Computers and Mathematics with Applications 212 (2026) 190–209

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Network inpainting via optimal transport

Enrico Facca [a,*], Jan Martin Nordbotten [b,c], Erik Andreas Hanson [d]

[a] *Department of Mathematics, University of Ferrara, Via Machiavelli, 30, Ferrara, 44121, Italy*
[b] *VISTA Center for Modeling of Coupled Subsurface Dynamics, Department of Mathematics, University of Bergen, Bergen, 5020, Norway*
[c] *NORCE Norwegian Research Center, Bergen, 5838, Norway*
[d] *Department of Computer Science, Electrical Engineering and Mathematical Sciences, Western Norway University of Applied Sciences, Inndalsveien 28, Bergen, 5063, Norway*

A R T I C L E I N F O

2000 MSC:
49-04
49K20
49N99
65N21
65N30
65Z05

Keywords:
Inpainting
PDE
Optimal transport
Image processing
Networks

A B S T R A C T

In this work, we present a novel tool for reconstructing transport networks from corrupted images. The reconstructed network is the result of a minimization problem that has a misfit term with respect to the observed data, and a physics-based regularizing term coming from the theory of optimal transport. Under the assumption that such physical information can be properly identified, we demonstrate through a range of numerical tests that our approach can effectively rebuild the primary features of damaged networks, even when artifacts are present.

1. Introduction

The problem of reconstructing images from corrupted data is a fundamental problem in science, arising in a wide range of problems, from medical imaging [1] to art restoration [2]. This reconstruction process is known in the literature of image processing as *inpainting* [3–5].

Inpainting problems may become particularly challenging when the image we are interested in represents a transport network and we want to recover specific characteristics. Here, for networks, we refer to structures where one dimension (length) is orders of magnitude larger than the others (thickness). For transport networks, we refer to those structures whose main purpose is to transport resources from one location into another, such as the blood vessels or river networks. An example of this kind of network is given in Fig. 1a, where it is easy to imagine a flow entering from below and exiting from the two branches on top.

Transport networks are often acquired via non-invasive/remote-sensing imaging techniques, and the resulting data may be corrupted by loss of data, noise, and/or the presence of artifacts, such as the example shown in Fig. 1b or in many real-world problems [6–8]. If we want to reconstruct the original network from the corrupted one, recovering the connectivity is a priority if we want to simulate a fluid passing through the network. If not properly reconstructed, the data cannot be used in numerical simulations, may require long manual preprocessing, or lead to wrong predictions.

* Corresponding author.
E-mail addresses: enrico.facca@unife.it (E. Facca), jan.nordbotten@uib.no (J.M. Nordbotten), erik.andreas.hanson@hvl.no (E.A. Hanson).

https://doi.org/10.1016/j.camwa.2026.04.003
Received 17 June 2025; Received in revised form 27 March 2026; Accepted 1 April 2026
Available online 13 April 2026

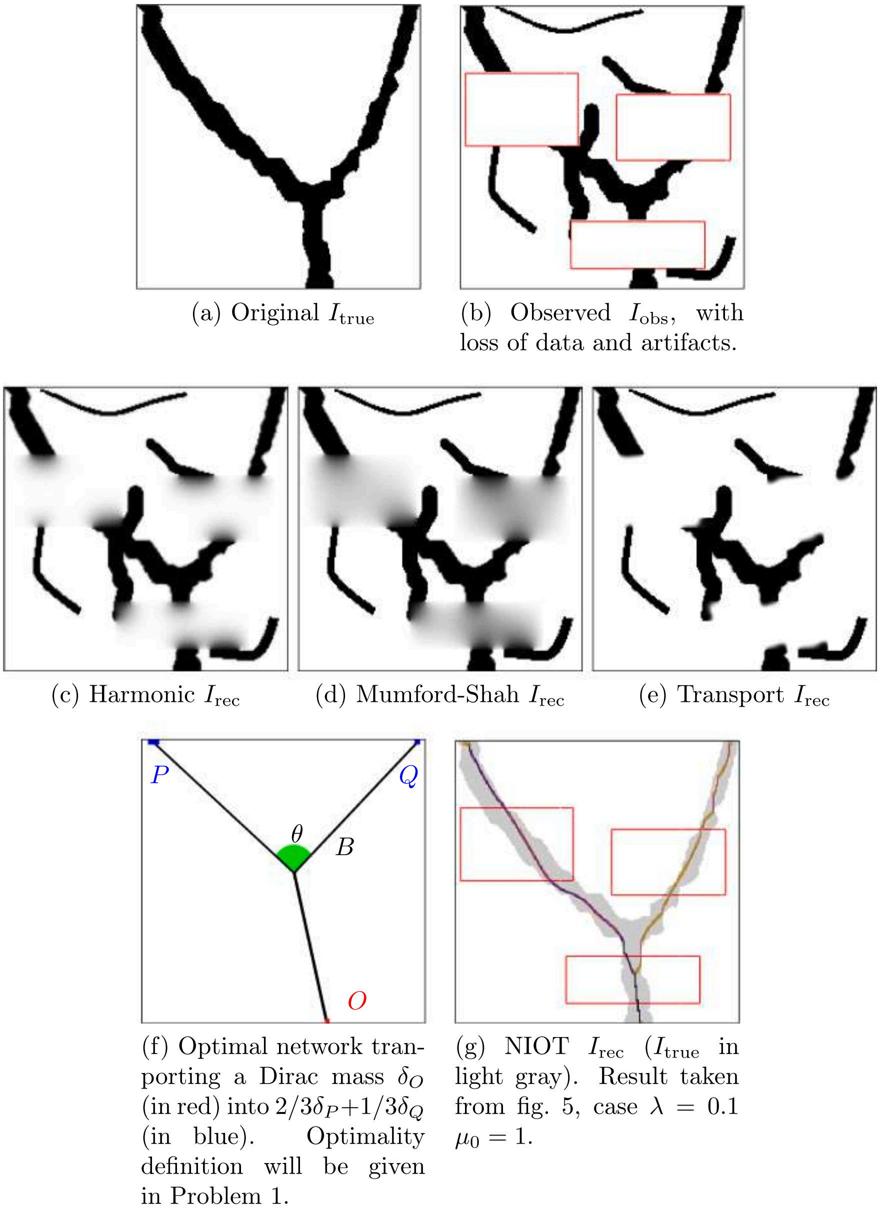


Fig. 1. Schematic representation of the network reconstruction problem. In the upper panels, we report an image I_{true} and a corrupted version I_{obs}. In the central panels, we report the images reconstructed I_{rec} using three inpainting algorithms implemented in [9], the harmonic, Mumford-Shah, and transport inpainting methods. In Fig. 1f we show the optimal transport network. Fig. 1g shows a network reconstructed with the method proposed in this paper.

The approaches used to solve inpainting problems can be divided into two main groups: machine learning-based methods and traditional methods. The former has shown versatile capabilities in recent years (see [10–12] for relatively recent reviews on the topic). However, these methods rely on the availability of a large amount of high-quality data, while in many applications a transparent algorithmic approach is desired.

Among traditional inpainting methods, the most commonly used is to recover missing information by solving PDEs [13–15]. In this setting, an image is seen as a function $I : \Omega \subset \mathbb{R}^d \to [0, \infty[$, in most problems $d = 2, 3$. In the Y-shaped network in Fig. 1a, the domain Ω is a 2d square and the image/function I_{true} is a binary map that indicates whether or not a point belongs to a network.

Traditional methods include the so-called variational inpainting methods, where a reconstructed image I_{rec} is obtained by solving, for a fixed $\lambda > 0$, the following minimization problem

$$\operatorname{argmin}_I \mathcal{J}(I) = \mathcal{R}(I) + \lambda \mathcal{D}(I, I_{\text{obs}}), \tag{1}$$

where different discrepancy functionals $\mathcal{D}$ and regularization functionals $\mathcal{R}$ can be used according to sought images. The parameter $\lambda > 0$ determines the balance between the two terms. Standard discrepancy measures are the weighted L^2 and H^{-1} norms. The most commonly used regularization functionals are the TV norm [15,16], the Cahn-Hilliard functional [17–20], and the Mumford-Shah-Euler functional [21–23].

However, when the image describes a network, this type of regularization functional may not preserve the network structure. In Fig. 1c to e we show how three inpainting algorithms (described in [4] and implemented in [9]) fail in reconstructing the main topological features of the network in I_{true}. This is because they capture only local information and do not express any physical concepts, like the presence of a flux entering from below and exiting from the two branches on top in Fig. 1b mentioned before. Furthermore, in the case of reconstructing transport networks, these traditional regularization functionals ignore the prior knowledge that several networks minimize the energy spent on transport. The most notable examples are natural transport networks such as blood vessels[24], rivers [25], and plant roots [26].

Even when information about the transport occurring within the network is readily available (or can at least be estimated), we still lack a mathematical framework that can simultaneously capture the network's essentially one-dimensional nature, encode the underlying transport physics, and assimilate observed data into the variational formulation.

The main idea of this paper is to address these issues thanks to recent advances in the optimal transport theory described in [27], including a physics-based regularization term $\mathcal{R}$ in Eq. (1). This functional is written with respect to a nonnegative function μ that can be seen as a conductivity. Hence, we rewrite the minimization problem in Eq. (1) in terms of this variable, solving the following problem

$$\mu_{\text{rec}} := \operatorname{argmin}_{\mu\geq 0}\mathcal{J}(\mu) := \mathcal{R}(\mu) + \lambda\mathcal{D}(\mathcal{I}(\mu), I_{\text{obs}}), \tag{2}$$

where $\mathcal{I}$ denotes a suitable map that transforms a conductivity μ into an image I. The reconstructed image/network will simply be $I_{\text{rec}} = \mathcal{I}(\mu_{\text{rec}})$.

The variational problem in Eq. (2) represents the core of the methodology proposed in this paper, which we call Network Inpainting via Optimal Transport (NIOT). The main contribution of this paper is showing how the NIOT approach is able to effectively reconstruct corrupted networks by encoding into equations our knowledge that the observed objects are transport networks.

The manuscript is organized as follows. In Section 2 we give a brief overview of the optimal transport problem and, in particular, of the branched transport problem that studies the ramified structures we are interested in. We also present the functional $\mathcal{R}$ used in Eq. (2) and how it allows us to include branched transport ideas in the framework of image inpainting problems. This integration and the details of its numerical implementation are described in Section 3. In Section 4 we present a series of numerical experiments that present the capabilities and limitations of the proposed approach.

2. Optimal and branched transport problems

The Optimal Transport (OT) problem deals with finding the most efficient way to move a nonnegative measure f^+ into another measure of equal mass f^-, in our problem both defined on $\Omega \subset \mathbb{R}^d$.

In most OT problems, a fixed cost of moving one unit of mass is assigned by a function $c : \Omega\times\Omega \to \mathbb{R}$. Typically, this cost is defined as $c(x, y) = |x - y|^p$, with $x, y \in \Omega$ and $p \geq 1$, for which there is a rich literature (see [28,29] for complete overviews of theoretical and numerical advances on the topic). The fundamental solution to this type of OT problems is that the mass travels along straight lines [30]. For example, if we consider the measures f^+ and f^- reported in Fig. 1f, the OT solution follows a V-shaped path connecting the point O to P and Q, sending $2/3$ of the mass to P and $1/3$ to Q.

However, one may want to model transport phenomena in which a cost of moving one unit of mass decreases for large volumes of goods being transported. In this setting, it becomes more convenient to move the mass together, following a path such as the Y-shaped network reported in Fig. 1f. The mass at the point O first moves to the point B and then splits along two different branches. The subarea of OT that studies this type of branching structures as minimal energy solutions is called the Branched Transport (BT) problem (see also [31] for a general overview of the topic).

The first BT formulation can be traced back to the work of Gilbert in [32], but in this paper we use the formulation described in [33]. Before presenting it, it is useful to state the following simple lemma on the mass balance equation on graphs.

Lemma 1. *Consider a graph $G = (V, E)$, with nodes V and edges E that contains no loops (this type of graph is said to be acyclic). Fix an orientation on the edges E and let $\vec{f} \in \mathbb{R}^{|V|}$ whose entries sum up to zero on each connected component of G. Then there exists a unique flux $\vec{v} \in \mathbb{R}^{|E|}$ that satisfies the following mass balance equation*

$$\sum_{e\in\delta^+(k)} \vec{v}_e - \sum_{e\in\delta^-(k)} \vec{v}_e = \vec{f}_k \quad \forall k = 1,\dots,|V|, \tag{3}$$

where $\delta^+(k)$ and $\delta^-(k)$ denote the set of edges entering and exiting the kth node (according to the fixed orientation).

The proof of this lemma can be found in [34]. It is based on the facts that on each connected component of acyclic graphs the number of vertices is one plus the number of edges. We can now present the BT problem as follows.

Problem 1. Let us denote with δ_x a Dirac measure centered at $x \in \Omega$. Consider two measures

$$f^+ = \sum_{i=1}^{n} f_i^+\delta_{x_i}, \quad f^- = \sum_{j=1}^{m} f_j^-\delta_{y_j}, \tag{4}$$

where:

- n, m are two positive integers;
- $(x_i)_{i=1}^n, (y_j)_{j=1}^m$ are distinct points in Ω;
- $(f_i^+)_{i=1}^n, (f_j^-)_{j=1}^m$ are strictly positive such that $\sum_{i=1}^n f_i^+ = \sum_{j=1}^m f_j^-$;

(in the example shown in Fig. 1f, we have $n = 1$ with $x_1 = O; f_1^+ = 1$ and $m = 2$ with $y_1 = P; f_1^- = 2/3$ and $y_2 = Q; f_2^- = 1/3$).

Consider the set $\mathcal{G}$ of those graphs $G = (V, E)$ such that:

- the graph G is a union of acyclic graphs;
- the set V contains all nodes x_i and y_j, and possibly extra nodes;
- to each edge $e \in E$ we associate a length $\vec{\ell}_e$ given by the Euclidean distance between the connected nodes. We denote by $\vec{\ell}_e(G)$ the vector in $\mathbb{R}^{|E|}$ with all these lengths.

Consider the following vector $\vec{f} \in \mathbb{R}^{|V|}$ with entries $\vec{f}_k$ given by

$$\vec{f}_k = \begin{cases} f_i^+ & \text{if } k\text{-th node} = x_i, \\ -f_j^- & \text{if } k\text{-th node} = y_j, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

for $k = 1, \dots, |V|$. If a graph $G \in \mathcal{G}$ falls under the assumptions of Lemma 1, then there exists a unique flux that satisfies the mass balance equation Eq. (3) with the forcing term $\vec{f}$, which we denote by $\vec{v}(G)$.

Now, fix an exponent $0 < \alpha < 1$. We want to find the optimal graph $G_{\text{opt}} \in \mathcal{G}$ that minimizes the energy

$$E^\alpha(G) := \sum_{e\in E} |\vec{v}_e(G)|^\alpha \vec{\ell}_e(G). \tag{6}$$

The presence of the concave and sub-additive term $|\cdot|^\alpha$ in Eq. (6) favors graphs with fewer edges but a higher flux rate, making the ramified structures optimal. Problem 1 admits a solution since there is a finite number of possible combinations of graphs that meet its requirements [35, Prop. 2.1, 2.2]. However, the number of these combinations increases exponentially with the total number of nodes x_i, y_j, making the BT problem NP-hard.

An explicit description of the optimal graph can be found only in simple cases, such as the problem in Fig. 1f, where there is an analytic formula for the branching point B given in [31]. This formula also shows that the branching angle θ between the edges PB and QB increases as the exponent α passes from one to zero, reaching the maximum value of 120°. The network reported in Fig. 1 corresponds to the case $\alpha = 0.5$. More generally, the minimal branching angle of the network is bounded from below by a quantity that decreases with the exponent α [36, Proposition 2.1].

Extensions of Problem 1 to general measures f^+, f^- have been proposed in recent years [33,37,38], with the latter being the most closely related to Problem 1.

The theory of BT provides a unified mathematical framework for various studies that model natural networks as optimal transport solutions. Examples include rivers [25,39], blood vessels [24,40], and plant roots [26]. However, two fundamental issues limit the use of this theory in applied problems and, more specifically, in inpainting problems considered in this paper.

The first issue is merely computational. Given the NP-hard nature of Problem 1 there is no hope of having an efficient numerical method to solve the BT problem, neither for f^+ and f^- given in Eq. (4) nor for general measures. All methods found in the literature work only for small numbers of nodes [41,42], converge toward local minima [27,43]), or require additional assumptions on the transported measures f^+ and f^- [44].

The second issue is the incompatibility between Problem 1 and the variational inpainting problem in Eq. (1). The first is written in terms of graphs, while in the second, the information on the observed image represents some form of intensity variation in space and is typically stored in pixels/voxels. There is no simple way to integrate these two problems. In the following section, we will introduce an optimization problem closely related to the BT problem, where this integration can be done more naturally.

2.1. A reformulation of the BT

In [27], the following problem has been proposed.

Problem 2. Consider f^+, f^- being two nonnegative densities in $L^2(\Omega)$ having the same mass. Fix $\mu_{\min} > 0$ and $\gamma \in]0, 1]$. Find an optimal "conductivity" $\mu_{\text{opt}} : \Omega \to [0, +\infty[$ solving

$$\operatorname{argmin}_{\mu\geq 0} \mathcal{R}(\mu) := \underbrace{\int_\Omega \frac{(\mu + \mu_{\min})|\nabla u(\mu)|^2}{2}\, dx}_{=:\mathcal{E}(\mu)} + \underbrace{\int_\Omega \frac{\mu^\gamma}{2\gamma}\, dx}_{=:\mathcal{M}_\gamma(\mu)}, \tag{7}$$

where, for any given $\mu : \Omega \to [0, \infty[$, the function $u(\mu)$ denotes the solution $u : \Omega \to \mathbb{R}$ of the equation

$$\begin{cases} -\operatorname{div}((\mu + \mu_{\min})\nabla u) = f^+ - f^-, \\ (\mu + \mu_{\min})\nabla u \cdot n_{\partial\Omega} = 0. \end{cases} \tag{8}$$

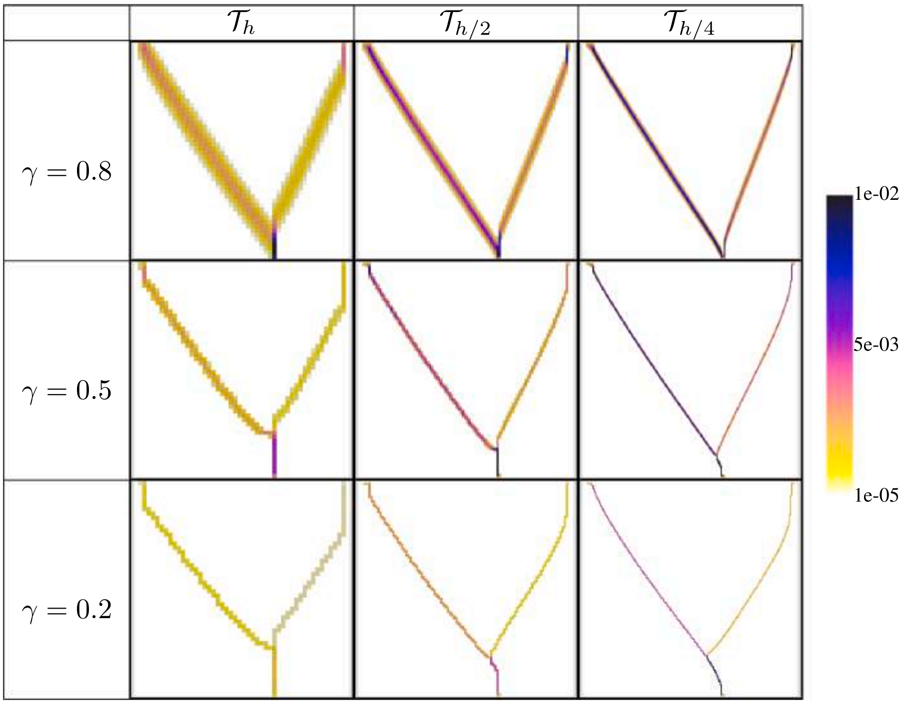


Fig. 2. Spatial distribution of DG^0-approximation of the $\mu_{\mathrm{opt},h}$ solving Problem 2 for the f^+ and f^- shown in Fig. 1f using $\gamma = 0.8, 0.5, 0.2$ (top to bottom) and refining the initial 50×50 grid $\mathcal{T}_h$ (left to right).

Here $\mu_{\min} > 0$ is a small scalar that ensures well-posedness of the PDE Eq. (8). In this paper, we fix $\mu_{\min} = 1.0e - 8$, while in [27] $\mu_{\min}$ was set to zero. Similar problem formulations can be found also in [45,46]. Problem 2 with $\gamma = 1$ and $\mu_{\min} = 0$ is equivalent to the OT problem with a cost equal to the Euclidean distance, as shown in [47].

Intuitively, the support of a function μ in Eq. (7) describes a "pathway" connecting f^+ to f^-. The divergence constraint ensures its linkage. An optimal solution μ_{opt} of Problem 2 provides the best trade-off between energy $\mathcal{E}(\mu)$, which is the Joule energy dissipated via potential flow, and cost $\mathcal{M}_\gamma(\mu)$, which represents the nonlinear cost of building the "pathway" μ. If $0 < \gamma < 1$, this cost is concave and encourages the concentration of the support of μ in small channels with high values of μ. Increasing γ from zero to one, this concentration effect is expected to weaken, passing from networks having few channels, wide branching angles, and high conductivities to networks with more channels, narrower branching angles, and reduced conductivities.

These intuitions are supported by a series of numerical experiments presented in [27] in which a discrete approximation of the minimization problem in Eq. (7) is obtained using finite elements, solution denoted hereafter by $\mu_{\mathrm{opt},h}$. In Fig. 2 we show the approximate $\mu_{\mathrm{opt},h}$ obtained for the measures $f^+ = \delta_O$ and $f^- = (2\delta_P + \delta_Q)/3$ (shown in Fig. 1f) using different values of γ and refining an initial grid $\mathcal{T}_h$ used to discretize the problem (see Section 3.3 for the details of the numerical solution of Problem 2).

We see that for $\gamma = 0.8$ a V-shaped network is obtained, while for $\gamma = 0.5$ and $\gamma = 0.2$ a Y-shaped network is obtained. We can also see that $\mu_{\mathrm{opt},h}$ depends on the mesh $\mathcal{T}_h$ used to discretize the problem. In fact, its support tends to concentrate on narrow "channels" composed of a sequence of single cells in the mesh $\mathcal{T}_h$, indicating convergence towards a solution having a support composed of a union of 1d lines. Similar results were shown in [27, Fig. 5-6-7]. Due to these phenomena, it is not clear how to recover the infinite-dimensional problem in Eq. (7) from its discrete counterpart. Therefore, there is no formal connection between Problem 2 and classical BT formulations.

However, this matter has a slight impact on the variational inpainting problem that we are focusing on. In such problems, the grid is given by the format used to store I_{obs}, and we do not concern ourselves with the issue of cells with size going to zero. More crucially, the mesh appears to have a minor influence on the overall network topology that is formed by the support of $\mu_{\mathrm{opt},h}$, as shown in Fig. 2 or in [27, Fig. 8]. For a given grid, Problem 2 offers the substantial advantage of defining an optimal network as the support of the function $\mu_{\mathrm{opt},h}$. This facilitates a seamless incorporation of BT principles into inpainting problems, where data are typically stored as pixels/voxels.

Another issue related to Problem 2, either in its continuous or discrete form, arises from the fact that the functional in Eq. (7) is non-convex. In this setting, there is no assurance that the gradient descent-based technique employed in [27] will discover a global minimum. As a result, any potential minimizer that we identify is merely local and depends on the initial data. Therefore, when we mention optimal conductivity $\mu_{\mathrm{opt},h}$, we will specifically refer to the local minimum achieved using the gradient descent method and will detail the initial data utilized.

3. Network inpainting via optimal transport

Two additional building blocks are required to define the NIOT problem in Eq. (2).

- A map $\mathcal{I}$ that transforms a nonnegative conductivity μ into a nonnegative image I, so that $I = \mathcal{I}(\mu)$.
- A discrepancy term suitable for the network inpainting problem.

The forthcoming two sections are devoted to defining the desired traits and building these two components.

The fundamental concept is that by minimizing the functional $\mathcal{R}(\mu)$, we ensure simultaneously that the support of μ connects f^+ and f^-, as a result of the divergence constraint in Eq. (8), and that it approximates a 1d structure. The allocation of f^+ and f^- is our indirect method to integrate our knowledge that the observed network is a structure that transports resources into the reconstruction problem.

On the other hand, by introducing the convex discrepancy term, we force the reconstructed network I_{rec} to closely match the observed data, ruling out those local minima for the BT problem that do not fit the data. In this way we are able to include in the reconstruction process other physics-based information encoded in the data but not encapsulated in the branched transport model in Problem 2.

This suggested method can be viewed as a model-driven inverse problem, similar to those described in [48].

Remark 1. Defining f^+ and f^- is essential for the proposed methodology, as this step embeds the transport physics into the inpainting formulation. These priors may be estimated from image data (see Section 4) or based on physical insights, such as the flow direction in river basins or vascular systems. It is vital to correctly identify the spatial support of f^+ and f^-, as mislocation results in topologically erroneous reconstructions. In our experience, however, the method is robust to moderate perturbations in the mass magnitudes, as long as the sources and sinks are correctly located.

3.1. Discrepancy term $\mathcal{D}$

The discrepancy term considered in this paper is the following weighted L^2-distance

$$\mathcal{D}(I, I_{\text{obs}}) := \frac{1}{2}\int_{\Omega}(I - I_{\text{obs}})^2 W \, dx, \tag{9}$$

where the weight $W : \Omega \to [0,1]$ measures our confidence in the observed data in each portion of the domain. In this paper, we consider two types of weights.

The first weight W can be defined only if there exists a region $D \subset \Omega$, hereafter denoted as a mask, where we know that the network is corrupted. For example, in Fig. 1b the mask D is the region within the red rectangles. Some inpainting algorithms (for example [13–15]) work only if this mask can be identified, inpainting the missed area only. Under these assumptions, it is rather natural to define the confidence weight as

$$W_{\text{mask}}(x) = \begin{cases} 1 \text{ if } x \in \Omega \setminus D \\ 0 \text{ if } x \in D \end{cases}, \tag{10}$$

in order to localize the reconstruction of the network only in the corrupted regions. However, in many inpainting problems, we may not have a priori information on where such regions are located and how the data are corrupted. In this case, a more agnostic discrepancy weight should be preferred. Our second candidate is simply $W = 1$.

3.2. Mapping μ into I

The transformation $\mathcal{I}$ has the scope of connecting two objects that may be different, the conductivity μ in the BT problem, which is a model-based variable, into an image I, which is instead an observation-based variable. This transformation must send nonnegative conductivity μ to nonnegative images I. Furthermore, it must be differentiable for all $\mu \geq 0$. This is necessary to compute the gradient of the functional $\mathcal{J}$ in Eq. (2). We do not require the map to be invertible, but it may be useful to infer information on μ from observations.

Any map $\mathcal{I}$ will necessarily include some multiplicative factor $c > 0$ such that $\mathcal{I}(\mu)$ and I_{obs} within the discrepancy term in Eq. (2) have the same order of magnitude. For example, when $\lambda = 0$, if the terms f^+ and f^- are both multiplied by a constant c, it is easy to see that the optimal conductivity $\mu_{\text{opt},h}$ will be multiplied by a factor $c^{\frac{2}{\gamma+1}}$. This multiplicative factor can also be deduced from the physical information encoded in the images.

Moreover, the map $\mathcal{I}$ should be designed such that $\mathcal{I}(\mu)$ falls within the range of values in which data I_{obs} are stored. For example, in the case of standard gray-scale images, it must be ensured that $\mathcal{I}(\mu) \in [0,1]$. Ultimately, the map $\mathcal{I}$ should be designed, having in mind that the optimization term in Eq. (7) tends to the conductivity $\mu_{\text{opt},h}$ on a union of 1d curves, as shown in Fig. 2.

In this paper, two maps $\mathcal{I}$ are considered. The first map is the identity map, that is, $\mathcal{I}(\mu) = c\mu$. This map is extremely simple and is suitable for inpainting problems where the supports of μ and I are expected to be similar, and I is proportional to the conductivity μ. Its main advantage is that it is linear and the gradient of the functional $\mathcal{J}$ in Eq. (2) can be computed without any additional cost involving the Jacobian $\mathcal{I}'(\mu)$.

The second map examined in this paper is designed to handle the case where some fluid is passing through the observed network. Under mild assumptions, this flow can be modeled using the Hagen-Poiseuille approximation, which states that the velocity of a fluid flowing through a pipe is proportional to the gradient of a pressure multiplied by a conductivity factor. For tubular pipes, this conductivity factor is expressed as a power law of the radius $r > 0$ of the channel, given by

$$\mu_{\mathrm{Pou}} = \kappa r^p \tag{11}$$

where κ is a constant factor, with $p = 3$ in $\mathbb{R}^2$, and $p = 4$ in $\mathbb{R}^3$. Note that in the BT formulation in Eq. (8) we implicitly assumed the Hagen-Poiseuille law to hold.

A map $\mathcal{I}$ adapted to this type of network should transform a conductivity μ concentrated on a union of 1d curves into an image I whose support satisfies the relation in Eq. (11). This implies mapping 1d channels with higher conductivities into tubular neighborhoods with wider support, making the map necessarily nonlinear.

In the variational setting of Eq. (2), we cannot take a simple tubular expansion of a 1d channel, possibly tuned according to μ, since this transformation is not differentiable. Moreover, it would be complicated to define this transformation, where the support of μ is not exactly supported on a 1 pixel wide curve.

We can define a map having the desired characteristics using the porous media (PM) equation [49,50], which can be formulated as follows. Given $m \geq 1$, $T > 0$, and a nonnegative initial condition ρ_0, we seek a function $\rho : [0,T] \times \Omega \to [0, +\infty]$ that satisfies the following equation

$$\begin{cases} \partial_t \rho - \mathrm{div}(\rho^{m-1} \nabla \rho) = 0, \\ \rho(0) = \rho_0, \\ \rho^{m-1} \nabla \rho \cdot n_{\partial\Omega} = 0. \end{cases} \tag{12}$$

Denoting by $\rho(t, m, \rho_0)$ the solution of the PDE Eq. (12) at time t with initial condition ρ_0, we can define a map $\mathcal{I}$ (hereafter referred to as the PM map) as follows

$$\mathcal{I}(\mu) := c\rho(t^*, m, \mu), \tag{13}$$

with t^* being a suitable time. In Section A.1 we explain how to deduce the exponent m and t^* from the exponent p and the constant κ in Eq. (11), with the formulas given in Eq. (A.3). When discretized on a grid, this map compensates the concentration of the support of $\mu_{\mathrm{opt},h}$ on 1 pixel-wide curves.

Note also that for $m = 1$, the equation in Eq. (12) reduces to the heat equation, and thus the maps would essentially act as a Gaussian filter. This map is suitable for the inpainting problem, with the advantage of being easier to compute. However, it has the disadvantage of infinite propagation of the support and, being linear, does not provide the "higher" to "wider" mapping needed to approximate the law in Eq. (11). For this reason, we did not consider it in this paper.

3.3. Discretization and optimization

The spatial discretization of Eq. (2) is performed using a finite-volume/finite-difference scheme based on a Cartesian grid $\mathcal{T}_h$ of the domain Ω. This discretization is particularly convenient in inpainting problems, where we mostly deal with images stored in pixels/voxels. We now look for the pair $(\mu_h, u_h) \in DG^0(\mathcal{T}_h) \times DG^0(\mathcal{T}_h)$, where $DG^0(\mathcal{T}_h)$ is the space of piecewise functions on $\mathcal{T}_h$. The discrete conductivity μ_h is the main unknown of the minimization problem, while u_h is the numerical solution of the elliptic equation in Eq. (8). This choice of discretization spaces requires the definition of a map from the cell-centered values of μ_h to the faces. We used the harmonic average for this purpose to ensure that the flux is properly conserved. Similarly to the notation introduced in Problem 2, we will denote by $u_h = u(\mu_h)$ the numerical solution of Eq. (8) for a given μ_h.

The gradient of the discretized functional $\mathcal{J}$ in Eq. (2) with respect to μ_h is given by

$$\nabla_{\mu_h} \mathcal{J}(\mu_h) = \frac{-|\nabla u_h|^2 + \mu_h^{\gamma-1}}{2} + \lambda(\mathcal{I}(\mu_h) - I_{\mathrm{obs}})W(x)\mathcal{I}'(\mu_h) \tag{14}$$

When the PM map is used, there is an additional cost to compute the term involving the Jacobian $\mathcal{I}'(\mu_h)$ coming from the discrepancy term in Eq. (14). To handle this, we employ adjoint-based methods [51]. In our case, these methods require solving the PM equation in Eq. (12) to obtain the term $\mathcal{I}(\mu_h)$. We use the backward Euler method combined with the Newton method for this purpose. Additionally, we need to solve a linear system with the Jacobian of the PM map $\mathcal{I}'(\mu_h)$ as the matrix. To improve accuracy, we divide the time t^* in Eq. (13) into five subintervals, gradually increasing the initial time step. Consequently, in the calculation of the gradient in Eq. (14), we solve five equations of porous media and five linear systems associated with the adjoint-based method.

The optimization part follows the approach used in [27]. Starting from an initial data $\mu_h^0 = \mu_0$, we update μ_h as follows

$$\mu_h^{k+1} = \mu_h^k - \Delta t^k (\mu_h^k)^{\frac{2}{\gamma}} \nabla_{\mu_h} \mathcal{J}(\mu_h^k), \quad k = 0, \dots, k_{\max}, \tag{15}$$

with $\Delta t^k > 0$ being a sequence of time steps chosen adaptively to ensure the stability of the scheme. The direction of update in Eq. (15) is (minus) the gradient of the functional $\mathcal{J}$ scaled by the factor $\mu_h^{\frac{2}{\gamma}}$. This scaling procedure ensures that μ_h remains positive and mimics a scheme known in the optimization community as the Mirror Descent approach [52]. We continue updating μ_h until the L^1-norm of the discrete right-hand side of Eq. (15) is below 10^{-5}.

The method we have introduced falls within the class of first-order methods, as it relies only on the gradient of $\mathcal{J}$. Approaches that incorporate second-order derivatives, such as a Newton-based scheme, appear to be not suitable for our problem. This is mainly due to the non-convex nature of the functional $\mathcal{J}$. Furthermore, these methods require specific linear algebra solutions, which are challenging to fit our circumstances, similar to those outlined in [53].

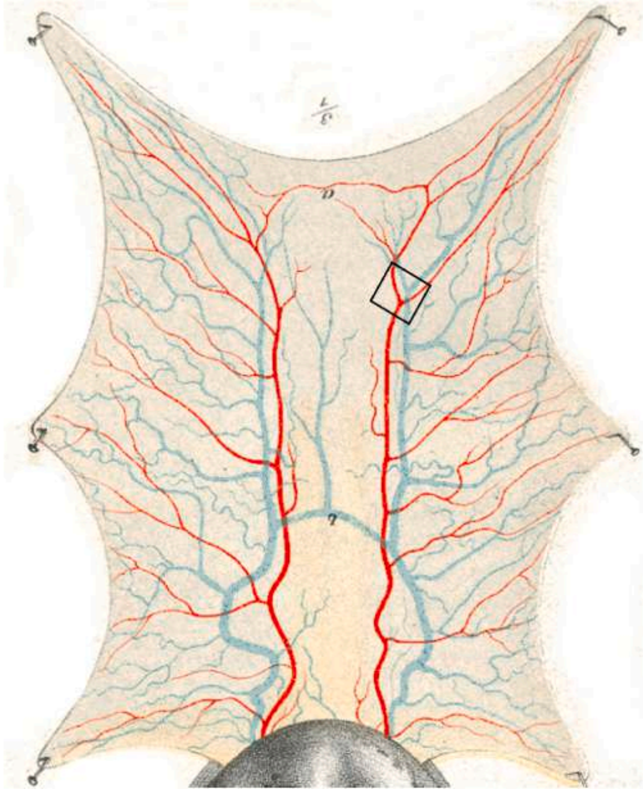


Fig. 3. Original drawing by Cohnheim with the arterial (red) and venal (blue) blood vessel network (picture taken from [62]). The Y-shaped red network within the black square is the one used in Section 4.1. Its binary map is shown in Fig. 1a.

3.4. Software and implementation details

The NIOT algorithm is implemented using Firedrake [54,55], a Python finite element library based on the Unified Form Language (UFL) introduced in [56]. Using these libraries, we can write our problem directly from the functional $\mathcal{J}$ in Eq. (7) and then use automatic differentiation to compute the gradient of the functional. The adjoint-based computations (required only when the PM map is used) are performed using the dolfin-adjoint library [57]. Firedrake is also built on top of the PETSc library [58,59], which provides a wide range of efficient linear and nonlinear solvers. PETSc ensures that the algorithm's main computational bottlenecks (namely, the solution of the linear systems associated with the elliptic PDE in Eq. (8) and the evaluation of the PM map) are handled using state-of-the-art parallelized software. Specifically, we employ iterative methods preconditioned with the Hypre BoomerAMG algebraic multigrid solver [60], which we found to be robust and efficient for our problem. These choices guarantee a smooth transition and scalability to large 3D problems, particularly in the context of blood vessel reconstruction.

4. Numerical experiments

We now present a set of experiments designed with a dual purpose. Our first objective is to address reconstruction problems in which data degradation is the result of data loss that causes network disconnection, and/or the existence of artifacts, as exemplified in Fig. 1b. Our second goal is to investigate the interaction between the parameters and functions of the NIOT model, summarized in Table 1. Our analysis is mostly focused on the fitting and optimization categories, as the modeling parameters f^+, f^-, and γ should be determined by the physical problem at hand, and should be known (or estimated) a priori by the user. With these two goals in mind, we examine, in different setups, two test cases.

One test case is an example of a real biological geometry with a reasonable 2d spatial representation. We turn to the classical work of Julius Cohnheim from 1872 in [61]. In his work, Cohnheim describes the vasculature of a frog tongue and includes a detailed colored drawing, shown in Fig. 3. As this geometry is 2d in nature, we can represent results without accounting for out-of-plane connections. The other test case is a Y-shaped network extracted from the first case. We start our numerical experiments with the second (simpler) case.

In both test cases, we corrupt an image and we try to recover it, in order to have the ground truth available for comparison. A reconstruction is considered good if the supports overlap and the main topological features, such as connectivity or subdivisions in subbranches, are similar. Unfortunately, finding quantitative estimates to compare network-like structures is a challenge in itself.

Table 1
Parameters and functions controlling the NIOT model.

Category	Parameter	Description
Modeling	f^+, f^-	Mass distribution densities in the BT problem
	γ	Scalar determining branching angles
Fitting	$\mathcal{I}$, c	Map from μ to I
	λ	Scalar measuring global trust in data
	W	Function measuring local trust in data
Optimization	μ_0	Initial data in Eq. (15)

Standard measures, like the L^2 distance or the Sørensen-Dice coefficient, do not provide meaningful estimates of discrepancies. They provide a meaningful comparison only if the supports overlap, which is unlikely for "thin" structures like networks. For these reasons, we restrict ourselves to a qualitative comparison with the ground truth to assess the accuracy of a reconstruction.

All images reported in the following experiments should be seen in their high-resolution colored version, which is available directly within the electronic version of the paper. These images were generated using the PyVista library [63].

4.1. Y-shaped

The first test case presented in this paper is the reconstruction of the Y-shaped network in Fig. 1a, which is the portion of the arterial network within the squared box shown in Fig. 3. It is corrupted using the mask made up of three rectangles shown in Fig. 1b. In this test case, the artifacts shown in Fig. 1b are removed.

First, we fix the resolution of the mesh used in our inpainting problem. We choose a mesh $\mathcal{T}_{h/4}$ with 200×200 pixels, used in the third column in Fig. 2. For this grid, the smallest branch is at least 4 pixels wide, while at coarser resolutions it is hard to see the effects of using the PM map used in the next experiments.

We also need to fix the exponent γ that influences the branching angle of the optimal network. Looking at the networks obtained in Fig. 2 we can see how the exponent $\gamma = 0.8$ does not provide a Y-shaped network, which is instead obtained for $\gamma = 0.2$ and $\gamma = 0.5$. We decided to fix $\gamma = 0.5$ for all experiments, because the network obtained for $\gamma = 0.2$ is not that different from the one obtained for $\gamma = 0.5$, while solving the BT problem with smaller values of γ is typically computationally more intense. This is because shorter time steps are required due to the factor $\mu^{\frac{2}{\gamma}}$ in Eq. (15), and therefore more time steps to reach convergence.

The transported measures are $f^+ = \delta_O$ and $f^- = (2\delta_P + \delta_Q)/3$. They are deduced by looking at the observed network and assuming that the flux on the left upper branch doubles the one on the right branch.

4.1.1. Images with lost data (case $W = W_{\text{mask}}$)

We now turn our attention to the discrepancy term in Eq. (2), which defines the fitting component of the NIOT model.

We consider two maps $\mathcal{I}$ described in Section 3.2, the identity map and the PM map with $m = 2$ and $t^* = 1.0e - 3$. The exponent m is the result of the formula in Eq. (A.3) with $p = 3$, while the time t^* was experimentally found to guarantee that the support of I_{obs} and the reconstructed image $\mathcal{I}(\mu)$ have approximately the same supports.

The identity map is scaled by the factor $c = 10$, while the PM map is scaled by $c = 50$. We choose these values by looking at the optimal $\mu_{\text{opt},h}$, reported in the rightmost column in Fig. 2, which correspond to the solution of Eq. (2) with $\lambda = 0$. We tune the scaling parameter c to have the same order of magnitude of the maximum values of the reconstructed image $\mathcal{I}(\mu)$ and the data term, which is a binary image. This calibration based on the maximum values presents some limitations. In fact, in full generality, we expect to have some mismatch between the observed data I_{obs} and the $\mathcal{I}(\mu)$ in the sub-branches of the network, where the conductivity is lower. Nonetheless, this method is rather straightforward and prevents the observed and reconstructed data from having values that vary by multiple orders of magnitude.

We fix the confidence term $W = W_{\text{mask}}$ given in Eq. (10) and the initial data $\mu_0 = 1$ (the influence of the second choice is discussed in Section 4.1.3). In the first and second rows of Fig. 4 we report the result obtained for the different values of λ and the two maps described above.

We can see how the proposed approach effectively restores the connectivity of the network for all values of λ and both maps. This result is the consequence of imposing a flux between the source term $f^+ = \delta_O$ and the sink term $f^- = (2\delta_P + \delta_Q)/3$. Having a disconnected network, which in our setting means having zero conductivity, would contradict the constraint given in Eq. (8).

Since $W = W_{\text{mask}}$, the discrepancy term has no influence within the rectangles that make up the mask, and the optimization of the regularization term $\mathcal{R}$ is the only factor driving the reconstruction. Thus, inside the mask, the reconstructed network is composed of (almost) straight channels or a union of them, reconnecting the ends of the original network cut by the mask. In this region, we can see some grid alignment artifacts.

Outside the mask, we see the influence of the discrepancy term. The network progressively fits to the observed data as λ increases. For both maps, the topology of the reconstructed network can change drastically even with a relatively small change of λ. For example, the bifurcation point in the middle of the domain is fitted by increasing λ from 5.0e − 2 to 1e − 1 when the PM map is used, while it is recovered only for $\lambda = 1.0$ when the identity map is used. This sudden transition can be attributed to the fact that the functional

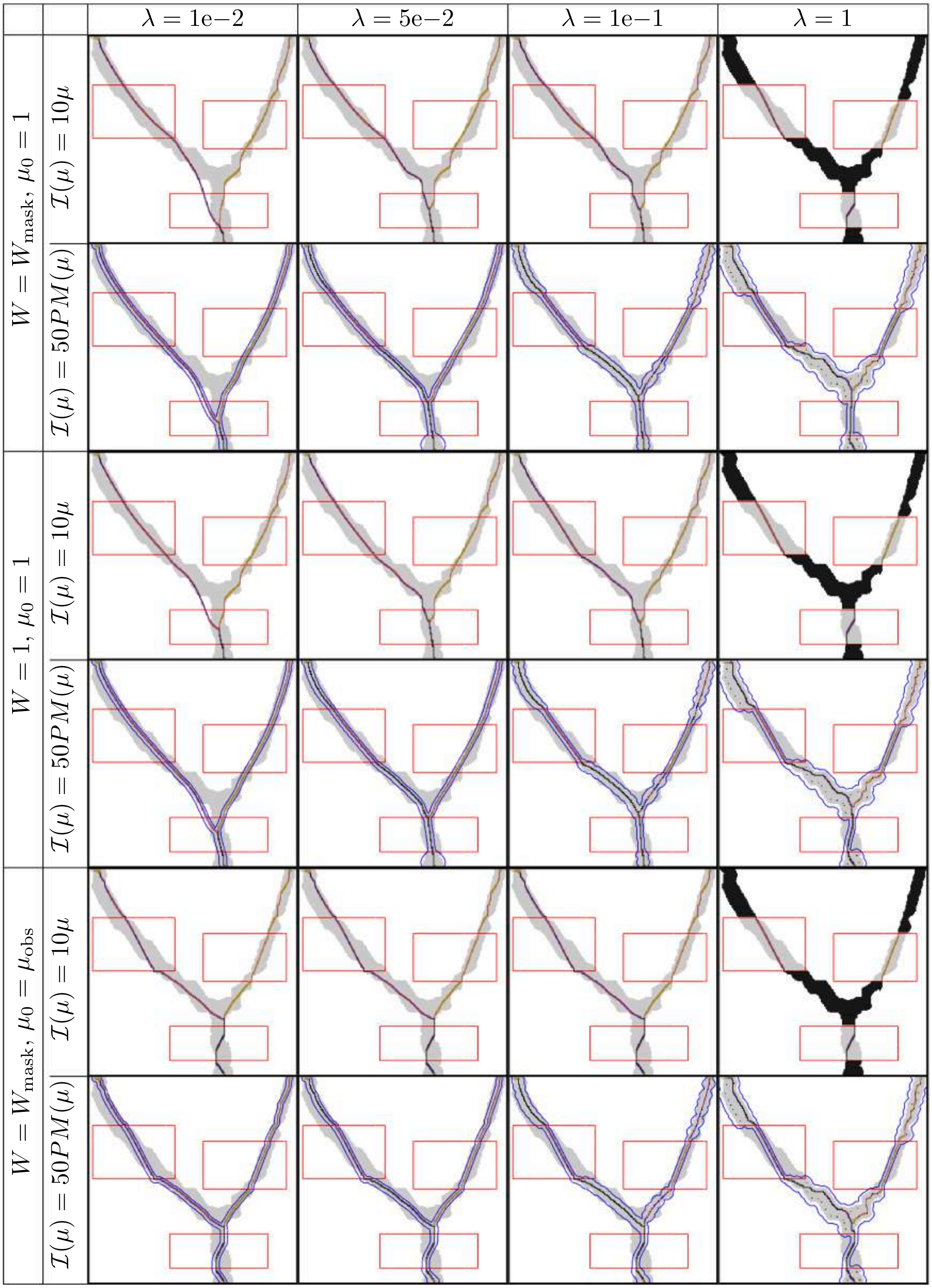


Fig. 4. Spatial distribution $\mu_{\mathrm{rec},h}$ for different combinations of λ, maps, and confidence W. When the PM map is used, the blue line defines the boundary for which $I_{\mathrm{rec}} = \mathcal{I}(\mu_{\mathrm{rec},h})$ is $1.0e-3$. The red rectangle defines the mask region. The ground truth image I_{true} is reported in light gray. .

$\mathcal{R}$ is not convex and that we are using gradient-based optimization methods. This leads to diverse convergence paths within the optimization landscape.

Both maps produced comparable results in terms of reconstructing the primary topological feature of the network, showing a similar change when λ is altered. However, the PM map provides a smoother transition of the reconstructed network I_{rec} at the mask boundary. Moreover, it generates both skeletonization and a reconstruction of the network. On the other hand, the PM map requires between 60 and 90 minutes to reconstruct an image, while the identity map requires between 20 and 30 minutes. These findings were derived by running our software on a single core, on a computer equipped with a 12th Gen Intel(R) Core(TM) i7-12700 processor and 64 Gb of RAM memory.

4.1.2. Images with lost data (case $W = 1$)

In this section, we propose a second experiment where we use the confidence term $W \equiv 1$, which is more suitable for those problems where we are not aware of the location of the corruption. In the third and fourth rows in Fig. 4 we report the spatial distribution of $\mu_{\text{rec},h}$ while changing the scaling parameter λ of the discrepancy and the map $\mathcal{I}$.

The results do not differ significantly from those obtained using $W = W_{\text{mask}}$. However, $\mu_{\text{rec},h}$ tends to have lower values and smaller support within the mask relative to the results obtained using $W = W_{\text{mask}}$. For example, when the identity map is used and $\lambda = 1.0\text{e}-2$ (first column, first and third row), the branching point within the lowest rectangle is moved upward. This effect can be attributed to the fact that the discrepancy term is now acting also within the mask D, where the data I_{obs} fitted is equal to zero, penalizing the presence of the network. If $\mu_{\text{rec},h}$ is still present within the mask, this penalization effect results in a network with wider channels, as a wider section is required to carry the same flux. This is more evident when comparing the case with the identity map and $\lambda = 1.0$.

4.1.3. Influence of initial data $However, in this \mu_0$

In Section 2.1 we already mentioned how the approximate solution $\mu_{\text{opt},h}$ of Problem 2 depends on the initial data. In this part, we illustrate the impact of this issue on the NIOT algorithm. So far, our consideration has been limited to uniform data with $\mu_0 = 1$. However, in this instance, we employ initial data that are extrapolated from the observed data. In particular, we set

$$\mu_0 = \mu_{\text{obs}} = c^{-1}(I_{\text{obs}}) + \mu_+, \tag{16}$$

where the lower bound μ_+ is required to avoid a strong penalization in the initial data where data are lost. In our experiment, we set it to $1.0\text{e}-5$ times the maximum value of I_{obs}. The results are reported in the last two rows of Fig. 4.

It is evident that the selection of initial data profoundly impacts the ultimate outcome. The magnitude of this influence is such that the selection of the map $\mathcal{I}$ and the scaling of the discrepancy λ become almost insignificant. This effect can be beneficial when images are corrupted only by data loss. However, if this assumption is not valid, using such initial data may result in incorrect outcomes, as shown in the next section.

4.1.4. Images with artifacts

In this section, we focus on a reconstruction problem in which false positive data are present in the image. Additional channels are included in Fig. 1a and then the mask shown in Fig. 1a is applied. The resulting I_{obs} can be seen in Fig. 1b.

In this case, we set the confidence measure $W = 1$, since we typically do not know the location of false positive data. We use the identity map with $c = 10$. In Fig. 5, we report the reconstructed networks using different $\lambda > 0$ and initial data $\mu_0 = 1, \mu_{\text{obs}}$.

The results show how large values of λ (rightmost panels) lead to overfitting of the data, including the artifacts that we would like to remove. On the other hand, a network more similar to the original Y-shaped network is obtained with $\lambda = 1.0\text{e}-3, 5.0\text{e}-2$ and unbiased initial data $\mu_0 = 1$. This suggests that the usage of smaller values of λ and uniform initial data should be preferred if false positive data are present.

The experiments in this section offer valuable guidelines on how to select the initial data μ_0. When the observed data I_{obs} are known to be free from false positive data, it is advantageous to use initial data derived from the observed data, as described in Eq. (16). Conversely, if there is a suspicion of false positive data, it is preferable to opt for unbiased initial data, such as $\mu_0 = 1$. In this setting, these experiments show an additional feature of the proposed method: its capability to remove artifacts from the corrupted data. This procedure is referred to in the literature on image processing as pruning [64]. In our model, this selection is the result of an optimization process, where the presence of the functional $\mathcal{R}$ removes those channels that do not carry enough flux. Similar ideas can be found in [65] in the context of graph pruning.

Remark 2. We considered only two types of initial data, uniform data $\mu_0 = 1$ and $\mu_0 = \mu_{\text{obs}}$, with the latter strongly biased by the observed data. However, the choice is not necessarily binary. In fact, it is easy to image a transition between the two configurations by progressively blurring μ_{obs} with a Gaussian filter. With such a procedure, the initial data are less biased by the observed data but still contain some information about the network that may be useful in the reconstruction process.

4.2. Frog tongue

In this section we consider the full arterial blood vessel network of the frog tongue in Fig. 3. We segmented it using a semi-automatic procedure used in [66], followed by a manual inspection to ensure spatial connectivity of the arterial vessel tree. The result is shown in Fig. 6a.

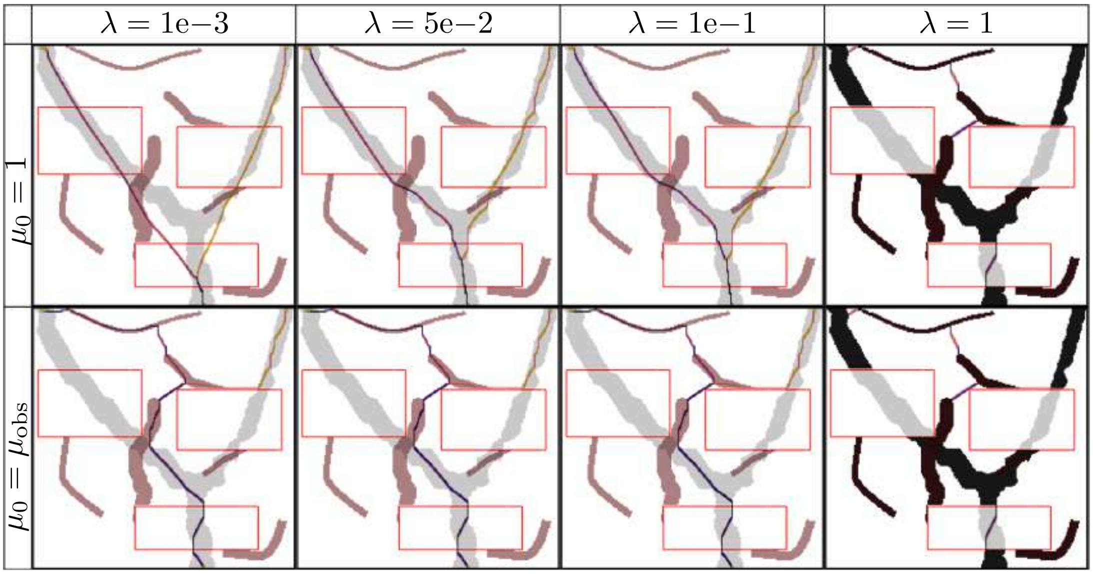


Fig. 5. Spatial distribution $\mu_{rec,h}$ for $\lambda = 1.0e-3, 5.0e-2, 1.0e-1, 1.0$ (left to right) and $\mu_0 = 1, \mu_{obs}$ (top to bottom). Confidence $W = 1$ and identity map. The observed data I_{obs} is obtained by adding additional artifacts (in brown) to the original network in Fig. 1a, and then removing all data within the mask (red rectangles). The ground truth I_{true} is reported in light gray.

We deduce the terms f^+ and f^- used in Problem 2 directly from Fig. 3. We set f^+ as the sum of two Dirac masses located at the root of the two main channels, having the same mass injection rate. The sink term f^- is set equal to one on the support of the frog tongue, modeling a uniform absorption of nutrients/oxygen. The two terms are balanced to ensure that the total mass injected is equal to the total mass absorbed. More accurate estimation of these terms would require experimental measurements that are not available in this case.

This configuration, characterized by few concentrated sources and a distributed sink (or vice versa), is observed in various natural systems, such as river basins and blood circulation. Indeed, this test case represents a simplified 2D analogue of our target application: the reconstruction of the brain vascular network from 3D MRI images, like those considered in Hodneland et al. [66].

Three experiments on the frog tongue data are presented. One experiment simulates the case where the source of corruption is loss of data that causes network disconnection, while the other simulates the presence of artifacts in the observed data. The observed network I_{obs} used in the second experiment is reported in Fig. 6c.

In the third experiment, we present a complementary approach to reconstruct the arterial network of the frog tongue. The idea is to estimate a conductivity μ_{obs} (shown in Fig. 6d) from the binary image I_{obs} by using the Hagen-Poiseuille law in Eq. (11). We use these enhanced data within Eq. (2) with the identity map, where the NIOT approach becomes a model-driven inverse problem written in terms of the conductivity μ.

In all experiments, each reconstruction requires between 20 and 60 minutes to be executed.

4.2.1. *Binary images with lost data*

The first experiment consists of reconstructing the binary map I_{true} shown in Fig. 6a corrupted with a mask composed of the seven circles, shown in Fig. 6c.

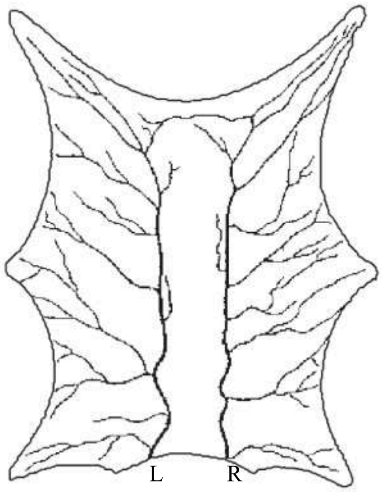
We first calibrate our model using $\lambda = 0$. The corresponding optimal $\mu_{opt,h}$, reported in Fig. 6b, is composed of a high-conductivity skeleton and a series of small channels that fill the domain of the sink term f^-. The maximum value of $\mu_{opt,h}$ is $2.0e-2$. So we set $c = 50$ to match the value of a binary representation of the arterial network.

In Fig. 7 we report the results obtained by fitting the observed data using different values of λ and $W = W_{mask}$ or $W = 1$. The initial data is set to $\mu_0 = 1$.

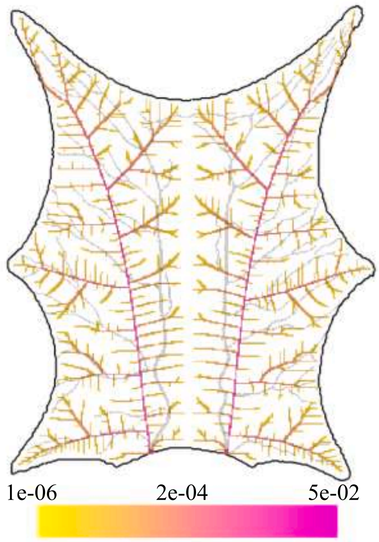
In all experiments, the curved portion of the network within circle (d) that branches from the main channel in the south-east direction is not consistently recovered. Neither is the channel oriented in the west direction in the upper part of circle (a). This can be attributed to the fact that within the mask, these configurations are not optimal for the $\mathcal{R}$ functional.

For $\lambda = 1.0e-3$, the discrepancy term already affects the resulting network. The main right branch is approximately fitted for both $W = W_{mask}$ and $W = 1$. However, despite the fact that these two confidence measures are different only within the mask, the left branch resulting networks are completely different. This can again be attributed to the non-convexity of the functional $\mathcal{R}$ and the usage of gradient-based optimization methods.

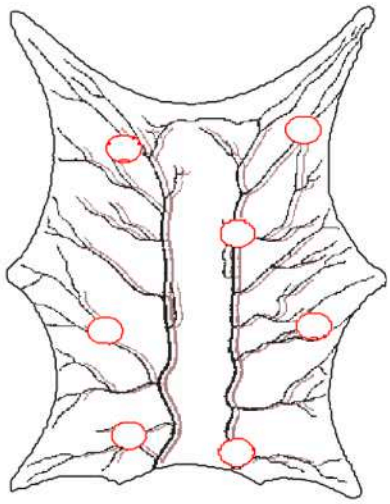
Within circle (g), note that a good approximation of the bifurcation point is created in all cases, even in the case $\lambda = 0$ shown in Fig. 6b. In fact, without any channel within the local mask, the flux cannot be carried from the (right) source to the sink.



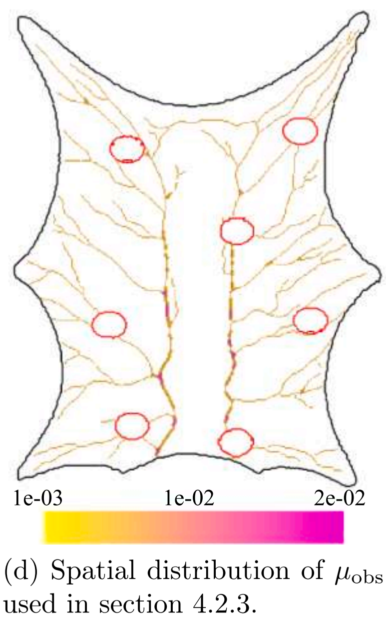
(a) Binary map I_{true} of the arterial network in fig. 3. The boundary of the frog tongue support is also shown. The points L and R are roots of the two main branches of the network.



(b) Spatial distribution of $\mu_{\text{opt},h}$ solving eq. (7) with $f^- \equiv 1$ within the support of the frog tongue and $f^+ = \delta_L + \delta_R$ (both normalized) and $\gamma = 0.5$. I_{true} is shown in the background.



(c) In black, the corrupted network I_{obs} used in section 4.2.1, with lost data within the red circles. In brown the artifacts added in section 4.2.2.



(d) Spatial distribution of μ_{obs} used in section 4.2.3.

Fig. 6. Frog tongue data. Figures adapted from Hodneland et al. [66].

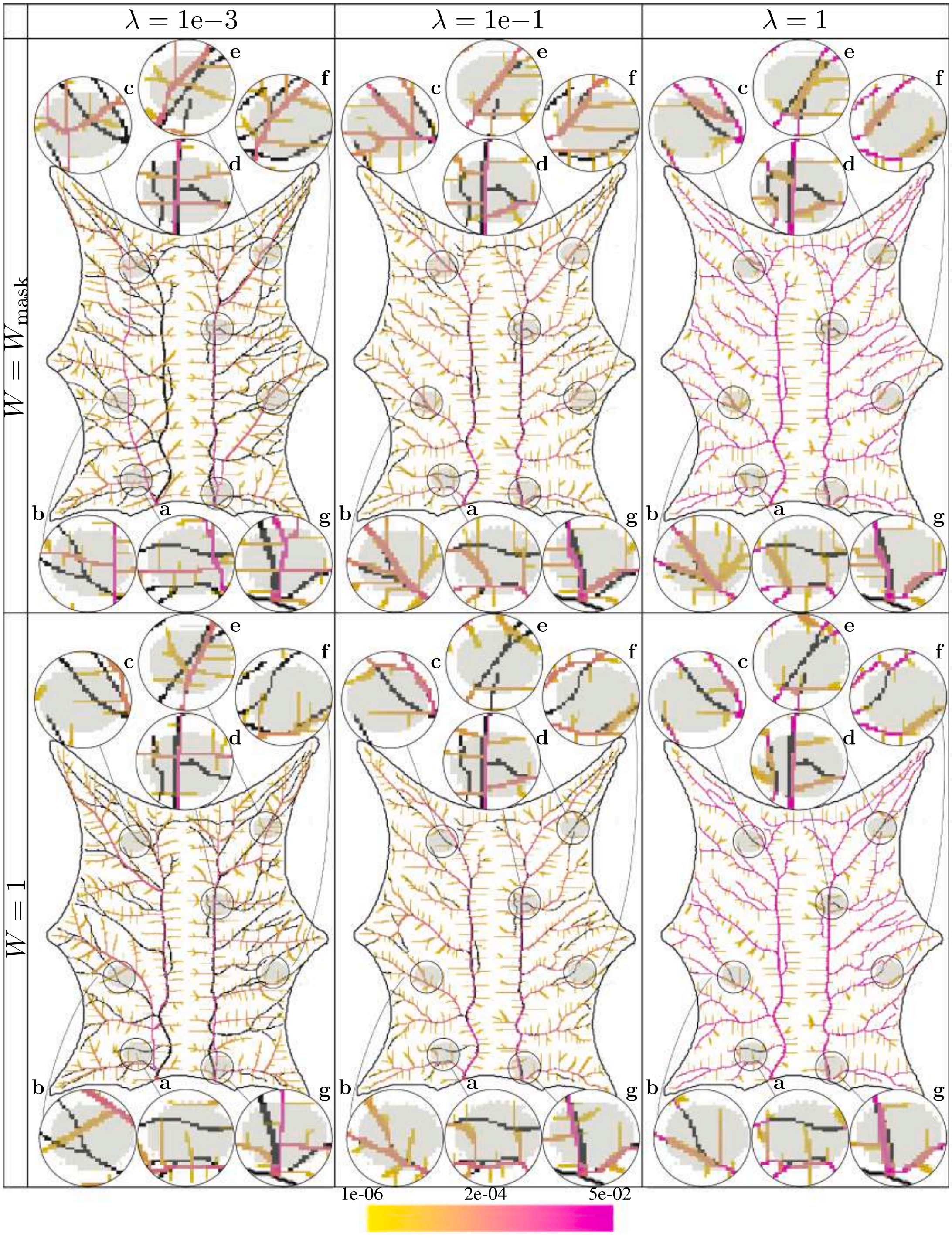


Fig. 7. Reconstruction with lost data. The ground truth I_{true} is reported in black for comparison. Circles where I_{true} was removed are marked in light gray. We report (color bar at the bottom) the spatial distribution of $\mu_{\text{rec},h}$ for $W = W_{\text{mask}}, 1$ and $\lambda = 1.0\text{e}-3, 1.0\text{e}-1, 1.0$.

At $\lambda = 1.0\text{e}-1$, the Y-shaped network within circle (b) is recovered for both $W = W_{\text{mask}}$ and $W = 1$. In the first case, the reconstructed network is not totally accurate, but the main structure of the network is recovered within all circles, excluding the inaccuracies already mentioned within circles (a) and (d). In the second case, the penalization imposed by having $W = 1$ within the mask removes the channels within circles (c), (e) and (f). However, outside the circular masks, the overall structure of the network is recovered with good quality.

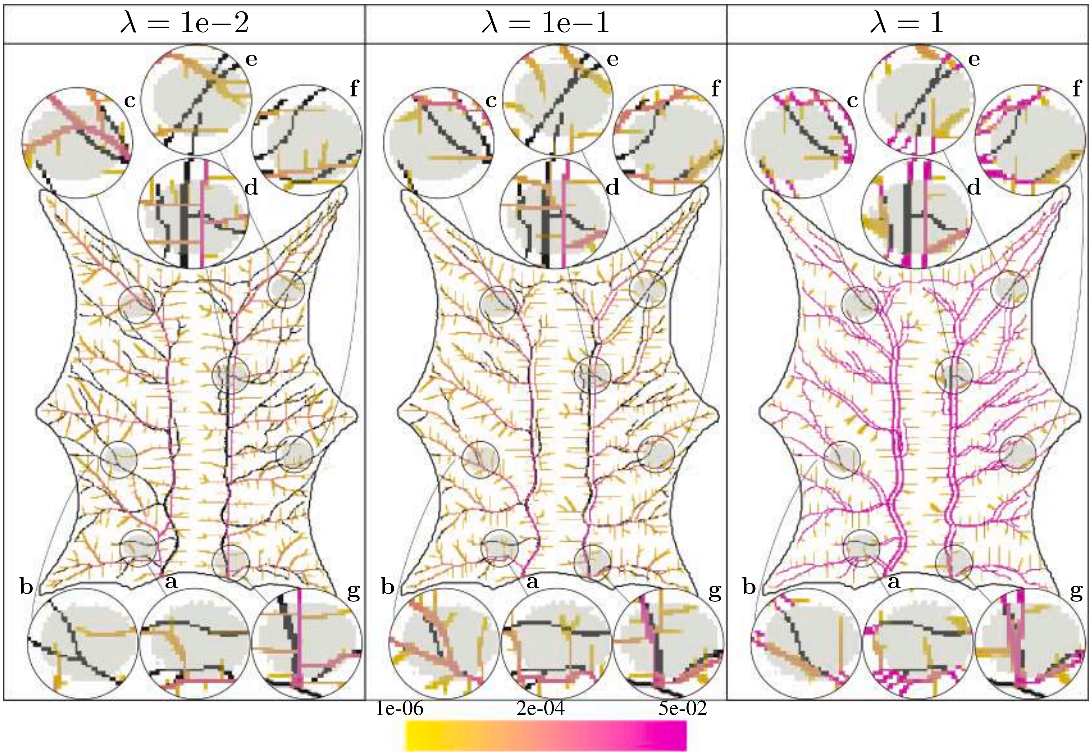


Fig. 8. Reconstruction with lost data and artifacts. The observed image I_{obs} is shown in Fig. 6c), while the ground truth image I_{true} is here reported in black in the background for comparison. Circles where I_{true} was removed are marked in light gray. We report (color bar at the bottom) the spatial distribution of $\mu_{\text{rec},h}$ for $\lambda = 1.0\text{e}-2, 1.0\text{e}-1, 1.0.$.

For $\lambda = 1e0$ and $W = W_{\text{mask}}$, the reconstructed network fits the observed data accurately outside the mask at the cost of having some errors within the circle (e).

For $\lambda = 1e0$ and $W \equiv 1$, an inconsistent network is obtained within circle (b). No network is created here, and the tail of the corrupted branch is "fed" by another sub-branch. This reconstructing error is not particularly satisfactory from a modeling point of view, because this would lead to flux flowing in the wrong direction. For $\lambda = 1e0$ and $W = W_{\text{mask}}$, the reconstructed network fits the observed data very well outside the mask at the cost of having some errors within circle (e).

We attribute these phenomena to the combination of different factors:

- Within the mask we have that $W \equiv 1$ and $I_{\text{obs}} = 0$, hence the discrepancy term is penalizing any presence of the network, similarly to what happened in the previous section. This effect can also be seen with circles (c), (e), and (f), where channels outside the mask are selected.
- The approach is currently blind to the direction of the flux, since this information is not encoded in our inpainting approach, neither in the data nor in the model.
- The fitted image is a binary map. It contains only information about the support of the network, but not its hierarchical structure, from the root to the leaves.

In Section 4.2.3, we will try to mitigate the effect of the last factor by extracting more information from the data.

4.2.2. Binary images with artifacts

We now present the experiment in which the data are both lost and corrupted by the presence of artifacts, where I_{obs} is shown in Fig. 6c. This image is obtained by shifting a portion of the original network and merging it with the original network. The data are also corrupted within the circular masks shown in Fig. 7.

In this false positive regime, the reconstruction of the network is also required to prune redundant channels and obtain the main structure of the network. We use $\mu_0 = 1$ and $W = 1$, and we show in Fig. 8 the results obtained for different λ.

At $\lambda = 1.0\text{e}-2$ we already recover a good approximation of the original network, with the main structure of the network being recovered but with a strong pruning of the redundant channels. At $\lambda = 1.0\text{e}-1$, we obtain even better results, like the reconstruction of the Y-shaped network in the circle (b), but we start seeing some overlay of small channels with the shifted network, in particular along the two main channels. For $\lambda = 1.0$ we fall into the overfitting regime.

In the context of pruning, let us also note how reducing λ cuts the vessel at the top of the image that connects the left and right basins of the frog arterial structure. In fact, this redundant structure cannot be taken into account in the current BT penalization.

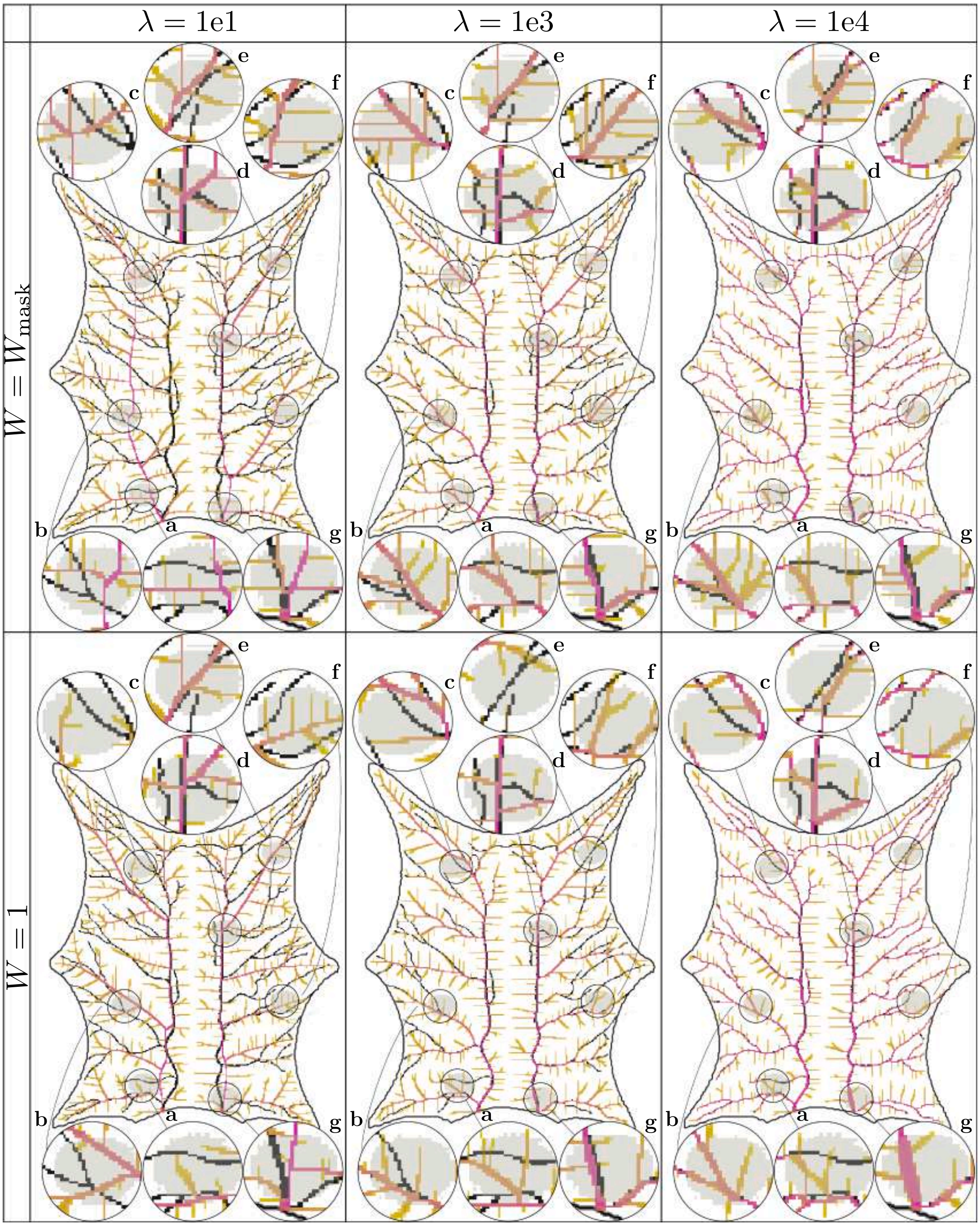


Fig. 9. Conductivity reconstruction. The observed conductivity μ_{obs} shown in Fig. 6d. The support of the original image I_{true} is reported in black in the background for comparison. We report the spatial distribution of $\mu_{\text{rec},h}$ for $W = W_{\text{mask}}, 1$ and $\lambda = 1.0\text{e} - 3, 1.0\text{e} - 1, 1.0$.

4.2.3. Enhanced image data

In this section, we present a complementary approach to reconstructing the arterial network of the frog tongue. In the previous pages we worked with the variables μ and the image I simultaneously, with the transformation $\mathcal{I}$ mapping the first into the second. In this instance, we focus solely on the conductivity parameter and examine the subsequent variational problem

$$\min_{\mu \geq 0} \mathcal{R}(\mu) + \lambda \mathcal{D}(\mu, \mu_{\text{obs}}), \tag{17}$$

where the term μ_{obs} is obtained from the observed data I_{obs}. This process may be seen as inverting the map $\mathcal{I}$, which may not be possible, as is the case for the map PM. However, using the relation $\mu_{\text{Pou}} = \kappa r^p$ in the Hagen-Poiseuille law, we define μ_{obs} through the following steps (we used ImageJ library [67] to perform the first two).

1. We compute the local thickness of the binary representation I_{true} of arterial network.
2. We get a skeleton s of I_{true}, which is a binary map inside the support of I_{true} that is only 1 pixel wide.
3. We define an approximate Dirac measure of the 1d skeleton s as $S = s/h$, where h is the width of 1 pixel.
4. We compute the conductivity $\mu_{\text{Pou}} = \kappa r^p S$ with $\kappa = 5.0e+02$ and $p = 3$.
5. Finally, we apply a PM map to μ_{Pou}, obtaining the data I_{obs} to be fitted. The exponent m and the time t^* that define the PM map are those given in Eq. (A.3). The value $\kappa = 5.0e+02$ was experimentally found so that I_{obs} and the optimal $\mu_{\text{opt},h}$ reported in the right panel of Fig. 6 have the same order of magnitude.

The last preprocessing step provides a smoother representation of the conductivity having the same support of the arterial network.

The corruption step involved in the application of the mask shown in Fig. 6d and the resulting conductivity μ_{obs} is reported in Fig. 6d. In Fig. 9, we report the spatial distribution of the conductivity $\mu_{\text{rec},h}$ using increasing values of λ and $W = W_{\text{mask}}, 1$. In this case, the values of λ are higher than those presented in Section 4.2.1 as the fitted data have lower values (the maximum value is 1.0e − 2).

The results are consistent with the previous experiments, the network is progressively reconstructed as λ increases. In this case, however, we see that the Y-shaped network with circle (b) is reconstructed even for $\lambda = 1e4$ and $W = 1$. This example shows that the process of retrieving the conductivity from the image can be beneficial for the inpainting process proposed in this paper.

5. Conclusions

We showed how using the branch-inducing functional $\mathcal{R}$ as the regularization functional in variational inpainting methods offers a natural way to encode physical information in the problem of reconnecting transport networks.

The proposed approach is particularly effective in restoring connectivity in corrupted networks. This is obtained by imposing a flux passing through the network, and it is possible only under the assumption that a priori knowledge of the transport phenomena passing through the network is available. Specifically, information on the source and sink terms f^+ and f^- is crucial for the success of the method. Identifying these terms (in particular their location) may not be straightforward in some applications and may require dedicated preprocessing strategies, such as automated or data-driven methods. Nevertheless, even the rough estimations presented in the examples (obtained both from the observed network or simple modeling deduction) led to a good reconstruction of the network.

Characterizing a network as a support of conductivity μ allows for easy merging of the modeling information and observed data. These two components of the model are linked by the map $\mathcal{I}$. In this paper, we presented only two maps, the simple identity map and the PM map. The latter is a useful tool in the case of the network for which the Hagen-Poiseuille law given in Eq. (11) holds.

Unfortunately, the approach requires the adjustment of several parameters, both for the discrepancy and the regularization/penalization term in Eq. (2), summarized in Table 1.

For the discrepancy term, in most of the examples presented, these parameters were estimated experimentally, but in general their tuning should be driven by the problem at hand. The experiments presented in the paper suggest that this tuning should also be driven by the way data are corrupted. If good data are available and the corruption is only due to data loss, high values of λ can be used. In this setting, it may be beneficial to choose the initial data μ_0 derived from the observed data and exploit the "memory" effect described in Section 4.1.3. On the other hand, if the corruption is due to the presence of false positive data, small values of λ, $W = 1$, and $\mu_0 = 1$ should be used. In this setting, the proposed approach has also shown interesting capabilities in removing artifacts from highly corrupted data, performing this selection according to an optimization principle, as shown in Sections 4.1.4, 4.2.2.

The parameter γ in the functional $\mathcal{R}(\cdot)$ was fixed at $\gamma = 0.5$ in most experiments. However, its choice should be driven by observed quantities (such as the typical branching angles in observed data) or from modeling deduction (such as Murray's law for blood vessels networks).

The software implementing the proposed methodology is designed to extend naturally to 3D, with a primary application being the reconstruction of blood vessel networks from MRI. Although the 3D weighted Poisson equation presents a higher computational load, our software is fully equipped to manage this complexity. The implementation is built upon PETSc, utilizing state-of-the-art parallel linear solvers that allow for efficient resolution of large-scale 3D systems.

On the other hand, the proposed model can be extended in several directions, enriching both its fitting and modeling components. In the first component, we can include further terms, like the flux or the pressure, when this information can be extracted from the data. Different discrepancy metrics can be used in place of the weighted L^2-norm considered here. Natural candidates are OT-based distances, similar to the unbalanced metrics described in [68], since these distances can deal with densities concentrated on support with a small width-to-length ratio. We did not explore this possibility in this paper, but it is a natural extension of the work presented here.

The regularizing component of the NIOT approach can be enriched, for example, by including a localized penalization of the term $\int_\Omega \mu^\gamma\,dx$ in Eq. (7) that favors (or disfavors) the presence of the network in certain regions, similarly to what was done in [69]. Networks having loops can be included using fluctuating source and sink terms as in [70].

Data availability

No data was used for the research described in the article.

Code availability statement

The source code used to produce the results in this paper is available at https://github.com/enricofacca/niot, together with the data and scripts to reproduce the results. The code and the data that support the findings of this study are openly available at the following DOI https://doi.org/10.5281/zenodo.13919718.

Acknowledgements

E.F. has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 104109101 (at the University of Bergen). E.F. was also partially supported by the Italian Ministry of University and Research (MUR) through the PRIN 2022 project 202277N5H9 “DEEP-GRAPH: Design and Theory of Deep Graph Learning” (at the University of Padova) and the FIS2 project “ADAMUS”, funded under Directorial Decree no. 1236 of 01/08/2023 “Fondo Italiano per la Scienza 2022-2023” (at the University of Ferrara).

Appendix A.

Here we present some well-known results on the porous media equation and how to set the exponent m and the time t^* to obtain the desired properties of the PM map introduced in Section 3.2.

A.1. Porous media equation and the Barenblatt solution

When the initial data of the porous media equation in Eq. (12) is given as $\rho_0(x) = M\delta_0(x)$ with $M > 0$, then there exists a closed-form solution of the problem, known as the Barenblatt solution [71]. It is given by

$$\rho(t,x) = t^{-\alpha}(A - B|x|^2 t^{-2\beta})_+^{\frac{1}{m-1}} \tag{A.1}$$

where α, β, A, and B are positive constants that can be expressed in terms of the dimension n, the exponent m, and the initial mass M. Following [50], they can be written as

$$\alpha = \beta n, \quad \beta = \frac{1}{2 + (m-1)n}, \quad A = \left(\frac{B^{n/2} M}{w_n z_{m,d}}\right)^{2\beta(m-1)}, \quad B = \beta \frac{m-1}{2m},$$

where w_n is the volume of the n sphere and $z_{m,n}$ is given by

$$z_{m,n} = \int_0^{\pi/2} [\cos(\theta)]^{\frac{m+1}{m-1}} [\sin(\theta)]^{n-1} d\theta.$$

The support of the Barenblatt solution $\rho(t,x)$ is a sphere expanding in time with radius

$$r(t) = t^\beta B^{-1/2} A^{1/2}. \tag{A.2}$$

A.2. Exponent m and time t^* for the PM map

We want to find the exponent m and the time t^* such that, given a singular measure of mass M centered at 0 at time $t = 0$, the evolution of this mass according to Eq. (12)

$$M = \kappa r(t^*)^p$$

for any M. This is equivalent to requiring that the following equality holds

$$r(t^*) = (t^*)^\beta B^{-1/2} A^{1/2} = (t^*)^\beta B^{-1/2} \frac{B^{n\beta(m-1)/2} M^{\beta(m-1)}}{(w_n, z_{m,n})^{\beta(m-1)}} = \kappa^{-1/p} M^{1/p} = r$$

for all M. This can be done by matching the exponents for M and proportionality constants, getting the following expression for the exponent m and the time t given by

$$m = \frac{2 + p - n}{p - n}, \quad t^* = \left(\kappa^{-1/p} \frac{(w_n, z_{m,n})^{\beta(m-1)}}{B^\beta}\right)^{1/\beta}. \tag{A.3}$$

Note that in our inpainting problem we need to set $n = d - 1$.

References

[1] H. Ben Yedder, B. Cardoen, G. Hamarneh, Deep learning for biomedical image reconstruction: a survey, Artif. Intell. Rev. 54 (1) (2021) 215–251.
[2] C. Guillemot, O. Le Meur, Image inpainting: overview and recent advances, IEEE Signal Process. Mag. 31 (1) (2013) 127–144.
[3] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., 2000, pp. 417–424.
[4] C.-B. Schönlieb, Partial Differential Equation Methods for Image Inpainting, 29, Cambridge University Press, 2015.
[5] O. Elharrouss, N. Almaadeed, S. Al-Maadeed, Y. Akbari, Image inpainting: a review, Neural Process. Lett. 51 (2) (2020).
[6] J.A. Kaufman, D. McCarter, S.C. Geller, A.C. Waltman, Two-dimensional time-of-flight MR angiography of the lower extremities: artifacts and pitfalls, AJR Am. J. Roentgenol. 171 (1) (1998) 129–135.
[7] R.R. Edelman, I. Koktzoglou, Noncontrast MR angiography: an update, J. Magn. Reson. Imaging 49 (2) (2019) 355–373.
[8] D. Li, B. Wu, B. Chen, C. Qin, Y. Wang, Y. Zhang, Y. Xue, Open-Surface river extraction based on sentinel-2 MSI imagery and DEM data: case study of the upper yellow river, Remote Sens. 12 (17) (2020).
[9] S. Parisotto, C.-B. Schönlieb, MATLAB/Python Codes for the Image Inpainting Problem, 2020. 10.5281/zenodo.4315173.
[10] X. Yi, E. Walia, P. Babyn, Generative adversarial network in medical imaging: a review, Med. Image Anal. 58 (2019) 101552.
[11] S. Moccia, E. De Momi, S. El Hadji, L.S. Mattos, Blood vessel segmentation algorithms - review of methods, datasets and evaluation metrics, Computer Methods Programs Biomed. 158 (2018) 71–91.
[12] J. Jam, C. Kendrick, K. Walker, V. Drouard, J.G.S. Hsu, M.H. Yap, A comprehensive review of past and present image inpainting methods, Comput. Vision Image Understanding 203 (May 2020) (2021) 103147.
[13] M. Nitzberg, D. Mumford, T. Shiota, Filtering, Segmentation and Depth, 662, Springer, 1993.
[14] S. Masnou, J.M. Morel, Level lines based disocclusion, in: Proceedings 1998 International Conference on Image Processing. ICIP98, 1998, pp. 259–263 vol.3.
[15] J. Shen, T.F. Chan, Mathematical models for local nontexture inpaintings, SIAM J. Appl. Math. 62 (3) (2002) 1019–1043.
[16] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1–4) (1992) 259–268.
[17] A.L. Bertozzi, S. Esedoglu, A. Gillette, Inpainting of binary images using the Cahn–Hilliard equation, IEEE Trans. Image Process. 16 (1) (2007) 285–291.
[18] M. Burger, L. He, C.-B. Schönlieb, Cahn–Hilliard inpainting and a generalization for grayvalue images, SIAM J. Imaging Sci. 2 (4) (2009) 1129–1167.
[19] A. Halim, B.V.R. Kumar, An anisotropic PDE model for image inpainting, Computers Math. Appl. 79 (9) (2020) 2701–2721.
[20] D. Jiang, M. Azaiez, A. Miranville, C. Xu, Nonlocal Cahn-Hilliard type model for image inpainting, Computers Math. Appl. 159 (2024) 76–91.
[21] S. Esedoglu, J. Shen, Digital inpainting based on the Mumford–Shah–Euler image model, Eur. J. Appl. Math. 13 (4) (2002) 353–370.
[22] C. Nan, Z. Qiao, Q. Zhang, Curvature-dependent elastic bending total variation model for image inpainting with the SAV algorithm, J. Sci. Comput. 101 (2) (2024) 29.
[23] C. Nan, Q. Zhang, Elastic bending total variation model for image inpainting with operator splitting method, Computers Math. Appl. 176 (2024) 150–164.
[24] C.D. Murray, The physiological principle of minimum work: I. The vascular system and the cost of blood volume, Proc. Natl. Acad. Sci. 12 (3) (1926) 207–214.
[25] I. Rodríguez-Iturbe, A. Rinaldo, Fractal River Basins: Chance and Self-Organization, Cambridge University Press, 2001.
[26] K.A. McCulloh, J.S. Sperry, F.R. Adler, Water transport in plants obeys Murray's law, Nature 421 (6926) (2003) 939–942.
[27] E. Facca, F. Cardin, M. Putti, Branching structures emerging from a continuous optimal transport model, J. Comput. Phys. 447 (2021) 110700.
[28] F. Santambrogio, Optimal Transport for Applied Mathematicians, Birkäuser, NY, 2015.
[29] G. Peyré, M. Cuturi, Computational optimal transport, Found. Trends Mach. Learn. 11 (5-6) (2019) 1–257.
[30] C. Villani, Optimal Transport: Old and New, 338, Springer, 2009.
[31] M. Bernot, V. Caselles, J.-M. Morel, Optimal Transportation Networks: Models and Theory, Springer, 2008.
[32] E.N. Gilbert, Minimum cost communication networks, Bell Labs Tech. J. 46 (9) (1967) 2209–2227.
[33] Q. Xia, Optimal paths related to transport problems, Commun. Contemp. Math. 5 (02) (2003) 251–279.
[34] R.B. Bapat, Graphs and Matrices, 27, Springer, 2010.
[35] Q. Xia, Motivations, ideas and applications of ramified optimal transportation, ESAIM Math. Modell. Numer. Anal. 49 (6) (2015) 1791–1832.
[36] Q. Xia, Boundary regularity of optimal transport paths, Adv. Calc. Var. 4 (2) (2011).
[37] F. Maddalena, S. Solimini, J.-M. Morel, A variational model of irrigation patterns, Interf. Free Boundaries 5 (2003) 391–415.
[38] L. Brasco, G. Buttazzo, F. Santambrogio, A Benamou–Brenier approach to branched transport, SIAM J. Math. Anal. 43 (2) (2011) 1023–1040.
[39] R. Rigon, A. Rinaldo, I. Rodriguez-Iturbe, R.L. Bras, E. Ijjasz-Vasquez, Optimal channel networks: a framework for the study of river basin morphology, Water Resour. Res. 29 (6) (1993) 1635–1646.
[40] J.R. Banavar, A. Maritan, A. Rinaldo, Size and form in efficient transportation networks, Nature 399 (6732) (1999) 130–132.
[41] Q. Xia, Numerical simulation of optimal transport paths, in: 2010 Second International Conference on Computer Modeling and Simulation, 1, 2010, pp. 521–525.
[42] P. Lippmann, E. Fita Sanmartín, F.A. Hamprecht, Theory and approximate solvers for branched optimal transport with multiple sources, Adv. Neural Inf. Process. Syst. 35 (2022) 267–279.
[43] E. Oudet, F. Santambrogio, A Modica–Mortola approximation for branched transport and applications, Arch. Ration. Mech. Anal. 201 (2011) 115–142.
[44] C. Dirks, B. Wirth, An adaptive finite element approach for lifted branched transport problems, Optim. Control Part. Differen. Equ. (2022) 189–236.
[45] J. Haskovec, P. Markowich, B. Perthame, Mathematical analysis of a PDE system for biological network formation, Commun. Part. Differen. Equ. 40 (5) (2015) 918–956.
[46] G. Albi, M. Burger, J. Haskovec, P. Markowich, M. Schlottbom, Continuum modeling of biological network formation, Active Particles, Volume (2017) 1.
[47] E. Facca, S. Daneri, F. Cardin, M. Putti, Numerical solution of Monge–Kantorovich equations via a dynamic formulation, J. Sci. Comput. 82 (3) (2020) 68.
[48] S. Arridge, P. Maass, O. Öktem, C.-B. Schönlieb, Solving inverse problems using data-driven models, Acta Numerica 28 (2019) 1–174.
[49] J.L. Vázquez, The Porous Medium Equation: Mathematical Theory, Oxford University Press, 2007.
[50] A. Fasano, M. Primicerio, Problems in Nonlinear Diffusion: Lectures Given at the 2nd 1985 Session of the Centro Internazionale Matematico Estivo (CIME) Held at Montecatini Terme, Italy, June 10-June 18, 1985, 1224, Springer, 2006.
[51] J. Cea, Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût, ESAIM: Modélisation mathématique et analyse numérique 20 (3) (1986) 371–402.
[52] A.S. Nemirovsky, D.B. Yudin, E.R. Dawson, Problem Complexity and Method Efficiency in Optimization, John Wiley & Sons, Inc New York, 1983.
[53] E. Facca, M. Benzi, Fast iterative solution of the optimal transport problem on graphs, SIAM J. Scient. Comput. 43 (3) (2021) A2295–A2319.
[54] F. Rathgeber, D.A. Ham, L. Mitchell, M. Lange, F. Luporini, A.T.T. Mcrae, G.-T. Bercea, G.R. Markall, P.H.J. Kelly, Firedrake: automating the finite element method by composing abstractions, ACM Trans. Math. Softw. 43 (3) (2016).
[55] D.A. Ham, P.H.J. Kelly, L. Mitchell, C.J. Cotter, R.C. Kirby, K. Sagiyama, N. Bouziani, S. Vorderwuelbecke, T.J. Gregory, J. Betteridge, D.R. Shapero, R.W. Nixon-Hill, C.J. Ward, P.E. Farrell, P.D. Brubeck, I. Marsden, T.H. Gibson, M. Homolya, T. Sun, A.T.T. McRae, F. Luporini, A. Gregory, M. Lange, S.W. Funke, F. Rathgeber, G.-T. Bercea, G.R. Markall, Firedrake User Manual, Imperial College London and University of Oxford and Baylor University and University of Washington, first edition edition, 2023.
[56] M.S. Alnæs, A. Logg, K.B. Ølgaard, M.E. Rognes, G.N. Wells, Unified form language: a domain-specific language for weak formulations of partial differential equations, ACM Trans. Math. Softw. 40 (2) (2014).
[57] S.K. Mitusch, S.W. Funke, J.S. Dokken, Dolfin-adjoint 2018.1: automated adjoints for FEniCS and Firedrake, J. Open Source Softw. 4 (38) (2019) 1292.
[58] S. Balay, S. Abhyankar, M.F. Adams, S. Benson, J. Brown, P. Brune, K. Buschelman, E. Constantinescu, L. Dalcin, A. Dener, V. Eijkhout, J. Faibussowitsch, W.D. Gropp, V. Hapla, T. Isaac, P. Jolivet, D. Karpeev, D. Kaushik, M.G. Knepley, F. Kong, S. Kruger, D.A. May, L.C. McInnes, R.T. Mills, L. Mitchell, T. Munson, J.E. Roman, K. Rupp, P. Sanan, J. Sarich, B.F. Smith, S. Zampini, H. Zhang, H. Zhang, J. Zhang, PETSc/TAO Users Manual, Technical Report ANL-21/39 - Revision 3.20, Argonne National Laboratory, 2023.

[59] S. Balay, W.D. Gropp, L.C. McInnes, B.F. Smith, Efficient management of parallelism in object oriented numerical software libraries, in: E. Arge, A.M. Bruaset, H.P. Langtangen (Eds.), Modern Software Tools in Scientific Computing, Birkhäuser Press, 1997, pp. 163–202.

[60] R.D. Falgout, J.E. Jones, U.M. Yang, The design and implementation of hypre, a library of parallel high performance preconditioners, in: Numerical Solution of Partial Differential Equations on Parallel Computers, Springer, 2006, pp. 267–294.

[61] J.F. Cohnheim, Untersuchungen über die embolischen Processe, A. Hirschwald, 1872.

[62] U.N.A. Qohar, A. Zanna Munthe-Kaas, J.M. Nordbotten, E.A. Hanson, A nonlinear multi-scale model for blood circulation in a realistic vascular system, R. Soc. Open Sci. 8 (12) (2021) 201949.

[63] C.B. Sullivan, A. Kaszynski, PyVista: 3D plotting and mesh analysis through a streamlined interface for the visualization toolkit (VTK), J. Open Source Softw. 4 (37) (2019) 1450. https://doi.org/10.21105/joss.01450

[64] J.C. Russ, The Image Processing Handbook, CRC Press, 2006.

[65] D. Baptista, D. Leite, E. Facca, M. Putti, C. De Bacco, Network extraction by routing optimization, Sci. Rep. 10 (1) (2020) 20806.

[66] E. Hodneland, E. Hanson, O. Sævareid, G. Nævdal, A. Lundervold, V. Ŝoltészová, A. Z. Munthe-Kaas, A. Deistung, J. R. Reichenbach, J.M. Nordbotten, A new framework for assessing subject-specific whole brain circulation and perfusion using MRI-based measurements and a multi-scale continuous flow model, PLoS Comput. Biol. 15 (6) (2019) 1–31.

[67] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al., Fiji: an open-source platform for biological-image analysis, Nat. Methods 9 (7) (2012) 676–682.

[68] G. Todeschi, L. Métivier, J.-M. Mirebeau, Unbalanced L^1-optimal transport for vector valued measures and application to Full Waveform Inversion, 2024. 2403.02764

[69] E. Facca, F. Cardin, M. Putti, Towards a stationary Monge–Kantorovich dynamics: the physarum polycephalum experience, SIAM J. Appl. Math. 78 (2) (2018) 651–676.

[70] A. Lonardi, E. Facca, M. Putti, C. De Bacco, Infrastructure adaptation and emergence of loops in network routing with time-dependent loads, Phys. Rev. E 107 (2023) 024302.

[71] G.I. Barenblatt, Y.B. Zel'Dovich, Self-similar solutions as intermediate asymptotics, Annu. Rev. Fluid Mech. 4 (1) (1972) 285–312.