

RESEARCH ARTICLE

Modeling full-scale leaf venation networks

Lars Erik J. Skjægstad¹, Julius B. Kirkegaard^{1,2*}¹ Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark, ² Department of Computer Science, Copenhagen, University of Copenhagen, Copenhagen, Denmark* juki@di.ku.dk

Abstract

The vascular network of leaves, comprising xylem and phloem, is a highly optimized system for the delivery of water, nutrients, and sugars. The design rules for these naturally occurring networks have been studied since the time of Leonardo da Vinci, who constructed a local rule for comparing the widths of in- and outgoing veins at branch points. Recently, physical models have been developed that seek to explain the full morphogenesis of leaf venial networks in which veins grow in response to local hydrodynamic feedback. Although these models go beyond simple local rules, they are challenging to compare to experimental data. Here, we extend these hydrodynamic models to a state where the direct comparison with images of full leaves becomes possible on the level of individual veins. We present a dataset of the venial networks of leaves that maintain full network topology and use this to discuss the benefits and drawbacks of such a direct comparison. We apply our approach to the direct estimation of a sink fluctuation parameter, demonstrating consistency within distinct leaf species. Finally, we utilize the ability of the model to run on full leaves to define and calculate exponents for a Murray's law that applies to reticulate venation networks.

OPEN ACCESS

Citation: Skjægstad LEJ, Kirkegaard JB (2025) Modeling full-scale leaf venation networks. *PLoS Comput Biol* 21(7): e1013292. <https://doi.org/10.1371/journal.pcbi.1013292>

Editor: Marc R. Birtwistle, Clemson University, UNITED STATES OF AMERICA

Received: October 25, 2024

Accepted: July 2, 2025

Published: July 21, 2025

Copyright: © 2025 Skjægstad, Kirkegaard. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: Source code is available from <https://github.com/kirkegaardlab/full-scale-leaves>. Data is provided at <https://doi.org/10.5281/zenodo.13990667>.

Funding: This work has received funding from the Novo Nordisk Foundation, Grant Agreement No. NNF200C0062047 (to JBK, funding LEJS). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

One of the most familiar characteristics of a leaf is its reticulate vein network, which, among other functions, is responsible for transporting and distributing water from the stem throughout the leaf. This network exhibits a significant number of loops, which in hydrodynamic models is often assumed to be the result of a fluctuating demand for water across the leaf. It has previously been shown in small theoretical networks that a single parameter controlling the varying degree of fluctuation can tune the loopiness of the emerging network structures. In this article, we apply this type of model to *full-scale* leaf images, and demonstrate that this single parameter can be fitted by directly comparing the vein widths output by the model to their image counterparts. This paves the way for research into data-driven modeling of leaf morphogenesis. Finally, we demonstrate an application of our model by using it to describe the relationship between vein

thicknesses at branchpoints in loopy networks, thus extending the well-known Murray's law.

1. Introduction

Leaf venation networks are intricate structures that are key to the overall health, growth, and adaptability of plants [1,2]. These veins are comprised of the xylem and phloem tissues. The xylem is primarily responsible for transporting water and minerals absorbed from the soil out to the leaves, where photosynthesis occurs, while the phloem distributes the photosynthetic products, such as sugars and other organic molecules, back throughout the plant. Additionally, the venial network can function as structural support for the leaf.

The venation structure typically exhibits self-similarity, dominated at the high level by a tree-like topology. However, the structure becomes more complex in environments where leaves are susceptible to damage, or where local changes to the demand of water within the leaves occur due to, e.g., variations in solar radiation. This complexity is characterized by the emergence of network loops, which are redundant structures that serve as adaptive mechanisms for the leaf to maintain functionality and resilience under changing or adverse conditions. These structural features ensure that the network can continue to satisfy the physiological demands of the leaves, even when external factors introduce variability.

Historically, one of the earliest documented explorations into leaf venation was made by Leonardo da Vinci [3]. His observation that the total cross-sectional area of all branches in a tree remains consistent from the base trunk to the leaves foreshadowed what is now known as Murray's law [4]. Since then, the understanding of leaf venation networks has advanced significantly on two distinct fronts. First, automated image analysis and network analysis have laid the foundation for studies on the statistical properties and topology of vein networks. Second, hydrodynamic models for transport networks have been developed, explaining not only the theory behind the local variations in vein thickness but also the overall morphogenesis of the network structure, particularly in small idealized networks. In this study, we bridge these two approaches by integrating automated image analysis and hydrodynamic modeling techniques. We develop methods to apply these theoretical models to networks extracted from images of actual leaves, allowing us to examine how the principles derived from the idealized models apply to complex, naturally occurring venation patterns [Fig 1].

In a seminal paper, Bohn & Magnasco [5] proposed a general framework for investigating the properties of optimal transport networks, which has been applied to a diverse set of problems, including leaves [6,7], coastal river networks [8], vascular blood systems [7,9], and traffic flow [10,11]. Their hydrodynamic model shows that in steady-state systems, i.e., where the current is constant, the optimal network configuration adopts a tree-like topology characterized by a complete absence of loops, which is highly different from the networks seen in many types of leaves, where loops are prevalent. However, when the system is subjected to fluctuating loads or probabilistic edge failures, redundancies become necessary, which leads to the emergence of more realistic reticulate networks as optimal structures [12–14].

The hydrodynamic model of network optimality [5] suffers from the existence of many local optima, and local adaptation rules [15] thus rarely lead to the globally optimal network. By taking into account the growth of the leaf [6], however, more optimal solutions can be reached from local rules, resulting in networks that better compare to real leaves. In its various formulations, the hydrodynamic model thus qualitatively captures many characteristics of leaf venation. However, two significant challenges arise when analyzing networks at realistic scales. First, the hydrodynamic model relies on graph-based computations that require



Fig 1. Sample image of *S. albus* (top half) with corresponding model output venation pattern (bottom half).

<https://doi.org/10.1371/journal.pcbi.1013292.g001>

the inversion of matrices whose dimensions correspond to the number of nodes or edges in the network. For optimization processes that necessitate a large number of iterations, this matrix inversion becomes computationally prohibitive in large-scale systems. Second, the number of local optima increases with network size, complicating the analysis of global optimality. To address these challenges, we employ two strategies to enable computations on full-size networks: (1) We partition the full-size networks into smaller subsections and optimize these substructures separately while preserving the influence of the surrounding network, and (2) we mitigate the issue of the numerous local optima by restricting our focus to the hydrodynamic optimality only in the vicinity of the network topology observed in the relevant leaf.

We demonstrate the utility of our model by applying it to images of leaves, showcasing the consistency and variability within leaf species by using several samples from each species. In particular, we exemplify on *Symphoricarpos albus*, *Lonicera xylosteum*, and *Crataegus monogyna*, which are native to the surroundings of the University of Copenhagen. We develop a hydrodynamic model with fluctuating sinks and evaluate it on networks extracted from the entire leaf, which results in data on each individual network node and edge. We subsequently compare the model output vein widths with those from the image data, discussing the possible reasons for discrepancies between them, and use the data to perform parameter estimations. Lastly, we do an empirical study of Murray's law on the full-scale networks that result from our simulations. We suggest an immediate extension to Murray's law to account for the reticulate structure in the networks and investigate its applicability.

2. Results

2.1. Model

We are interested in comparing theoretical models on the formation of complete venial networks in leaves directly to leaf data, a process which relies on the availability of networks describing these full-scale systems. To that end, we employ a custom pipeline for extracting the network topology and edge attributes from leaf images. The process, illustrated in Fig 2a–2d, relies on the deep learning segmentation of veins [16], as well as a custom network extraction algorithm. Details can be found in the Methods section, as well as in the associated source code.

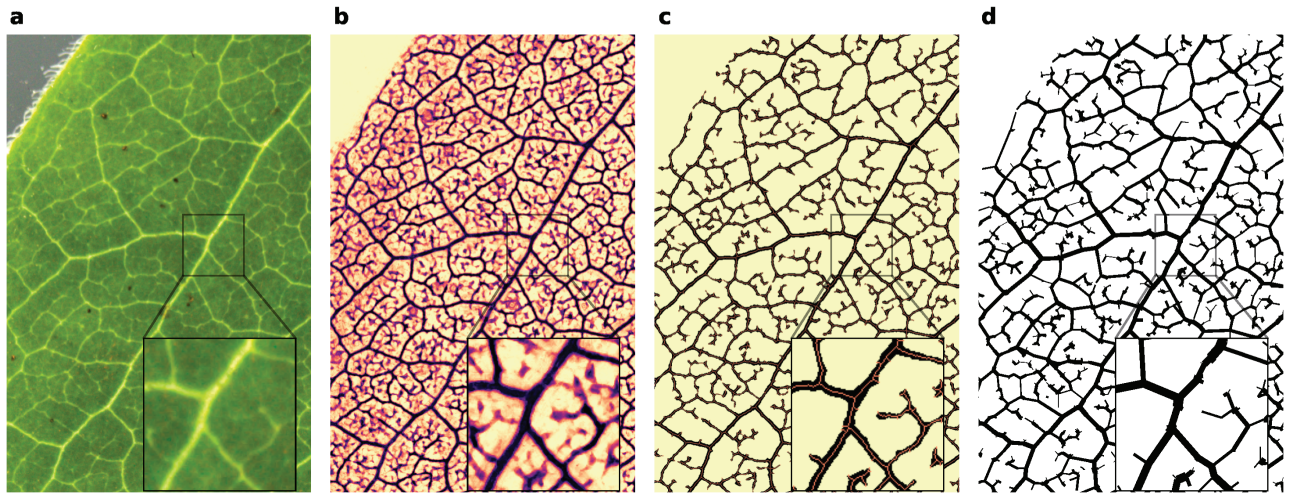


Fig 2. Network extraction. (a) Section of input image. (b) Semantic image segmentation of the veins from background. (c) Binary segmentation and the network skeleton. (d) The final graph extracted, represented by the coordinates and incidences of all nodes and edges, as well as the lengths and widths of each edge.

<https://doi.org/10.1371/journal.pcbi.1013292.g002>

The network structure is represented by a directed, planar graph consisting of a total of n_n nodes and n_e edges, each edge being associated with a length ℓ_i and width w_i . The venial network has two roles: it transports and distributes fluid from the petiole to the rest of the leaf (xylem) and it distributes the photosynthetic products from the leaf cells to the plant (phloem). Mathematically, this leads to dual formulations with identical end results, so here we formulate the problem in terms of xylem only. Thus, to model the fluid flow through the network, we place a fluid source at the petiole, and sinks (which account for fluid evaporation and photosynthetic consumption) at all other nodes in the network.

Assuming laminar fluid flow, we model each edge as a cylindrical pipe, such that the flux f_i over an edge i is given by the Hagen-Poiseuille law,

$$f_i = \frac{c_i}{\ell_i} (\Delta p)_i, \quad (1)$$

where Δp is the pressure difference between the two nodes connected by the edge, and c is the conductivity of the edge, which obeys proportionality $c \sim w^4$. We further define the conductance $c^{\text{eff}} = c/\ell$.

Defining \mathbf{B} as the $n_n \times n_e$ incidence matrix for the network allows us to write Eq (1) for the entire network as

$$\mathbf{f} = \mathbf{C}^{\text{eff}} \mathbf{B}^T \mathbf{p}, \quad (2)$$

where \mathbf{f} and \mathbf{p} are flux and pressure vectors, respectively, and \mathbf{C}^{eff} is a matrix with the conductances on its diagonal.

Combining this with Kirchhoff's law

$$\mathbf{B}\mathbf{f} = \mathbf{s}, \quad (3)$$

where \mathbf{s} is the source-sink vector, gives us

$$\mathbf{BC}^{\text{eff}}\mathbf{B}^T\mathbf{p} = \mathbf{Ap} = \mathbf{s}, \quad (4)$$

where $\mathbf{A} = \mathbf{BC}^{\text{eff}}\mathbf{B}^T$ is the weighted graph Laplacian, a symmetric matrix of rank $n_n - 1$. Since \mathbf{s} is in the column space of \mathbf{A} , we can solve for the pressures as

$$\mathbf{p} = \mathbf{A}^\dagger \mathbf{s}, \quad (5)$$

where \dagger denotes the pseudo-inverse (see Sect 2 in SI Text 1 for an efficient computation of the above). Lastly, inserting Eq 5 into Eq 2, we can compute the fluxes as

$$\mathbf{f} = \mathbf{C}^{\text{eff}}\mathbf{B}^T\mathbf{A}^\dagger\mathbf{s} \equiv \mathbf{Gs}, \quad (6)$$

where we have defined the matrix \mathbf{G} for convenience.

We focus our attention on energy-efficient networks, which have previously been shown to be a suitable model to describe leaf venial systems [5,6,12,13,15]. Thus we minimize the energy dissipation

$$P = \sum_i \frac{f_i^2}{C_{ii}^{\text{eff}}}, \quad (7)$$

subject to constant material cost $\sum_i \ell_i c_i^\gamma$. Here γ is a parameter that determines the cost of building and maintaining the (biological) network. The higher the value of γ , the flatter the hierarchy of vein widths, as thicker widths become more costly. Using the method of Lagrange multipliers, it can be shown that the optimal conductivities must scale with the fluxes as [5]

$$c_i \propto |f_i|^{2/(1+\gamma)}. \quad (8)$$

This equation can henceforth be used to find the optimal conductivities by fixed-point iteration.

We have now fully specified the theoretical foundation for the calculation of fluxes and optimal conductivities for a given source-sink vector \mathbf{s} . A standard choice is $\mathbf{s} = [1, \frac{-1}{n-1}, \frac{-1}{n-1}, \dots, \frac{-1}{n-1}]$, where each sink is constant and of equal magnitude. To enable computations on the non-uniform node distribution that emerges from leaf-image network extraction, we assign to each node j a source or a sink term proportional to the area a_j of the Voronoi cell associated with the node. This is motivated by the assumption that the drainage around a node is proportional to the surface area surrounding that node.

Power minimization with a constant source-sink vector \mathbf{s} leads to networks with tree topologies [5]. However, the loopless optimal structures emerging from this are vastly different from the highly loopy networks seen in many species of leaves. Several models have been developed to mitigate this issue, e.g., the cut-bond and the moving sink models [12,13]. Here, we generalize the moving sink model in order to account for the variations in the node-associated drainage area. In the original model [12], the squared flux of Eqs (7) and (8) is averaged over a distribution of source-sink vectors, parameterized in terms of sink nodes, i.e.,

$$\langle f_i^2 \rangle = \frac{1}{n_n - 1} \sum_k f_i^2(\mathbf{s}(k)), \quad (9)$$

where the sum in k runs over all sink nodes. Note that the average is over the squares of flows, which is what distinguishes a model using fluctuating sinks from a static sink model.

In our approach, the distribution of \mathbf{s} is constructed as a combination of constant overall network load and locally fluctuating sink sizes. Thus we define source-sink vectors with excess fluid removal at the k -th node, making sure that the sink sizes are proportional to the corresponding drainage areas:

$$s_j(k) = \begin{cases} 1 & j = \text{source index} \\ -a_j(s_{\text{avg}} + \delta_{jk}\sigma)C(k) & \text{otherwise.} \end{cases} \quad (10)$$

Here, $s_{\text{avg}} = \frac{1-\sigma}{n_n-1}$ is the same constant value for all sinks, and $\delta_{jk}\sigma$ represents the influence of the moving sink, the relative amplitude of which is controlled by the sink fluctuation parameter $\sigma \in [0, 1]$. In the limit $\sigma = 0$, which corresponds to no sink fluctuation, we have that the fluid removal is solely dependent on the drainage area. Conversely, in the limit $\sigma = 1$, we have full sink fluctuation, i.e., $s_k = -1$, and $s_{j \neq k} = 0$ (excluding the source node). This limit corresponds to the original moving sink model [12]. Lastly, the normalization constant $C(k) = (\sigma a_k + s_{\text{avg}} \sum_{j \in \text{sinks}} a_j)^{-1}$ ensures that for any k , $\sum s_j(k) = 0$, and is necessitated by the fact that the amount of fluid entering the leaf through the source always equals the amount escaping through the sinks. Keeping the source term constantly equal to unity, and only letting the sink values vary, corresponds to constant load on the system level. We note that the moving sink model that we employ here is only one of many possible choices, and that the influence of fluctuations on loop formation is generally very complex [14].

2.1.1. Optimization on extracted graphs. The theoretical model specified above can in principle be run on any connected graph. Our model uses networks extracted from images of real leaves as input. This raises some interesting questions. For example, how can we find the parameters that best reproduce the vein widths observed in leaf data, and is it possible to define new relevant statistics based on the knowledge of the fluid flow in the venial network?

To perform computations on the large networks of real leaves, we are faced with two complications: (1) It is computationally expensive to run models on networks with the number of nodes and edges in the thousands, and (2) the loss function of a complex transport network model of natural leaves is, in general, highly non-convex [6], which leads to issues in choosing which optima to compare to.

Regarding computational efficiency, we detail in the Methods section our approach of dividing the large networks into subsystems, as well as our use of hardware acceleration to speed up computations within each system. As for non-convexity, consider the small example system in Fig 3a. Here, we initialize two networks, each with a uniform distribution of edge widths, with small (but different) perturbations added. Optimizing the power [Eq (7)] results in two entirely different solutions, the networks of which are topologically distinct. The two minima have very similar power usage, illustrating the highly non-convex nature of the problem. This makes comparisons with data challenging, as there is a plethora of candidates for the correct local minimum to use.

Our solution to this problem is simple: As our aim is to find the local minimum closest to that of the input data taken from the leaf image, we choose the leaf image data widths as initial values for the iteration process. This should ideally constitute the best starting point for reaching the desired optimal model output. The procedure is illustrated in Fig 3b, where the initial state (pink) is located relatively close to the desired local minimum (blue) in the non-convex landscape.

2.1.2. The optimal model network. We now have the necessary ingredients to find the optimal network that most closely models an experimental leaf. In particular, this enables

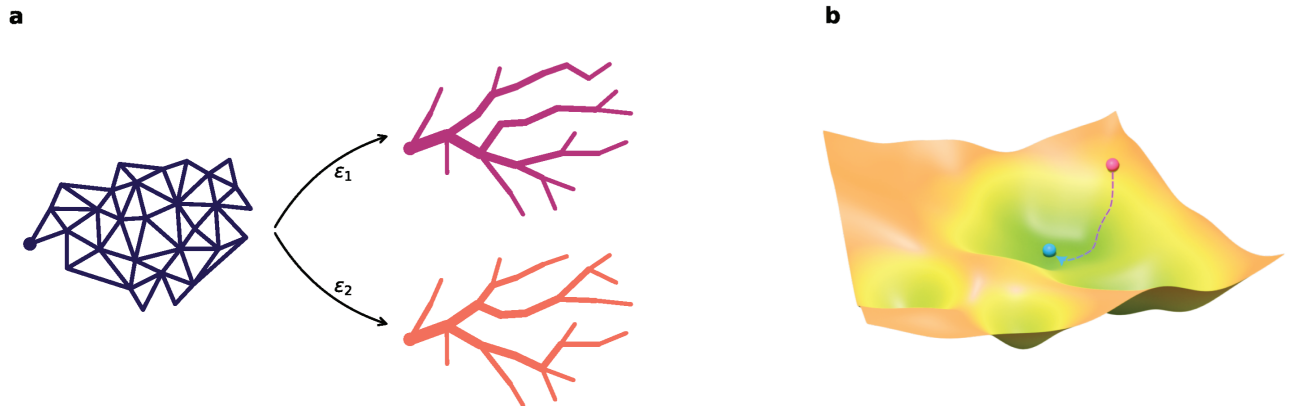


Fig 3. Local optimization. (a) Applying two different, but small perturbations, starting from a uniform edge width distribution, leads to very different topologies in the final minima. (b) The final solution can be biased by starting near the desired local minimum.

<https://doi.org/10.1371/journal.pcbi.1013292.g003>

the direct comparison between the edges from leaf images and model output. Specifically, this is achieved by solving for the edge conductivities using fixed-point iteration and then converting them back to edge widths. We thus initialize $c_i^{(0)} = (w_i^{\text{data}})^4$, and iterate

$$f_i^{(n+1)}(k) = \sum_j G_{ij}^{(n)} s_j(k),$$

$$c_i^{(n+1)} = \left\langle (f_i^{(n+1)}(k))^2 \right\rangle_k^{1/(1+\gamma)}$$

towards convergence in **c**. The second equation follows from Eq (8) combined with the fluctuating sink average of Eq (9). Finally, we calculate the edge widths, $w_i^{\text{model}} = c_i^{1/4}$.

We note that in the above, we have worked in arbitrary units of conductivity. The scale of the conductivity (and of the output widths) is set by the material constraint, which in turn serves to set the scale of the output. Instead of finding the optimal value for this budget, we simply normalize the output after optimization:

$$\hat{w}_i^{\text{model}} = \frac{\langle w_i^{\text{data}} \rangle}{\langle w_i^{\text{model}} \rangle} \cdot w_i^{\text{model}}. \quad (11)$$

2.2. Applications

Having fully defined the model, we apply it to a set of leaf images from the three species *S. albus*, *L. xylosteum*, and *C. monogyna*. For each leaf, we extract a network representation of the observable venation pattern and use this pattern directly as input to our hydrodynamic model to obtain output data on each individual edge (example model output is shown in Fig 1, bottom). The fact that we can access output data for individual edges allows us to perform a number of experiments: First, we fit the sink fluctuation amplitude σ to the image data from each of the three species. This will reveal that each species exhibits a distinct optimal value for this parameter. Second, we investigate the use of Murray's law for reticulate venation networks generated by the model.

2.2.1. Sink fluctuation amplitude. We now consider whether direct comparisons of the model output to the extracted leaf networks can be used to fit model parameters. This contrasts with previous approaches [7,17], where summary statistics typically serve as the metrics for comparison. While this can often be sufficient, the approach ignores the specificities of the individual edges during comparison. Here, we use the full-scale leaf image network structure as input data, allowing for individual edge-to-edge comparisons with our model. We hence ask the question of whether this approach can be used to fit theoretical venation models. Thus, we define a loss function as an average over each individual edge,

$$L(\sigma) = \left\langle \frac{(w_i^{\text{data}} - \hat{w}_i^{\text{model}}(\sigma))^2}{w_i^{\text{data}}} \right\rangle_i, \quad (12)$$

which averages the squared relative difference, weighted by the widths (which corresponds to assigning every vein pixel in the original images equal weight). We minimize this loss function with respect to the sink fluctuation parameter σ , while fixing $\gamma = \frac{1}{2}$ (corresponding to a fixed-volume budget).

The results for the three species can be seen in Fig 4a. The different leaves within a species exhibit the same qualitative behavior over the range of sink fluctuation values, and each leaf has a well-defined optimum value for σ . Furthermore, the optimal values for a given species are located within a distinct range, so we use their means (indicated by dashed vertical lines) to represent an optimal value for the species as a whole.

The distributions of the optimal values can be seen in Fig 4b. Here we find that both the means and the spreads are markedly different between the species, indicating variability not only between species but also within species. The distributions appear to be both distinct and well-defined, which suggests sink fluctuation as a relevant parameter for species classification based on the venation pattern [17].

2.2.2. Individual vein comparisons. In order to explore how well the model fits the data on an edge-to-edge level, rather than only relying on summary statistics, we now choose one representative leaf from each species, and study the distributions of the edge widths. In Fig 4c–4e, we compare the distributions between the image data and the model output, using the species-specific fit values as parameters.

The overall shape of the vein width distribution is captured well by the model. Nonetheless, it is clear that there are discrepancies between the real leaves and model outputs. More concretely, the most pronounced difference between the two distributions occurs at the smallest widths (note the logarithmic axes). In general, the model underestimates the number of small veins in the data, but conversely produces a significant number of veins of size zero. Note that by construction, there are no zero-width edges in the data. Lastly, a histogram for model output with sink fluctuation set to zero can be seen for all three leaves. This distribution appears both to have a longer tail and to have a higher number of zero-width veins than the corresponding optimal-sink fluctuation distribution. This is in line with the fact that the mean vein width is independent of sink fluctuation (due to normalization), so a higher number of zero-width veins must be compensated by a higher number of wider veins. The relatively higher number of veins of size zero is also consistent with less sink fluctuation, as expected.

In our approach, we can move beyond the comparison of histograms and compare each individual vein width of the model output with the corresponding vein width in the leaf image data. This is shown in Fig 4f–4h. This use of individual edges in the calculations gives a more detailed insight into the model and provides more information than histograms, which describe only the global scale. This insight reveals that the apparent overlap between data and

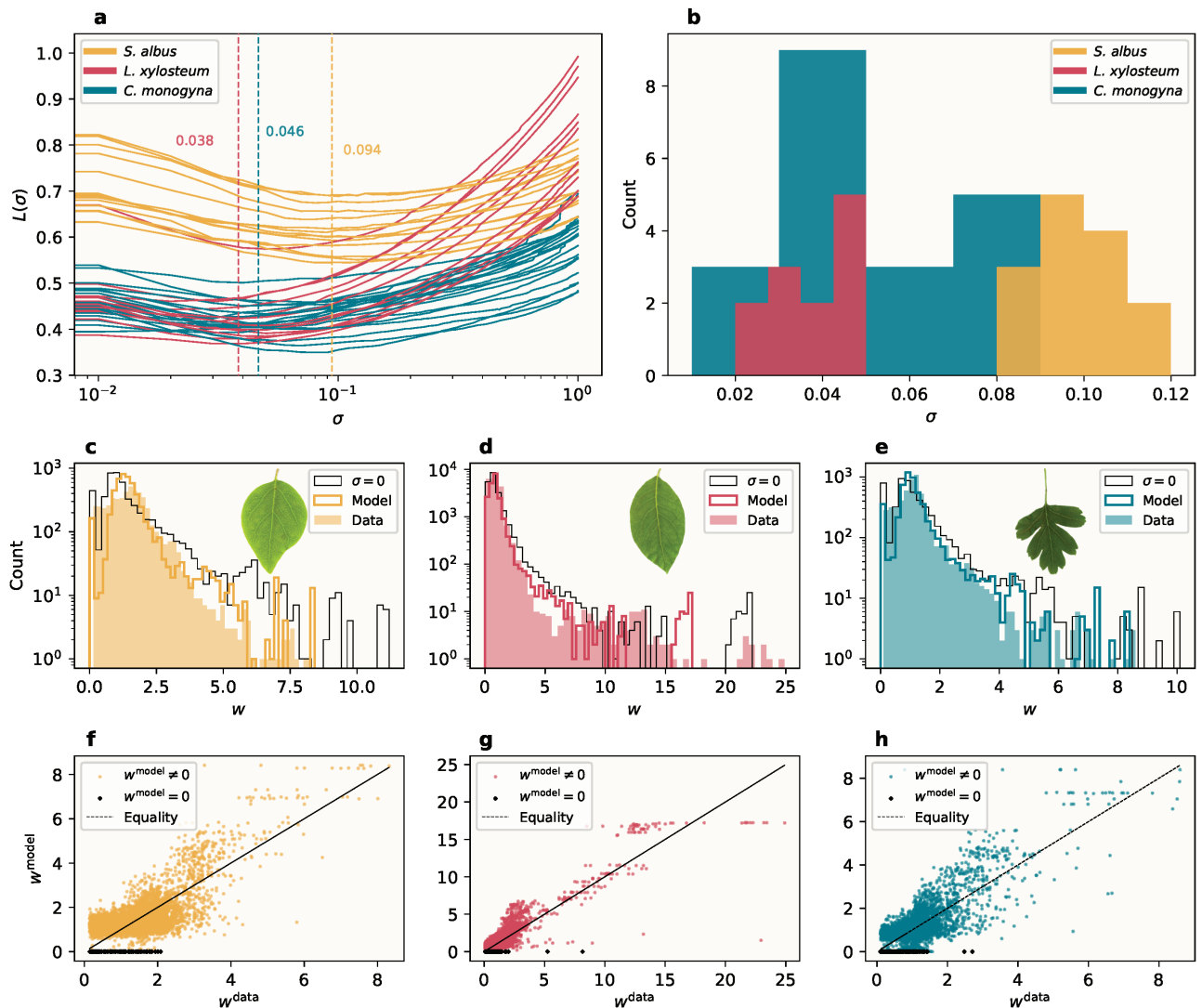


Fig 4. (a) Loss values as a function of sink fluctuation amplitude σ for leaves from the three different species. The average minimum loss values are indicated for each species. **(b)** Distributions of the optimal values of σ for each species. **(c–e)** Vein width distributions for specific leaf samples, one from each species. The distributions are plotted for image data and model output, using the species-specific optimal value of σ , as well as zero sink fluctuation. Note that the histograms are shown with logarithmic axes. (Insets) The actual leaves analyzed. **(f–h)** Model output vein widths as a function of the corresponding image data vein widths.

<https://doi.org/10.1371/journal.pcbi.1013292.g004>

model does not occur on an individual edge level, but is a global phenomenon. We can see the same trend in all three leaves: Although the points are distributed around the line of equality, the varying degrees of spread seem to be significant. Here we can see clearly how the model overestimates the widths of the smallest veins, for which the model has a preferred size. This overestimation is partly due to the normalization and compensation for the significant number of model output vein widths with zero size (black dots). We note that at the level of individual edges, inaccuracies that originate in the preprocessing of the data become more pronounced and that the discrepancies arise due to a combination of model-data mismatch and preprocessing errors.

In conclusion, we find that a single-variable explanation is insufficient and that more complex physical models are needed to fully capture the individual variations in the venation patterns.

2.2.3. Murray's law. Reticulate network formulation. We now exploit the fact that we can run the venation model on full-scale leaves to extend Murray's law [4,18] so that it takes into account the loopy structure in the networks. Murray's law states that the relationship between the radius r_p of the parent vein and the radii $r_{d,i}$ of the daughter veins at a branch point is $r_p^\alpha = \sum_i r_{d,i}^\alpha$, with the original law setting $\alpha = 3$, though different values of α have been found in other types of networks [19,20]. Notably, the law was originally restricted to tree-topology networks, and attempts at applying it to reticulate networks typically involve simplifying the network structure (such as reducing it to tree topologies) before performing the calculations [7,21]. To make the law well-defined for loopy networks, we here present a natural extension to it by including sink fluctuation in its definition.

In the following we calculate the expression $R^{in/out}$ for in- or outgoing veins, respectively (initially ignoring the assumption that the network be leak-free for simplicity). For a given branch node b with degree >2 , we take the sum of in- or outgoing vein radii, each term to the power of the Murray exponent α , and then average over possible moving sink positions k in the leaf. Thus, we define

$$R_b^{in/out}(\alpha) \equiv \left\langle \sum_i I_{b,i}^{in/out}(k) \cdot r_{b,i}^\alpha \right\rangle_k, \quad (13)$$

where $I_{b,i}^{in/out}(k)$ is an indicator function, the value of which depends on whether the flow of vein i at branch point b is in- or outgoing, given the moving sink position k , respectively. The average over different flow directions is motivated by the fact that the reticulate leaf network has been optimized to accommodate varying flow conditions. A consequence of this is that the number of terms in the sum can change as the sink distribution changes.

The variations in flux in a venation network due to sink fluctuations are illustrated in Fig 5a. Here we show a close-up of a leaf image, where the directions of flow are indicated with arrows, along with the model output edge widths. The arrow sizes indicate the average fraction of time when the flow moves in the indicated direction.

As mentioned, the flow direction in a given edge depends on the moving sink position k . Consequently, when calculating Murray sums, we have to account for these changing directions by averaging over all possible sink positions. The extended version of Murray's law can thus be defined as

$$R_b^{in}(\alpha) = R_b^{out}(\alpha). \quad (14)$$

This further allows us to construct a loss function for the system. Here we take the squared difference of the sums and in turn average the resulting residuals over the number of branch nodes in the leaf,

$$L(\alpha) = \left\langle \left(R_b^{out}(\alpha) - R_b^{in}(\alpha) \right)^2 \right\rangle_b, \quad (15)$$

the minimum of which defines an optimal value for α .

Optimizing α . In the following, we apply the extended Murray's law of Eq (14) to the model output networks. We first search for an optimal value for α for each of the three species of leaf, using the loss function in Eq (15). Fig 5c–5d show the loss values for a range of different exponents α for sink fluctuation $\sigma = 0$ and the best-fit value for each leaf (see Fig 4a–4b), respectively. The losses are shown for all leaf samples, with the mean loss for each species

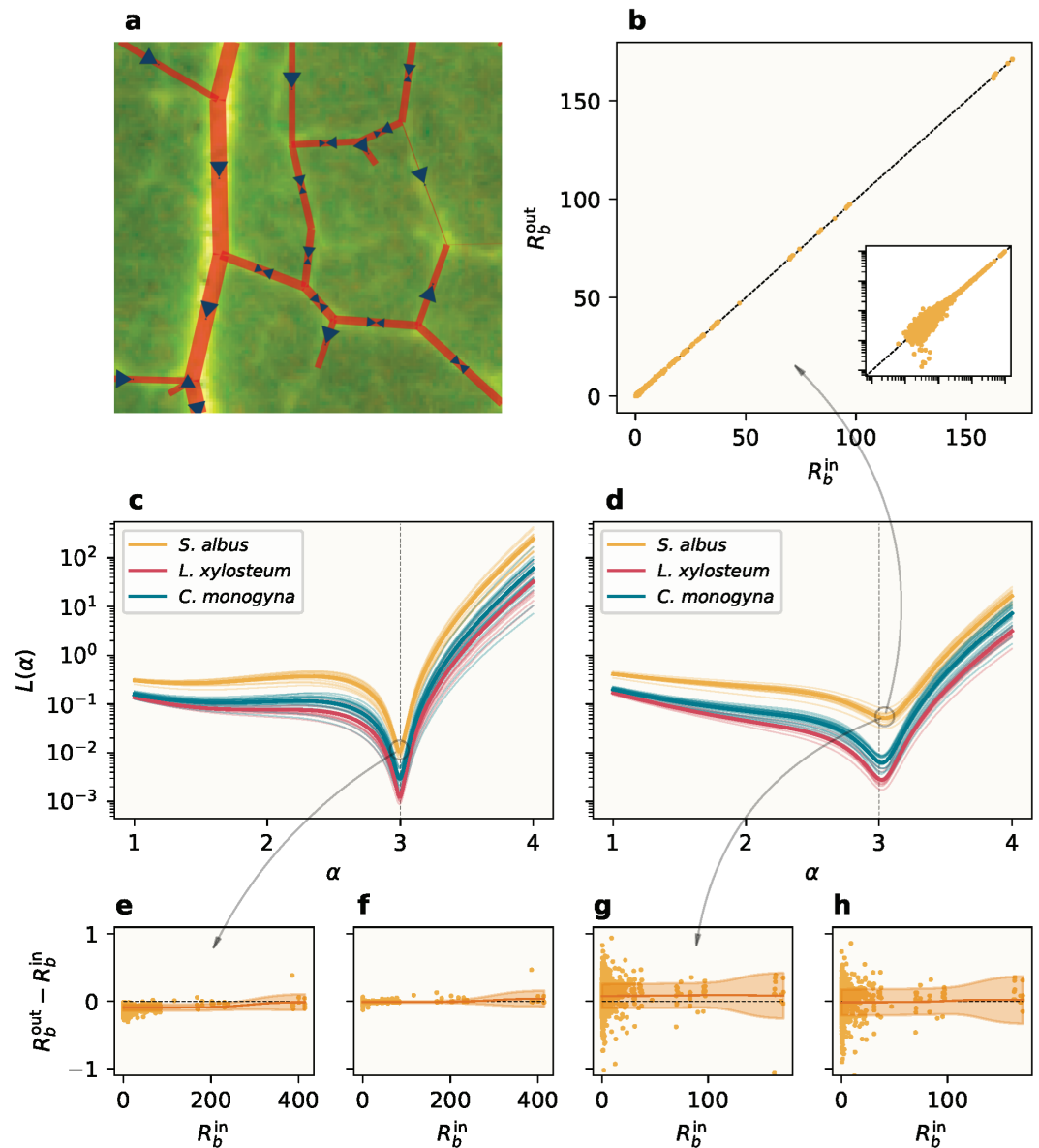


Fig 5. Murray's law for loopy networks. (a) Subsection of a leaf image, with model output edge widths (red lines) and fluxes (blue arrows) indicated. The arrow sizes indicate the relative average time the flow moves in the respective directions given by the arrows. (b) R_b^{out} vs. R_b^{in} for all nodes in a sample of *S. albus*, for the optimal $\alpha \approx 3.03$ and $\sigma \approx 0.094$. The points closely follow the line of equality. (Inset) The same data with both axes logarithmic. (c-d) Loss values for individual leaf samples (thin lines), with the mean value for each species (thick lines). The dashed vertical line is located at $\alpha = 3.0$. (c) $\sigma = 0$. The optimal value of $\alpha = 3$ is predicted for all samples. (d) Best-fit values of σ [Fig 4a-4b]. The optimal values have been shifted in the positive direction. (e-h): Residual plots with Gaussian weighted means and standard deviations (SD). (e-f) $\sigma = 0$. The residuals are (e) biased negatively, and (f) subsequently corrected by a term proportional to the associated sink area. (g-h) $\sigma = 0.094$. The residuals are (g) biased positively, have a larger SD compared to the case of no sink fluctuation, and are (h) corrected by a term proportional to the associated sink area.

<https://doi.org/10.1371/journal.pcbi.1013292.g005>

indicated in bold. In both cases, we see that the branching observed in the model output leads to well-defined Murray exponents that are consistent on the species level. With no sink fluctuation, we find optima at $\alpha = 3$, as expected, and with the inclusion of sink fluctuation, all

species have optima near $\alpha = 3$ with small variations that are larger between species than within species. This indicates that the optimal values are significant.

We note that setting $\sigma = 0$ when running the model leads to output networks with tree topologies, for which the original definition of Murray's law should hold exactly. Indeed, in Fig 5c, we see that the original exponent $\alpha = 3$ is predicted for every individual leaf sample. However, we also observe that the loss values are not exactly zero for any of the samples. This non-zero loss is due to the presence of sinks in our model, where the original model assumes leak-free flow. This leads to a bias in the assumed outgoing fluxes at the branch points, a phenomenon that should likewise be present in real data. Thus, unless the dissipation of fluid is taken into account, a direct application of Murray's law cannot be expected to perfectly fit data.

In Fig 5d, where we have non-zero sink fluctuation, the predicted optimal exponents have shifted in the positive direction for all samples. This shift occurs due to the flattening of the hierarchy of vein widths caused by the fluctuating sinks, and the magnitude of the effect is influenced by both the fluctuating sink amplitude and the fact that the network topologies in leaves from different species are fundamentally different. The latter results in outgoing veins that are relatively wider than in the case of no sink fluctuation, which in turn implies that the minimum of Eq (15) will occur at a higher value of α . We still observe species-level consistency, as well as greater variability between species than within species, which supports that the predicted exponents are significant.

We now explore the data in more detail by looking at the individual branch nodes. In Fig 5b we plot individual R_b^{out} versus R_b^{in} for all branch nodes in a representative sample of *S. albus*, using the observed optimal value of $\alpha \approx 3.03$, as well as the best-fit value for σ . The points very closely follow the line of equality. This indicates that Murray's law is obeyed over all ranks of veins, although the inset figure (which shows the same data with logarithmic axes) reveals a larger variation for lower-rank veins, where the effect of sink fluctuation is the highest. Lastly, Fig 5e–5h show residual plots for the same leaf sample and value of α . We use Gaussian weighted means and standard deviations (SD) to illustrate the discrepancies over vein widths. In Fig 5e, which corresponds to $\sigma = 0$, we see that all residuals are negative, with small variations across vein widths. The SD is more or less constant across scales, but the residuals also reveal a negative bias. This suggests that the sink values have to be taken into account in order for Murray's law to be valid for models that employ sinks at the branch nodes. We investigate this by correcting the residual of each branch node by adding a value proportional to the area of the associated sink. The resulting residuals can be seen in Fig 5f. All values have been shifted such that the mean is equal to zero for most branch nodes, while the remaining discrepancies are due to the finite convergence threshold of our simulations. Fig 5g–5h show the corresponding plots for the best-fit value of σ . Now, the mean values are positive, and we see a much higher variability, particularly for small-vein branches. The positive bias occurs due to the aforementioned flattening of the vein width hierarchy. This means that there are now two effects acting in opposite directions: The flattening of the vein width hierarchy, and fluid leakage through sinks, which lead to positive and negative bias in the residuals, respectively. In the same fashion as before, we attempt to correct the residuals by values proportional to the sink areas. Although the mean values have been shifted towards zero, the effect on the variability is only barely noticeable in this case.

3. Discussion

In this paper, we have shown that hydrodynamic models can be run on full-scale leaves, performing analyses on the level of individual network nodes and edges. This approach stands

apart from those of earlier studies and brings with it both advantages and disadvantages. As an example, looking at the distributions in Fig 4c–4e, it is initially difficult to dismiss the notion that the model faithfully reproduces the image data. However, in the corresponding plots of individual vein widths in Fig 4f–4h we observe significant deviations from the image data, and even a clear bias in some regimes. When applied to high-quality data, this aspect of our approach thus provides a more fine-grained method of analysis for leaf venation modeling, in that it exposes the insufficiencies of using mere distributions to compare theoretical venation models to experimental leaf data.

The hydrodynamic model is adapted to the morphology of experimental leaf images by using, e.g., area-weighted, fluctuating sinks. Including sink fluctuation in the model facilitates the reproduction of loopy structures [12,13], and is in our case controlled by the amplitude parameter σ . This parameter was used to perform fits, which yielded species-dependent parameter distributions. However, the use of a single parameter for this purpose is almost certainly insufficient, as we are fitting tens of thousands of edge widths. On the other hand, the data is extremely systematic in its variations. It is therefore pertinent to ask whether the observed discrepancies are caused by either systematic differences or stochastic fluctuations during the venation network formation. For instance, in our study we set $\gamma = \frac{1}{2}$, which defines a volume-based material budget, but this parameter could also turn out to be species-dependent, just as we found for sink fluctuation amplitude. More generally, being able to compare model and data on an edge-to-edge level allows for an inverse approach to the problem: instead of merely fitting parameters of predefined models, local feedback models can be parameterized (for example, by neural networks) and then optimized. This would allow data-driven model discovery for leaf venation feedback models. In particular, we have implemented our approach using JAX, which has enabled hardware acceleration in this study, and has the potential to further facilitate automatic differentiation in our simulations. It would therefore be feasible to train such neural networks end-to-end, which is an intriguing avenue for further research. While our study does not answer the fundamental question of whether the morphogenesis of leaf venation results from local hydrodynamic feedback, it opens the door for a more data-driven approach for further investigation on this topic.

There is a number of caveats to our approach. Crucially, we rely on an accurate preprocessing pipeline. The venation models are compared to the networks extracted from segmented images, thus any bias in this procedure will be reflected in the analyses performed. Furthermore, we want to demonstrate that both tree-topology networks, as well as fully loopy networks can be reproduced by the model, which implies that output edge widths can take a value of zero. When mapping the leaf image to the planar graph, we only include the nodes and edges that represent the actual leaf, so there are, in other words, no edges of zero width in the input data. Consequently, the model can only be correct if it sets all output edge widths to non-zero values. According to Fig 4c–4h, however, there is apparently a trade-off: in order to match the image data, the best-fit model does set a significant number of output edge widths to zero. This behavior is fundamentally caused by inherent differences between our model and the leaf data. Conversely, due to the fact that the input network topology is completely determined by the experimental images, we cannot at present observe whether the model would output network edges that do not exist in the input data. This issue could perhaps be mitigated by adding additional, artificial edges to the leaf input data before running the model, which could further constrain a data-driven model discovery.

Finally, we showed how our approach enables a novel definition of Murray's law that applies to loopy networks [Eq (14)]. Thus we circumvent the need to, e.g., transform the graph to an approximate tree-structure [21]. We find this definition highly applicable to reticulate networks, although with minor discrepancies, the nature of which we explore further: When

calculating the Murray exponent, we initially assume leak-free flow through the branches, which contradicts the fact that fluid is removed from the xylem caused by evaporation and photosynthesis. Correcting for this, we find that this effect is only important for loopless networks, and thus that the insufficiency of Murray's original law applied to reticulate networks far outweighs the inclusion of fluid dissipation.

We also assume that the reticulate networks are purely optimized for fluid flow, and not for structural support. This does not happen in plant xylem in general, although the assumption arguably holds for leaves [18]. Additionally, there are indications that the optimization of structural support and fluid flow velocity primarily occur in the larger veins [21]. Since leaves are dominated in number by small veins, structural support can mostly be neglected in our case.

4. Methods

4.1. Extracting the network from leaf images

The leaf vein structure is extracted from the images using the *Ilastik* application [16]. A 2D representation of the venation structure is made by classifying the image pixels into veins and non-veins [Fig 2]. This binary image is then skeletonized using a medial axis transform. End-points and branch-points of the skeleton are found using simple binary hit-miss operations. A graph is then constructed by recursively walking between these node points on the skeleton, simultaneously calculating the width of the edges using the distance from the skeleton to the background. The final graph is pruned by merging nodes that are very close to each other, as well as removing all 2-degree nodes.

Furthermore, to each node is associated a drainage area, which we take to be the size of the corresponding surface in the vicinity of the node. These areas are calculated by using Voronoi tessellation and employing alpha shapes for the leaf boundary.

For algorithmic details, the code can be found at <https://github.com/kirkegaardlab/full-scale-leaves>.

4.2. Hardware acceleration

We mitigate some of the computational bottlenecks by implementing our models in *JAX*, which enables hardware acceleration on GPUs. This allows us to execute large matrix operations far more efficiently than on CPUs, given that GPUs are optimized for such parallel tasks. The approach does however introduce a trade-off: While the use of GPUs yields significantly faster computations, it necessitates more careful memory management, as GPUs generally have less memory than CPUs.

4.3. Subsystems

Hydrodynamic models of leaf networks are typically run on artificial networks, i.e., small, regular grids [5,12,13,15]. The size of the networks emerging from the real leaves used for this study is very large, and for most leaves we have $n_n, n_e \sim 10^4-10^5$. This sets the scale of the relevant matrices on which we must compute, e.g., the matrix inverse, which scales as $\mathcal{O}(n^3)$. Clearly, it would be infeasible to perform these computations at every step of an optimization procedure. In addition, the use of GPUs imposes a relatively low upper bound on the size of the matrices that can fit in memory at any time. We thus adapt our model to these large networks by dividing them into subsystems, and run the model on each of these separately.

Because we initialize the system using the image data widths, we can assume that the changes in a subsystem only affect the surrounding network to a minor degree. In other

words, the effects of a local change have limited range within the entire leaf. Concretely, we assume that the conductivities outside of a subnetwork remain constant during the optimization of the subnetwork itself. We use k -means clustering to divide the network into a number of subsystems (~ 10), and calculate fluxes by

$$\tilde{\mathbf{f}} = \tilde{\mathbf{G}}\tilde{\mathbf{s}}, \quad (16)$$

where the tilde indicates that only the matrix rows and columns corresponding to the nodes and edges from the subsystem are included. A source-sink vector of the full system is translated to a source-sink vector for the subsystem by adding the (possibly negative) net flow crossing into the subsystem from the surrounding full system,

$$\tilde{s}_j = s_j + \sum_{i \in \text{crossing}, j} f_i^{\text{full}}, \quad (17)$$

where the sum is over the incident edges of node j (if any) that are also incident to nodes outside of the subsystem. Technically, as detailed below, \tilde{s}_j is calculated using predefined translation matrices that only need to be calculated once for the full system, and do not require updating during optimization.

An example of this process can be seen in Fig 6. The demarcation of one hypothetical subsystem within a full leaf is shown in Fig 6a, and 6b illustrates how the sink values at the border nodes are affected by the flow in the full leaf. Nodes closer to/further from the leaf petiole act as pseudo-sources/-sinks for the subsystem, respectively. In Fig 6c the resulting network after performing the optimization on the indicated subsystem is shown, assuming no sink fluctuation for clarity. With no sink fluctuation, the stark contrast between the leaf input and model output becomes clearly visible.

Calculating $\tilde{\mathbf{s}}$. Given a subnetwork, let Q be the set of all n_Q nodes within this subnetwork. Let $P(j,k)$ be the statement “Node $j \in Q$, and edge k is incident to node j and node $l \notin Q$.” Then

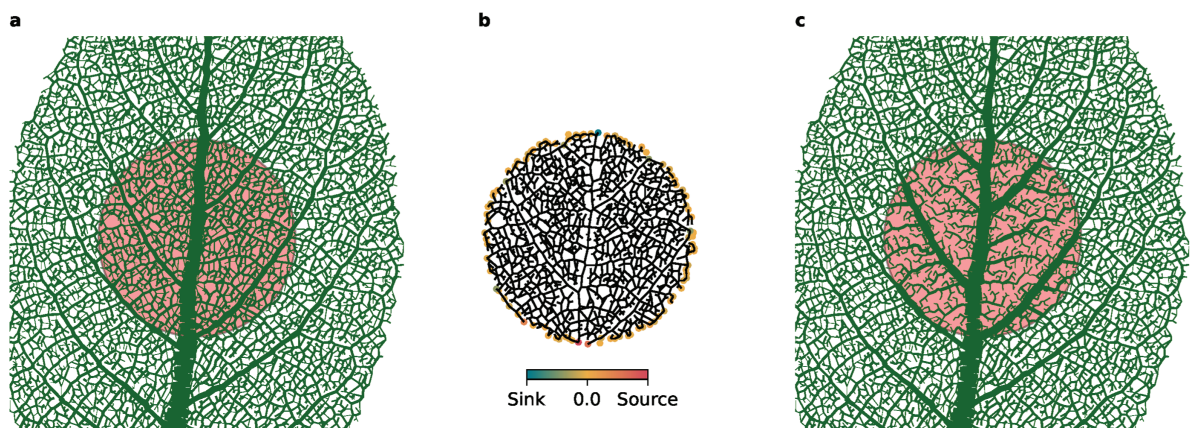


Fig 6. Running the model on a subsystem. (a) Original network with the subnetwork indicated. (b) The subnetwork with redefined sources/sinks, the amplitudes of which are indicated by colors. (c) Output from model with sink fluctuation = 0. Note that most of the reticulate structure has disappeared.

<https://doi.org/10.1371/journal.pcbi.1013292.g006>

the $(n_n \times n_e)$ matrix J can be defined as

$$J_{jk} = \begin{cases} B_{jk} & \text{if } P(j, k) \\ 0 & \text{otherwise,} \end{cases} \quad (18)$$

where B is the incidence matrix of the full system. Furthermore, let $f(j)$ be the enumeration of the nodes in Q . Then we can define the $(n_Q \times n_n)$ matrix M that maps nodes in the large system to the small system, i.e.,

$$M_{ij} = \begin{cases} 1 & \text{if } j \in Q \text{ and } i = f(j) \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

Given source/sinks, s , for the full system, we can calculate the corresponding vector \tilde{s} for the small system by

$$\tilde{s} = M(I - JG)s \equiv Zs, \quad (20)$$

where I is the identity matrix, and $G = C^{\text{eff}}B^T(BC^{\text{eff}}B^T)^\dagger$ as defined in the Model subsection. (Note that G is independent of the subsystem, and thus only needs to be calculated once before the optimization. This is where the method of subsystems facilitates significant computational savings.) This entails that the second term represents the flow between the subsystem and the rest of the leaf, whereas the first term accounts for the nodes within the subsystem only.

Finally, the flow rates for the subsystem edges can be calculated as

$$\tilde{F} = \tilde{C}^{\text{eff}}\tilde{B}^T(\tilde{B}\tilde{C}^{\text{eff}}\tilde{B}^T)^\dagger\tilde{s} = \tilde{G}\tilde{s} = \tilde{G}Zs \quad (21)$$

Note also that Z only needs to be calculated once for each subsystem.

Supporting information

S1 Text. Effects of perturbing the segmentation, Faster computation of A^\dagger , Flow bi-directionality.
(PDF)

Acknowledgments

We acknowledge helpful discussions with Louise Isager Ahl.

Author contributions

Conceptualization: Lars Erik J. Skjægstad, Julius B. Kirkegaard.

Data curation: Julius B. Kirkegaard.

Formal analysis: Lars Erik J. Skjægstad, Julius B. Kirkegaard.

Funding acquisition: Julius B. Kirkegaard.

Investigation: Lars Erik J. Skjægstad.

Methodology: Lars Erik J. Skjægstad.

Software: Lars Erik J. Skjeggstad.

Supervision: Julius B. Kirkegaard.

Validation: Lars Erik J. Skjeggstad.

Visualization: Lars Erik J. Skjeggstad.

Writing – original draft: Lars Erik J. Skjeggstad, Julius B. Kirkegaard.

Writing – review & editing: Lars Erik J. Skjeggstad, Julius B. Kirkegaard.

References

1. Sack L, Scoffoni C. Leaf venation: structure, function, development, evolution, ecology and applications in the past, present and future. *New Phytol.* 2013;198(4):983–1000. <https://doi.org/10.1111/nph.12253> PMID: 23600478
2. Katifori E. The transport network of a leaf. *Comptes Rendus Physique.* 2018;19(4):244–52.
3. Da Vinci L. The notebooks of Leonardo da Vinci. Courier Corporation; 2012.
4. Murray CD. The physiological principle of minimum work: I. The vascular system and the cost of blood volume. *Proc Natl Acad Sci U S A.* 1926;12(3):207–14. <https://doi.org/10.1073/pnas.12.3.207> PMID: 16576980
5. Bohn S, Magnasco MO. Structure, scaling, and phase transition in the optimal transport network. *Phys Rev Lett.* 2007;98(8):088702.
6. Ronellenfitch H, Katifori E. Global optimization, local adaptation, and the role of growth in distribution networks. *Phys Rev Lett.* 2016;117(13):138301.
7. Katifori E, Magnasco MO. Quantifying loopy network architectures. *PLoS One.* 2012;7(6):e37994. <https://doi.org/10.1371/journal.pone.0037994> PMID: 22701593
8. Konkol A, Schwenk J, Katifori E, Shaw JB. Interplay of river and tidal forcings promotes loops in coastal channel networks. *Geophys Res Lett.* 2022;49(10):e2022GL098284.
9. Kirkegaard JB, Sneppen K. Optimal transport flows for distributed production networks. *Phys Rev Lett.* 2020;124(20):208101. <https://doi.org/10.1103/PhysRevLett.124.208101> PMID: 32501061
10. Lonardi A, Facca E, Putti M, De Bacco C. Designing optimal networks for multicommodity transport problem. *Phys Rev Res.* 2021;3(4):043010. <https://doi.org/10.1103/physrevresearch.3.043010>
11. Lonardi A, De Bacco C. Bilevel optimization for traffic mitigation in optimal transport networks. *Phys Rev Lett.* 2023;131(26):267401. <https://doi.org/10.1103/PhysRevLett.131.267401> PMID: 38215368
12. Katifori E, Szöllösi GJ, Magnasco MO. Damage and fluctuations induce loops in optimal transport networks. *Phys Rev Lett.* 2010;104(4):048704. <https://doi.org/10.1103/PhysRevLett.104.048704> PMID: 20366746
13. Corson F. Fluctuations and redundancy in optimal transport networks. *Phys Rev Lett.* 2010;104(4):048703. <https://doi.org/10.1103/PhysRevLett.104.048703> PMID: 20366745
14. Waszkiewicz R, Shaw JB, Lisicki M, Szymczak P. Goldilocks fluctuations: dynamic constraints on loop formation in scale-free transport networks. *Phys Rev Lett.* 2024;132(13):137401. <https://doi.org/10.1103/PhysRevLett.132.137401> PMID: 38613264
15. Hu D, Cai D. Adaptation and optimization of biological transport networks. *Phys Rev Lett.* 2013;111(13):138701. <https://doi.org/10.1103/PhysRevLett.111.138701> PMID: 24116821
16. ilastik - the interactive learning and segmentation toolkit. <https://www.ilastik.org/>
17. Ronellenfitch H, Lasser J, Daly DC, Katifori E. Topological phenotypes constitute a new dimension in the phenotypic space of leaf venation networks. *PLoS Comput Biol.* 2015;11(12):e1004680. <https://doi.org/10.1371/journal.pcbi.1004680> PMID: 26700471
18. McCulloh KA, Sperry JS, Adler FR. Water transport in plants obeys Murray's law. *Nature.* 2003;421(6926):939–42. <https://doi.org/10.1038/nature01444> PMID: 12607000
19. Sherman TF. On connecting large vessels to small. The meaning of Murray's law. *J Gen Physiol.* 1981;78(4):431–53. <https://doi.org/10.1085/jgp.78.4.431> PMID: 7288393
20. Williams HR, Trask RS, Weaver PM, Bond IP. Minimum mass vascular networks in multifunctional materials. *J R Soc Interface.* 2008;5(18):55–65. <https://doi.org/10.1098/rsif.2007.1022> PMID: 17426011
21. Price CA, Knox S-JC, Brodribb TJ. The influence of branch order on optimal leaf vein geometries: Murray's law and area preserving branching. *PLoS One.* 2013;8(12):e85420. <https://doi.org/10.1371/journal.pone.0085420> PMID: 24392008