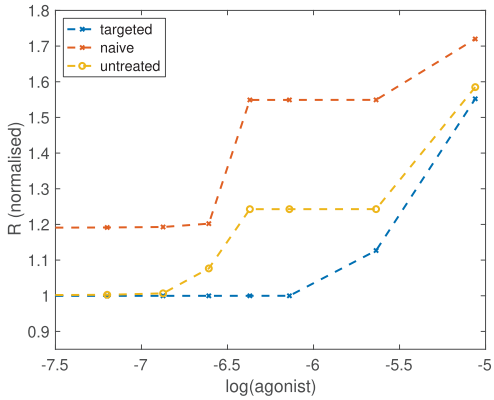
It is worth reiterating the particular aspects which might favour this approach over alternatives. First and foremost, this is only relevant to situations in which we have access only to the action of the forward operator; if either the adjoint operator is available, or the entire matrix can be formed, other methods are almost certainly preferred.

Moreover, we are envisioning situations in which the input and output spaces are both of relatively high dimension; otherwise, conventional optimization methods are probably favoured. For the BT problem in particular, the combinatorial optimization problem is completely infeasible on computational cost grounds: the search space is enormous ($\sim \binom{300}{40} \approx 10^{50}$) and the computational cost of each forward evaluation is significant ($\sim$ 1 cpu-day), though significant savings may be possible via reduction techniques (Whitfield *et al.*, 2020).

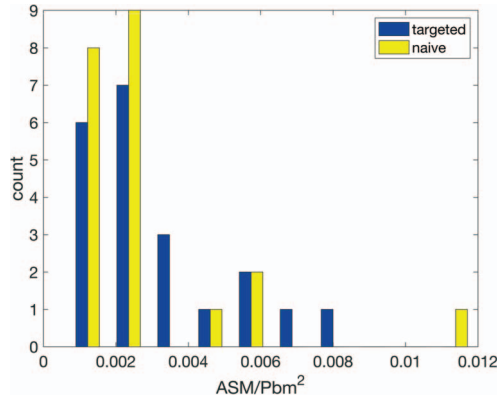
Moreover, it is also important that one has some idea of how to search the input space efficiently. If it becomes necessary to sample the entire input space, we have gained little over the naïve approach of simply reconstructing A by a complete sampling of the input space (e.g. example 3).

There are connections between this approach and uncertainty quantification (e.g. Smith, 2013). Although we have not explicitly formulated the method as such, perturbations to the input space could be viewed not as treatment effects but rather as measurement error. In such a case, the input space would need to be viewed as the full parameter set.

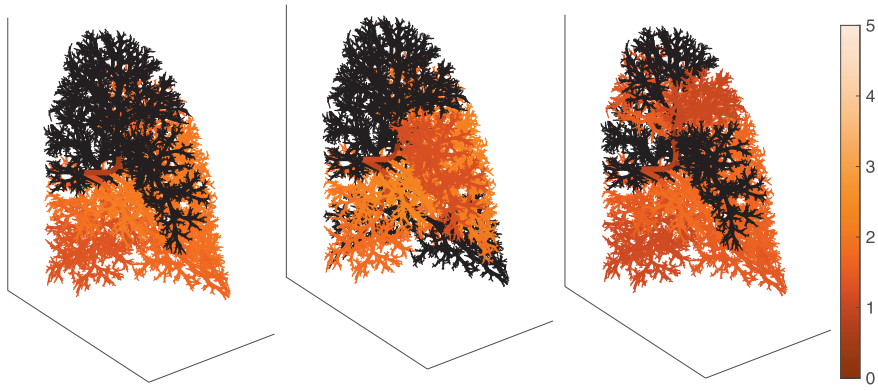
One additional aspect that we have not so far discussed is the choice of reference configuration x_0 and the size of the linearizing perturbation δ (see Eq. 2.6). Of course, if M is a linear operator, then the choice is unimportant, but in the more likely situation that M is nonlinear then these choices are



(a) Dose-response curve, airways resistance (normalised) as a function of contractile agonist dose (log scale). For clarity, lower agonist doses are not shown, though note that (normalised) $R = 1$ for very low doses in all three cases.



(b) Comparison of ASM content in targeted and naïve treatment sets by histogram of treated airways.



(c) Flow (q) distribution, untreated.

(d) Flow (q) distribution, naïve treatment.

(e) Flow (q) distribution, targeted treatment.

FIG. 8. Targeted versus naïve airway treatment sets in the BT model. In (a), resistance (R) is normalized to nominal; similarly in (c)–(d) flow is normalized to nominal (Donovan, 2017) with a single common colourbar on the right for all panels.

important. In the BT case it is straightforward—the reference configuration is the pre-treatment lung, and the linearizing perturbation is binary: each airway is either treated or untreated. However, in other cases, it may be less obvious how to make these choices. Similarly, the method of snapshots requires the choice of distribution from which to draw the snapshots ω_j. The idea here is to efficiently sample the range of M, so that the choice of distribution of ω is informed by what we know about the input domain. One might expect that the standard deviation of ω should be comparable to δ, though this could be applications-specific deviations from this.

There are a few subtle points that have not been previously discussed that bear mention. The first is that our choice of $\mathcal{E}$ (Eq. 2.11) is by no means definitive. Other choices could most certainly be made, depending on the specifics of the problem at hand. A related problem is the stability of the mode ordering as the iterations progress: there is no guarantee that the ordering of the modes (by singular value) remains the same as iterations are added to the truncated approximation. In such a case, it may be necessary to

apply other methods to disambiguate the modes. The consequences are two-fold: (1) if the error measure incorporates the leading mode, switching of the leading mode will strongly effect the error measure, and (2) we have explicitly shown the leading mode in the figures as the iterations proceed. As with the modal mixing of example 2, modal switching can confuse the interpretation.

It is also worth considering the practicalities of sampling limitations and the utility of $\mathcal{E}$. In the examples given here, we have used a fixed number of iterations (and tracked $\mathcal{E}$) rather than using a convergence criterion *per se* in order to demonstrate the method more fully; however, the algorithms of Sec. 2 are formulated to stop at a fixed convergence criterion and this is how they would be best used in practice.

In summary, we have demonstrated the proposed method on a sequence of examples in order to demonstrate both its strengths and limitations. In all cases, we assume linearity of the underlying operator, and this can be a significant assumption. Selection of the sampling basis is also key—if the sampling basis is such that we must sample the entire space in order to achieve an adequate estimate, then we have gained little. Indeed, the sampling resources used for the method of snapshots in such a case are in some sense wasted, except perhaps for their utility in demonstrating the need to continue sampling. There can also be issues with modal mixing and switching, as shown in example 2. However, if these conditions are met, this approach offers a new avenue for otherwise difficult problems, as demonstrated by our ability to find a targeted BT treatment set that significantly improves upon the untargeted approach, in a computationally feasible manner.

References

Alencar, A. M., Arold, S. P., Buldyrev, S. V., Majumdar, A., Dimitrije Stamenović, H., Stanley, E. & Suki, B. (2002) Dynamic instabilities in the inflating lung. *Nature*, **417**, 809–811.

Anafi, R. C. & Wilson, T. A. (2001) Airway stability and heterogeneity in the constricted lung. *J. Appl. Physiol.*, **91**, 1185–1192.

Brunton, S. L., Nathan, J. & Kutz. (2022) *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*. Cambridge University Press.

Chernyavsky, I. L., Russell, R. J., Saunders, R. M., Morris, G. E., Berair, R., Singapuri, A., Chachi, L., Mansur, A. H., Howarth, P. H., Dennison, P., et al. (2018) In vitro, in silico and in vivo study challenges the impact of bronchial thermoplasty on acute airway smooth muscle mass loss. *Eur. Respir. J.*, **51**, 1701680.

Donovan, G. M. (2016) Clustered ventilation defects and bilinear respiratory reactance in asthma. *J. Theor. Biol.*, **406**, 166–175.

Donovan, G. M. (2017) Inter-airway structural heterogeneity interacts with dynamic heterogeneity to determine lung function and flow patterns in both asthmatic and control simulated lungs. *J. Theor. Biol.*, **435**, 98–105.

Donovan, G. M. (2018) Biological version of Braess' paradox arising from perturbed homeostasis. *Phys. Rev. E (3)*, **98**, 062406.

Donovan, G. M., Elliot, J. G., Boser, S. R., Green, F. H. Y., James, A. L. & Noble, P. B. (2019) Patient-specific targeted bronchial thermoplasty: predictions of improved outcomes with structure-guided treatment. *J. Appl. Physiol.*, **126**, 599–606.

Donovan, G. M., Elliot, J. G., Green, F. H. Y., James, A. L. & Noble, P. B. (2018) Unraveling a clinical paradox: why does bronchial thermoplasty work in asthma? *Am. J. Respir. Cell Mol. Biol.*, **59**.

Donovan, G. M. & Kath, W. L. (2011) An iterative stochastic method for simulating large deviations and rare events. *SIAM J. Appl. Math.*, **71**, 903–924.

Donovan, G. M., Wang, K. C. W., Elliot, J. G., James, A. L. & Noble, P. B. (2023) Quantifying airway remodelling for research or clinical purposes: how should we normalize for airway size? *Respirology*, **28**, 223–225.

Hackmann, M. J., Elliot, J. G., Green, F. H. Y., Cairncross, A., Cense, B., McLaughlin, R. A., Langton,

D., James, A. L., Noble, P. B. & Donovan, G. M. (2022) Requirements and limitations of imaging airway smooth muscle throughout the lung in vivo. *Respir. Physiol. Neurobiol.*, **301**, 103884.

Halko, N., Martinsson, P.-G. & Tropp, J. A. (2011) Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Rev.*, **53**, 217–288.

Hall, C. S., Quirk, J. D., Goss, C. W., Lew, D., Kozlowski, J., Thomen, R. P., Woods, J. C., Tustison, N. J., Mugler III, J. P., Gallagher, L., et al. (2020) Single-session bronchial thermoplasty guided by 129xe magnetic resonance imaging. A pilot randomized controlled clinical trial. *Am. J. Respir. Crit. Care Med.*, **202**, 524–534.

Hoffman, E. A. & McLennan, G. (1997) Assessment of the pulmonary structure–function relationship and clinical outcomes measures: quantitative volumetric ct of the lung. *Acad. Radiol.*, **4**, 758–776.

Lambert, R. K., Wilson, T. A., Hyatt, R. E. & Rodarte, J. R. (1982) A computational model for expiratory flow. *J. Appl. Physiol.*, **52**, 44–56.

Leary, D., Winkler, T., Braune, A. & Maksym, G. N. (2014) Effects of airway tree asymmetry on the emergence and spatial persistence of ventilation defects. *J. Appl. Physiol.*, **117**, 353–362.

Michael, J. (2021) What do we mean when we talk about "structure/function" relationships? *Adv. Physiol. Educ.*, **45**, 880–885.

Postlethwaite, C. M. (2022) *Private communication*.

Shakarian, P., Bhatnagar, A., Aleali, A., Shaabani, E., Guo, R., Shakarian, P., Bhatnagar, A., Aleali, A., Shaabani, E. & Guo, R. (2015) *The SIR Model and Identification of Spreaders*. Springer.

Sinkin, O. V., Holzlöhner, R., Zweck, J. & Menyuk, C. R. (2003) Optimization of the split-step Fourier method in modeling optical-fiber communications systems. *J. Light. Technol.*, **21**, 61.

Sirovich, L. (1987) Turbulence and the dynamics of coherent structures. i. Coherent structures. *Quart. Appl. Math.*, **45**, 561–571.

Skloot, G. S. (2017) The effects of aging on lung structure and function. *Clin. Geriatr. Med.*, **33**, 447–457.

Smith, R. C. (2013) *Uncertainty Quantification: Theory, Implementation, and Applications*, vol. 12. SIAM.

Stewart, P. S. & Jensen, O. E. (2015) Patterns of recruitment and injury in a heterogeneous airway network model. *J. R. Soc. Interface*, **12**, 20150523.

Stoll, M. (2012) A Krylov–Schur approach to the truncated SVD. *Linear Algebra Appl.*, **436**, 2795–2806.

William Thorpe, C. & Bates, J. H. T. (1997) Effect of stochastic heterogeneity on lung impedance during acute bronchoconstriction: a model analysis. *J. Appl. Physiol.*, **82**, 1616–1625.

Venegas, J. G., Winkler, T., Musch, G., Vidal, M. F., Melo, D. L., Tgavalekos, N., Fischman, A. J., Callahan, R. J., Bellani, G., Scott, R. & Harris. (2005) Self-organized patchiness in asthma as a prelude to catastrophic shifts. *Nature*, **434**, 777.

Watts, D. J. & Strogatz, S. H. (1998) Collective dynamics of "small-world" networks. *Nature*, **393**, 440–442.

Whitfield, C. A., Latimer, P., Horsley, A., Wild, J. M., Collier, G. J. & Jensen, O. E. (2020, 2020) Spectral graph theory efficiently characterizes ventilation heterogeneity in lung airway networks. *J. R. Soc. Interface*, **17**, 20200253.