

Connected Components on Lie Groups and Applications to Multi-Orientation Image Analysis

Nicky Janneke van den Berg

`n.j.v.d.berg@tue.nl`

Eindhoven University of Technology

Olga Mula

Eindhoven University of Technology

Leanne Vis

Eindhoven University of Technology

Remco Duits

Eindhoven University of Technology

Research Article

Keywords: Connected Components, Lie Groups, Vessel Tree Identification, Image Analysis, Morphological Operators, Medical Image Analysis

Posted Date: November 19th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-5389923/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: Competing interest reported. R. Duits is a member of the editorial board of JMIV.

Connected Components on Lie Groups and Applications to Multi-Orientation Image Analysis

Nicky J. van den Berg · Olga Mula · Leanne Vis · Remco Duits

the date of receipt and acceptance should be inserted later

Abstract We develop and analyze a new algorithm to find the connected components of a compact set I from a Lie group G endowed with a left-invariant Riemannian distance. For a given $\delta > 0$, the algorithm finds the largest cover of I such that all sets in the cover are separated by at least distance δ . We call the sets in the cover the δ -connected components of I . The grouping relies on an iterative procedure involving morphological dilations with Hamilton-Jacobi-Bellman kernels on G and notions of δ -thickened sets. We prove that the algorithm converges in finitely many iteration steps, and we propose a strategy to find an optimal value for δ based on persistence homology arguments. We also introduce the concept of affinity matrices. This allows for grouping of δ -connected components based on their local proximity and alignment.

Among the many different applications of the algorithm, in this article, we focus on illustrating that the method can efficiently identify (possibly overlapping) branches in complex vascular trees on retinal images. This is done by applying an orientation score transform to the images that allows us to view them as functions from $\mathbb{L}_2(G)$ where $G = SE(2)$, the Lie group of rotations. By applying our algorithm in this Lie group, we illustrate that we obtain δ -connected components that differentiate between crossing structures and that group well-aligned, nearby structures. This contrasts standard connected component algorithms in $\mathbb{R}^2$.

Keywords Connected Components · Lie Groups · Vessel Tree Identification · Image Analysis · Morphological Operators · Medical Image Analysis

1 Introduction

1.1 Context and Goals

In image analysis, the task of connecting components plays an important role in numerous applications such as image segmentation [14, 36, 16, 53], object recognition [45, 11, 54, 28, 30], motion tracking [28, 30, 38, 27, 40], and data compression [42, 52]. In this article, we develop a novel method for identifying connected components in images with line structures that possibly intersect with each other, as in Fig. 1, 2. When applying a naive connected component algorithm in $\mathbb{R}^2$ on these images, the algorithm finds only one or two connected components, as it cannot differentiate between the different structures at crossings (cf. Fig. 1b, 2b).

We present a strategy that does not automatically merge different crossing structures. The image data is lifted from $\mathbb{R}^2$ to the space of positions and orientations, where crossing structures are disentangled based on their local orientation (cf. Fig. 1d, 2d).

After the lifting step, we apply our connected component algorithm on the Lie group $SE(2)$. This algorithm is built for any Lie group G and uses theory involving (logarithmic approximations of) Riemannian distances. This results in the practical advantage of grouping different, possibly overlapping, anisotropic connected components based on their local alignment (measured by a Riemannian distance).

Once all connected components have been identified, the output is projected back onto the input image in $\mathbb{R}^2$,

Nicky J. van den Berg · Olga Mula · Leanne Vis · Remco Duits

Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands

E-mail: n.j.v.d.berg@tue.nl, o.mula@tue.nl, l.vis@student.tue.nl, r.duits@tue.nl

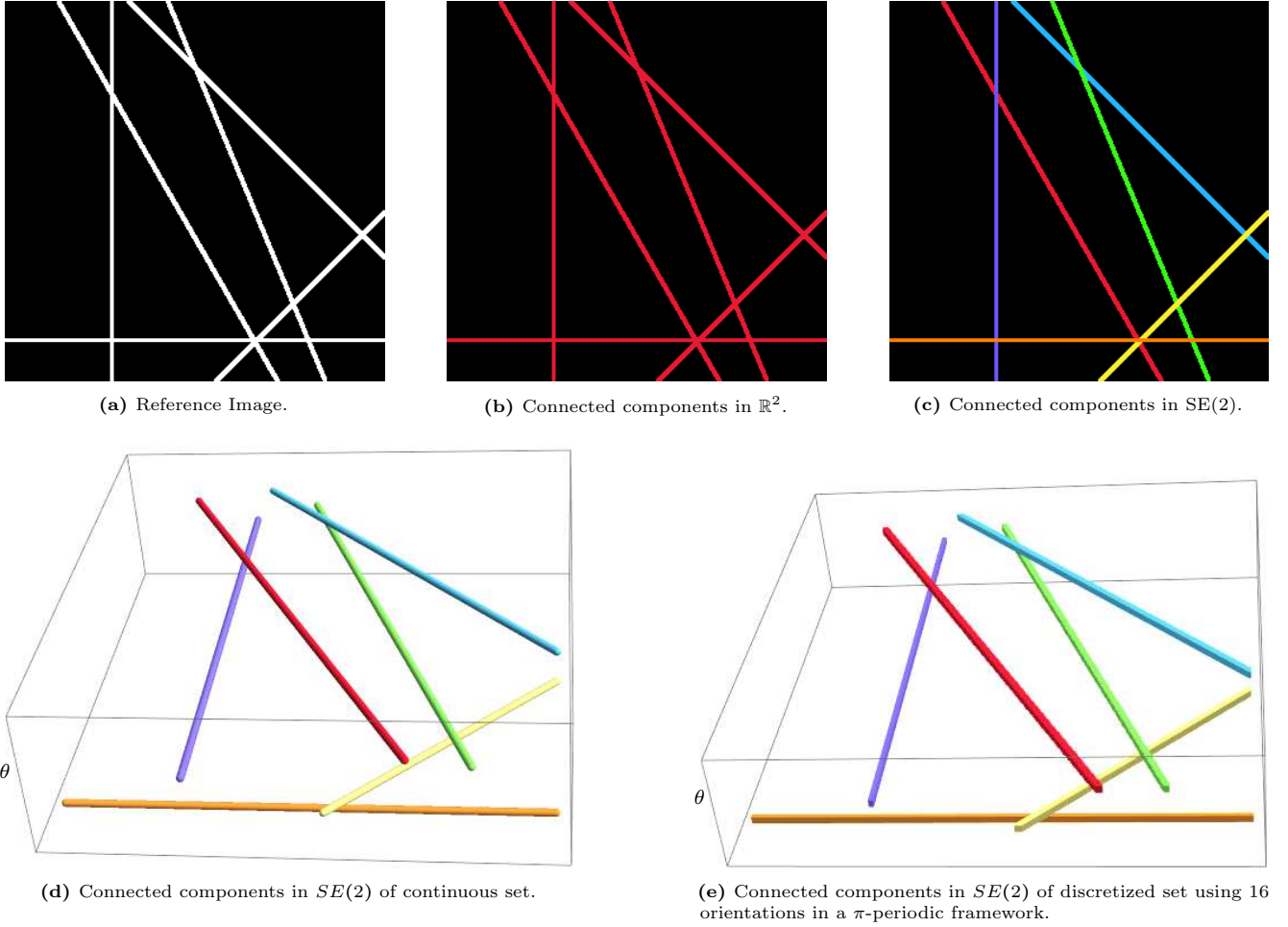


Fig. 1 Visualization of the connected component algorithm in $\mathbb{R}^2$ and in the Lie group $SE(2)$ on an image of straight lines. The classical connected component algorithm in $\mathbb{R}^2$ cannot differentiate between the different line structures, unlike the extension to $SE(2)$. Here we applied the algorithm in Sec. 5.2 using parameters $(w_1, w_2, w_3) = (0.1, 1, 4)$ for the left-invariant metric (4).

as visualized in Fig. 1c, 2c. The new algorithm cannot only deal with overlapping structures, but can also deal with small interruptions of lines, still assigning them to the same connected component, as shown in Fig. 3.

1.2 Fundamentals of Connected Component Algorithms

Now that we have illustrated the main geometric idea on basic examples, let us zoom out, and consider the vast literature on connected components analysis. There are typically two main approaches to define connectivity: 1) the graph-theoretic approach and 2) the topological approach. In the first, one considers images as a set of pixels connected through a graph [17, 41, 48, 33]; in the second, images are rather seen as continuous functions on a smooth manifold [19, 3, 50, 13]. This work belongs to the second category. There are two main subclasses of topology-based algorithms: grey-level mathematical morphology [32, 51] and binary morphology [10, 15, 6,

2]. In this article, we use techniques from grey-level mathematical morphology, such as dilation, to solve binary morphology applications.

Connected components have received a lot of attention over the years, resulting in the development of various techniques and algorithms.

- **Pixel connectivity:** Initially, people were interested in identifying connected components in images and sets. They tackled the problem by looking at the direct neighbors of a pixel or voxel in the set, i.e. either sharing an edge or corner in 2D, or sharing a face, edge, or corner in 3D [43]. This results in a very local analysis of a given input.
- **Distance inclusion:** This local connectivity approach is not always desirable depending on the application. Therefore, techniques were developed where the connectivity was defined in terms of the distances between neighboring points, using mathematical morphology [31]. This kind of connectivity allows

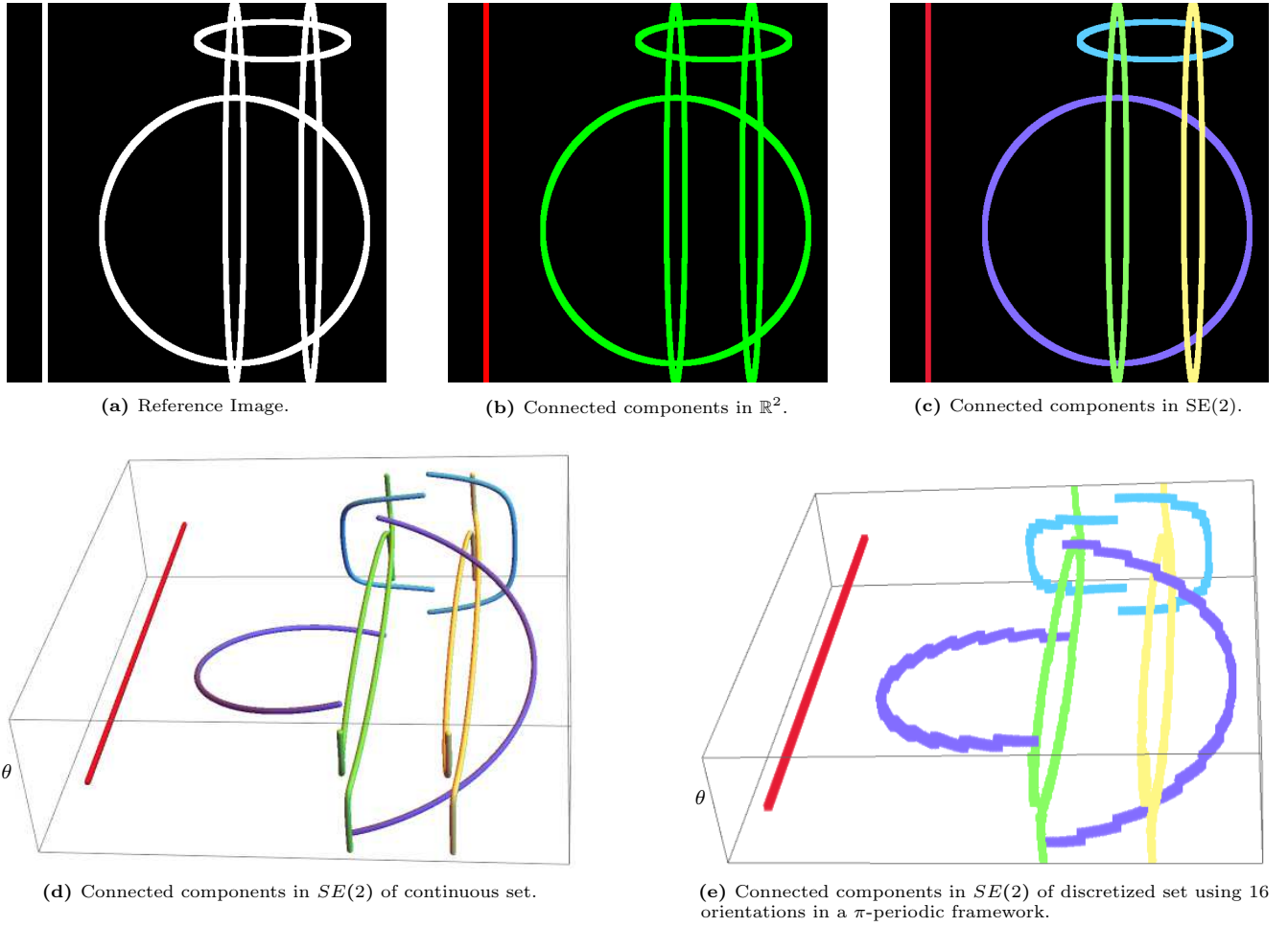


Fig. 2 Visualization of the effect of performing the connected component algorithm in $\mathbb{R}^2$ and in the Lie group $SE(2)$ on an image of ovals and lines. The classical connected component algorithm in $\mathbb{R}^2$ is not able to differentiate between the different line structures, unlike the extension to $SE(2)$. Here we applied the algorithm in Sec. 5.2 using parameters $(w_1, w_2, w_3) = (0.2, 1, 4)$ for the left-invariant metric (4).

to assign non-adjacent pixels to the same connected component. This is useful when one wants to identify multiple, previously separate connected components, as one, e.g. in the identification of paw prints or words [39].

- **Symmetry inclusion:** The grouping of line elements should be equivariant under roto-translations. This is guaranteed if the metric is left-invariant. Note that for isotropic distances, the equivariance under roto-translations is always satisfied.

1.3 Connectivity on the Lifted Space $SE(2)$

In this article, we focus on dealing with overlapping structures. We highlight some of the works that have focussed on this in the past.

Van de Gronde et al [48] suggested a graph-based approach. There, a local orientation tensor was identified

for each vertex in the graph. This allowed for the identification of all possible paths containing every vertex while satisfying a set of constraints (local orientations of two adjacent pixels sufficiently aligned). The resulting set of possibly connected vertices represented one connected component. Since this method only considers adjacent pixels for connectivity, this is a local approach in the lifted space.

Besides the graph-based approach in the lifted space, different algorithms based on perceptual grouping (grouping of regions and parts of the visual scene to get higher order perceptual units such as objects or patterns [12]) have been introduced to identify vascular trees (in the lifted space) [6, 2]. In [2], the connectivity is learned from retinal images, after which the learned convection-diffusion kernel is used to determine the connectivity of the vessel fragments. On the other hand, in [6], a set of key points located on blood vessels are

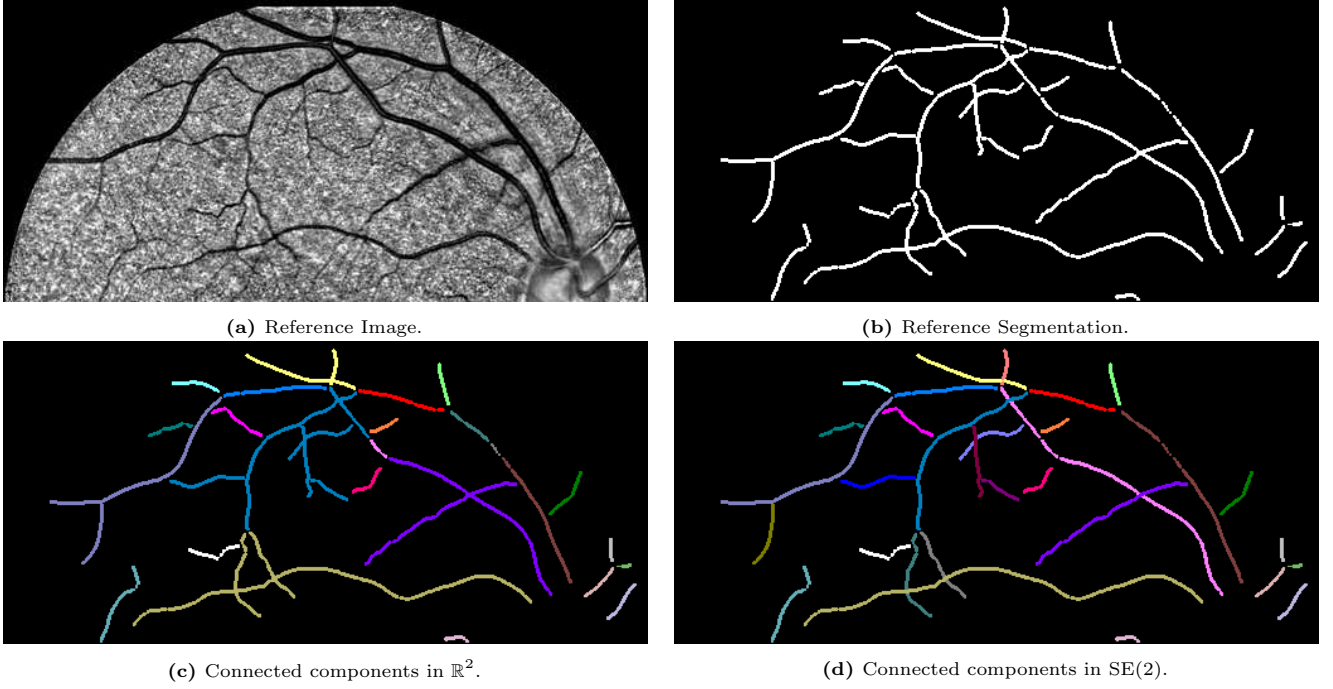


Fig. 3 Visualization of the effect of the connected component algorithm in $\mathbb{R}^2$ and in the Lie group $SE(2)$. The classical connected component algorithm is not able to differentiate between crossing vessels, and additionally breaks up not perfectly connected vessels, into different components. The algorithm presented in Sec. 5.2 using parameters $(w_1, w_2, w_3) = (0.2, 1, 4)$ for the left-invariant metric (4), can better differentiate between different crossing structures and can additionally group well-aligned structures resulting in a more intuitive result.

used as input for the algorithm. Then, the key points are connected based on their mutual distance. It is important to note that these algorithms do not rely on mathematical morphology.

1.4 Our approach: Connected Components by PDE-Morphology in the Lifted Space

We aim for an approach that relies on mathematical morphology techniques to identify connected components. We consider Lie groups G equipped with a left-invariant metric tensor field $\mathcal{G}$. This metric tensor field induces a left-invariant distance d_G on the Lie group. The precise formulas for the metric d_G and metric tensor field $\mathcal{G}$ will follow in Section 2 (cf. (3),(4)).

In the specific case $G = SE(2)$, the group of roto-translations, the metric tensor field $\mathcal{G}$ is determined by three parameters w_1, w_2 , and w_3 . Intuitively, these parameters represent tangential, lateral, and angular movement costs respectively. The influence of different choices for these parameters is illustrated in Fig. 4, where the reference image is lifted to the space of roto-translations. One single Riemannian ball is plotted in each figure (lifted to $SE(2)$ and projection on $\mathbb{R}^2$) to illustrate the influence of the metric tensor field.

The introduced algorithm will identify, for a given threshold $\delta > 0$, the largest cover of I such that all sets are separated by at least distance δ . Every set in this cover is called a δ -connected component. The threshold distance $\delta > 0$ is a parameter to be tuned by the user. In calculating these δ -connected components, we rely on formal morphological convolutions via Hamilton-Jacobi-Bellman (HJB) equations and Riemannian distance approximations, allowing for fast, parallelizable morphological convolution algorithms.

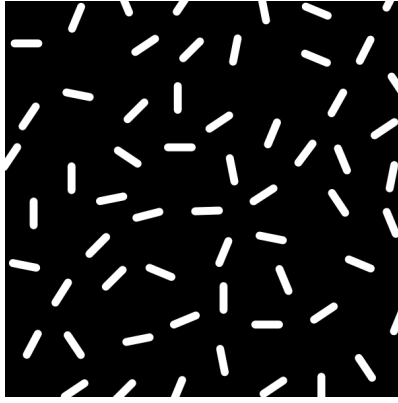
In the experimental section of this article, we have restricted ourselves to the Lie group $G = SE(2)$. We consider two-dimensional images which we view as functions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ belonging to the space $\mathbb{L}_2(\mathbb{R}^2, \mathbb{R})$. We next lift the images to the space of roto-translations $SE(2)$ with a mapping

$$W_\phi : \mathbb{L}_2(\mathbb{R}^2, \mathbb{R}) \rightarrow \mathbb{L}_2(SE(2), \mathbb{R}),$$

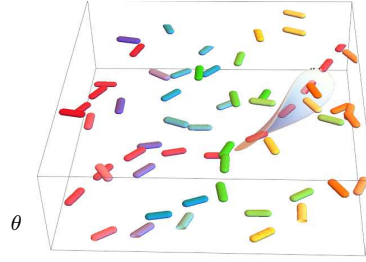
where, for every $f \in \mathbb{L}_2(\mathbb{R}^2, \mathbb{R})$,

$$(W_\phi f)(g) := \int_{\mathbb{R}^2} \phi(R_\theta^{-1}(\mathbf{y} - \mathbf{x})) f(\mathbf{y}) d\mathbf{y} \quad \forall g = (\mathbf{x}, \theta) \in SE(2), \quad (1)$$

and where $\phi \in \mathbb{L}_1(\mathbb{R}^2) \cap \mathbb{L}_2(\mathbb{R}^2)$ is a rotating anisotropic wavelet. In this article, we use the real-valued cake



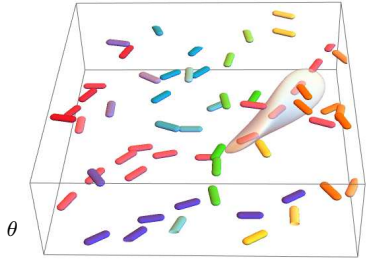
(a) Reference Image containing 50 line segments.



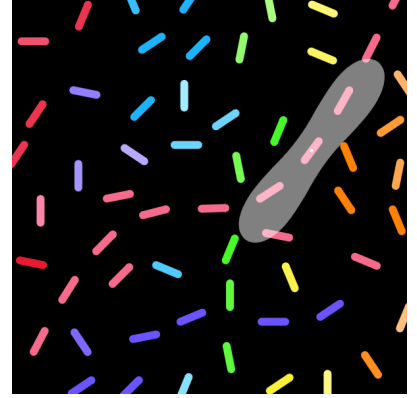
(b) Connected components when using an anisotropic metric tensor field with parameters $(w_1, w_2, w_3) = (0.05, 1, 2)$, anisotropy factor 20. 1-connected component count: 38.



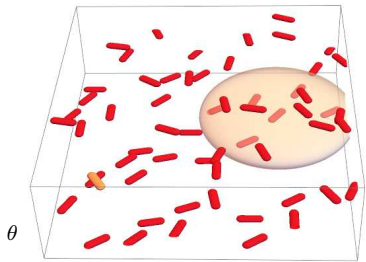
(c) Connected components when using an anisotropic metric tensor field with parameters $(w_1, w_2, w_3) = (0.05, 1, 2)$. 1-connected component count: 38.



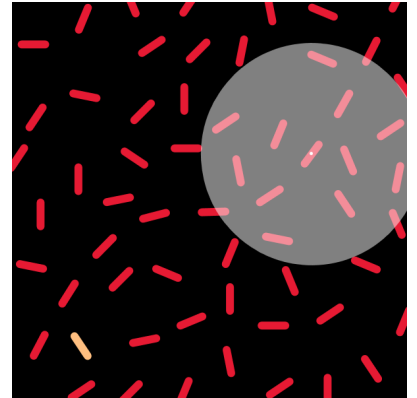
(d) Connected components when using an anisotropic metric tensor field with parameters $(w_1, w_2, w_3) = (0.05, 0.3, 2)$, anisotropy factor 6. 1-connected component count: 32.



(e) Connected components when using an anisotropic metric tensor field with parameters $(w_1, w_2, w_3) = (0.05, 0.3, 2)$. 1-connected component count: 32.



(f) Connected components when using an isotropic metric tensor field with parameters $(w_1, w_2, w_3) = (0.05, 0.05, 2)$, anisotropy factor 1. 1-connected component count: 2.



(g) Connected components when using an isotropic metric tensor field with parameters $(w_1, w_2, w_3) = (0.05, 0.05, 2)$. 1-connected component count: 2.

Fig. 4 Visualization of the effect of the metric tensor field on the connected component algorithm in the Lie group $SE(2)$. The output heavily relies on the chosen distance metric d_G introduced in (3), with metric tensor parameters as defined in (5).

wavelets (for more information, see [21,24]). The resulting function $W_\phi f$ is usually called the orientation

score of the image f , and the operator W_ϕ is called the orientation score transform.

The δ -connected components are calculated on suitably binarized orientation scores $W_\phi f$, after which the resulting components are “projected back” onto $\mathbb{L}_2(\mathbb{R}^2, \mathbb{R})$ by inverting W_ϕ [20].

In the experimental section, we see that the δ -connected component algorithm allows us to identify aligned structures in images. In the experiments, we look at both artificial images (cf. Figs. 1,2,4,8,9) and images of the human retina (Figs. 3,12-16). On both datasets, we can identify connected components that do not suffer from the crossing structures. We have also performed a stability analysis on the δ -connected components as illustrated in Fig. 8,9. Lastly, we look at grouping different components in images of the human retina based on their mutual affinity (linked to perceptual grouping). By doing this, we can find even more complete vascular trees.

As the approach generalizes to other Lie groups and applications, cf. [5,25,35,29,22], we formulate the theory beyond the $SE(2)$ setting. Crucial ingredients are a good logarithmic norm approximation (valid for $G = H(d), SO(d), SE(d)$) and a lift via a unitary group representation so that a left action on the image is a left action on the score [20].

1.5 Main Contributions

The main contributions of this paper are:

1. We introduce the notion of δ -connected components in compact metric spaces, which is a relaxation of the definition of connected components.
2. We develop a method to find the connected components of a binary function defined on any Lie group G equipped with a left-invariant metric tensor field using morphological convolutions, resulting in δ -connected components that are identified in a group equivariant way.
3. We mathematically analyze our method:
 - (a) We prove that the provided algorithm always converges to the correct δ -connected component in a finite number of steps in Theorem 1.
 - (b) We show reflectional symmetries of our connected component algorithm (Cor. 2 in App. B) and how this generalizes to other Lie groups (Lemma 4).
4. We introduce the concept of affinity matrices, allowing us to measure alignment (both spatial and in orientation) between different δ -connected components. We provide bounds for the affinity in Proposition 4 and use it to group different components based on their proximity.
5. We introduce the idea of “persistence homology” of δ -connected components, allowing us to analyze the

stability of all components, depending on the choice of δ .

6. Notebooks for all experiments are available via [8] that allow for easy reproduction of all results.

1.6 Structure of the Article

First, we briefly introduce Lie groups and left-invariant distances in Sec. 2. In Sec. 3, we introduce our notion of δ -connected components and elaborate on our approach towards identifying those components. Then, we introduce the morphological convolutions (Sec. 4) that are used in the presented δ -connected component algorithm (Sec. 5.2). We prove it converges to the desired components in a finite number of steps (Sec. 5.3). Additionally, we introduce the idea of persistency of δ -connected components in Sec. 6. We define the affinity between different δ -connected components in Sec. 7, which allows us to quantify how well different components are aligned. Once all these theoretical concepts have been introduced, we move on to Sec. 8, where we present the experimental part of the article. We summarize and conclude in Sec. 9.

2 Background on Lie Groups and Left-Invariant Distances

The theory presented in this document applies to all finite-dimensional Lie groups G of dimension $n \in \mathbb{N}$, with a left-invariant Riemannian distance d . We recall that a Lie group is a smooth manifold G equipped with a group structure such that the group multiplication $\mu : G \times G \rightarrow G$ given by $\mu(x, y) = xy$ and the inverse map $\nu : G \rightarrow G$ given by $\nu(x) = x^{-1}$ are smooth maps.

The tangent space at $h \in G$ of Lie group G is the vector space of all tangent vectors of curves passing through $h \in G$, which we denote by $T_h(G)$.

Let $g \in G$. The left-action on G is denoted by $L_g : G \rightarrow G$ and defined as $L_g(h) = gh$, and its corresponding push-forward $(L_g)_* : T_h(G) \rightarrow T_{gh}(G)$ is defined as $(L_g)_* A_h u = A_h(u \circ L_g)$ for all smooth functions $u : G \rightarrow \mathbb{C}$, where $A_h \in T_h(G)$. In particular if we take $h = e$, the unit element, and a coordinate basis $\{\partial_{x^i}|_e\}_{i=1}^n$ at $T_e(G)$, then

$$(L_g)_* \partial_{x^i}|_e u = \partial_{x^i}|_e (u \circ L_g).$$

The Riemannian metric tensor field is a field of inner products $\mathcal{G}_g : T_g(G) \times T_g(G) \rightarrow \mathbb{R}$ for each $g \in G$. We always consider a left-invariant Riemannian metric tensor field $\mathcal{G}$, i.e. for all $g, h \in G$ and for all $\dot{h} \in T_h(G)$

$$\mathcal{G}_{gh}((L_g)_* \dot{h}, (L_g)_* \dot{h}) = \mathcal{G}_h(\dot{h}, \dot{h}). \quad (2)$$

Then, the Riemannian left-invariant distance d_G is defined by

$$d_G(g, h) := \inf \left\{ \int_0^1 \sqrt{\mathcal{G}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))} dt \mid \gamma \in \Gamma_1, \gamma(0) = g, \gamma(1) = h \right\}, \quad (3)$$

where $\Gamma_1 = \text{PC}([0, 1], G)$ denotes the family of piecewise continuously differentiable curves. Then from (2), it follows readily that $d_G(gh_1, gh_2) = d_G(h_1, h_2)$ for all $g, h_1, h_2 \in G$. Intuitively, this means that the distance is invariant to left-actions, and one has $d_G(g, h) = d_G(h^{-1}g, e)$, where $e \in G$ is the unit element.

Remark 1 The distance d_G always relies on a metric tensor $\mathcal{G}$. Henceforth, we omit the label $\mathcal{G}$, and write $d := d_G$ for the distance whenever we do not need to stress the metric tensor used to calculate the distances.

Remark 2 (Explanation weights of metric tensor field) From (2), one can deduce that all left-invariant (recall Eq. (2)) metric tensor fields are given by

$$\mathcal{G}_g(\dot{g}, \dot{g}) = \sum_{i,j=1}^n c_{ij} \dot{g}^i \dot{g}^j \quad (4)$$

with *constant* coefficients $c_{ij} \in \mathbb{R}$, and where $\dot{g} = \sum_{i=1}^n \dot{g}^i \mathcal{A}_i|_g \in T_g(G)$ for a basis $\{\mathcal{A}_i\}_{i=1}^n$ of left-invariant vectors $G \ni g \mapsto \mathcal{A}_i|_g \in T_g(G)$, $i = 1, \dots, n$. Note that in general the matrix $[c_{ij}]$ is positive symmetric definite, but in this article we constrain ourselves to the diagonal case

$$c_{ij} = w_i \delta_{ij}, \quad (5)$$

with positive weights $w_i > 0$.

Remark 3 The theoretical results in this article also hold when the distance denotes a sub-Riemannian distance, though this requires different logarithmic approximations than (8) below, for details see [47, 23, 7].

We have a particular interest in the cases of the roto-translation group

$$G = SE(2) := \mathbb{R}^2 \rtimes SO(2)$$

and the rotation group

$$G = SO(3) := \{X \in \mathbb{R}^{3 \times 3} \mid X^T X = I, \det(X) = 1\}.$$

The group product of $G = SE(2)$ is given by

$$g_1 g_2 = (x_1, R_1)(x_2, R_2) = (x_1 + R_1 x_2, R_1 R_2) \quad (6)$$

for all $(x_1, R_1), (x_2, R_2) \in \mathbb{R}^2 \rtimes SO(2)$, and the left-actions are given by roto-translations. Similarly, the group product $R_1, R_2 \in SO(3)$ is given by the ordinary matrix product $R_1 R_2 \in SO(3)$, where we note that $(R_1 R_2)(R_1 R_2)^T = R_1 R_2 R_2^T R_1^T = R_1 I R_1^T = I$ and $\det(R_1 R_2) = \det(R_1) \det(R_2) = 1$.

Remark 4 In the case where $G = SE(2)$, we have in (4), $\dot{g} = \sum_{i=1}^3 \dot{g}^i \mathcal{A}_i|_g \in T_g(G)$, where

$$\mathcal{A}_1 = \cos \theta \partial_x + \sin \theta \partial_y, \mathcal{A}_2 = -\sin \theta \partial_x + \cos \theta \partial_y, \mathcal{A}_3 = \partial_\theta, \quad (7)$$

are left-invariant vector fields and where we identify

$$\theta \in \mathbb{R}/(2\pi\mathbb{Z}) \leftrightarrow R_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \in SO(2).$$

In that case, w_1, w_2 and w_3 in Eq. (4) and (5) describe costs for forward, sideways and angular movement in $SE(2)$ respectively.

The groups $G = SE(2)$ and $G = SO(3)$ have the special property that the exponential map $\text{Exp} : T_e(G) \rightarrow G$ is surjective and we have a complete logarithmic norm approximation

$$d_G(g, h) = d_G(h^{-1}g, e) \approx \|\log h^{-1}g\|_G \quad (8)$$

for $g, h \in G$ close enough to each other, cf. Fig. 5. For details on the quality of this approximation for the $G = SE(2)$ case, see [7]. For a concise self-contained summary and explicit formulas for the logarithmic norm approximation in (8) in coordinates, see Eq. (34) for $G = SE(2)$ and Eq. (37) for $G = SO(3)$ in Appendix A.

Remark 5 (Inclusion of Symmetries) It is well-known in Lie group theory that left-invariant Riemannian metrics have a constant Gram matrix in any left-invariant frame. So in (4), $c_{ij} \in \mathbb{R}$ independent of $g \in G$ if $(\mathcal{A}_i)_{i=1}^n$ is a left-invariant frame, i.e. a basis of left-invariant vector fields. The associated distance maps of such metrics and their logarithmic norm estimates carry $2^{\dim(G)}$ reflectional symmetries. In the specific Lie group case of roto-translations $G = SE(2)$ there are 2^3 reflectional symmetries (for intuitive illustration see App. B, Fig. 17) which also become reflectional symmetries of our connected component algorithm.

Remark 6 The Lie group $SE(2)$ can be identified with the homogeneous manifold of positions and orientations in two dimensions $\mathbb{M}_2$. However, this is not true for $n > 2$: $SE(n) \not\subset \mathbb{M}_n$.

3 An Algorithm to Find the δ -Connected Components of the Compact Set I

To explain our algorithm, we first introduce the notion of δ -connectedness of a compact set I .

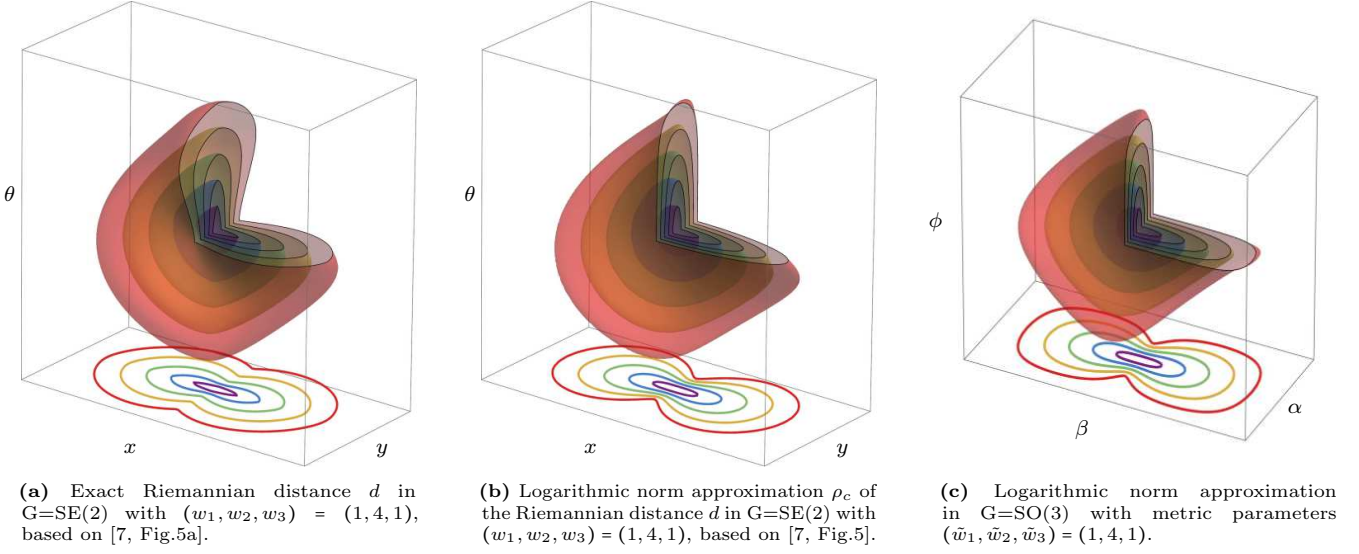


Fig. 5 Visualization of distance balls and their logarithmic approximations in $G=SE(2)$ and $G=SO(3)$. We see isocontours of $d(p_0, \cdot)$ in G , and on the bottom, we see the min-projection over the orientation θ of these contours. The visible contours are $d = 0.5, 1, 1.5, 2, 2.5$ and the metric parameters are $(w_1, w_2, w_3) = (\tilde{w}_1, \tilde{w}_2, \tilde{w}_3) = (1, 4, 1)$. Parameter w_1 controls costs for tangential motion, w_2 controls costs for lateral motion on the base manifold (resp. $\mathbb{R}^2$ and S^2) and w_3 controls costs for changing the orientation along a geodesic. Thereby, together, they control the anisotropic shape of the Riemannian ball and the identification of connected components, recall white balls in column 2 of Fig. 4.

3.1 The Notion of δ -Connectedness

We start by recalling the notion of the covering number of I , for which we use the definition of a ball.

Definition 1 (Riemannian Ball) The ball around $g \in G$ with radius $\delta > 0$ is given by

$$B(g, \delta) = \{h \in G \mid d(g, h) < \delta\}.$$

Definition 2 (δ -covering) A set $C \subset I$ is a δ -covering of I if $I \subseteq \bigcup_{g \in C} B(g, \delta)$.

Definition 3 (Covering number) The covering number, denoted by $n_\delta(I)$, is the smallest cardinality of all possible δ -coverings of the set I , i.e.

$$n_\delta(I) := \min \{|C| : C \text{ is a } \delta\text{-covering of } I\}. \quad (9)$$

Since we assume that the set I is compact, its covering number $n_\delta(I)$ is finite. Next, we introduce the notion of δ -connectedness between points $g, h \in I$.

Definition 4 (δ -connectedness of points) Let $\delta > 0$. We say that two elements g and $h \in I$ are δ -connected, and we denote it as $g \tilde{\sim} h$, if and only if there exists a finite number $m \in \mathbb{N}$ of elements $\{q_i\}_{i=0}^m \subset I$ such that $q_0 = g, q_m = h$ and

$$d(q_{i+1}, q_i) \leq \delta, \quad \forall i \in \{0, \dots, m-1\}.$$

Remark 7 Note that $m \in \mathbb{N}$ may depend on $g, h \in G$ and δ , so $m := m_\delta(g, h) \in \mathbb{N}$.

One can readily check that $\tilde{\sim}$ is an equivalence relation on I , and the equivalence class of a $g \in I$ under $\tilde{\sim}$, denoted by $[g]$, is defined as

$$[g] := \{q \in I \mid q \tilde{\sim} g\}.$$

These preliminary concepts allow us to give the following definition to the notion of ‘ δ -connected components’.

Definition 5 (δ -connected components) The set of δ -connected components of a given set I is defined as its equivalence class $\tilde{I}^\delta := I / \tilde{\sim}$, and each element $[g] \in \tilde{I}^\delta$ is called a ‘ δ -connected component’.

The following result guarantees that there are at most $n_\delta(I)$ δ -connected components to search for when I is a compact set.

Lemma 1 For every compact set I one has that the number of δ -connected components is bounded by the covering number, i.e. $|\tilde{I}^\delta| \leq n_\delta(I)$.

Proof We argue by contradiction. Suppose that the covering number is $n_\delta(I)$, and suppose there are $m > n_\delta(I)$ connected components in I . Then there are m points $\{g_i\}_{i=1}^m$ in I such that $d(g_i, g_j) > \delta$ for all $1 \leq i, j \leq m$. Therefore, we need at least m balls of radius δ to cover the set I , so $n_\delta(I) \geq m$. However, by assumption $m > n_\delta(I)$. $\square$

Remark 8 It follows from Lemma 1 that the distance between the closest pair of points between two δ -connected

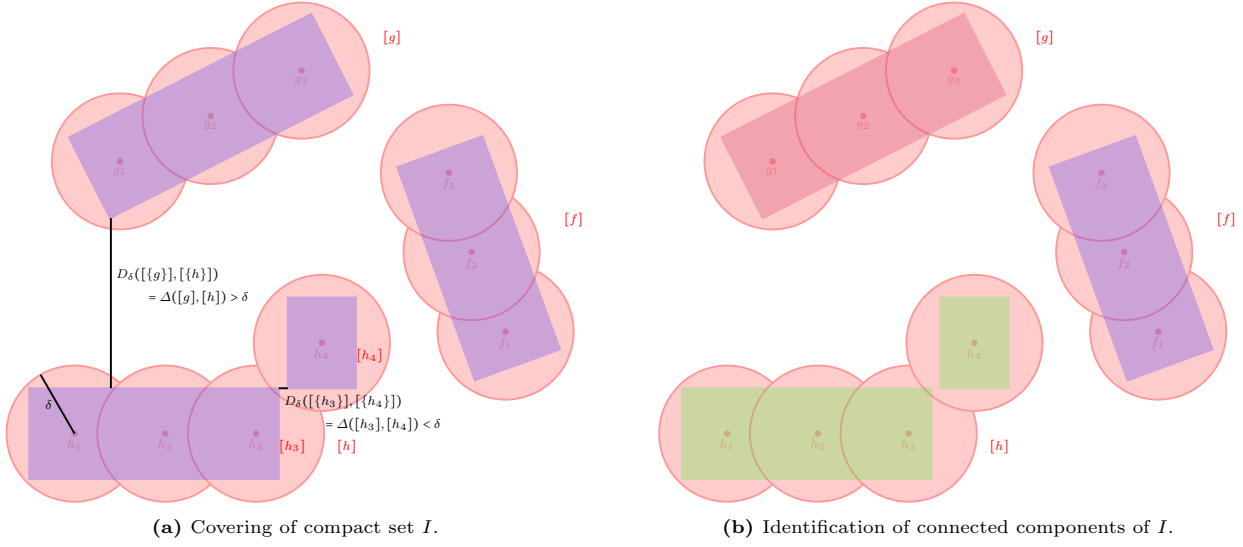


Fig. 6 The set of connected components I (in blue in Fig. 6a) has a finite covering. In this case, it is covered by 10 balls with radius δ (in red), and has covering number $n_\delta(I) = 10$. The distance between different connected components is larger than the radius of a ball δ . The example shows three different δ -connected components, indicated by color in Fig. 6b.

components is larger than δ , that is,

$$\inf_{(g,h) \in [g] \times [h]} d(g,h) > \delta \quad \text{if } [g] \neq [h].$$

We next define several notions of distances that will be needed in subsequent developments.

Definition 6 (Distance from point to set) The distance from a point $g \in G$ to a non-empty set $A \subset G$ is given by

$$d(g, A) := \inf_{a \in A} d(g, a).$$

Definition 7 (Distance between two sets) The distance from a non-empty set $A \subset G$ to a non-empty set $B \subset G$ is given by

$$d(A, B) := \inf_{a \in A, b \in B} d(a, b).$$

Additionally, we will need the concept of an ϵ -thickened set when we explain the δ -connected component algorithm:

Definition 8 (ϵ -thickened set) Let $\epsilon > 0$. For every closed subset $A \subset G$, we define its ϵ -thickened version A_ϵ as the set

$$A_\epsilon := \{g \in G \mid d(g, A) \leq \epsilon\} = \bigcup_{a \in A} \overline{B(a, \epsilon)}.$$

3.2 A General Algorithm for (Disjoint) Connected Components

Given a threshold $\delta > 0$, the following strategy finds the δ -connected components of a compact set I . We defer

the discussion on how to select an optimal value of δ to Sec. 6, and focus here on explaining the algorithm once this parameter is fixed. The strategy consists of a main algorithm `find_all_components` which identifies all δ -connected components relative to the compact set I . It relies on an algorithm that identifies the δ -connected component $[g] \in \tilde{I}^\delta$ of a given $g \in I$ (algorithm `find_full_component`).

Algorithm `find_all_components`

Find all δ -Connected Components $[g] \in \tilde{I}^\delta$
for a Compact $I \subset G$

To find all $[g_i] \in \tilde{I}^\delta$ in I , we proceed iteratively:

- $k = 1$:
 1. Pick any $g_1 \in I$.
 2. Compute $[g_1]$ with `find_full_component` to compute δ -connected components, which is introduced in the next box.
 3. Set $\tilde{I}^\delta(1) := \{[g_1]\}$.
 4. Define $I^\delta(1) := \{g \in G \mid [g] \in \tilde{I}^\delta(1)\}$.
- $k > 1$: Given $\tilde{I}^\delta(k-1) = \{[g_1], \dots, [g_{k-1}]\}$ and its associated set $I^\delta(k-1)$.
 1. Search an element $g_k \in I \setminus I^\delta(k-1)$
 - (a) if g_k exists: Update

$$\tilde{I}^\delta(k) := \tilde{I}^\delta(k-1) \cup \{[g_k]\}$$
 and $I^\delta(k)$ accordingly.
 - (b) if g_k does not exist: all k δ -connected components are identified, i.e. $\tilde{I}^\delta(k-1) = \tilde{I}^\delta$. The algorithm ends.

Algorithm find_full_component

Find δ -Connected Component $[g]$ for a $g \in G$
 To find $[g] \in \tilde{I}^\delta$ for a given g , in practice, we again proceed iteratively:

- $n = 0$: We set $C(g, 0) = \{g\}$.
- $n \geq 1$: Given $C(g, n-1)$, find

$$C(g, n) = C_\delta(g, n-1) \cap I, \quad (10)$$

where $C_\delta(g, n-1) := (C(g, n-1))_\delta$ is the δ -thickened set of $C(g, n-1)$ (see definition 8).

As we prove next in Lemma 2, if $C(g, n) = C(g, n-1)$, then $C(g, n) = [g]$ and the algorithm terminates. Otherwise, we go to step $n+1$.

In practice, $C(g, n)$ is computed by means of

$$C(g, n) = C_\delta(g, n-1) \cap I = \left(\bigcup_{c \in C(g, n-1)} B(c, \delta) \right) \cap I.$$

As a result, it suffices to compute $B(c, \delta) \cap I$ for every $c \in C(g, n-1)$ to find the new set $C(g, n)$.

In practice, the thickened sets $C_\delta(g, n)$ in **find_full_component** are computed with a strategy involving morphological convolutions which we explain in Sec. 4.

Lemma 2 *Let $g \in G$ and $n \geq 1$. If $C(g, n) = C(g, n-1)$, then $C(g, n) = [g]$.*

Proof We prove the statement by contradiction: Suppose $C(g, n) = C(g, n-1)$ and $C(g, n-1) \not\subseteq [g]$. Then, there exists a $h \in [g]$ such that $h \notin C(g, n-1)$ and $d(h, C(g, n-1)) \leq \delta$ (due to the definition of the equivariance classes $[g]$), so $h \in C_\delta(g, n-1)$. Since we also have that $h \in I$, it follows that $h \in C_\delta(g, n-1) \cap I = C(g, n)$. Hence $h \in C(g, n-1)$, which is a contradiction. $\square$

3.3 How the Algorithm Works in an Example

An example of finding connected components in (a discrete subset of) $\mathbb{R}^2$ is given in Figure 7. Here, we consider an image of a person and a tree. The set I consists of the center points of the pixels with width ρ . Then, we execute the algorithm with $\delta = \rho\sqrt{2}$.

The first δ -connected component is identified in row 2 to 4 of Fig. 7. We initiate the algorithm by picking $g_1 = C(g_1, 0) \in I$, depicted by the green pixel in the first column. Then, the set $C_\delta(g_1, 0)$ is identified in transparent green and contains all pixels whose center point has at most distance δ to the point g_1 . Then, set $C(g_1, 1)$ is shown in non-transparent green in the 2nd row, 3rd column. These steps are repeated until

stabilization is reached, so until $C(g_1, n) = C(g_1, n-1)$ (for $n = 7$ in Fig. 7). The final two rows depict the identification of the second connected component in red.

4 Computing a Thickened Set A_ϵ with Hamilton-Jacobi Equations

The above algorithm hinges on computing δ -thickened versions of the sets $C(g, n)$ in the sense of Def. 8. In this section, we give a practical algorithm based on morphological dilations on the Lie group G . We carry the discussion for a generic set A first. The transformations used to create a thickened set A_ϵ rely on morphological convolutions as defined next.

Definition 9 (Morphological convolution) Let G be a Lie group. Let $f_1, f_2 : G \rightarrow \mathbb{R}$ be two lower semi-continuous functions, bounded from below. Then the morphological convolution $(f_1 \square f_2) : G \rightarrow \mathbb{R}$ is given by

$$(f_1 \square f_2)(g) = \inf_{h \in G} \{f_1(h^{-1}g) + f_2(h)\}.$$

To create an ϵ -thickened set A_ϵ , we rely on morphological dilations. The concept of morphological dilations is connected to the notion of viscosity solutions of Hamilton-Jacobi-Bellmann (HJB) equations. For a given $\alpha > 1$ and initial condition $f \in C(G)$, we search for a $W : G \times [0, T] \rightarrow \mathbb{R}$ that satisfies the Hamilton-Jacobi-Bellmann equation

$$\begin{cases} \frac{\partial W}{\partial t}(g, t) = \frac{1}{\alpha} \|\nabla_G W(g, t)\|^\alpha, & t \in (0, T] \\ W(g, 0) = f(g), & g \in G. \end{cases} \quad (11)$$

The above equation has a unique viscosity solution W which is given by

$$W(g, t) = -(k_t^\alpha \square -f)(g), \quad (12)$$

where

$$k_t^\alpha(g) := \frac{t}{\beta} \left(\frac{d(g, e)}{t} \right)^\beta =: \kappa_t^\alpha(d(g, e)) \quad (13)$$

is called a morphological kernel with $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, $\alpha, \beta > 1$ and where $\kappa_t^\alpha(x) = \frac{t}{\beta} (x/t)^\beta$, $x \in \mathbb{R}$ denotes the 1D morphological kernel. Equation (13) shows that the morphology on G relates to the 1D morphology on the distance map. For more details on this relation, cf. Appendix C.1.

Definition 10 (Morphological dilation) We say that $W(\cdot, t)$ is a morphological dilation of f with kernel k_t^α given by (12) if W is the viscosity solution of Eq. (11).

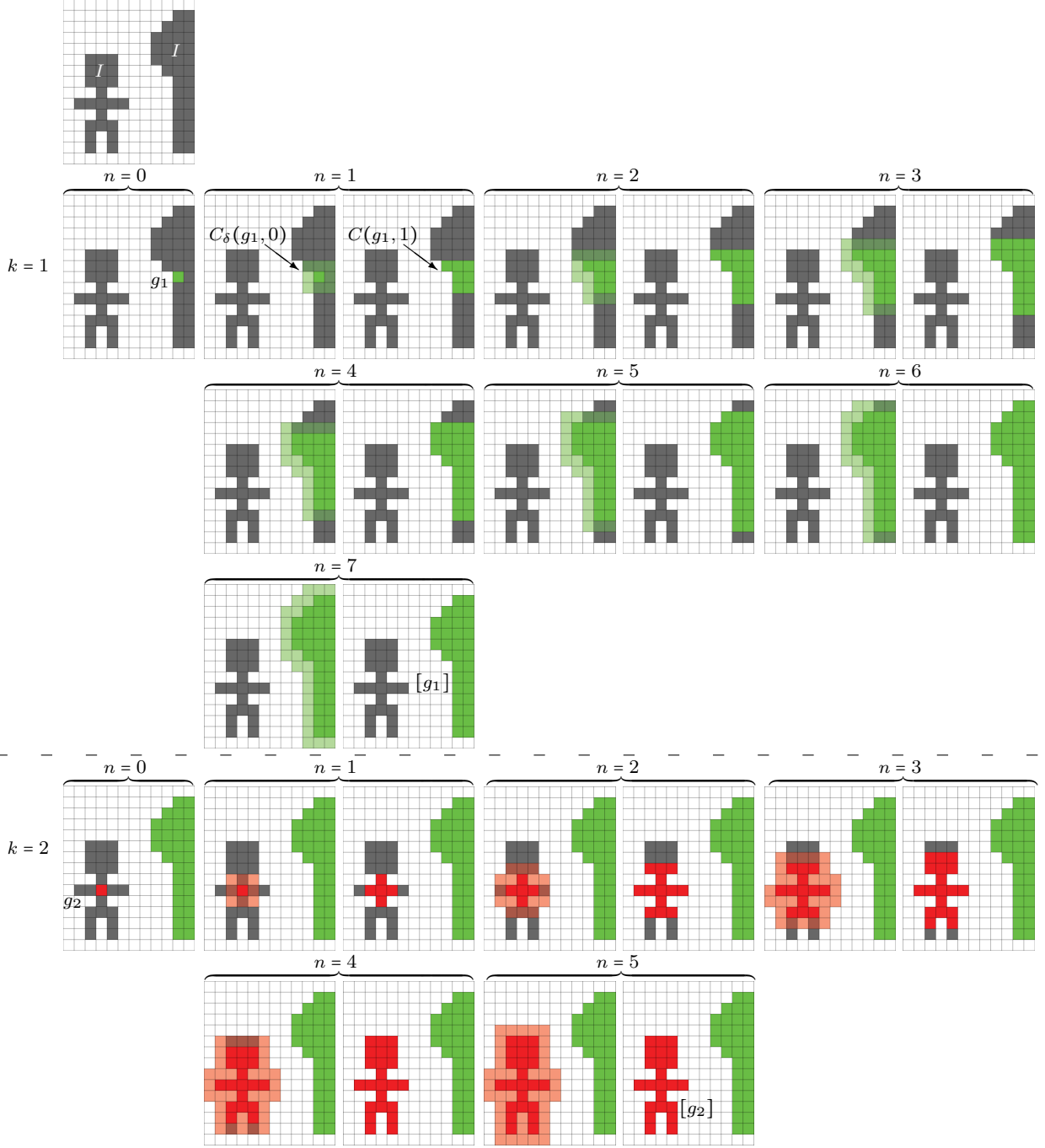


Fig. 7 A visualization of a connected component algorithm on a set $\Omega = \{1, \dots, 12\} \times \{1, \dots, 15\}$. Here, one sees in the top figure $I = \{\text{grey pixel locations}\}$. There are two separate connected components which are identified in two steps: $k = 1$ shows the process of identifying $[g_1] := \{\text{green pixel locations}\}$ and $k = 2$ shows the process of finding the second connected component $[g_2] := \{\text{red pixel locations}\}$.

For subsequent developments, it will be useful to view the viscosity solution $W(\cdot, t)$ given in (12) as the image of f under the flow map $\varphi_t^\alpha : C(G) \rightarrow \mathbb{R}$ given by

$$\varphi_t^\alpha(f) := W(\cdot, t) = -(k_t^\alpha \square - f). \quad (14)$$

The semi-group property of the flow implies that

$$\varphi_t^\alpha \circ \varphi_s^\alpha(f) = \varphi_{t+s}^\alpha(f), \quad \forall t, s \in \mathbb{R}, \quad \forall f \in C(G). \quad (15)$$

Remark 9 Note that the semi-group property of the flow implies that $\varphi_t^\alpha \circ \varphi_t^{\alpha(n-1)} \varphi_t^\alpha f = \varphi_{nt}^\alpha f$ for all $t > 0$. This is

due to the well-posedness of (11) in terms of viscosity solutions, but can also be seen more explicitly by looking at the corresponding kernels, see Lemma 5 in Appendix C.1.

We next list some useful properties connected to the semi-group property of the HJB equation, where we use the indicator function and the support:

Definition 11 (Indicator function) The indicator function $\mathbb{1}_S : G \rightarrow \{0, 1\}$ of some set $S \subset G$ is defined as

$$\mathbb{1}_S(x) := \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{else.} \end{cases}$$

Definition 12 (Support of a function) The support of the function $f : G \rightarrow \mathbb{R}$ is given by the closure:

$$\text{supp}(f) := \overline{\{g \in G \mid f(g) \neq 0\}}.$$

Lemma 3 Let $\alpha \geq 1$. Let $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. For a given closed set $A \subset G$ and a function $f : G \rightarrow [0, 1]$, one has

$$\varphi_t^\alpha(f)(g) \in [0, 1], \quad \forall g \in G, \quad (16)$$

$$A_{\varepsilon(t, \alpha)} = \text{supp}(\varphi_t^\alpha(\mathbb{1}_A)), \quad \forall t \geq 0, \quad (17)$$

$$\mathbb{1}_{A_t} = (\varphi_t^1(\mathbb{1}_A)), \quad (18)$$

with

$$\varepsilon(t, \alpha) = t \sqrt[\beta]{\frac{\beta}{t}}, \quad (19)$$

and where A_t , $A_{\varepsilon(t, \alpha)}$ are t - and $\varepsilon(t, \alpha)$ -thickened sets of A , as defined in Def. 8.

Proof We prove the statements one by one, starting with (16): Let $f : G \rightarrow [0, 1]$. Then, as the kernels are positive (13), we have for all $g \in G$

$$\begin{aligned} \varphi_t^\alpha(f)(g) &\stackrel{(14)}{=} -(k_t^\alpha \square -f)(g) = \sup_{h \in G} \{f(h) - k_t^\alpha(h^{-1}g)\} \\ &\leq \|f\|_\infty = 1, \end{aligned}$$

and

$$\begin{aligned} \varphi_t^\alpha(f)(g) &\stackrel{(14)}{=} -(k_t^\alpha \square -f)(g) = \sup_{h \in G} \{f(h) - k_t^\alpha(h^{-1}g)\} \\ &\geq \sup_{h \in G} \{-k_t^\alpha(h^{-1}g)\} = 0. \end{aligned}$$

The two estimates above prove (16).

We next prove (17): We are interested in the support of $\varphi_t^\alpha(\mathbb{1}_A)$. From (16), we know

$$\varphi_t^\alpha(\mathbb{1}_A) = -(k_t^\alpha \square -\mathbb{1}_A) \in [0, 1].$$

By definition of the support and the positivity of $\varphi_t^\alpha(\mathbb{1}_A)$, we have $\text{supp}(\varphi_t^\alpha(\mathbb{1}_A)) = \overline{\{g \in G \mid \varphi_t^\alpha(\mathbb{1}_A)(g) > 0\}}$,

so we aim to find the set B , such that for all $g \in B$, we have

$$0 < \varphi_t^\alpha(\mathbb{1}_A)(g) = \sup_{h \in G} \left\{ \mathbb{1}_A(h) - \frac{t}{\beta} \left(\frac{d(h^{-1}g, e)}{t} \right)^\beta \right\} \leq 1, \quad (20)$$

Note that $\overline{B} = \text{supp}(\varphi_t^\alpha(\mathbb{1}_A))$ and that the upper boundary is always satisfied as the kernel k_t^α , recall Eq. (13), is positive and $\mathbb{1}_A \leq 1$. The supremum is only strictly larger than zero if the supremum is reached for $h \in A$, so $A \subset \overline{B}$. Then, (20) gives

$$0 < 1 - \frac{t}{\beta} \left(\frac{d(g, A)}{t} \right)^\beta \Leftrightarrow d(g, A) < t \sqrt[\beta]{\frac{\beta}{t}} \text{ for } g \in B.$$

Taking $g \in \overline{B}$ gives the condition $d(g, A) \leq t \sqrt[\beta]{\frac{\beta}{t}}$. Therefore, the set $\overline{B} = A_{\varepsilon(t, \alpha)}$, where $A_{\varepsilon(t, \alpha)}$ is a thickened set of A as defined in Def. 8. This proves Eq. (17).

The last statement (18) follows immediately from (17) by taking the limit of $\alpha \downarrow 1$, combined with

$$\begin{aligned} k_t^1(g) &= \lim_{\alpha \downarrow 1} k_t^\alpha(g) = \lim_{\beta \rightarrow \infty} \frac{t}{\beta} \left(\frac{d(g, e)}{t} \right)^\beta \\ &= \begin{cases} 0 & \text{if } d(g, e) \leq t, \\ \infty & \text{else.} \end{cases} \end{aligned} \quad (21)$$

□

We conclude this section by connecting one last time to the main goal in this part of the article which is to explain how to create an ϵ -thickened set A_ϵ . The main message is that A_ϵ can be obtained through Eqs. (17) and (18), which crucially rely on morphological dilations. In practice, it is useful to rely on Eq. (18) to determine the ϵ -thickened set A_ϵ in the δ -connected component algorithm, as we will see in the next section.

5 Computing δ -Connected Components using Iterative Morphological Convolutions

In Sec. 3.2, we explained the δ -connected component algorithm in a general setting. In Sec. 5.2, we will explain how the δ -connected component algorithm is constructed using morphological dilations, but first, we discuss the choice of the parameter α in the morphological dilation kernel k_t^α in Sec. 5.1. We finish with a convergence analysis for the algorithm in Sec. 5.3.

5.1 Fixing $\alpha = 1$ in the δ -Connected Component Algorithm

We aim to construct a δ -connected component algorithm that relies on morphological dilations. Before we introduce the algorithm, we discuss the choice of the parameter $\alpha \geq 1$ in the morphological dilation kernel k_t^α (13).

We recall that identifying the δ -connected components is an iterative procedure as introduced in Sec. 3.2. The identification of one single δ -connected component is mainly described by Eq. (10). In every iteration step n , we create a δ -thickened set $C_\delta(g, n-1)$ and take the intersection with a reference set $I \subset G$, resulting in an updated set $C(g, n)$. This is repeated until the updated set is the same as the initial set, i.e. $C(g, n) = C(g, n-1)$.

Hence, the set $C(g, n)$ is thickened iteratively, which motivates the identification of the maximum radius of influence of one single point after n morphological dilations, as will be introduced in Def. 13. The maximum radius of influence follows immediately from the semi-group property in Remark 9 and the support after applying the flow to an indicator function in Equation (17).

Definition 13 (Maximum radius of influence) The maximum radius of the region of influence of a specific point after n applications of a morphological dilation of time t and kernel steepness α is denoted by

$$d_{n,t}^\alpha := nt \left(\frac{\beta}{nt} \right)^{\frac{1}{\beta}} = \varepsilon(nt, \alpha),$$

where $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ and $\varepsilon(t, \alpha)$ was defined in Eq. (19).

We recall that an ϵ -thickened set was produced with Eq. (17) and (18). We aim to identify δ -connected components. Therefore, we want the support of the set to grow with rate δ in every iteration step while applying Eq. (17). That means we would like to have that $(A_{\varepsilon(t, \alpha)})_{\varepsilon(t, \alpha)} = A_{\varepsilon(2t, \alpha)}$ which is only satisfied if $\beta \rightarrow \infty \Leftrightarrow \alpha = 1$ as is clear from Eq. (19).

If $\alpha > 1$, the algorithm cannot always identify the full δ -connected component, while using the same kernel k_t^α in every set, as is visualized in Fig. 18 in App. D. Therefore, in our computations of the δ -connected components, we set $\alpha = 1$ and apply morphological convolutions with k_t^1 , given by (21), with $\delta := t$.

5.2 δ -Connected Component Algorithm

Next, we detail how we implement the general δ -connected component algorithm as discussed in Sec. 3.2 using the morphological dilations that were discussed in

Sec. 4. More specifically, we explain how the δ -thickened set C_δ in `find_full_component` is determined using the properties of morphological dilations as stated in Lemma 3.

First, we discuss how to identify one single δ -connected component $[g_i]$ for a $g_i \in I$. We initialize

$$C(g_i, 0) = \{g_i\} = \text{supp}(\mathbb{1}_{\{g_i\}}).$$

During iteration step n of identifying one single δ -connected component $[g_i]$, the set $C(g_i, n-1)$ from the previous iteration is expanded with a radius δ using morphological dilations and Eq. (18) in Lemma 3.

Therefore, we define a function $U : G \times \mathbb{N} \rightarrow \{0, 1\}$ which is the indicator function of the computed δ -connected component initialized in $g_i \in I$ at step n , i.e.

$$U_{g_i}(\cdot, n-1) := \mathbb{1}_{C(g_i, n-1)}(\cdot).$$

This allows us to calculate the set $C(g_i, n)$ using morphological dilations on the function $U_{g_i}(\cdot, n-1)$

$$U_{g_i}(\cdot, n) := \mathbb{1}_I(\cdot) \varphi_\delta^1(U_{g_i}(\cdot, n-1)), \quad (22)$$

where

$$C(g_i, n) := \text{supp}(U_{g_i}(\cdot, n)). \quad (23)$$

These steps are repeated until $C(g_i, n) = C(g_i, n-1)$. Then, we update $I = I \setminus C(g_i, n)$ and pick a new $g_{i+1} \in I$ to identify the next component as explained in `find_all_components` in Sec. 3.2. The complete algorithm is also stated in Algorithm 1: `find_deltaCC` and produces δ -connected components as stated in the next proposition.

Proposition 1 *Algorithm 1: `find_deltaCC` produces δ -connected components, i.e. $[g_1], \dots, [g_K]$ are all δ -connected.*

The result follows immediately from Eq. (22), using Eq. (18) in Lemma 3 where $t = \delta$.

Prop. 1 shows that `find_deltaCC` finds δ -connected components. It remains to be shown that the algorithm is indeed a stable algorithm which we will address next.

5.3 Convergence Analysis

In this section, we prove in Theorem 1 that Algorithm 1: `find_deltaCC` computes the δ -connected components of any compact set I on any Lie group G in a finite number of calculation steps for $\alpha = 1$.

For this, we need to have an explicit non-recursive expression of $U_{g_0}(\cdot, n)$ because this describes the state of the δ -connected component algorithm in the n -th

Algorithm 1: find δ CC

Input: Binary map $\mathbb{1}_I : G \rightarrow \{0, 1\}$, where $I \subset G$ the region of interest for the connected components, kernel k_δ^α with $\alpha = 1$, $\delta > 0$
 set $k = 0$
 set $\mathcal{I} = I$
 while $\mathcal{I} \neq \emptyset$:
 update $k = k + 1$
 pick $g_k \in \mathcal{I}$
 initialize $U_{g_k}(\cdot, 0) = \mathbb{1}_{\{g_k\}}$
 initialize connected component sets
 $C(g_k, -1) := \emptyset$, $C(g_k, 0) = \text{supp}(U_{g_k}(\cdot, 0)) = \{g_k\}$.
 set $n := 0$
 while n such that $C(g_k, n) \neq C(g_k, n - 1)$:
 do a morphological convolution calculating $U_{g_k}(\cdot, n + 1)$ as described in (22).
 create set $C(g_k, n + 1) = \text{supp}(U_{g_k}(\cdot, n + 1))$ (see Eq. (23))
 update $n := n + 1$
 end
 set $[g_k] := C(g_k, n)$
 update $\mathcal{I} = \mathcal{I} \setminus [g_k]$
 end
 $K = k$
Output: Connected components $[g_1], \dots, [g_K]$.

iteration step. In Proposition 2, we find such a non-recursive expression for $U_{g_i}(\cdot, n)$, where Definition 13 and Lemma 3 are key ingredients in proving the convergence of our δ -connected component algorithm (Theorem 1).

Proposition 2 *Let $\alpha = 1$, $t = \delta > 0$, $g \in G$, $n \in \mathbb{N}$. Let $g_0 \in G$ be the reference point. Set $\delta = d_{1,t}^1$. Then, we have, for $U_{g_0}(g, n)$ given by Eq. (22), that $0 \leq U_{g_0}(g, n) \leq 1$ where*

$$U_{g_0}(g, n) = \begin{cases} 1 & \text{if } g \stackrel{\delta}{\sim} g_0 \wedge m_\delta(g, g_0) \leq n \\ 0 & \text{else.} \end{cases},$$

where $m_\delta(g, g_0)$ was defined in Def. 4 and Remark 7.

For the details of the proof of Proposition 2, see Appendix C.2. Intuitively, the condition $m_\delta(g, g_0) \leq n$ ensures that the n steps of the connected component algorithm are sufficient to reach the point g starting from g_0 .

Knowing the algorithm's state after n iterations, allows us to confidently say that the algorithm can identify all δ -connected components in a finite number of steps, as proved in the next theorem.

Theorem 1 (Convergence of Algorithm 1: find δ CC) *Let $\delta > 0$, $\alpha = 1$, and assume that I has K connected components $[g_1], \dots, [g_K]$ such that $I = [g_1] \cup \dots \cup [g_K] \subset G$ and $[g_i] \cap [g_j] = \emptyset$ for all $i \neq j \in \{1, \dots, K\}$. Then Algorithm 1: find δ CC correctly finds all the δ -connected components in a finite number of steps.*

Proof Consider $\alpha = 1$. Let $i \in \{1, \dots, K\}$ be given. Note that in this case, one has the dilated volume $U_{g_i}(\cdot, n) : G \rightarrow \{0, 1\}$ and the connected components

$$C(g_i, n) = \text{supp}(U_{g_i}(\cdot, n)) = \{g \in G \mid U_{g_i}(g, n) \neq 0\}.$$

In the proof of Prop. 2, we defined the set $\tilde{B}(g_i; n)$ by

$$\tilde{B}(g_i; n) := \{g \in I \mid m_\delta(g_i, g) \leq n\}. \quad (24)$$

Then, by Proposition 2 where we set g_i instead of g_0 , one has

$$U_{g_i}(g, n) = \mathbb{1}_{[g_i]}(g) \cdot \mathbb{1}_{\tilde{B}(g_i; n)}(g)$$

$$C(g_i, n) = \tilde{B}(g_i; n).$$

From the compactness of I and the fact that $I = \bigcup_i [g_i]$, $[g_i] \cap [g_j] = \emptyset$, one can deduce that $[g_i]$ is closed. Therefore, also $[g_i]$ is compact. Consequently, its covering number $n_\delta([g_i])$, as introduced in (9), is finite and thus $U_{g_i}(\cdot, n)$ converges in at most $n = n_\delta([g_i]) + 1 < \infty$ steps (see Corollary 1) to $[g_i]$ (Prop. 1), i.e. $U_{g_i}(\cdot, n_\delta([g_i]) + 1) = U_{g_i}(\cdot, n_\delta([g_i]) + 2)$ and thus $C(g_i, n_\delta([g_i]) + 1) = C(g_i, n_\delta([g_i]) + 2)$. $\square$

As a direct consequence of Theorem 1, we can give an upper bound for the maximum number of iterations that find δ CC needs to identify all δ -connected components.

Corollary 1 (Upper bound maximum number of iterations) *Let $I = [g_1] \cup \dots \cup [g_K] \subset G$ compact, where the δ -connected components $[g_i] \cap [g_j] = \emptyset$ for all $i \neq j \in \{1, \dots, K\}$. The δ -connected component algorithm given in Algorithm 1: find δ CC finishes in at most $n = n_\delta(I) + K$ steps.*

Proof We start by noticing that $n_\delta(I) = \sum_{i=1}^K n_\delta([g_i])$ because $d([g_i], [g_j]) > \delta$ for $i \neq j$.

To identify δ -connected component $[g_i]$, we need at most 1 step to get from the starting point $g_i \in I$ to a point $\tilde{g}_i \in C$, where C is the minimal δ -covering of I . Then, by Lemma 1, we need at most $n_\delta([g_i])$ steps to identify the full δ -connected component $[g_i]$.

Hence, identifying all δ -connected components in $I = \bigcup_{i=1}^K [g_i]$ is done in at most $\sum_{i=1}^K 1 + n_\delta([g_i]) = K + n_\delta(I)$ steps. $\square$

6 Choosing δ with Persistence Homology Arguments

In the previous section, we have discussed how to find δ -connected components using morphological convolutions. We chose a fixed δ at the start of the algorithm, and

use `find_δCC` to identify all δ -connected components. Consequently, the algorithm's output heavily depends on the choice of δ (recall Fig. 6).

To gain some insight into the behavior of the δ -connected components, we will create 'persistence diagrams' to study the stability of the components. We visualize the persistence in two different ways. Firstly, we create a figure where we plot the total number of δ -connected components for different values of δ . When the distance between two components is smaller than the considered δ , they merge and hence the number of connected components decreases. Secondly, we visualize, for sets of points, how many points are grouped in each δ -connected component. These plots (for instance in Fig. 8) allow us to verify the persistence of the δ -connected components. Depending on the application, we want them to be as persistent as possible, e.g. when looking at vascular images, or we want to choose δ small enough to differentiate between the smallest possible δ -connected components. Note that using a very small value of δ results in a high number of δ -connected components due to noise being assigned its own δ -connected component.

In Fig. 8, we show the persistence diagrams for two examples in $\mathbb{R}^2$. The first concerns a set of points (7 in total) that belong to our set I_0 , depicted in Fig. 8a. Then, we determine the δ -connected components for a range of values of δ . We plot the number of connected components per value of δ (cf. Fig. 8b) and the number of vertices that are part of every δ -connected component (cf. Fig. 8c). One can see from the plots that the most stable connected components (at least one connection, but not everything is connected) occur when $\delta \in (2.2, 4.1)$.

We do the same for the second example, a different set I_1 , where 20 points are placed in a circle, and with two outliers (see Fig. 8d). Then, the persistence graphs in Fig. 8e and Fig. 8f show a stabilization for $\delta \in (0.3, 2.6)$.

In Fig. 9, we do the same, but now on the Lie group $G = SE(2)$. Again, we see persistence of the δ -connected components for $\delta \in (0.57, 1.13)$ and $\delta \in (1.13, 1.92)$ with 6 and 4 δ -connected components respectively. Therefore, choosing δ in these intervals results in the most persistent δ -connected components.

7 Affinity Matrices between δ -Connected Components

Now that we have described a way to choose the parameter δ we can calculate the δ -connected components by Algorithm 1: `find_δCC`. Then we would like to quantify how well-aligned the δ -connected components are as

a whole. For this we introduce and analyze affinities between connected components in this section. Intuitively, 'affinity' measures the proximity and alignment of δ -connected components.

We define the affinity matrix $A = (a_{ij})$ by

$$a_{ij} = \sup \left\{ \left(\frac{1}{\mu([g_i])} \int_{[g_i]} \varphi_t^\alpha \left(W_{[g_j]}^{(0)} \right) (h)^p dh \right)^{1/p}, \right. \\ \left. \left(\frac{1}{\mu([g_j])} \int_{[g_j]} \varphi_t^\alpha \left(W_{[g_i]}^{(0)} \right) (h)^p dh \right)^{1/p} \right\}, \quad (25)$$

with measure normalisation $\mu([g_i]) := \int_{[g_i]} dg$, and $\alpha > 1$, $t > 0$, $p \geq 1$ fixed, and where $W_{[g_i]}^{(0)}$ is given by

$$W_{[g_i]}^{(0)}(g) = \frac{D(g) \mathbb{1}_{[g_i]}(g)}{\sup_{h \in [g_i]} \{D(h)\}}, \quad g \in G, \quad (26)$$

where $D : G \rightarrow [0, 1]$ is a (non-zero) data term that can be freely chosen. In the experimental section, we chose the data term D to be the absolute value of the orientation score as defined in Eq. (1), rescaled to be between 0 and 1. Let us also recall that the flow operator φ_t^α was given by (14) and implemented by the morphological convolution in (12). We say a_{ij} denotes the affinity between δ -connected components $[g_i]$ and $[g_j]$.

In the computation of the affinity matrices we fix the parameter t , based on the choice of α and the compact set I , such that the dilation of any δ -connected component $[g_i]$ has non-zero values for all $h \in I$. How this can be achieved is explained in the next proposition.

Proposition 3 (Non-zero affinity between δ -connected components) *Let $\alpha > 1$ be given and*

$\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Let $I = \bigcup_{k=1}^K [g_k] \subset G$ compact be given, $[g_i] \neq [g_j]$ for all $i \neq j$. Then

$$\varphi_t^\alpha \left(W_{[g_i]}^{(0)} \right) (g) \neq 0 \text{ for all } g \in I, [g_i] \in \tilde{I}^\delta$$

$$\text{at least when } t > \left(\sup_{q_1, q_2 \in I} d(q_1, q_2) \right)^\alpha \beta^{1-\alpha}.$$

Proof We need to show that one dilation step with morphological kernel k_t^α on $W_{[g_i]}^{(0)}(g)$ results in all non-zero values for all $g \in I$. That means that we start with calculating:

$$0 \leq \varphi_t^\alpha \left(W_{[g_i]}^{(0)} \right) (g) = - \left(k_t^\alpha \square - W_{[g_i]}^{(0)} \right) (g) \\ = \sup_{h \in G} \left\{ W_{[g_i]}^{(0)}(h) - \frac{t}{\beta} \left(\frac{d(g, h)}{t} \right)^\beta \right\} \leq 1,$$

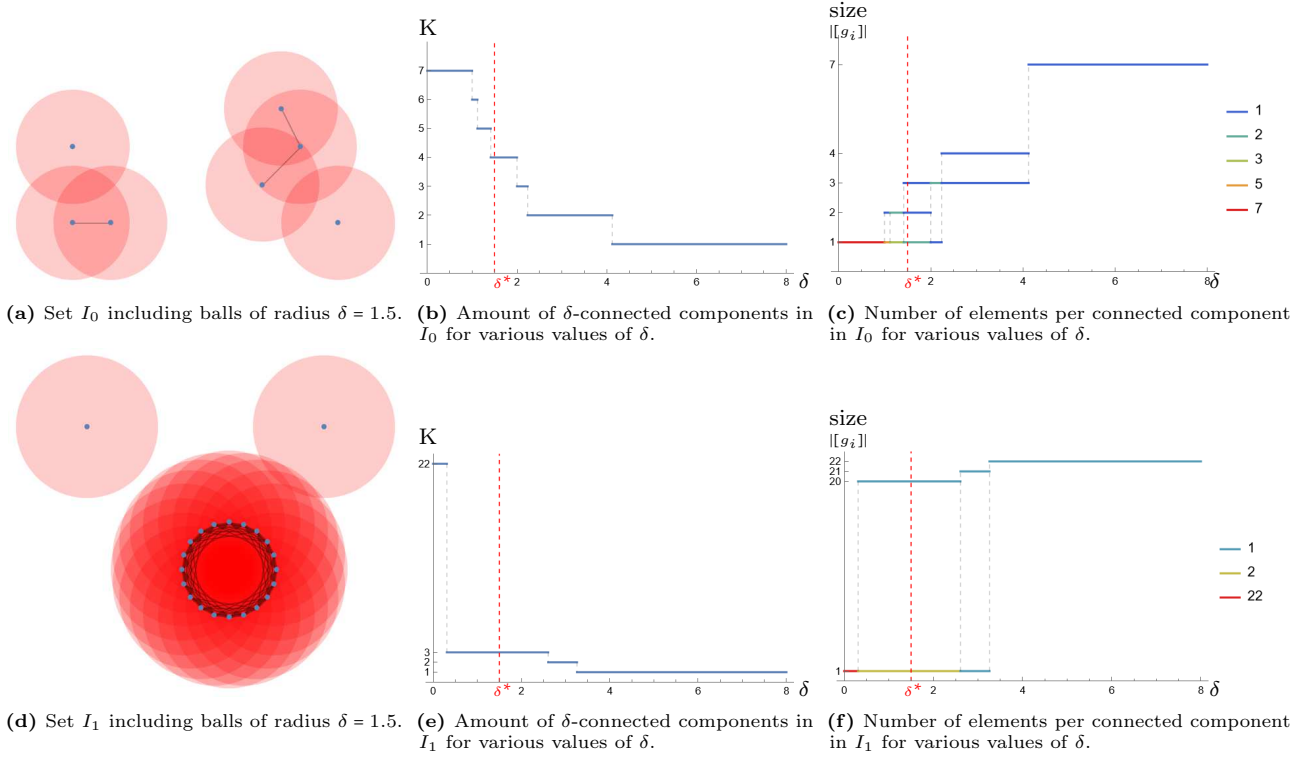


Fig. 8 Persistence plots for different sets $I \subset \mathbb{R}^2$. The longest lines (most values of δ) denote the most persistent connected components. Depending on the application, the optimal choice of δ is where the δ -connected components are the most stable/persistent (i.e. long lines in Fig. 8b,8c,8e,8f) while aiming for the smallest possible choice of δ (to be as distinctive as possible).

where we find the inequalities by Lemma 3. Then, we calculate the value of $t > 0$ such that

$$0 < \sup_{h \in I} \left\{ W_{[g_i]}^{(0)}(h) - \frac{t}{\beta} \left(\frac{d(g, h)}{t} \right)^\beta \right\} \leq 1.$$

By definition of $W_{[g_i]}^{(0)}$, for any element in the connected component $q \in [g_i]$ where the data does not vanish ($D(q) \neq 0$), we have

$$\begin{aligned} W_{[g_i]}^{(0)}(q) - \frac{t}{\beta} \left(\frac{d(g, q)}{t} \right)^\beta \\ \leq \sup_{h \in I} \left\{ W_{[g_i]}^{(0)}(h) - \frac{t}{\beta} \left(\frac{d(g, h)}{t} \right)^\beta \right\} \leq 1. \end{aligned}$$

We are sure that all affinities are nonzero when

$$W_{[g_i]}^{(0)}(q) - \frac{t}{\beta} \left(\frac{d(g, q)}{t} \right)^\beta > 0, \text{ for all } g \in I.$$

By definition of $W_{[g_i]}^{(0)}$ in (26), there exists a $q \in [g_i]$ such that $W_{[g_i]}^{(0)}(q) = 1$, so that for all $g \in I$

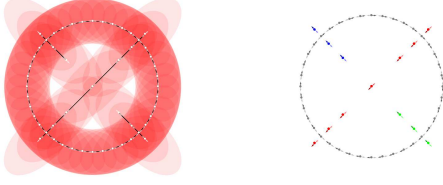
$$\begin{aligned} W_{[g_i]}^{(0)}(q) - \frac{t}{\beta} \left(\frac{d(g, q)}{t} \right)^\beta &= 1 - \frac{t}{\beta} \left(\frac{d(g, q)}{t} \right)^\beta > 0 \\ \Leftrightarrow t &> \left(\frac{d(g, q)^\beta}{\beta} \right)^{\frac{1}{\beta-1}} = (d(g, q) \beta^{-1/\beta})^\alpha = (d(g, q))^\alpha \beta^{1-\alpha} \end{aligned}$$

for all $g \in I$, and where $\frac{1}{\alpha} + \frac{1}{\beta} = 1$ with $\alpha, \beta \geq 1$ and the last estimate holds because the distance $d(g, q)$ is positive. Then, since the inequality has to hold for all $q \in I$ and should be independent of the origin, we have

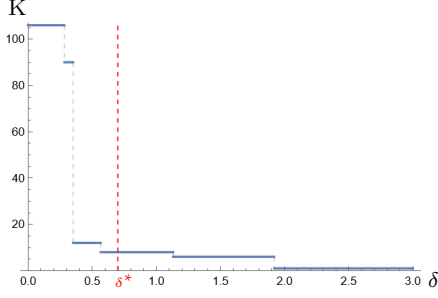
$$t > \left(\sup_{q_1, q_2 \in I} d(q_1, q_2) \right)^\alpha \beta^{1-\alpha} \geq \left(\sup_{\substack{q_1 \in I, \\ q_2 \in [g_i]}} d(q_1, q_2) \right)^\alpha \beta^{1-\alpha}. \quad \square$$

Algorithm 2: **find_affinity** describes the algorithm to find the affinity matrices between the δ -connected components of any set $I \subset G$.

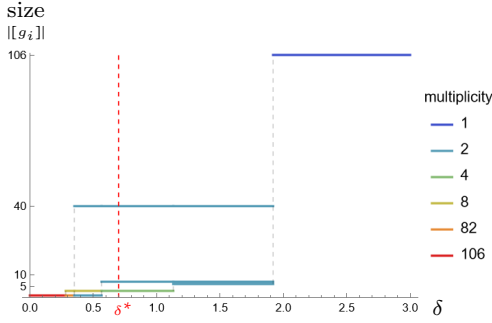
If the data $D(\cdot)$ is constant and non-zero in the connected components, we can give an upper and lower bound for the affinity of $[g_i]$ on $[g_j]$.



(a) Set $I_0 \subset SE(2)$ including balls of radius $\delta^* = 0.7$. (b) Different δ -connected components identified with $\delta^* = 0.7$.



(c) Amount of δ -connected components in I_0 for various values of δ .



(d) Number of elements per connected component in I_0 for various values of δ .

Fig. 9 Persistence plots for different sets $I \subset G = SE(2)$ using metric tensor field parameters $(w_1, w_2, w_3) = (1, 2, 2)$. The longest lines (most values of δ) denote the most persistent connected components. Depending on the application, the optimal choice of δ is where the δ -connected components are the most stable/persistent (i.e. long lines in Fig 9c,9d) while aiming for the smallest possible choice of δ (to be as distinctive as possible).

Proposition 4 Let the data $D : G \rightarrow [0, 1]$ be such that $D(g) = c_{[g]} \neq 0$ for all $g \in I$, where $c_{[g]}$ is constant in the same δ -connected component. Let $\alpha > 1$ be given and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Let $t > \sup_{q_1, q_2 \in I} d(q_1, q_2)^\alpha \beta^{1-\alpha}$ fixed.

Then the affinities on and off-diagonal satisfy respectively

$$a_{ii} = 1 \quad \text{and} \quad 0 < 1 - \frac{t}{\beta} \left(\frac{r}{t} \right)^\beta \leq a_{ij} < 1 - \frac{t}{\beta} \left(\frac{\delta}{t} \right)^\beta < 1$$

for all $[g_i] \neq [g_j] \subset I$, with diameter $r = \sup_{q_1, q_2 \in I} d(q_1, q_2)$.

Proof First, we note that since the data term $D(g) = c_{[g]} \neq 0$ for all $g \in I$, with $c_{[g]}$ constant in the same

Algorithm 2: find_affinity

Input: δ -connected components $[g_1], \dots, [g_K]$

determined by **find_δCC**, $I = \bigcup_{k=1}^K [g_k] \subset G$ region of interest for the connected components, kernel k_t^α

with $\alpha > 1$, $t > 0$, data $D : G \rightarrow [0, 1]$, parameter $p \geq 1$.

set $t > \left(\sup_{q_1, q_2 \in I} d(q_1, q_2) \right)^\alpha \beta^{1-\alpha}$ fixed

set $k = 1$

for $k \leq K$

initialize $W_{[g_k]}^{(0)}(g) := \frac{D(g) \cdot \mathbb{1}_{[g_k]}(g)}{\sup_{h \in [g_k]} \{D(h) \cdot \mathbb{1}_{[g_k]}(h)\}}$

update

$W_{[g_k]}^{(1)}(g) = \varphi_t^\alpha \left(W_{[g_k]}^{(0)} \right) (g) = - \left(k_t^\alpha \square - W_{[g_k]}^{(0)} \right) (g)$

for $1 \leq l \leq K$

set $\tilde{a}_{kl} := \frac{\left(\int_{[g_l]} \varphi_t^\alpha \left(W_{[g_k]}^{(0)} \right) (h)^p dh \right)^{1/p}}{\left(\int_{[g_l]} dh \right)^{1/p}}$

update $k := k + 1$

end

for $i, j = 1, \dots, K$

set $a_{ij} = \sup \{ \tilde{a}_{ij}, \tilde{a}_{ji} \}$

end

Output: Affinity Matrix $A = (a_{ij})$ for all

$i, j = 1, \dots, K$.

δ -connected component, we have

$$W_{[g_i]}^{(0)}(g) = \mathbb{1}_{[g_i]}(g).$$

Then, the morphological dilation of the initialization yields

$$W_{[g_i]}^{(1)}(g) := \varphi_t^\alpha \left(W_{[g_i]}^{(0)} \right) (g) = 1 - \inf_{h \in [g_i]} \frac{t}{\beta} \left(\frac{d(g, h)}{t} \right)^\beta \quad (27)$$

where we used the relation $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, with $\alpha > 1$. The distance satisfies two conditions:

1. If $g \in [g_i]$, then $\inf_{h \in [g_i]} d(g, h) = 0$.
2. If $g \notin [g_i]$, then $\delta < \inf_{h \in [g_i]} d(g, h) \leq r := \sup_{g, h \in I} d(g, h)$.

The upper bound follows from $[g_i] \subset I$. The lower bound follows from g not being part of the δ -connected component $[g_i]$. Recall that by Definition 5, we know that the distance between two δ -connected components is larger than δ .

Using this information on the distances, we can further simplify the expression in (27) to

$$W_{[g_i]}^{(1)}(g) = 1 \text{ if } g \in [g_i], \quad (28)$$

and

$$1 - \frac{t}{\beta} \left(\frac{r}{t} \right)^\beta \leq W_{[g_i]}^{(1)}(g) < 1 - \frac{t}{\beta} \left(\frac{\delta}{t} \right)^\beta \text{ if } g \notin [g_i], \quad (29)$$

using the distance estimates in items 1 and 2. Then, the affinity (25) equals $a_{ij} = \sup\{\tilde{a}_{ij}, \tilde{a}_{ji}\}$ with

$$\tilde{a}_{ij} := \left(\frac{1}{\mu([g_j])} \int_{[g_j]} \left(W_{[g_i]}^{(1)}(h) \right)^p dh \right)^{1/p}.$$

Clearly if $i = j$ then $a_{ij} = a_{ii} = 1$ by (28). For $i \neq j$, Eq. (29) provides an upper bound

$$\begin{aligned} \tilde{a}_{ij} &< \left(\frac{1}{\mu([g_j])} \int_{[g_j]} \left(1 - \frac{t}{\beta} \left(\frac{\delta}{t} \right)^\beta \right)^p dh \right)^{1/p} \\ &= 1 - \frac{t}{\beta} \left(\frac{\delta}{t} \right)^\beta < 1, \end{aligned} \quad (30)$$

and a lower bound

$$\begin{aligned} \tilde{a}_{ij} &\geq \left(\frac{1}{\mu([g_j])} \int_{[g_j]} \left(1 - \frac{t}{\beta} \left(\frac{r}{t} \right)^\beta \right)^p dh \right)^{1/p} \\ &= 1 - \frac{t}{\beta} \left(\frac{r}{t} \right)^\beta > 0, \end{aligned} \quad (31)$$

where $1 - \frac{t}{\beta} \left(\frac{r}{t} \right)^\beta > 0$ follows from the choice of t as was explained and proven in Prop. 3. Thereby one has

$$0 < 1 - \frac{t}{\beta} \left(\frac{r}{t} \right)^\beta \leq a_{ij} = \max\{\tilde{a}_{ij}, \tilde{a}_{ji}\} < 1 - \frac{t}{\beta} \left(\frac{\delta}{t} \right)^\beta < 1,$$

and the result follows. $\square$

8 Experiments

In the previous section, we introduced several concepts and algorithms, such as a formal algorithm to identify connected components (`find_δCC`) and an algorithm to calculate affinity matrices (`find_affinity`). Here, we will show the results of these algorithms applied to several images of the STAR-dataset [55, 1]. All experiments are performed on the Lie group $G = SE(2)$, and the Mathematica-notebooks are available via [8].

We start with discussing the experimental set-up in Sec. 8.1 and our approach to identifying the δ -connected components and the corresponding results in Sec. 8.2. Then, we discuss the results of the affinity matrices on some of the δ -connected component experiments in Sec. 8.3.

8.1 Experimental Set-Up

In the experiments, we will identify the δ -connected components (flowchart in Fig. 10) in retinal images from the STAR-dataset [55, 1] on the Lie group $G = SE(2)$.

The retinal images in this dataset are standard 2D-images. Therefore, we first explain how we prepare the images for processing in the Lie group $SE(2)$.

We consider an input image $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. We lift it to the space of positions and orientations $SE(2)$ (by creating an orientation score) to disentangle crossing structures. The orientation score $W_\phi f$ is calculated by a convolution with a rotating anisotropic wavelet ϕ and is given by Eq. (1). For ϕ , we use real-valued cake wavelets (see [21, 24]). We use 32 orientations for all experiments.

From the orientation score $W_\phi f$, we calculate a crossing-preserving vesselness $\mathcal{V}^{SE(2)}(W_\phi f)$, which is done as described in [49, Appendix D]. In all experiments, we use parameter settings $\xi = 1$, $\zeta = 1$, $s = \{1.5, 2, 2.5, 3\}$, $\sigma_{s,Ext} = 0$ and $\sigma_{a,Ext} = 0$.

From this crossing-preserving vesselness, we compute a cost function $\mathcal{C} = 1/(1 + \lambda \mathcal{V}^{SE(2)}(W_\phi f)^p)$ with parameters $\lambda > 0$ and $p > 0$. Then, we binarize the cost function using Otsu's method to identify the threshold value. Because the width of the structures varies within the same image, we choose to identify the centerline of the binarized structure, using the approach proposed in [34]. Since this is done in an isotropic way, we compensate for that by slightly dilating the structures in the $\mathcal{A}_1$ -direction; the principal direction of the local structure (see Eq. (7) for the explicit formula of $\mathcal{A}_1$). The support of the resulting binarized function on $SE(2)$ together with the lifted bifurcation information (for all orientations at bifurcation position equal to 1) is then used for the reference set I in the δ -connected component algorithm, more specifically, we identified $\mathbb{1}_I$.

Then, we use the connected component algorithm as introduced in `find_δCC` to identify all connected components in the lifted image. To visualize the output of the δ -connected component algorithm, we define the function $f^{CC} : SE(2) \rightarrow \mathbb{N}$ by

$$f^{CC}(x, y, \theta) = \begin{cases} i & \text{if } (x, y, \theta) \in [g_i] \\ 0 & \text{else.} \end{cases}$$

The δ -connected components are visualized by projecting them back onto $\mathbb{R}^2$ (taking per location the maximum over all orientations), i.e.

$$f_{\text{out}}^{CC}(x, y) = \max_{\theta \in [-\pi, \pi]} f^{CC}(x, y, \theta).$$

Moreover, we also calculate the affinity matrices (flowchart in Fig. 11). Because the algorithm used to compute them uses the δ -connected components as input. Therefore, one first follows the steps for computing the δ -connected components described in the previous paragraph. The projection step to visualize the results

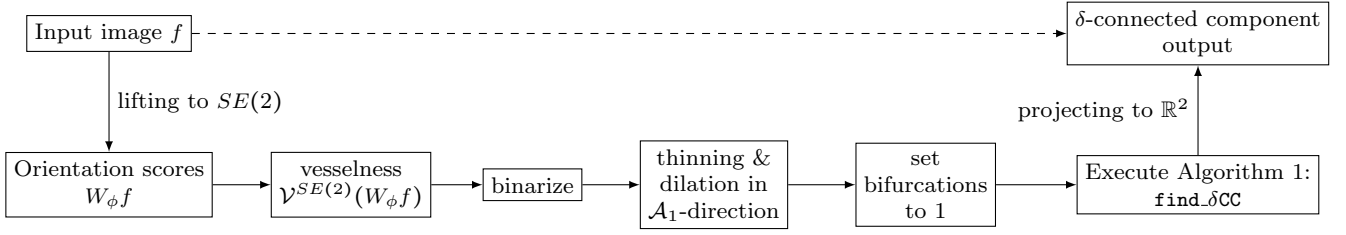


Fig. 10 Flowchart of the executed steps in the connected component experiments shown in Fig. 12-16.

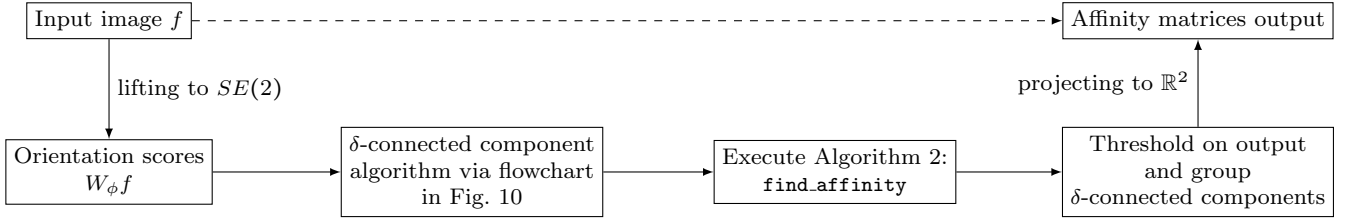


Fig. 11 Flowchart of the executed steps in the affinity matrices experiments shown in Fig. 14-16.

can be skipped as the identification of the affinity matrices also happens in $SE(2)$. Instead, we perform the affinity matrices algorithm as given in `find_affinity` to identify all affinity matrices. Here, we use the orientation score data $W_\phi f$ as the data term $D = |W_\phi f| / \sup |W_\phi f|$. To visualize the results of the output of the affinity matrices algorithm, we determine a threshold value of T and group the δ -connected components having an affinity higher than this threshold value T , i.e.

$$\tilde{C}_i = \bigcup_{j \in \mathbb{I}_i} [g_j],$$

with $\mathbb{I}_i$ the set containing the vertices that belong to the same connected component of the adjacency graph $A > T$, i.e. $\mathbb{I}_i = \{j \in \{1, \dots, K\} \mid \exists k_1, \dots, k_n; k_1 = i, k_n = j \text{ s.t. } a_{k_l k_{l+1}} > T\}$. Consequently, multiple δ -connected components will be grouped in the new visualization

$$f^{AM}(x, y, \theta) = \begin{cases} i & \text{if } (x, y, \theta) \in \tilde{C}_i \\ 0 & \text{else.} \end{cases}$$

To visualize the results, we again project the affinity matrices results back onto $\mathbb{R}^2$ by taking per location the maximum over all orientations

$$f_{\text{out}}^{AM}(x, y) = \max_{\theta \in [-\pi, \pi)} f^{AM}(x, y, \theta)$$

The additional parameter settings used to process all images can be found in Table 1. Additionally, the *Mathematica*-notebooks are publicly available via [8].

8.2 δ -Connected Components

We identify the δ -connected components using the method described in Sec. 8.1 and visualized in the

flowchart in Fig. 10. We identify the components for five different images from the STAR-dataset [55, 1], and show the results in Fig. 12-16.

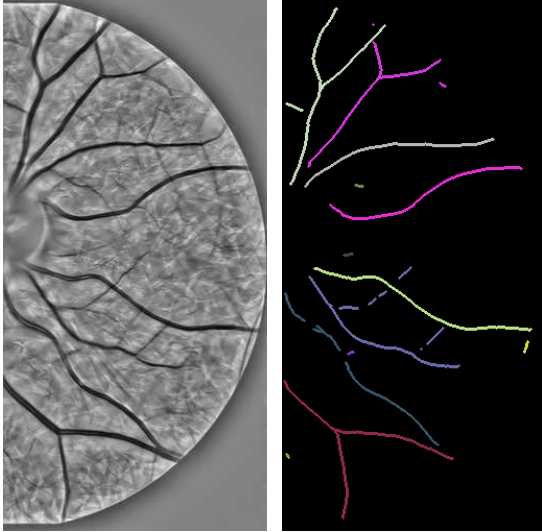
We start with the retinal image STAR48 in Fig. 12. Due to the binarization, we see that not all vascular structures are nicely connected. However, the δ -connected component algorithm can compensate for interrupted vascular structures by choosing the right threshold value for δ , and the correct distance parameters w_1 , w_2 and w_3 . We see that the algorithm has correctly grouped most of the segments that belong to the same vessel in the underlying image.

The δ -connected components of the retinal image STAR13 are visualized in Fig. 13. The vascular structure in this image does not contain a lot of bifurcations or crossing structures. The algorithm is good at identifying components that correspond to the structures in the underlying image. We also see that the choice of the threshold in the binarization step (found with Otsu's method) has a big influence on the input in the δ -connected component algorithm; many vascular structures are interrupted, and therefore the algorithm needs to compensate for that as well. The results are good as long as the gaps are smaller than a certain threshold value. If the gaps are bigger, the δ -connected component algorithm is not able to identify them as belonging to the same vascular structure.

In Fig. 14, we applied the algorithm to STAR34. The results are shown in Fig. 14b. The vessels in the image are more sinuous. We see that this causes some challenges if the vascular structure gets interrupted in the binarization. This is because we need to choose the metric parameters in the δ -connected component algorithm. We chose these parameters, cf. column 7&8 in

Figure	Image	Cost		Dilation		δ -connected Components		Affinity Matrices			
		λ	p	(w_1, w_2, w_3)	α	$(\tilde{w}_1, \tilde{w}_2, \tilde{w}_3), w_i = k\tilde{w}_i$	α	(w_1, w_2, w_3)	α	T	p
Fig. 12	STAR48	100	3	(0.2, 1.5, 50)	1.3	(0.06, 0.7, 2)	1	-	-	-	-
Fig. 13	STAR13	50	3.5	(0.1, 1.5, 50)	1.3	(0.1, 0.7, 4)	1	-	-	-	-
Fig. 14	STAR34	200	3	(0.2, 1.5, 50)	1.3	(0.2, 1, 2)	1	(0.1, 4, 4)	2	0.99957	2
Fig. 15	STAR37	100	3	(0.2, 1.5, 50)	1.3	(0.08, 0.7, 4)	1	(0.5, 2, 0.5)	2	0.9985	2
Fig. 16	STAR38	100	3	(0.2, 1.5, 50)	1.3	(0.2, 1, 2)	1	(0.5, 2, 0.1)	2	0.993	2

Table 1 Parameter settings for preprocessing of data and calculation of δ -connected components and affinity matrices. The parameter k in the δ -connected components is given by $k = (13/3)^{6/13}$.



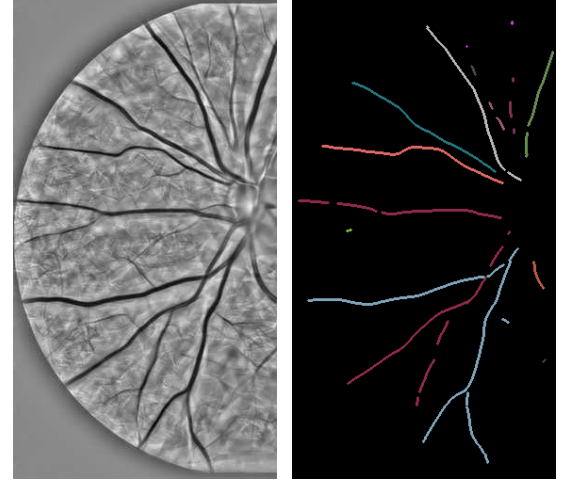
(a) Underlying Image (STAR48). (b) Connected components calculated by the flowchart in Fig. 10 and parameter settings in Table 1.

Fig. 12 The output of the δ -connected component algorithm executed on a thinning of the calculated vesselness of the given image. The parameter settings can be found in Table 1.

Table 1, such that forward movement is allowed, but sideways movement and changing orientation are not, to avoid crossing structures being connected. However, one does need to change orientation and forward movement at interrupted tortuous structures. Consequently, the δ -connected component algorithm has trouble connecting the vessel segments that are interrupted at highly tortuous parts.

In Fig. 15, we performed the algorithm on the retinal image STAR37. The algorithm groups the vessel segments correctly using the chosen parameters. However, it does not identify full vascular trees, but only parts of it. This is due the relatively large spatial gaps between vessel parts, often also changing orientation.

Last, we look at STAR38 in Fig. 16b. The δ -connected component algorithm can identify the large vascular structures correctly. Some small vessels are not correctly connected to the main vessel at the



(a) Underlying Image (STAR13). (b) Connected components calculated by the flowchart in Fig. 10 and parameter settings in Table 1.

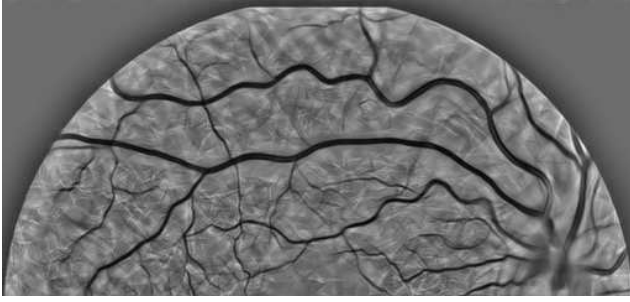
Fig. 13 The output of the δ -connected component algorithm executed on a thinning of the calculated vesselness of the given image. The parameter settings can be found in Table 1.

bifurcations, but all vessel segments are correctly connected.

We conclude that this δ -connected component algorithm allows us to identify parts of vascular trees. Additionally, it can differentiate between different structures when they are crossing. However, when the vascular structure is interrupted at a very tortuous part, the algorithm can have difficulties connecting the right parts, depending on the chosen metric parameters.

8.3 Affinity Matrices

In the previous section, we discussed the experimental results of the δ -connected component algorithm. We saw that the algorithm can identify parts of vessels belonging to the same vascular structure, but often does not identify full vascular trees. We will use the affinity matrices to group different δ -connected components that are most likely to belong to the same vascular structure based on their local alignment and proximity. We have performed the affinity matrices algorithm on



(a) Underlying Image (STAR34).



(b) Connected components calculated by the flowchart in Fig. 10 and parameter settings in Table 1.

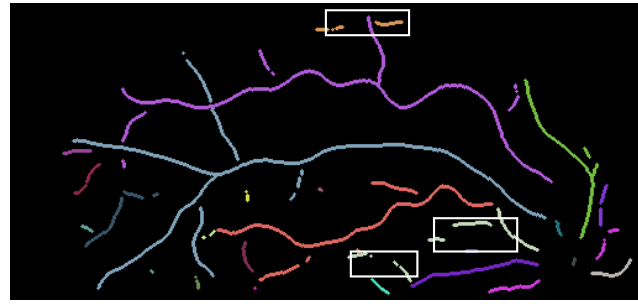
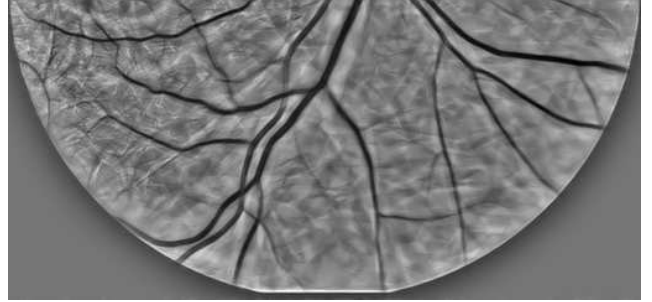
(c) Affinity Matrices calculated on the connected components in Fig. 14b. Newly grouped δ -connected components are indicated by a white box.

Fig. 14 The output of the δ -connected component algorithm and affinity matrices algorithm executed on a thinning of the calculated vesselness of the given image. The parameter settings can be found in Table 1.

the δ -connected component results of Fig. 14-16. We use the orientation score data $|W_{\phi}f|$ as the data term in the affinity matrices algorithm given in `find_affinity`.

First, we calculated the affinity matrices on the δ -connected component output in Fig. 14c. Thresholding on the affinities allows us to group different components, which results in slightly more complete vascular trees. We cannot pick parameter settings such that all parts that belong to the same vessel are connected without connecting them to other structures. However, the output in Fig. 14c is a more complete vascular tree classification than the δ -connected component output in Fig. 14b. We have marked the correctly grouped vascular structures with a white box.



(a) Underlying Image (STAR37).



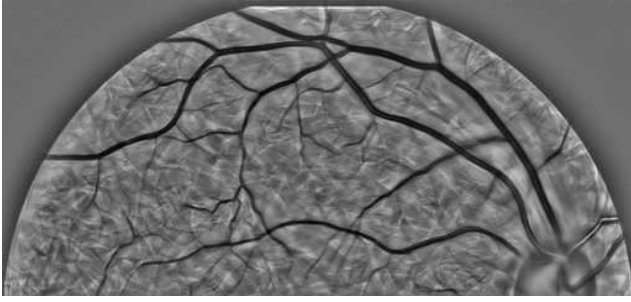
(b) Connected components calculated by the flowchart in Fig. 10 and parameter settings in Table 1.

(c) Affinity Matrices calculated on the connected components in Fig. 15b. Newly grouped δ -connected components are indicated by a white box.

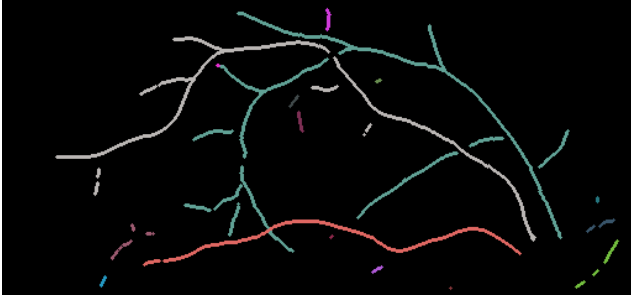
Fig. 15 The output of the δ -connected component algorithm and affinity matrices algorithm executed on a thinning of the calculated vesselness of the given image. The parameter settings can be found in Table 1.

In Fig. 15c, we have also applied the affinity matrices algorithm to the output of the δ -connected component algorithm in Fig. 15b. The thresholded output of the affinity matrices have correctly grouped different δ -connected components that belong to the same vascular structure. The newly grouped structures have been indicated with a white box. We note that the output is much closer to the real underlying vascular structure. Still, some cases have not been grouped correctly (due to the imperfect data term D).

Last, we have calculated the affinity matrices for the retinal image of STAR38 (cf. Fig. 16). The δ -connected component output in Fig. 16b has been used as input for `find_affinity`. The output of this algorithm has



(a) Underlying Image (STAR38).



(b) Connected components calculated by the flowchart in Fig. 10 and parameter settings in Table 1.

(c) Affinity Matrices calculated on the connected components in Fig. 16b. Newly grouped δ -connected components are indicated by a white box.

Fig. 16 The output of the δ -connected component algorithm and affinity matrices algorithm executed on a thinning of the calculated vesselness of the given image. The parameter settings can be found in Table 1.

been thresholded and the δ -connected components having an affinity higher than the threshold are grouped. The output is visualized in Fig. 16c. Here, full vascular trees are identified completely and correctly. The white boxes indicate the changes compared to the δ -connected component output.

We found that the thresholding on the affinity matrices improves the results; δ -connected components are grouped such that more complete vascular trees are identified. It is important to note that the choice of the metric parameters of Table 1 and threshold value T are important to the output.

9 Conclusion and Future Work

In this article, we have introduced a way to identify so-called δ -connected components on a Lie group G . These components consist of sets (of points) with a maximum distance δ from each other. First, we introduced the general idea behind the algorithm and some theoretical background on morphological dilations. Then, we connected this theory to the general algorithm, resulting in the δ -connected component algorithm, stated in `find_δCC`.

We studied the convergence of the δ -connected component algorithm in Theorem 1. We proved that the algorithm always finishes in a finite number of iteration steps. Subsequently, we discussed the choice of the parameter δ . We suggested using persistence homology arguments to choose the optimal value for δ and illustrated it with a few examples.

Once the δ -connected components are introduced and calculated, we aim to determine a hierarchy between the different components. Therefore, we introduce the concept of affinity matrices. They describe a way to group components based on their proximity and local alignment. To account for subtleties in the data, we include a data term in the initialization of the affinity matrices. The full algorithm can be found in `find_affinity`.

To show the performance of the δ -connected component algorithm and the affinity matrices algorithm, we have tested both algorithms on several 2D images of the retina. All experiments show that the δ -connected component algorithm can differentiate between different structures at crossings. Additionally, the δ -connected components groups well-aligned structures, resulting in more complete vessels in the output. However, it cannot always identify full vascular trees. Therefore, we calculated the affinity matrices for several results of the δ -connected component algorithm. We see that this leads to more complete vascular trees, where different δ -connected components are grouped that belong to the same vascular tree.

The algorithms are challenged by gaps at highly bending parts in the vascular structure. The algorithm cannot connect these structures without connecting different vascular trees at crossings. Large spatial gaps also form challenges: choosing δ too high results in connecting vessel segments belonging to different vascular trees.

For future work, it would be interesting to train the metric tensor weights (w_1, w_2, w_3) via a PDE-G-CNN on $G = SE(2)$, cf. [46], to improve results further, and to see if optimal transport on $SE(2)$ [9] would be a viable

alternative to determine the affinity between different δ -connected components.

Acknowledgments

We gratefully acknowledge Gijs Bellaard for the code underlying Figure 5.

A Logarithmic Norm Approximation

Calculating the exact distances in a Lie group G is expensive. Therefore, we approximate the norm with a so-called ‘logarithmic norm approximation’. This section elaborates on how this is done, as explained in [46, 7].

Let us consider any Lie group G , equipped with a left-invariant metric tensor field $\mathcal{G}$, Lie group elements $g \in G$ and identity element $e \in G$. We consider an exponential curve $\gamma : [0, 1] \rightarrow G$, connecting e to g where $\|\dot{\gamma}(0)\|$ is minimal, i.e. $\gamma(0) = e$, $\gamma(1) = g$ with the property $\gamma(s+t) = \gamma(s) \cdot \gamma(t)$. We want to be able to calculate distances between different elements in the Lie group, e.g. the distance between e and g , denoted by $d_G(e, g)$. This is quite an expensive operation. Therefore, we approximate the exact distance with the distance of a logarithm, as explained in this section:

$$d_G(e, g) \leq \text{Length}_{\mathcal{G}}(\gamma) \quad (32.1)$$

$$= \int_0^1 \|\dot{\gamma}(t)\|_{\mathcal{G}} dt \quad (32.2)$$

$$= \int_0^1 \|(L_{\dot{\gamma}(t)})_* \dot{\gamma}(0)\|_{\mathcal{G}} dt \quad (32.3)$$

$$= \int_0^1 \|\dot{\gamma}(0)\|_{\mathcal{G}} dt \quad (32.4)$$

$$= \|\dot{\gamma}(0)\|_{\mathcal{G}} \quad (32.5)$$

$$= \|\log g\|_{\mathcal{G}}, \quad (32.6)$$

where the equality in (32.6) holds because by definition $\log g := \dot{\gamma}(0) \in T_e(G)$, (32.5) holds because the metric tensor field is left-invariant, and where (32.4) holds because

$$\begin{aligned} \dot{\gamma}(s) &= \frac{d}{dt} \gamma(s+t) \Big|_{t=0} = \frac{d}{dt} (\gamma(s) \cdot \gamma(t)) \Big|_{t=0} \\ &= \left(\frac{d}{dt} L_{\gamma(s)} \gamma(t) \right) \Big|_{t=0} = (L_{\gamma(s)})_* \dot{\gamma}(0), \end{aligned}$$

with $L_g : G \rightarrow G$ the left action defined by $L_g h = g \cdot h$ for all $h \in G$, and where the last equality follows directly by the definition of the push-forward. Hence, we have found an upper bound for the distance from $e \in G$ to a point $g \in G$ in any Lie group G .

In this article, we consider two Lie groups; the special Euclidean group $G = SE(2)$ and the special orthogonal group $G = SO(3)$. For both, we will calculate $\log g$ for $g \in G$. Let $R_G(g)$ denote the matrix representation of the element $g \in G$, such that $R_G(g_1)R_G(g_2) = R_G(g_1g_2)$. Then, the associated Lie algebra is denoted by $\text{span}(A_i)$, where

$$A_i = \frac{\partial R_G(g)}{\partial g^i} \Big|_e, \quad (33)$$

with $e \in G$ the identity element and where $\{A_i\}$ denotes the left-invariant frame, with dual $\{\omega^i\}$ such that $\langle \omega^i, A_j \rangle = \delta_j^i$. Then,

$$\sum_{i=1}^n g^i A_i \leftrightarrow g \in G.$$

A.1 Logarithmic Norm Approximation in $SE(2)$

We start with calculating $\log g$ for the special Euclidean group $G = SE(2)$, where $g = (x, y, \theta) \in G$. The generating matrix $R_{SE(2)}$ is given by

$$R_{SE(2)}(x, y, \theta) = \begin{pmatrix} \cos \theta & -\sin \theta & x \\ \sin \theta & \cos \theta & y \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, the associated Lie algebra $\mathfrak{se}(2) = \text{span}(A_1, A_2, A_3)$ is spanned by

$$A_1 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

calculated by (33) using identity element $e = (0, 0, 0)$. Hence,

$$\begin{aligned} SE(2) \ni (x, y, \theta) &\leftrightarrow \\ \begin{pmatrix} \cos \theta & -\sin \theta & x \\ \sin \theta & \cos \theta & y \\ 0 & 0 & 1 \end{pmatrix} &= \exp(xA_1) \exp(yA_2) \exp(\theta A_3) \\ &= \exp(c^1 A_1 + c^2 A_2 + c^3 A_3) \\ &= \begin{pmatrix} \cos c^3 & -\sin c^3 & \left(c^1 \cos \frac{c^3}{2} - c^2 \sin \frac{c^3}{2} \right) \text{sinc} \frac{c^3}{2} \\ \sin c^3 & \cos c^3 & \left(c^1 \sin \frac{c^3}{2} + c^2 \cos \frac{c^3}{2} \right) \text{sinc} \frac{c^3}{2} \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Hence, one has the following relations:

$$\begin{cases} x = \left(c^1 \cos \frac{c^3}{2} - c^2 \sin \frac{c^3}{2} \right) \text{sinc} \frac{c^3}{2} \\ y = \left(c^1 \sin \frac{c^3}{2} + c^2 \cos \frac{c^3}{2} \right) \text{sinc} \frac{c^3}{2} \\ \theta = c^3 \pmod{2\pi} \end{cases}, \quad \begin{cases} c^1 = \frac{x \cos \frac{\theta}{2} + y \sin \frac{\theta}{2}}{\text{sinc} \frac{\theta}{2}} \\ c^2 = \frac{-x \sin \frac{\theta}{2} + y \cos \frac{\theta}{2}}{\text{sinc} \frac{\theta}{2}} \\ c^3 = \theta, \end{cases} \quad (34)$$

where $c^1 \partial_x|_e + c^2 \partial_y|_e + c^3 \partial_\theta|_e \in T_e(SE(2))$ represent the logarithmic coordinates.

A.2 Logarithmic Norm Approximation in $SO(3)$

We follow the same approach as above for the special orthogonal group $G = SO(3)$, with group elements $(\alpha, \beta, \varphi) \in SO(3)$. The generating matrix $R_{SO(3)}$ is given by

$$R_{SO(3)}(\alpha, \beta, \varphi) = \begin{pmatrix} c\alpha c\beta & s\beta c\varphi - s\alpha c\beta s\varphi & -s\alpha c\beta c\varphi - s\beta s\varphi \\ -c\alpha s\beta & s\alpha s\beta s\varphi + c\beta c\varphi & s\alpha s\beta c\varphi - c\beta s\varphi \\ s\alpha & c\alpha s\varphi & c\alpha c\varphi \end{pmatrix},$$

where $c\alpha = \cos \alpha$, $c\beta = \cos \beta$, $c\varphi = \cos \varphi$, $s\alpha = \sin \alpha$, $s\beta = \sin \beta$ and $s\varphi = \sin \varphi$. The associated Lie algebra $\mathfrak{so}(3)$ is spanned by

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

calculated by (33) using identity element $e = (0, 0, 0)$. Hence,

$$\begin{aligned} SO(3) \ni (\alpha, \beta, \varphi) &\leftrightarrow R_{SO(3)}(\alpha, \beta, \varphi) \\ &= \exp(-\beta A_3) \exp(-\alpha A_2) \exp(\varphi A_1) \\ &= \exp(c^1 A_1 + c^2 A_2 + c^3 A_3). \end{aligned} \quad (35)$$

We need to find the relation between (α, β, φ) and (c^1, c^2, c^3) such that Eq. (35) holds. To this end we note that we can interpret the rotation described by $R_{SO(3)}$ as a counterclockwise rotation around an axis $\mathbf{a}$ with an angle ϕ . Both can be identified from the matrix representation of the group elements of $SO(3)$ given by $R_{SO(3)}$. The rotation axis, notated by the vector $\mathbf{a}$, is the eigenvector corresponding to the eigenvalue 1. If the rotation matrix $R_{SO(3)}$ is not symmetric, the eigenvector $\mathbf{a}$ and the corresponding rotation angle ϕ are given by

$$\mathbf{a} = \frac{1}{2 \sin \phi} \begin{pmatrix} R_{32} - R_{23} \\ R_{13} - R_{31} \\ R_{21} - R_{12} \end{pmatrix}, \quad \phi = \arccos \left(\frac{\text{Tr}(R_{SO(3)}) - 1}{2} \right). \quad (36)$$

The expression for the rotation angle ϕ is readily verified, as for a rotation matrix $\tilde{R}$ rotating a vector with an angle $\tilde{\phi}$ around the z -axis. Then, using the conjugation invariance of the trace operator and the fact that there exists a matrix P such that $P^{-1} \tilde{R} P = R_{SO(3)}$, Eq. (36) holds for the rotation with angle ϕ around the axis $\mathbf{a}$.

Then, one has, by Rodrigues' rotation formula and the well-known Lie group exp formula in $SO(3)$ that

$$\begin{aligned} \exp(\phi(a^1 A_1 + a^2 A_2 + a^3 A_3))(\mathbf{x}) &= \exp(\phi \cdot \mathbf{a})(\mathbf{x}) \\ &= \mathbf{x} \cos \phi + (\mathbf{a} \times \mathbf{x}) \sin \phi + \mathbf{a} (\mathbf{a} \cdot \mathbf{x}) (1 - \cos \phi) = R_{SO(3)}(\mathbf{x}). \end{aligned}$$

Therefore, we can now conclude that the logarithmic coordinates in $SO(3)$ are given by

$$\begin{aligned} (c^1, c^2, c^3)^T &= \phi \cdot \mathbf{a} = \frac{1}{2 \text{sinc } \phi} \begin{pmatrix} R_{32} - R_{23} \\ R_{13} - R_{31} \\ R_{21} - R_{12} \end{pmatrix} \\ \Leftrightarrow \Omega_{\mathbf{a}} &= a^1 A_1 + a^2 A_2 + a^3 A_3 = \frac{1}{2 \text{sinc } \phi} (R - R^T), \end{aligned} \quad (37)$$

where $\Omega_{\mathbf{a}} \in \mathfrak{so}(3) = T_e(SO(3))$ is the antisymmetric matrix associated to $\Omega_{\mathbf{a}}(\mathbf{x}) = \mathbf{a} \times \mathbf{x}$. The formula above has an (innocent) removable singularity at $\phi = \pm\pi$ where the matrix R becomes symmetric (with 1,-1,-1 on the diagonal after diagonalisation). At $\phi = 0$ we have $\text{sinc}(0) = 1$.

B Reflectional Symmetries

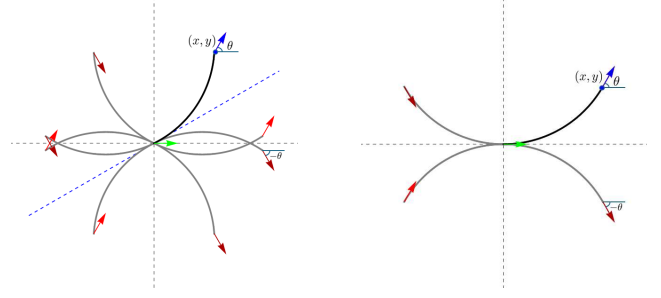
Note that by (8), we have left-invariance $d_{\mathcal{G}}(g, h) = d_{\mathcal{G}}(h^{-1}g, e) \approx \|\log h^{-1}g\|_{\mathcal{G}}$. Furthermore, this distance's (reflectional) symmetries carry over to the δ -connected component algorithm as we will show next.

Lemma 4 *Set $q = h^{-1}g$. Then $\log q = \sum_{i=1}^n c^i(q) \mathcal{A}_i|_e \in T_e(G)$. Now its norm $\|\log q\|_{\mathcal{G}}$ is invariant under $c^i \mapsto \pm c^i$, $i = 1, \dots, n$, which gives rise to 2^n symmetries.*

Proof We recall from App. A that $\|\log q\|_{\mathcal{G}} = \sqrt{\sum_{i=1}^n w_i (c^i)^2}$. Then, its invariance under $c^i \mapsto \pm c^i$ follows immediately. $\square$

See Fig. 17 for what this means on $SE(2)$ where $n = \dim SE(2) = 3$. For further details, see [7, 37].

Corollary 2 *If the connected component algorithm is applied to a point cloud $\{q_1, \dots, q_n\} \subset SE(2)$, then it produces the same reflected connected components as when it is applied to $\{\tilde{\varepsilon}^i(q_1), \dots, \tilde{\varepsilon}^i(q_n)\}$, where $\tilde{\varepsilon}^i(q) = \exp(\varepsilon^i \log(q))$, $i = 0, \dots, 2^n - 1$, denotes the symmetry given by the reflection of the logarithmic coordinates as denoted in [7, Table 5].*



(a) Reflectional symmetries with c^2 unconstrained (general Riemannian case [7]).

(b) Reflectional symmetries with $c^2 = 0$ (sub-Riemannian case [37]).

Fig. 17 Reflectional symmetries of the blue point $(x, y, \theta) \equiv (x, y, R_\theta) \in SE(2)$, reflected in the three reflection axes. Depending on the axis, the orientation is updated to $-\theta$ or stays the same. The blue reflection axis is only for the endpoints of the geodesics.

Proof The statement follows immediately from Lemma 4 applied to $\|\log \tilde{\varepsilon}^i(g)\|_{\mathcal{G}}$:

$$\|\log \tilde{\varepsilon}^i(q)\|_{\mathcal{G}} = \|\varepsilon^i(\log q)\|_{\mathcal{G}} \stackrel{\text{Lemma 4}}{=} \|\log q\|_{\mathcal{G}}.$$

Thus, the distances between elements are the same, and consequently, the δ -connected components are also the same. $\square$

The precise formulas of these reflections can be found in [37, 7]. For a quick intuition, see Fig. 17. Essentially, they are found by line reflections in the half-angle axis, the x - and y -axes, possibly negating the angle. Intuitively, they are depicted in Fig. 17a.

C Proofs of different Lemma's and Propositions

C.1 Multiple consecutive morphological dilations

We will show what happens when multiple consecutive morphological dilations are executed in Lemma 5. In the statement and proof, we use the morphological delta which we define first.

Definition 14 (Morphological delta) We define the morphological delta $\delta_e^M : G \rightarrow \mathbb{R} \cup \{\infty\}$ as

$$\delta_e^M(g) = \begin{cases} 0 & \text{if } g = e \\ \infty & \text{else.} \end{cases} \quad (38)$$

Lemma 5 *Let $1 \leq \alpha \leq \infty$. Consider the morphological kernel k_t^α . Let $g \in G$ be given. When defining¹ $k_0^\alpha = \delta_e^M$ for $\alpha < \infty$, and $k_0^\infty = d(\cdot, e)$, we have:*

$$\forall t, s \geq 0 : (k_t^\alpha \square k_s^\alpha)(g) = k_{t+s}^\alpha(g) = (\kappa_t^\alpha \square_{\mathbb{R}} \kappa_s^\alpha)(d(g, e)),$$

where $\kappa_t^\alpha : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, is defined as $\kappa_t^\alpha(d(g, e)) = k_t^\alpha(g)$.

Proof For $t = 0$, $\alpha < \infty$, one has $k_t^\alpha = \delta_e^M$, and $\delta_e^M \square f = f$, so let us examine $1 \leq \alpha < \infty$, $t, s > 0$. Let $\kappa_t^\alpha(r) := \kappa_t^\alpha(|r|)$ for all

¹ These pointwise limits arise from (13).

$r \in \mathbb{R}$. Then, one can write

$$(k_t^\alpha \square k_s^\alpha)(g) = \inf_{h \in G} k_t^\alpha(h^{-1}g) + k_s^\alpha(h) \quad (39.1)$$

$$= \inf_{h \in G} \kappa_t^\alpha(d(h^{-1}g, e)) + \kappa_s^\alpha(d(h, e)) \quad (39.2)$$

$$= \inf_{h \in \gamma_{e,g}^{\min}} \kappa_t^\alpha(d(g, e) - d(h, e)) + \kappa_s^\alpha(d(h, e)) \quad (39.3)$$

$$= \inf_{0 \leq r \leq d(g, e)} \kappa_t^\alpha(d(g, e) - r) + \kappa_s^\alpha(r) \quad (39.4)$$

$$= \inf_{r \in \mathbb{R}} \kappa_t^\alpha(d(g, e) - r) + \kappa_s^\alpha(r) \quad (39.5)$$

$$= (\kappa_t^\alpha \square_{\mathbb{R}} \kappa_s^\alpha)(d(g, e)) \quad (39.6)$$

$$= \kappa_{t+s}^\alpha(d(g, e)) = k_{t+s}^\alpha(g), \quad (39.7)$$

where the third equality (39.3) holds because one has $\geq$ due to the triangle inequality and κ_t^α being monotonous. The $\leq$ holds because $\gamma_{e,g}^{\min} \subset G$, where $\gamma_{e,g}^{\min}$ denotes the minimizing geodesic connecting e and g . To show why the fifth equality (39.5) holds, we define $G(r) := \kappa_t^\alpha(d(g, e) - r) + \kappa_s^\alpha(r) = \frac{t}{\beta} \left(\frac{d-r}{t} \right)^\beta + \frac{s}{\beta} \left(\frac{r}{s} \right)^\beta$, where we omit the dependence on the elements g and e in the notation for the distance $d = d(g, e)$. Then

$$G'(r) = -\left(\frac{d-r}{t} \right)^{\beta-1} + \left(\frac{r}{s} \right)^{\beta-1} = 0 \iff r = \frac{s}{t+s}d.$$

Since $s, t > 0$, r attains a global minimum between 0 and distance $d = d(g, e)$. The sixth equality (39.6) is by definition of the morphological convolutions on $\mathbb{R}$. The seventh equality (39.7) follows by well-known results for morphological PDEs on $\mathbb{R}$ [44].

Note that the case $\alpha = \infty$, $t = 0$, $s > 0$ is the same as $\alpha = \infty$, $t, s > 0$ since $k_0^\infty(g) = k_t^\infty(g)$ for all $t \geq 0$, and where

$$(k_t^\infty \square k_s^\infty)(g) = \inf_{h \in G} d(h^{-1}g, e) + d(h, e) = d(g, e). \quad \square$$

C.2 Proof of Proposition 2

Proof First of all, note that $m_\delta(g, g_0) \leq n$ implies that $d(g, g_0) \leq n\delta$ by applying the triangular inequality to the sequence of intermediate points in Def. 4.

Then, note that the inequalities follow directly from Lemma 3, Eq. (16). We will prove the rest of this statement by induction w.r.t. n . Without loss of generality in view of left-invariance, we prove the statement for $g_0 = e$. This is analogous to doing a roto-translation with g_0^{-1} of the given binary map $\mathbb{1}_I$, and afterwards pushing forward the output $U_e(n, \cdot)$ to $U_{g_0}(n, \cdot)$. First, we show that the statement is true for the first step of the algorithm. We recall Eq. (21), which stated

$$k_t^1(g) := \begin{cases} 0 & \text{if } d(g, e) < t \\ \infty & \text{else.} \end{cases}$$

We can easily express the first step of the connected component algorithm $n = 1$ by calculating $U_e(1, \cdot)$:

$$\begin{aligned} U_e(1, g) &= \left(-\left(k_\delta^1 \square -U_e(0, \cdot) \right)(g) \right) \mathbb{1}_I(g) \\ &= \begin{cases} \sup_{h \in G} \{U_e(0, g) - k_\delta^1(h^{-1}g)\} & \text{if } g \in I, \\ 0 & \text{else;} \end{cases} \\ &= \begin{cases} 1 & \text{if } g \in I \text{ and } d(e, g) < \delta \\ 0 & \text{else,} \end{cases} \end{aligned}$$

where $g \in I$ and $d(e, g) < \delta$ is equivalent to $g \stackrel{\delta}{\sim} e$ and $m_\delta(e, g) \leq 1$. Note that $g \stackrel{\delta}{\sim} e$ automatically implies $g \in I$.

Next, we assume that the statement is true for a given $n \in \mathbb{N}$, meaning that one has

$$\begin{aligned} U_e(n, g) &= \begin{cases} 1 & \text{if } g \stackrel{\delta}{\sim} e \wedge m_\delta(e, g) \leq n \\ 0 & \text{else.} \end{cases} \\ &= \mathbb{1}_{I_{n,\delta}^e}(g), \end{aligned}$$

where

$$I_{n,\delta}^e = [e] \cap \tilde{B}(e; n),$$

where $[e]$ denotes the set of points that are δ -connected to e and where $\tilde{B}(e; n)$ is defined by

$$\tilde{B}(e; n) := \{g \in I \mid m_\delta(e, g) \leq n\}. \quad (40)$$

Then, we aim to show that the statement is also true for iteration $n + 1$. That means that one investigates whether one can write $U_e(n + 1, \cdot)$ in the same form as $U_e(n, \cdot)$. Let $g \in I$, then

$$\begin{aligned} U_e(n + 1, g) &= -\left(-k_\delta^1 \square -U_e(n, \cdot) \right)(g) \\ &= -\left(-k_\delta^1 \square -\mathbb{1}_{I_{n,\delta}^e} \right)(g) \\ &= \sup_{h \in G} \left\{ \mathbb{1}_{I_{n,\delta}^e}(h) - k_\delta^1(h^{-1}g) \right\}. \end{aligned} \quad (41)$$

Then, in order to reach the supremum in (41), we can simplify the expression to

$$U_e(n + 1, g) = \begin{cases} 1 & \text{if } g \in I_{n,\delta}^e \text{ or if } \exists h \in I_{n,\delta}^e \text{ s.t. } d(h, g) < \delta \\ 0 & \text{else.} \end{cases}$$

We will reformulate the condition for $U_e(n + 1, g)$ to equal 1. First, it is important to note that $I_{n,\delta}^e \subseteq I_{n+1,\delta}^e$. Let us assume $g \notin I_{n,\delta}^e$, but there exists a $h \in I_{n,\delta}^e$ such that $d(h, g) < \delta$. Then, the first condition gives us

$$m_\delta(e, g) > n. \quad (42)$$

Simultaneously, there exists a $h \in I_{n,\delta}^e$ such that $d(h, g) < \delta$, so

$$m_\delta(e, h) \leq n \text{ and } d(h, g) < \delta, \text{ so } m_\delta(e, g) \leq n + 1. \quad (43)$$

Combining (42) and (43) tells us that $m_\delta(e, g) = n + 1$. Additionally, the value of $U_e(n + 1, g)$ is equal to 1 when $g \in I_{n,\delta}^e$, i.e., $m_\delta(e, g) \leq n$. Hence, $U_e(n + 1, g)$ can be reformulated to

$$U_e(n + 1, g) = \begin{cases} 1 & \text{if } g \in I_{n+1,\delta}^e, \text{ i.e. } m_\delta(e, g) \leq n + 1 \text{ and } g \stackrel{\delta}{\sim} e \\ 0 & \text{else.} \end{cases}$$

This is the same expression as was assumed to be true for some $n \in \mathbb{N}$ but now for $n + 1 \in \mathbb{N}$, and hence, the statement is proven. $\square$

D For $\alpha > 1$ the equivalence relation breaks

For an example where the equivalence relation $\stackrel{\delta}{\sim}$ breaks down for $\alpha > 1$; see Fig. 18. It motivates why in connected component algorithms we choose $\alpha \downarrow 1$ (or $\alpha = 1$) and why we used the evolutions for $\alpha > 1$ only for the affinity measure experiments (where $\alpha > 1$ shrinks the wavefront propagation and softens the max-pooling). Some theoretical results on viscosity solutions for HJB equations [7, 4, 18, 26], do require $\alpha > 1$. In the algorithms for the experiments (e.g. find δ CC) one can either take $\alpha = 1$ or $\alpha \downarrow 1$ which boils down to the same when taking the pointwise limit in the morphological convolutions (12) that solve the HJB PDE system (11) almost everywhere.

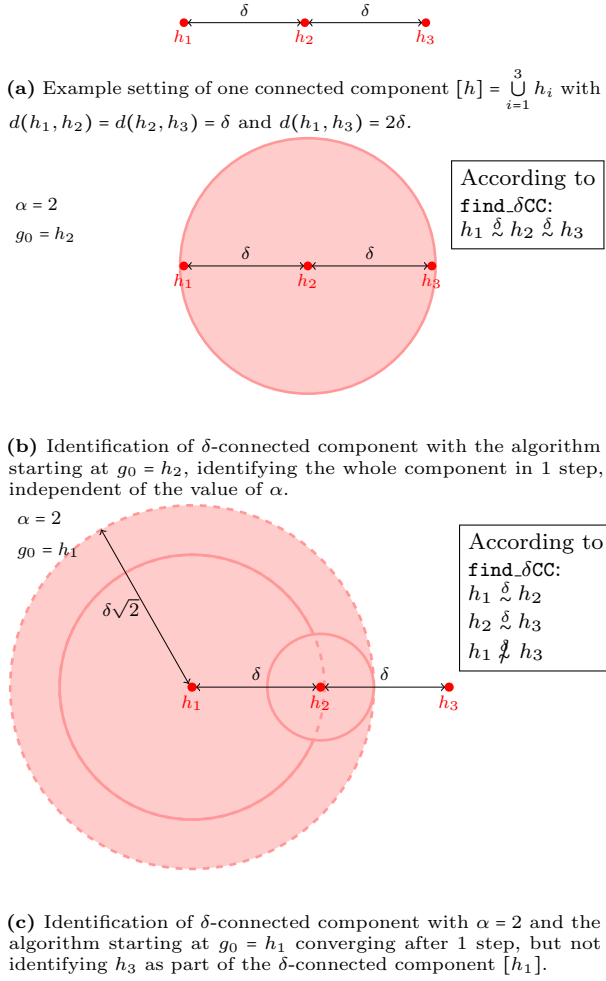


Fig. 18 The δ -connected component $[h]$ (in red in Fig. 18a) is not always completely identified with `find_δCC` when not choosing $\alpha = 1$. In the example, we have used $\alpha = 2$ to show that the algorithm returns different δ -connected components for different starting points that may differ from the full δ -connected component. The example uses starting points h_2 and h_1 for results in Fig. 18b and 18c respectively.

Declarations

Competing Interests

R. Duits is a member of the editorial board of JMIV.

Authors' Contributions

N.J. van den Berg developed the used model, wrote the first draft of the manuscript, created the base code for all the experiments, and did most of the experiments, and wrote the code for the experimental section (available via GitHub [8]). She also created most of the theoretical results in Section 5. O. Mula is one of the project supervisors, provided relevant input and suggestions to the presented theory, and helped with the careful structuring of the manuscript and polished the mathematics of the manuscript. L. Vis helped with the

experiments involving δ -connected component experiments in the experimental section, and she helped in pre-development of the connected component algorithm (`find_δCC`). R. Duits is one of the project supervisors, provided relevant input and setup for the theory and advice on the experiments, and he polished the mathematics of the manuscript.

All authors collaborated closely, wrote parts of the manuscript, and reviewed the manuscript.

Author contributions per section (in order of appearance): Sec. 1 - N.J. van den Berg & O. Mula & R. Duits; Sec. 2 - N.J. van den Berg & R. Duits; Sec. 3 - N.J. van den Berg & O. Mula & R. Duits; Sec. 4 - N.J. van den Berg & O. Mula & R. Duits; Sec. 5 - N.J. van den Berg & L. Vis & O. Mula & R. Duits; Sec. 6 - N.J. van den Berg & O. Mula & R. Duits; Sec. 7 - N.J. van den Berg & O. Mula & R. Duits; Sec. 8 - N.J. van den Berg & L. Vis; Sec. 9 - N.J. van den Berg; App. A - N.J. van den Berg & R. Duits; App. B - N.J. van den Berg & R. Duits; App. C - N.J. van den Berg, L. Vis and R. Duits; App. D - N.J. van den Berg and R. Duits;

Funding

We gratefully acknowledge the Dutch Foundation of Science NWO for its financial support by Talent Programme VICI 2020 Exact Sciences (Duits, Geometric learning for Image Analysis, VI.C. 202-031).

The EU is gratefully acknowledged for financial support through the REMODEL - MSCA-SE 101131557 project which supported the research directions and results of this work.

Olga Mula gratefully acknowledges support from the DesCartes project.

Availability of Data and Materials

The STAR-dataset [55,1] is, together with the *Mathematica* notebooks, available via [8].

References

1. S. ABBASI-SURESHJANI, I. SMIT-OCKELOEN, J. ZHANG, AND B. TER HAAR ROMENY, *Biologically-inspired supervised vasculature segmentation in SLO retinal fundus images*, in Image Analysis and Recognition, M. Kamel and A. Campilho, eds., Cham, 2015, Springer International Publishing, pp. 325–334.
2. S. ABBASI-SURESHJANI, J. ZHANG, G. SANGUINETTI, R. DUTS, AND B. T. HAAR ROMENIJ, *Geometric connectivity analysis based on edge co-occurrences in retinal images*, in Proceedings of the Ophthalmic Medical Image Analysis International Workshop 3(2016), 2016, pp. 154–155.
3. J. ANGULO AND S. VELASCO-FORERO, *Riemannian mathematical morphology*, Pattern Recognition Letters, 47 (2014), pp. 93–101. Advances in Mathematical Morphology.
4. D. AZAGRA, J. FERRERA, AND F. LÓPEZ-MESAS, *Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds*, Journal of Functional Analysis, 220 (2005), pp. 304–361.
5. D. BARBIERI, G. CITTI, G. COCCI, AND A. SARTI, *A cortical-inspired geometry for contour perception and motion integration*, JMIV, 49 (2014), pp. 511–529.

6. E. J. BEKKERS, D. CHEN, AND J. M. PORTEGIES, *Nilpotent Approximations of Sub-Riemannian Distances for Fast Perceptual Grouping of Blood Vessels in 2D and 3D*, Journal of Mathematical Imaging and Vision, 60 (2018), pp. 882–899.
7. G. BELLAARD, D. L. J. BON, G. PAI, B. M. N. SMETS, AND R. DUTS, *Analysis of (sub-)Riemannian PDE-G-CNNs*, Journal of Mathematical Imaging and Vision, (2023).
8. N. V. D. BERG, *Image analysis applications of connected components on lie groups in $SE(2)$ in Mathematica*. <https://github.com/NickyvdBerg/ConnectedComponentsOnLieGroups>, 2024.
9. D. BON, G. PAI, G. BELLAARD, O. MULA, AND R. DUTS, *Optimal transport on the lie group of roto-translations*, 2024.
10. R. V. D. BOOMGAARD AND R. V. BALEN, *Methods for fast morphological image transforms using bitmapped binary images*, CVGIP: Graphical Models and Image Processing, 54 (1992), pp. 252–258.
11. T. BRENDDEL, J. SCHWANKE, P. F. JENSCH, AND R. MEGNET, *Knowledge-based object recognition for different morphological classes of plants*, in Optics in Agriculture, Forestry, and Biological Processing, vol. 2345, SPIE, 1995, pp. 277–284.
12. J. L. BROOKS, *57 Traditional and new principles of perceptual grouping*, in The Oxford Handbook of Perceptual Organization, Oxford University Press, 08 2015.
13. B. BURGETH, M. WELK, C. FEDDERN, AND J. WEICKERT, *Morphological operations on matrix-valued images*, in Computer Vision - ECCV 2004, T. Pajdla and J. Matas, eds., Berlin, Heidelberg, 2004, Springer Berlin Heidelberg, pp. 155–167.
14. D. CHUDASAMA, T. PATEL, S. JOSHI, AND G. PRAJAPATI, *Image segmentation using morphological operations*, International Journal of Computer Applications, 117 (2015), pp. 16–19.
15. F. G. B. DE NATALE AND G. BOATO, *Detecting morphological filtering of binary images*, IEEE Transactions on Information Forensics and Security, 12 (2017), pp. 1207–1217.
16. T. S. DEEPTHI MURTHY AND G. SADASHIVAPPA, *Brain tumor segmentation using thresholding, morphological operations and extraction of features of tumor*, in 2014 International Conference on Advances in Electronics Computers and Communications, 2014, pp. 1–6.
17. R. DIESTEL, *Graph Theory*, Springer-Verlag New York, 1997.
18. E. H. S. DIOP, A. MBENGUE, B. MANGA, AND D. SECK, *Extension of mathematical morphology in Riemannian spaces*, in Scale Space and Variational Methods in Computer Vision, Cham, 2021, Springer International Publishing, pp. 100–111.
19. J. DUGUNDJI, *Topology*, Allyn and Bacon, inc., 1966.
20. R. DUTS, *Perceptual Organization in Image Analysis*, PhD thesis, Eindhoven University of Technology, 2005.
21. R. DUTS, M. DUTS, M. V. ALMSICK, AND B. T. HAAR ROMENY, *Invertible Orientation Scores as an Application of Generalized Wavelet Theory*, Pattern Recognition and Image Analysis, 17 (2007), pp. 42–75.
22. R. DUTS, M. FELSBERG, G. GRANLUND, AND B. T. HAAR ROMENY, *Image Analysis and Reconstruction using a Wavelet Transform Constructed from a Reducible Representation of the Euclidean Motion Group*, International Journal of Computer Vision, 72 (2007), pp. 79–102.
23. R. DUTS AND E. M. FRANKEN, *Left invariant parabolic evolution equations on $SE(2)$ and contour enhancement via invertible orientation scores*, Quarterly of Applied mathematics, AMS, 68 (2010), pp. 255–331.
24. R. DUTS AND E. M. FRANKEN, *Left-Invariant Parabolic Evolutions on $SE(2)$ and Contour Enhancement via Invertible Orientation Scores Part I: Linear Left-Invariant Diffusion Equations on $SE(2)$* , Quarterly of Applied Mathematics, 68 (2010), pp. 255–292.
25. R. DUTS, H. FUEHR, B. JANSSEN, L. FLORACK, AND H. VAN ASSEN, *Evolution equations on Gabor transforms and their applications*, ACHA, 35 (2013), pp. 483–526.
26. A. FATHI AND E. MADERNA, *Weak KAM theorem on non compact manifolds*, Nonlinear Differential Equations and Applications NoDEA, 14 (2007), p. 1–27.
27. P. J. FIGUEROA, N. J. LEITE, AND R. M. BARROS, *A flexible software for tracking of markers used in human motion analysis*, Computer Methods and Programs in Biomedicine, 72 (2003), pp. 155–165.
28. G. L. FORESTI, *Object recognition and tracking for remote video surveillance*, IEEE Transactions on circuits and systems for video technology, 9 (1999), pp. 1045–1062.
29. P.-E. FORSSÉN, *Sparse representations for medium level vision*, PhD thesis, Linköping University, 2001.
30. C. GU AND M.-C. LEE, *Semiautomatic segmentation and tracking of semantic video objects*, IEEE Transactions on Circuits and Systems for Video Technology, 8 (1998), pp. 572–584.
31. A. HAAS, G. MATHERON, AND J. SERRA, *Morphologie mathématique et granulométries en place*, in Annales des mines, vol. 11, 1967, pp. 736–753.
32. H. HEIJMANS, *Theoretical aspects of gray-level morphology*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 13 (1991), pp. 568–582.
33. H. J. HEIJMANS, *Connected morphological operators for binary images*, Computer Vision and Image Understanding, 73 (1999), pp. 99–120.
34. T. LEE, R. KASHYAP, AND C. CHU, *Building skeleton models via 3-d medial surface axis thinning algorithms*, CVGIP: Graphical Models and Image Processing, 56 (1994), pp. 462–478.
35. A. MASHTAKOV, R. DUTS, Y. SACHKOV, E. J. BEKKERS, AND I. BESCHASTNYI, *Tracking of Lines in Spherical Images via Sub-Riemannian Geodesics in $SO(3)$* , Journal of Mathematical Imaging and Vision, 58 (2017), pp. 239–264.
36. F. MEYER AND S. BEUCHER, *Morphological segmentation*, Journal of Visual Communication and Image Representation, 1 (1990), pp. 21–46.
37. I. MOISEEV AND Y. L. SACHKOV, *Maxwell strata in sub-Riemannian problem on the group of motions of a plane*, ESAIM: Control, Optimisation and Calculus of Variation, 16 (2010), pp. 380–399.
38. D. MURRAY AND A. BASU, *Motion tracking with an active camera*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 16 (1994), pp. 449–459.
39. G. K. OUZOUNIS, *Generalized Connected Morphological Operators for Robust Shape Extraction*, PhD thesis, Rijksuniversiteit Groningen, 2009.
40. B. PAWAR, V. T. HUMBE, AND L. KUNDNANI, *Morphology based moving vehicle detection*, in 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), 2017, pp. 217–223.
41. J. B. T. M. ROERDINK AND A. MEIJSTER, *The watershed transform: Definitions, algorithms and parallelization strategies*, Fundamenta informaticae, 41 (2000), pp. 187–228.
42. P. SALEMBIER, P. BRIGGER, J. CASAS, AND M. PARDAS, *Morphological operators for image and video compression*, IEEE Transactions on Image Processing, 5 (1996), pp. 881–898.
43. P. SALEMBIER AND M. H. WILKINSON, *Connected operators*, IEEE Signal Processing Magazine, 26 (2009), pp. 136–157.
44. M. SCHMIDT AND J. WEICKERT, *Morphological Counterparts of Linear Shift-Invariant Scale Spaces*, Journal of Mathematical Imaging and Vision, 56 (2016), pp. 352–366.

45. F. Y. SHIH, *Object representation and recognition using mathematical morphology model*, Journal of Systems Integration, 1 (1991), pp. 235–256.
46. B. M. N. SMETS, J. PORTEGIES, E. J. BEKKERS, AND R. DUIJS, *PDE-Based Group Equivariant Convolutional Neural Networks*, Journal of Mathematical Imaging and Vision, (2023).
47. A. F. M. TER ELST AND D. W. ROBINSON, *Weighted sub-coercive operators on Lie groups*, Journal of Functional Analysis, 157 (1998), pp. 88–163.
48. J. J. VAN DE GRONDE, M. LYSSENKO, AND J. B. T. M. ROERDINK, *Path-based mathematical morphology on tensor fields*, in Visualization and Processing of Higher Order Descriptors for Multi-Valued Data, I. Hotz and T. Schultz, eds., Cham, 2015, Springer International Publishing, pp. 109–127.
49. N. J. VAN DEN BERG, B. M. N. SMETS, G. PAI, J.-M. MIREBEAU, AND R. DUIJS, *Geodesic tracking via new data-driven connections of cartan type for vascular tree tracking*, Journal of Mathematical Imaging and Vision, 66 (2024), pp. 198–230.
50. S. VELASCO-FORERO AND J. ANGULO, *On nonlocal mathematical morphology*, in Mathematical Morphology and Its Applications to Signal and Image Processing, C. L. L. Hendriks, G. Borgefors, and R. Strand, eds., Berlin, Heidelberg, 2013, Springer Berlin Heidelberg, pp. 219–230.
51. J. WEICKERT, *Efficient image segmentation using partial differential equations and morphology*, Pattern Recognition, 34 (2001), pp. 1813–1824.
52. C. S. WON, *A block-based map segmentation for image compressions*, IEEE Transactions on Circuits and Systems for Video Technology, 8 (1998), pp. 592–601.
53. Y. WU, X. PENG, K. RUAN, AND Z. HU, *Improved image segmentation method based on morphological reconstruction*, Multimedia Tools and Applications, 76 (2017), pp. 19781–19793.
54. Z. YU-QIAN, G. WEI-HUA, C. ZHEN-CHENG, T. JING-TIAN, AND L. LING-YUN, *Medical images edge detection based on mathematical morphology*, in 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, 2005, pp. 6492–6495.
55. J. ZHANG, B. DASHTBOZORG, E. J. BEKKERS, J. P. PLUIM, R. DUIJS, AND B. M. T. HAAR ROMENY, *Robust Retinal Vessel Segmentation via Locally Adaptive Derivative Frames in Orientation Scores*, IEEE Transactions on medical imaging, 35 (2016), pp. 2631–2644.