Generalized Morse theory of distance functions to surfaces for persistent homology ☆

Anna Song [a,b], Ka Man Yim [c,*], Anthea Monod [a]

[a] *Department of Mathematics, Imperial College London, London, UK*
[b] *Haematopoietic Stem Cell Laboratory, The Francis Crick Institute, London, UK*
[c] *School of Mathematics, Cardiff University, Cardiff, UK*

A B S T R A C T

This paper brings together three distinct theories with the goal of quantifying shape textures with complex morphologies. Distance fields are central objects in shape representation, while topological data analysis uses algebraic topology to characterize geometric and topological patterns in shapes. The most well-known and widely applied tool from this approach is persistent homology, which tracks the evolution of topological features in a dynamic manner as a barcode. Morse theory is a framework from differential topology that studies critical points of functions on manifolds; it has been used to characterize the birth and death of persistent homology features. However, a significant limitation to Morse theory is that it cannot be readily applied to distance functions because distance functions lack smoothness, which is required in Morse theory. Our contribution to addressing this issue is two fold. First, we generalize Morse theory to Euclidean distance functions of bounded sets with smooth boundaries. We focus in particular on distance fields for shape representation and we study the persistent homology of shape textures using a sublevel set filtration induced by the signed distance function. We use transversality theory to prove that for generic embeddings of a smooth compact surface in $\mathbb{R}^3$, signed

distance functions admit finitely many non-degenerate critical points. This gives rise to our second contribution, which is that shapes and textures can both now be quantified and rigorously characterized in the language of persistent homology: signed distance persistence modules of generic shapes admit a finite barcode decomposition whose birth and death points can be classified and described geometrically. We use this approach to quantify shape textures on both simulated data and real vascular data from biology.

Contents

1. Introduction

Information on the shape and texture of objects captured in images can provide valuable insight to the context of study. For example, in a study to predict the sur-

vival of patients diagnosed with glioblastoma multiforme—an extremely aggressive brain cancer—it has been shown that both the shape and texture of tumor images is even more informative in the prediction task than gene expression [37], which is a type of molecular data obtained from surgery following biopsy and sequencing and is a mainstream type of data used to study cancer. However, the task of feasibly quantifying shape and texture information for computational, statistical, and machine learning tasks in the most informative yet interpretable manner remains challenging. This is especially true when considering complex morphologies entailing porosity and branching behavior, which are not easily represented by meshes [107]. Quantifying shape and texture in an interpretable manner is the driving motivation and practical task of interest in our work. To answer this practical question, we contribute novel, fundamental theoretical results in differential topology with respect to distance functions.

Distance functions (or fields) are fundamental objects in geometry and topological data analysis, [99,85,30,8], PDEs [5], as well as non-smooth analysis and singularity theory [31,13]. They have become popular tools to represent geometry in multiple fields of applications, ranging from computer graphics to shape analysis [81,87,105,100]. Given a subset $A \subset \mathbb{R}^n$, the unsigned distance function assigns to each point in space its distance to A. The *signed* distance function for a set A with non-empty interior is defined by modifying the distance function to ∂A, distinguishing points in the interior from those outside by attributing a negative sign or a positive sign respectively. Signed distance fields are useful in shape representation and analysis [99]; since an object can be implicitly represented as the zero sublevel set of its corresponding signed distance field, an equivalent amount of geometric and topological information is contained in both the shape and the field. Signed distance fields are useful tools to represent complex porous and branching structures, which is a particular application interest in this work.

In this paper, we focus on a method combining signed distance fields and persistent homology—*signed distance persistent homology* (SDPH) [43,44,68,93]—to rigorously define and quantify the texture of materials. Persistent homology is the leading algebraic method in topological data analysis that measures multiscale features in data and fundamentally depends on a distance function [46,56,25]. By measuring the scale at which components, cycles, voids, appear or disappear, persistent homology intrinsically captures heuristic geometric notions such as "shape" and "size" as topological (specifically, homological) features. The lifetimes and algebraic relations between these features are summarized in a *barcode* or *persistence diagram*, a collection of birth–death intervals where each interval corresponds to a topological feature. Moreover, barcodes can be further processed as features in statistical analysis, using either a variety of vectorization methods available or methodology developed to intake persistence diagrams.

For the case of SDPH barcodes, there is a need for a precise interpretation of birth and death events in terms of the geometry of the shape from which they are derived. This lack of interpretability is a major obstacle to using SDPH in applications where interpretability and geometric understanding are key. To achieve interpretability of birth and death events of the barcode, we appeal to *Morse theory* [92], which is a framework

relating the gradient dynamics of a smooth function to the topology of the underlying manifold. For smooth functions with only non-degenerate critical points—*Morse functions*—Morse theory implies that birth and death values in their persistence diagrams correspond to critical values of the function. Furthermore, critical points of a Morse function are classified by their *Morse index*, λ: crossing a critical point with index λ is topologically equivalent to attaching a λ-dimensional "handle," which in terms of persistence, corresponds to the creation or destruction of a homology generator, i.e., endpoints of a birth–death interval. Moreover, smooth functions are generically Morse, in the sense that Morse functions form a dense subset of smooth functions [69, Chapter 6].

An immediate impasse to applying Morse theory to achieve interpretability in our setting is that distance functions are not smooth. This leads to our main theoretical contribution, which is a generalization of Morse theory of smooth functions to the class of general Euclidean distance functions f to smooth compact boundaries in $\mathbb{R}^n$. In particular, our results also adapt to the case of signed distance functions and we leverage these results to systematically interpret SDPH diagrams, giving rise to an interpretable quantifier of both shape and texture of complex morphologies.

Contributions. We generalize Morse theory to Euclidean distance functions (signed or unsigned) of smooth compact boundaries in $\mathbb{R}^n$. To do this, we first show that the fundamental results of smooth Morse theory—the isotopy and handle attachment lemmas—can also be generalized to topological Morse functions [95]. Such functions f admit either a regular form (up to a homeomorphism) at topological regular points,

$$f \underset{\text{homeo}}{\sim} \text{const.} + x_n,$$

or a normal form at topological critical points with index λ,

$$f \underset{\text{homeo}}{\sim} \text{const.} - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2.$$

Similar to the smooth case, we obtain the topological counterparts of the two fundamental Morse lemmas, the isotopy lemma (Theorem 4) and the handle attachment lemma (Theorem 5). They state that the topology of the sublevel sets changes exactly when the level crosses a critical value, at which a λ-cell is attached.

Distance functions are, in fact, special cases of topological Morse functions, under suitable geometric assumptions on the surface that we provide in Lemma 9. The essential idea is to express distance functions locally as Min-type functions [55], for which the notion of non-degenerate critical points is well-defined and provides a topological normal form to distance functions (Theorem 7). This sequence of ideas can be figuratively summarized as

$$\text{distance functions} \xrightarrow[\text{conditions}]{\text{geometric}} \text{Min-type functions} \xrightarrow[\text{form}]{\text{normal}} \text{topological Morse functions}.$$

As a key result, we prove that signed distance functions of generically embedded compact surfaces in $\mathbb{R}^3$ are topological Morse functions with finitely many critical points (Theorem 13). We use transversality theory [1] to prove this result which echoes the fact that smooth functions are generically Morse.

This theoretical investigation has important consequences on the interpretation of SDPH in applications to data, which we present as our applied contribution. Our previous theoretical contributions imply that, for generic shapes in $\mathbb{R}^3$, the signed distance persistence modules are tame and thus admit a finite barcode decomposition. Furthermore, we give a geometric description of the Min-type index of critical points and obtain a classification of critical points as well as of birth–death points in persistence diagrams. We demonstrate how the SDPH approach quantifies texture in shapes by applying it to both simulated and real data: in particular, realizations of Gaussian random fields; synthetic porous shapes generated by "curvatubes" [107]; and confocal images of bone marrow vessels. From these case studies, we derive a general interpretation of SDPH diagrams that is useful for shape texture analysis.

Outline. In Section 2, we provide the necessary background and prove results associated with foundational concepts necessary for our work; we recall the essential facts linking persistent homology to Morse theory and introduce the Min-type framework. In particular, we focus on describing non-degenerate Min-type critical points of functions. Section 3 is the theoretical pillar of this work and contains most of our theoretical contributions. In Section 3.1, we first show that the fundamental lemmas of smooth Morse theory also hold in the more general setting of topological Morse functions. Then, in Section 3.2, we describe in Proposition 1 the geometric conditions where signed distance functions are topological Morse functions. Concluding with Section 3.3, we prove that signed distance functions induced by *generic* embeddings of surfaces are topological Morse functions in Theorem 13 by satisfying the conditions of Proposition 1.

Section 3 lays the foundations for Section 4, where we define the SDPH approach rigorously. In particular, Theorem 13 allows us to show for generic embeddings of surfaces, the persistence module of the signed distance filtration admits a barcode decomposition into finitely-many interval modules (Corollary 5). Moreover, as the topological critical points of the signed distance filtration are shown to be generically non-degenerate, we can classify and geometrically interpret birth–death pairs in the barcode using the index of the critical points. We use this fact in Section 5, where we apply SDPH to analyze synthetic and real data describing textures and derive a general interpretation of the SDPH diagrams for applications. Readers solely interested in the practical applications of SDPH may go to Sections 4 and 5 directly.

2. Background and foundational concepts

In this section, we focus on outlining foundational concepts in persistent homology, smooth Morse theory, and Min-type Morse theory. The ideas and results described in this

section motivate our study of topological Morse functions in Section 3.1, and give us the theoretical framework for analyzing signed distance persistent homology in Section 3.2.

Morse theory [92] is an important tool in differential topology for studying topological properties of manifolds, grounded on a fundamental relationship between gradient dynamics of smooth functions and the homology of the underlying manifold. A thorough account of applications of Morse theory to topology, differential geometry, and mathematical physics is given by [20]. Within the scope of topological data analysis, Morse theory also has foundational consequences; first, for theoretically understanding sublevel sets filtrations; second, for deriving new computational algorithms; and third, for generating new methods for data analysis. We outline below how Morse theory describes the persistence modules of filtrations of smooth functions in terms of relations between critical points (see also Chapter 7 of [57]). *Discrete* Morse Theory—an extension of smooth Morse theory to monotone functions on simplicial complexes—underpins efficient algorithms for computing the barcodes of filtered simplicial complexes [67] and other topological invariants [39]. Beyond the manifold setting, Morse theory has been generalized to describe functions on subsets of positive reach in Euclidean space [51], and a related class of subsets called complementary regular sets [34]. Morse-theoretic concepts have also been applied to methods for clustering data [28], shape analysis [26], image analysis and partitioning [47], and topographical function simplification [12,111].

Min-type Morse theory—a variant of Morse theory for *Min-type functions* [55]—has been instrumental in quantifying and bounding the homological behavior of geometric complexes on point clouds, a natural setting in many applications that topological data analysis is geared towards. A rich vein of research in this topic has been contributed by [16,18,17,41], among others. In contrast to those approaches, which cast distance functions to point clouds as Min-type functions, we focus on distance functions to surfaces which may not be globally Min-type. Nevertheless, we show in Section 3.2 that the Min-type Morse theory framework is a key tool for describing the critical points of distance functions to generic surfaces. We give a summary of the relevant ideas from Min-type theory in this section to build up to that result. However, we note that for a full Morse theory for distance functions, we will require the more general framework of *topological Morse theory* described in Section 3.1.

2.1. Persistent homology and Morse theory

We begin our discussion with a brief summary of basic notions in persistent homology, the leading framework in topological data analysis. In particular, we highlight how classical Morse theory allows us to relate the barcode of a persistence module with critical points of a smooth function, setting the stage for generalizations to persistence modules induced by filtrations of Min-type or topological Morse functions.

We present how persistent homology relates to Morse theory in the smooth setting. To do this, we introduce some essential facts and basic terminology related to persistent homology and Morse theory in the smooth setting.

Sublevel set filtrations and persistence diagrams. Given a topological space X and a continuous function $f : X \to \mathbb{R}$, the *sublevel set filtration* of f is the sequence of inclusion of sublevel sets $X_a = f^{-1}(-\infty, a]$, given by

$$X_\bullet : \quad X_a \subseteq X_b, \quad \forall\, a \le b \in \mathbb{R}.$$

The inclusion maps $\iota_{a,b} : X_a \to X_b$ induce by functoriality a sequence of homology groups

$$\mathrm{PH}_k(f) : \quad \mathrm{H}_k(X_a) \xrightarrow{\mathrm{H}_k(\iota_{a,b})} \mathrm{H}_k(X_b), \quad \forall\, a \le b \in \mathbb{R}$$

that describes the changes in the topology of sublevel sets. Here, $k \ge 0$ refers to the homology dimension. If the homology groups are taken with coefficients in a field $\mathbb{F}$, this sequence is called the *persistence module* of f [114,46]. Furthermore, if the vector spaces $\mathrm{H}_k(f^{-1}(-\infty, a])$ are finite-dimensional for all $a \in \mathbb{R}$, the persistence module is said to be *pointwise finite dimensional* (p.f.d.); in this case, the persistence module can be specified up to isomorphism by a *barcode*, which is a multiset of (open–open, closed–closed, closed–open, or open–closed) intervals in the extended real number line $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$,

$$\mathrm{Bar}_k(f) = \{\langle b, d \rangle \subset \bar{\mathbb{R}}\},$$

where $\mathbb{R} \cup \{-\infty\} \ni b < d \in \mathbb{R} \cup \{+\infty\}$. Alternatively, the module can be characterized by a *persistence diagram*, which is a multiset of *birth–death points* in the upper-diagonal extended half-plane

$$\mathrm{Dgm}_k(f) = \{(b, d) \in (\bar{\mathbb{R}})^2\}.$$

Each interval (or bar, or birth–death point) describes a k-dimensional homology class being born at filtration value b and being trivialized at value d. This characterization is possible due to the interval decomposition theorem of [38] for p.f.d. persistence modules: algebraically, a persistence module can be uniquely decomposed into a direct sum of *interval modules*, where each interval module is in one-to-one correspondence with an interval in the barcode. Details can be found in [52] and [38].

Stability. Persistence diagrams exhibit a property which is fundamental to their use in data analysis, namely, their *stability* with respect to perturbations of the input data, in terms of the bottleneck distance. Let $\mathrm{Dgm}^{(1)}$ and $\mathrm{Dgm}^{(2)}$ be two persistence diagrams, and Π denote the set of all possible bijections $\pi : \mathrm{Dgm}^{(1)} \to \mathrm{Dgm}^{(2)}$ between them, where it is assumed that any point on the diagonal $\{x = y\}$ in the diagrams has infinite multiplicity. The *bottleneck distance* between two persistence diagrams is

$$d_B(\mathrm{Dgm}^{(1)}, \mathrm{Dgm}^{(2)}) = \inf_{\pi \in \Pi} \sup_{x \in \mathrm{Dgm}^{(1)}} \|x - \pi(x)\|_\infty.$$

Here, we use the convention that $\infty - \infty = 0$ to deal with infinite death times.

The stability result holds for a broad class of functions called tame. A function $f : X \to \mathbb{R}$ on a topological space X is *tame* if the persistent homology of the sublevel set filtration $\mathrm{PH}(f) = \mathrm{PH}(X_\bullet)$, $X_t = \{f \leq t\}$, has *finite type*. Equivalently, the persistence module is p.f.d. and has finitely many *homological critical values*, at which, by definition, there exists k such that for all sufficiently small $\epsilon > 0$, the map $\mathrm{H}_k(f^{-1}(-\infty, t - \epsilon]) \to H_k(f^{-1}(-\infty, t + \epsilon]))$ is *not* an isomorphism.

We suppose that X is *triangulable*, i.e., there is a (finite) simplicial complex K whose underlying space $|K|$ is homeomorphic to X.

Theorem 1 *(Stability Theorem [33]). Let X be a triangulable space with continuous tame functions $f, g : X \to \mathbb{R}$. Then the persistence diagrams of the sublevel set filtrations they induce on X satisfy*

$$d_B(\mathrm{Dgm}(f), \mathrm{Dgm}(g)) \leq \|f - g\|_\infty.$$

The Stability Theorem states that, if two input functions are geometrically close to each other in L_∞ norm, then necessarily their diagrams are also close to each other in bottleneck distance d_B. This is a crucial result for applicability of persistent homology to real data: if the input is perturbed by noise, sampling error, and so on, the output diagram will not be too different from the actual diagram. There are many other stability results in different settings; an account of these works can be found in [24, Section 2.3].

Smooth Morse theory and persistent homology. If f is a Morse function [92], there is a correspondence between the critical values of f (values where $\nabla f = 0$), and the end points of the intervals in the barcode of f. A *Morse function* is a smooth, proper function on a smooth manifold M, where any critical point q is non-degenerate (i.e., the Hessian evaluated at q has full rank). The correspondence between the end points of intervals in the barcode of f and its critical values stem from the following fundamental theorems of Morse theory.

Recall that a λ-cell is a closed ball of dimension λ, $e^\lambda = \{y \in \mathbb{R}^\lambda \mid \|y\| \leq 1\}$, and its boundary is denoted $\partial e^\lambda = \{y \in \mathbb{R}^\lambda \mid \|y\| = 1\}$. Attaching a λ-cell to a topological space Y consists of taking first the disjoint union of Y and e^λ, and then identifying points $y \in \partial e^\lambda$ to $\Phi(y) \in Y$ via a continuous map $\Phi : \partial e^\lambda \to Y$. The resulting space is endowed with the quotient topology and denoted by $Y \cup_\Phi e^\lambda$.

Theorem 2 *([92]). Let $f : M \to \mathbb{R}$ be a smooth, proper function on an n-dimensional manifold M and let $\mathrm{Crit}(f)$ denote the set of critical points of f.*

(M0) If q is a non-degenerate critical point, there is a diffeomorphism $\phi : \mathbb{R}^n \to U$ onto an open neighborhood U of q, such that $\phi(0) = q$, and

$$f \circ \phi(x) = f(q) - \sum_{i=1}^{\lambda(q)} x_i^2 + \sum_{i=\lambda(q)+1}^{n} x_i^2. \tag{1}$$

The number of negative coefficients $\lambda(q)$ is the index of q.

(M1) If $f^{-1}[a,b]$ contains no critical points of f, then $f^{-1}(-\infty, a]$ is a deformation retract of $f^{-1}(-\infty, b]$ to; thus they are homotopy equivalent.

(M2) Furthermore, if $\mathrm{Crit}(f) \cap f^{-1}[a,b] = \{q_1, \dots, q_k\}$ is a finite set of non-degenerate critical points, and $\mathrm{Crit}(f) \cap f^{-1}[a,b] \subset f^{-1}(c)$, then $f^{-1}(-\infty, b]$ has the homotopy type of $f^{-1}(-\infty, a]$ with cells $e^{\lambda(q_i)}$ of dimension $\lambda(q_i)$ attached along the boundaries of the cells:

$$f^{-1}(-\infty, b] \simeq f^{-1}(-\infty, a] \cup e^{\lambda(q_1)} \cup \dots \cup e^{\lambda(q_k)}.$$

The two statements (M1) and (M2) are informally known as the "isotopy lemma" and the "handle attachment lemma". Equation (1) is known as the normal form of the function at the critical point q.

As a consequence of the description of topological changes in sublevel sets of Morse functions in Theorem 2, we can characterize the persistence modules of Morse functions where the following finiteness conditions hold. Note that the conditions always hold if the underlying manifold is compact.

Corollary 1. *Let $f : M \to \mathbb{R}$ be a smooth, proper Morse function on a manifold M with finite dimensional homology groups $\mathrm{H}_\bullet(M)$ with field coefficients. If f only has finitely many critical points, then:*

(i) *The persistence module of the sublevel set filtration f is pointwise finite dimensional (p.f.d.);*

(ii) *For all k, the barcode $\mathrm{Bar}_k(f)$ is a finite multiset of intervals $\{[b,d) \subset \bar{\mathbb{R}}\}$;*

(iii) *A critical point with index λ either corresponds to the birth of an interval in $\mathrm{Bar}_\lambda(f)$ with homology dimension λ, or the death of an interval in $\mathrm{Bar}_{\lambda-1}(f)$ with homology dimension $\lambda - 1$.*

Proof. Since f only has finitely many critical points, (M1) and (M2) together imply $\mathrm{H}_\bullet(f^{-1}(-\infty, t]) \to \mathrm{H}_\bullet(M)$ only has finite rank and kernel. Therefore $\mathrm{H}_\bullet(f^{-1}(-\infty, t])$ is finite dimensional, and $\mathrm{PH}_\bullet(f)$ are p.f.d. persistence modules. As each critical point p of index λ corresponds to attaching a λ-cell in the sublevel set filtration, the filtration passing p will either correspond to the birth of an interval in $\mathrm{Bar}_\lambda(f)$ with homology dimension λ, or the death of an interval in $\mathrm{Bar}_{\lambda-1}(f)$ with homology dimension $\lambda - 1$. Finally (M1) and (M2) together with finite $\mathrm{Crit}(f)$ imply that homological features can only be born and destroyed at finitely many critical values; thus any interval in the barcode of f can only be of the form $[b,d)$. $\square$

2.2. Min-type functions and their critical points

Two major components are needed here to generalize Morse theory to Euclidean distance functions: *criticality* and *index*. [65] introduced a definition of *critical points* for distance functions d_p w.r.t. a single point on a Riemannian manifold, precisely to generalize Morse's isotopy lemma (see [31,64]). However, no explicit notion of index was declared. Later, [54] and [55] generalized Morse theory to the broad class of *Min-type* functions, which allowed for the definition of an appropriate notion of *index* and *non-degeneracy* of critical points.

Essentially, Min-type functions are functions that can be expressed as the minimum of a finite number of smooth functions. [55] studied Min-type distance functions d_p on non-positively curved Riemannian manifolds M and showed that for compact M, the d_p are even Morse-Min-type (i.e., have non-degenerate regular and non-degenerate critical points only) for generic non-positively curved Riemannian metrics on M. Later, [73] obtained similar results by rephrasing Min-type properties of the distance function d_p using more direct geometric arguments.

In topological data analysis, the Min-type framework had an important application to distance functions $d_{\mathcal{P}}$ defined w.r.t. a finite point cloud $\mathcal{P}$ in $\mathbb{R}^n$ [16]. In particular, the ϵ-Čech complexes of $\mathcal{P}$ are homotopy equivalent to the sublevel sets $\{d_{\mathcal{P}} \leq \epsilon\}$ of the distance function, by the Nerve Theorem [19]. Therefore, if $\mathcal{P}$ has a generic configuration, the critical points of $d_{\mathcal{P}}$ are non-degenerate and determine the exact homological changes in the Čech filtration. Yet, to the best of our knowledge, there has been no generalization of Morse theory for Euclidean distance fields w.r.t. smooth boundaries, especially in the context of persistent homology.

2.2.1. Min-type functions and their representations

We begin with an introduction to Min-type functions. For the sake of clarity and conciseness, we only cover the necessary content from the Min-type function theory for our goals; moreover, we rephrase it in simpler terms for the Euclidean case only. See [55] for complete details of the Min-type framework in full generality. In what follows, we borrow from [55] and provide some proofs with the goal of making this work self-contained.

As mentioned previously, a Min-type function is a function which can be locally written as the minimum of a finite number of smooth functions. As such, a Min-type function itself may not be smooth even though it can be constructed from a collection of smooth functions.

As Min-type functions are locally constructed, we first review some terminology regarding local equivalences of sets and functions before formally giving the definition of Min-type functions. The properties stated below are only concerned with the behavior of functions and sets in arbitrarily small neighborhoods of a point. We use the framework of *germs* to describe such local behavior; we refer the reader to [21, I.6.9-10] for a detailed exposition. Given a topological space X, we say two maps $f, g : X \to \mathbb{R}$ are *germ equiv-*

alent at $x \in X$ if there is a neighborhood V of x, such that the restrictions $f|_V = g|_V$ are equal. Similarly, if A and B are two subsets of X, then they are germ equivalent at x if there is a neighborhood V of x such that $A \cap V = B \cap V$. Germ equivalence at a given point x is an equivalence relation (on maps or sets), and we call the equivalence classes *germs* (of maps or of sets). We slightly abuse notation and use $f = g$ "as germs of functions" to denote germ equivalence of functions, and, similarly, $A = B$ "as germs of sets" to denote germ equivalence of sets. Given a function f or subset A, "the germ of f" or "the germ of A" refers to the class of functions or subsets represented by f or A respectively.

As a notational convenience, hereafter we use $\{f = c\}$, $\{f < c\}$ and $\{f \leq c\}$ to denote level sets, sublevel sets, and strict sublevel sets of a function f at level c.

Definition 1 *(Min-type functions [55]).* Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. f is a C^k-smooth *Min-type function at* q if there exists a finite family of C^k-smooth functions $\alpha_1, \ldots, \alpha_m$, called a *representation* of f, such that

$$f = \min\{\alpha_1, \ldots, \alpha_m\}$$

as germs of functions at q.

That is, there is a neighborhood N of q on which f can be expressed as the minimum of $\alpha_1, \ldots, \alpha_m$. Given f is min-type and admits a representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$, the *active set* of α_i at q is defined to be the germ of the following set at q:

$$A_i = \{x \mid f(x) = \alpha_i(x)\}.$$

That is, the equivalence class of subsets of $\mathbb{R}^n$ whereby there is a neighborhood of q on which such subsets coincide with the level set $f - \alpha_i = 0$. Since $f = \min\{\alpha_1, \ldots, \alpha_m\}$, we remark that the germ of A_i can be equivalently expressed as the germ of

$$A_i = \bigcap_{j \neq i} \{\alpha_i - \alpha_j \leq 0\}.$$

The germ of the interior A_i° of A_i is called the *open active set* of α_i.

The statement that f is a Min-type function at q is equivalent to saying there exists a neighborhood N of q on which $\forall\, x \in N, f(x) = \min_i \alpha_i(x)$. In particular, $N = \cup_i A_i$ (as germs of sets) and f is C^k-smooth on A_i°.

In general, a Min-type function can have more than one representation $\{\alpha_i\}$. A "most concise" representation that uses the fewest constituent functions α_i is said to be *efficient*.

Definition 2 *(Efficient representation).* The representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$ as germs of functions is *efficient at* q if m is the minimal number of functions necessary to represent f around q.

Efficient representations were originally referred to as "minimal representations" by [55]. To avoid confusing terminology, we use the term "efficient" instead of minimal in this work.

While efficient representations may not be unique, it can be shown that if an efficient representation admits *linearly-independent* gradients (LIG), then the efficient-LIG representation is unique as a set of functions [55].

Recall that a set of vectors $v_1, \ldots, v_m \in \mathbb{R}^n$ is said to be in *general position* in $\mathbb{R}^n$ if the dimension of the affine subspace spanned by $\{v_1, \ldots, v_m\}$ is equal to $m - 1$. Equivalently, this means the collection $v_1, \ldots, v_m$ are in general position if $v_2 - v_1, \ldots, v_m - v_1$ are linearly independent. Notice that the origin may or may not belong to the affine subspace (depending on whether $v_1, \ldots, v_m$ are linearly dependent or not).

Definition 3. A representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$ at q is *efficient-LIG* if it is efficient and if the gradients $\{\nabla \alpha_1, \ldots, \nabla \alpha_m\}$ are in general position at q.

Note that the general position condition implies $m \leq n + 1$. The condition also implies $\nabla(\alpha_i - \alpha_j) \neq 0$ for all $i \neq j$ on an open neighborhood U of q. Since zero is a regular value of $\alpha_i - \alpha_j$ on U, the interior of the sublevel set $\alpha_i - \alpha_j \leq 0$ is the strict sublevel set $\alpha_i - \alpha_j < 0$. Thus, we can express the open active site of α_j as follows:

$$A_i^\circ = \bigcap_{i \neq j} \{\alpha_i - \alpha_j < 0\}.$$

We can thus characterize the open active sites at q by considering the functions $\alpha_i - \alpha_j$.

Lemma 1. *Let $f = \min\{\alpha_1, \ldots, \alpha_m\}$ be an efficient-LIG representation at q. For any $i \in 1, \ldots, m$, there is a C^k map $F_i : (\mathbb{R}^n, q) \to (\mathbb{R}^n, 0)$, and a connected neighborhood U of q, such that the restriction $F_i|_U : U \to F_j(U)$ is a diffeomorphism, and for any $j \neq i$, the function $\alpha_i - \alpha_j$ on U is equivalent as germs to $y_j \circ F_i(x)$, where y_j is the jth Cartesian coordinate function on the codomain.*

Proof. Without loss of generality consider $i = m$, and set $f_j(x) := \alpha_m(x) - \alpha_j(x)$ for $j = 1, \ldots, m - 1$. If $m \leq n$, since $\nabla f_1(q), \ldots, \nabla f_{m-1}(q)$ are linearly independent, let $x_m, \ldots, x_n$ be a set of orthogonal coordinates for the vector subspace orthogonal to that spanned by $\nabla f_1(q), \ldots, \nabla f_{m-1}(q)$. Consider then $F_m(x) = (f_1(x), \ldots, f_{m-1}(x), x_m - q_m, \ldots, x_n - q_n)$ (here q_k denotes the kth coordinate of q). By construction, the Jacobian $dF_m(q)$ is invertible; thus the inverse function theorem [82, Theorem 4.5] implies that F_m is a diffeomorphism on a connected neighborhood U of q. By construction, for $j = 1, \ldots, m - 1$, and $x \in U$, the function $(\alpha_m - \alpha_j)(x) = f_j(x)$ yields the jth coordinate of $F_m(x)$. $\square$

As a consequence we can characterize active sets via the diffeomorphism F_i on a neighborhood U of q, and Cartesian coordinates y_j. Because $\alpha_i(x) < \alpha_j(x)$ on U iff $y_j \circ F_i(x) < 0$, we have the equivalence as germs at q

$$A_i = F_i^{-1}\left(F_i(U) \cap \bigcap_{j \neq i}\{y_j \leq 0\}\right) \quad \text{and} \quad A_i^\circ = F_i^{-1}\left(F_i(U) \cap \bigcap_{j \neq i}\{y_j < 0\}\right). \quad (2)$$

The following two properties of open active sets follow directly from this representation.

Lemma 2. *Let $f = \min\{\alpha_1, \ldots, \alpha_m\}$ be an efficient-LIG representation at q. Then $A_i^\circ \neq \emptyset$ as germs of sets.*

We say a subset $Z \subset \mathbb{R}^n$ is locally connected at $q \in \mathbb{R}^n$ if there is a neighborhood V of q such that $V \cap Z$ is non-empty and connected. Note that in this particular setting we do not require $q \in Z$. Because this is a local property of Z in the neighborhood of q, local connectedness at q is a property shared across subsets with the same germ at q. The following lemma states that germs of open active sets of an efficient-LIG representation are locally connected at q in this sense.

Lemma 3. *Given an efficient-LIG representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$ at q, open active sets A_i° are locally connected at q.*

Proof. Let U, F_i, and y be as described in Lemma 1 and eq. (2). We choose a sufficiently small ball $B \subseteq F(U_i)$ around the origin, and consider we note that $B \cap \bigcap_{j \neq i}\{y_j < 0\}$ is connected. Since

$$F_i^{-1}\left(B \cap \bigcap_{j \neq i}\{y_j < 0\}\right) = F_i^{-1}(B) \cap F_i^{-1}\left(F_i(U) \cap \bigcap_{j \neq i}\{y_j < 0\}\right) = F_i^{-1}(B) \cap A_i^\circ,$$

because F_i is a diffeomorphism on $U \supset F_i^{-1}(B) \ni q$, we have found an open neighborhood $V = F_i^{-1}(B)$ of q such that $V \cap A_i^\circ$ is connected. Thus, the germ of A_i° is locally connected at q. $\square$

Lemma 4. *Let $f = \min\{\alpha_1, \ldots, \alpha_m\}$ be an efficient-LIG representation at q. Denote by Ξ the set of germs at q of connected open subsets on which f is C^k. Then $\{A_i^\circ\}$ are the maximal elements of Ξ.*

Proof. The germ of set A_i° is locally connected (Lemma 3) and f is smooth on it, so $A_i^\circ \in \Xi$. There also exists $a \in A_i^\circ \neq \emptyset$ (Lemma 2), with $\alpha_i(a) = f(a)$. Let $V \in \Xi$ be such that $A_i^\circ \subset V$. We need to show that $V = A_i^\circ$. Suppose not, then $V \setminus A_i \neq \emptyset$ and there exists $b \in V \setminus A_i$ and $j \neq i$ such that $\alpha_i(b) > f(b) = \alpha_j(b)$. Then since V is connected, consider a continuous path $\gamma(t) \in V$ joining $a \in A_i^\circ$ to b. f coincides with α_i

on a but then α_i becomes strictly greater than f around b. By continuity and definition of Min-type functions, we can then find $\ell \neq i$ (which is not necessarily j) and a point y on the path such that $\alpha_\ell(y) = \alpha_i(y) = f(y)$, so $y \in A_\ell \cap A_i \cap V$.

The previous arguments apply to any neighborhood of q. By extracting an index $\ell_0 \neq i$, we can build a sequence $(y_\nu) \in V \cap A_i \cap A_{\ell_0}$ converging to q. But since f, α_i, and α_{ℓ_0} are C^k-smooth on V and $\forall\, j$, $A_j = \overline{A_j^\circ}$, we get $\nabla f(y_\nu) = \nabla \alpha_i(y_\nu) = \nabla \alpha_{\ell_0}(y_\nu)$. At the limit, we obtain $\nabla \alpha_i(q) = \nabla \alpha_{\ell_0}(q)$, which contradicts the LIG property. Therefore $V = A_i^\circ$.

To show that the sets A_i° are the only maximal elements of Ξ, consider $V \in \Xi$. There exists i such that $V \cap A_i \neq \emptyset$ as germs of sets. We can re-use the previous argument to show that $V \setminus A_i = \emptyset$, hence $V \subset A_i^\circ$.   □

It then follows that germs of efficient-LIG representations are uniquely defined up to permutation since they induce the same maximal connected open sets on which f is C^k-smooth.

Corollary 2 *(Uniqueness). Let $f = \min\{\alpha_1, \ldots, \alpha_m\} = \min\{\beta_1, \ldots, \beta_m\}$ be two efficient-LIG representations at q. Then there exists a permutation of indices σ such that the germs of active sets and functions $A_i = B_{\sigma(i)}$ and $\alpha_i = \beta_{\sigma(i)}$ are equal.*

In subsequent discussions, we will find it useful to characterize the intersection of active sets $A_i = \{f(x) = \alpha_i\}$ given an efficient-LIG representation.

Definition 4. Given an efficient-LIG representation $f = \min\{\alpha_1, \ldots \alpha_m\}$ at q, $G_f(q)$ is defined as the *germ of set*

$$G_f(q) = \{\alpha_1 = \ldots = \alpha_m = f\} = \bigcap_i A_i,$$

where recall that $A_i = \{f = \alpha_i\}$.

Notice that $q \in G_f(q)$ and that this definition does not depend on the choice of representation due to Corollary 2.

The form of $G_f(q)$ is locally a smooth submanifold of dimension $n - m + 1$ due to the LIG property. In particular, if the number of functions m over which the minimum is taken is equal to $n + 1$, then $G_f(q)$ may reduce to the point $\{q\}$ itself.

Lemma 5. *Given an efficient-LIG representation $f = \min\{\alpha_1, \ldots \alpha_m\}$ at q, the germ of set $G_f(q)$ is an $(n - m + 1)$-dimensional manifold.*

Proof. Let U be a neighborhood of q where $f = \min\{\alpha_1, \ldots \alpha_m\}$. As the representation is efficient, $G_f(q) = \cap_{i=1,\ldots,m}\{\alpha_i = f\} = \cap_{i=2,\ldots,m}\{\alpha_i - \alpha_1 = 0\}$. Thus, we can express $G_f(q)$ as an intersection of $(m-1)$ level sets of smooth functions $\cap_{i=2,\ldots,m}\{\alpha_i - \alpha_1 = 0\}$.

Since the representation is LIG, the set of $(m-1)$ $\{\nabla(\alpha_i - \alpha_1)\}_{i=2,\ldots,m}$ consists of linearly independent vectors. Thus, due to the implicit function theorem, there is an open neighborhood $V \subset U$ on which $G_f(q)$ is a $(n - (m-1))$-dimensional submanifold. $\square$

We also remark that the germ of the restriction $f_{|G_f(q)}$ is a smooth function, since it coincides with the restriction of α_i to $G_f(q)$ for any i.

2.2.2. Non-degenerate critical points and index

We now define the critical points of Min-type functions.

Definition 5 *(Min-type critical point).* Let f be a C^k Min-type function that admits an efficient-LIG representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$ at q. If 0 is contained in the convex hull of gradients $\{\nabla\alpha_1, \ldots \nabla\alpha_m\}$ at q, then q is said to be a *Min-type critical point* of f.

We use the terminology of "critical points" for points described in Definition 5 because by restricting to the submanifold on which f is smooth, those points are indeed critical points of the restricted smooth function.

Lemma 6. *If q is a Min-type critical point of f then $\nabla f_{|G_f(q)}|_q = 0$ on the submanifold $G_f(q)$.*

Proof. Recall $G_f(q)$ is the germ of the $(n - m + 1)$ dimensional submanifold (Lemma 5) locally defined by the solution set $\alpha_i - \alpha_1 = 0$ for $i = 2, \ldots, m$. Since $f_{|G_f(q)} = \alpha_i \,_{|G_f(q)}$ and α_i is smooth, the restriction $f_{|G_f(q)}$ of f to $G_f(q)$ is smooth.

If $\nabla\alpha_i|_q = 0$, then $\nabla f_{|G_f(q)} = 0$ (notice that since the representation is LIG at most one of $\nabla\alpha_i$ vanish at q). Thus, consider the case where none of α_i are critical at q. Then q is a critical point of $f_{|G_f(q)}$ if and only if $\nabla\alpha_i|_q \perp T_q G_f(q)$. Since $G_f(q)$ is locally defined by the solution set $\alpha_i - \alpha_1 = 0$ and the representation is LIG, the set of linearly independent vectors $\{\nabla(\alpha_i - \alpha_1)|_q\}_{i=2,\ldots,m}$ is a basis of $N_q G_f(q)$. Therefore, q is a critical point of $f_{|G(q)}$ if and only if $\nabla\alpha_1|_q$ is in the span of $\{\nabla(\alpha_i - \alpha_1)|_q\}_{i=2,\ldots,m}$.

Now assume 0 is contained in the convex hull of gradients $\{\nabla\alpha_1, \ldots \nabla\alpha_m\}$ at q: i.e., there is a set of non-negative coefficients t_i that sum to one, such that

$$0 = \sum_{i=1}^{m} t_i \nabla\alpha_i|_q = \sum_{i=2}^{m} t_i \nabla(\alpha_i - \alpha_1)|_q + \nabla\alpha_1|_q.$$

Hence we have $\nabla\alpha_1|_q \in N_q G(q)$ and thus q is a critical point of the smooth function $f_{|G(q)}$ on $G_f(q)$. $\square$

We now turn to the notions of non-degeneracy and index in the context of Min-type functions.

Definition 6 *(Non-degenerate critical point [55])*. Let f be a C^k-Min-type function. If q is a Min-type critical point with an efficient-LIG representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$ at q, such that

1. The origin lies strictly in the *interior* of the convex hull of the gradients $\nabla\alpha_1, \ldots, \nabla\alpha_m$ at q; and
2. The restriction $f_{|G_f(q)}$ is a Morse function.

Then q is said to be a *non-degenerate critical point* (in the Min-type sense) of f.

Notice that the definition of non-degenerate critical points does not depend on the choice of efficient representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$, due to Corollary 2. Moreover, a non-degenerate critical point q of f is a non-degenerate critical point for the restriction $f_{|G_f(q)}$ in the sense of smooth Morse theory.

Definition 7 *(Index of a Min-type function)*. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a C^k-Min-type function with efficient representation $f = \min\limits_{i=1,\ldots,m} \alpha_i$ at q, and suppose that q is a non-degenerate critical point. The *index* of f at q is

$$\text{index}(q; f) := (m - 1) + \text{index}(q; f_{|G_f(q)}).$$

The index has a geometric interpretation: it is the maximal dimension of a submanifold in a neighborhood of q such that the restriction of f on it has a strict maximum at q [55, Section 3.1, Proposition 3].

The Min-type framework is interesting for our purpose of extending Morse theory to distance functions because it allows us to write Min-type functions into a normal form around non-degenerate critical points, similar to the setting of smooth Morse theory (see Theorem 2). In classical Morse theory [92], smooth functions f have simple forms at regular points and non-degenerate critical points. These forms are on the same orbit as f under the action of the group of local diffeomorphisms of $\mathbb{R}^n$, and we either have $f \sim \text{const.} + x_n$ at regular points, or $f \sim \text{const.} - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2$ at non-degenerate critical points. However, for Min-type functions, another group of transformations is needed [89,54].

Definition 8 *(Group of almost smooth homeomorphisms)*. A homeomorphism $\phi : U \to V$, where U and $V = \phi(U)$ are open neighborhoods of 0 in $\mathbb{R}^n$ and $\phi(0) = 0$, is said to be *almost smooth* if ϕ and ϕ^{-1} are smooth except on a submanifold of codimension ≥ 1.

The following normal form theorem for Min-type functions is due to [55] (Section 3.2, Theorem 1) and [54] (Theorem 1.3).

Theorem 3 *(Normal form [55])*. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Min-type function with efficient representation $f = \min\{\alpha_1, \ldots, \alpha_m\}$ at q, and suppose that q is a non-degenerate critical*

point. Then there exists an almost smooth homeomorphism $\phi : U \to V$, where U and $V = \phi(U)$ are open neighborhoods of 0 and q in $\mathbb{R}^n$, such that $\phi(0) = q$ and for all $x \in U$,

$$f \circ \phi(x) = f(q) - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2,$$

where $\lambda = \mathrm{index}(q; f)$.

This implies that non-degenerate critical points of a Min-type function are *isolated*.

Remark 1. [55] also defined a notion of *non-degenerate regular points*. The definition is similar to Definition 6, where Condition 2 must be replaced by 2': the origin does not belong to the convex hull of the gradients $\nabla \alpha_1, \dots, \nabla \alpha_m$ at q, and a Condition 4' must be added, which states that: any $m - 1$ gradients among $\nabla \alpha_1, \dots, \nabla \alpha_m$ are linearly independent at q. Then they defined *Morse Min-type functions* to be Min-type functions that only admit non-degenerate regular points and non-degenerate critical points. These refinements will not be needed for our purpose, as we are only concerned by non-degenerate critical points.

Remark 2. It can be shown that Min-type critical points (Definition 5) are also critical points of the *Clarke subgradient* [32,4], a generalization of gradients for functions that need only be locally Lipschitz continuous. For a compact interlevel set $f^{-1}[a, b]$ that contains only regular points of the Clarke subgradient, it can be shown that $f^{-1}(-\infty, a]$ is a strong deformation retract of $f^{-1}(-\infty, b]$ using the theory of *weak slopes* [35] (another related generalization of gradients).

3. Morse theory for signed distance functions

In this section, we generalize Morse theory to *signed* distance functions $d : \mathbb{R}^n \to \mathbb{R}$ to smooth compact surfaces bounding an open set in $\mathbb{R}^n$. With the goal of computing SDPH to analyze shapes, on the theoretical front, we investigate conditions under which the persistence module $\mathrm{PH}_k(d)$ is p.f.d. (see Section 2.1), and aim to find a correspondence between homological births or deaths and geometric features of the surface. Since d is not continuously differentiable on all of its domain, it is not a Morse function in the classical sense. Therefore we cannot apply Theorem 2. However, we will show that d can be framed as a *topological Morse function* that shares properties of smooth Morse functions described by classical Morse theory. Topological Morse functions either have the form $f \sim \mathrm{const.} + x_n$ at topological regular points, or $f \sim \mathrm{const.} - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2$ at non-degenerate topological critical points, up to a homeomorphism between open neighborhoods.

We first show in Section 3.1 that the fundamental Morse lemmas formulated in the smooth setting also hold in the more general setting of topological Morse functions. This

allows us to generalize the Morse lemmas to signed distance functions later on in Section 3.2: the idea is to successively reframe them as C^k-Min-type functions, under some geometric assumptions, and then as topological Morse functions. Finally, in Section 3.3, we show that these geometric assumptions are not restrictive and hold for generic signed distance functions.

3.1. Topological Morse functions

The main obstruction to generalizing classical Morse theory to continuous functions is the requirement of continuous gradients and gradient flows that underpin the results of Theorem 2. This motivated the study of continuous functions that carry over generalized notions of critical and regular points from smooth functions. One such family are *topological Morse functions* (originally called topological non-degenerate functions) introduced by [95].

Definition 9 *(Topological regular point [95]).* Let $f : X \to \mathbb{R}$ be a continuous function. A point $y \in X$ is a *topological regular point* of f if there is a homeomorphism $\varphi : M_1 \to M_2$ between open neighborhoods M_1 of 0 in $\mathbb{R}^n$ and M_2 of y with $\varphi(0) = y$, such that for all $x = (x_1, \ldots, x_n) \in M_1 \subset \mathbb{R}^n$,

$$f \circ \varphi \ (x) = f(y) + x_n. \tag{3}$$

For notational convenience, we write $f \overset{\varphi}{\sim} f(y) + x_n$ where $\varphi : (M_1, 0) \to (M_2, y)$.

Definition 10 *(Topological critical point [95]).* Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. A point $y \in \mathbb{R}^n$ is said to be a *topological critical point* of f if y is not a topological regular point of f. Furthermore, y is a *non-degenerate topological critical point* of f if there exists an integer $0 \le \lambda \le n$ and a homeomorphism $\phi : N_1 \to N_2$ where N_1 and $N_2 = \phi(N_1)$ are open neighborhoods of 0 and y in $\mathbb{R}^n$, such that $\phi(0) = y$ and for all $x \in N_1$,

$$f \circ \phi \ (x) = f(y) - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2.$$

For notational convenience, we write $f \overset{\phi}{\sim} f(y) - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2$ where $\phi : (N_1, 0) \to (N_2, y)$. Notice that a non-degenerate topological critical point cannot be a topological regular point [95]. We also note that non-degenerate topological critical points are *isolated*.

Definition 11 *(Topological Morse function [95]).* A continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is a *topological Morse function* if all the points $y \in \mathbb{R}^n$ are either topological regular points or non-degenerate topological critical points of f.

The literature dealing with topological Morse functions appears to be sparse (see for instance [23,94,79]); in particular, there does not currently exist full counterparts to the isotopy and handle attachment lemmas in the same form as that given by [92]. Essay III by [77] asserts that the handle attachment lemma is still true for topological Morse functions, by mentioning that a proof by [104] based on mutually transverse foliations exists. However some details of the proofs appear to be missing to allow their complete reconstruction. For completeness, we provide the full proofs based on purely topological arguments; some of them are similar to those found in [23] (see Lemma 3.1) and [36] (see Theorem 2.8).

3.1.1. The isotopy lemma

We now show that the isotopy lemma that describes deformation retractions between sublevel sets of smooth functions over intervals of regular values (Theorem 2(M1)) generalizes to topological Morse functions.

Unlike the smooth setting, we do not have gradient vector fields that generate a flow to deform sublevel sets onto one another. To create a deformation retraction from $f^{-1}(-\infty, b]$ onto $f^{-1}(-\infty, a]$ without a vector field, we form an open cover of the interlevel set $f^{-1}[a, b]$, and perform successive f-decreasing deformations on the individual cover elements, on each of which f is simply the projection onto a Euclidean axis as specified in eq. (3) of Definition 9.

Recall that a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is *proper* if for any compact subset $K \subset \mathbb{R}$, $f^{-1}(K) \subset \mathbb{R}^n$ is compact.

Theorem 4 *(Isotopy Lemma, topological version). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous proper function, and $a < b$. Suppose that all points $y \in f^{-1}[a, b]$ are topological regular points for f.*

Then there exists a strong deformation retraction sending $f^{-1}(-\infty, b]$ onto $f^{-1}(-\infty, a]$. As a consequence, $f^{-1}(-\infty, b]$ and $f^{-1}(-\infty, a]$ are homotopy equivalent.

Proof. The strategy of the proof is to consider arbitrary level sets $f^{-1}(c)$, small enough ϵ, and then apply successive deformations of $\mathbb{R}^n$ that only move points inside finitely many subsets that form an open cover of the interlevel set $f^{-1}[c - \epsilon, c + \epsilon]$. The aim of these deformations is to deform the interlevel set $f^{-1}[c - \epsilon, c + \epsilon]$ into the level set $f^{-1}(c - \epsilon)$ by lowering the values of f. To facilitate this construction, we consider the *open solid cylinder* $CL(b, s) \subset \mathbb{R}^n$, which is the open set bounded by the cylinder whose base disk is $D^{n-1}(0, b)$ along the $n - 1$ first coordinates and side is $D^1(0, s)$ along the last coordinate:

$$CL(b, s) = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \ldots + x_{n-1}^2 < b^2 \text{ and } x_n^2 < s^2\}.$$

Finite open and compact covers of $f^{-1}[c - \epsilon, c + \epsilon]$. Let $c \in [a, b]$ and consider the compact set $f^{-1}(c)$. It can be covered by open sets U_y that are neighborhoods of points

$y \in f^{-1}(c)$ where f can be written in a regular form over U_y: $f \overset{\varphi_y}{\sim} c + x_n$ with $\varphi_y : (V_y, 0) \to (U_y, y)$. By choosing small enough $\epsilon_y > 0$ and $b_y > 0$, we may suppose that $V_y = CL(b_y, \epsilon_y)$ are of open solid cylinder form. In particular, we have $f(U_y) = f(\varphi_y(V_y)) = (c - \epsilon_y, c + \epsilon_y)$.

By compactness, $f^{-1}(c)$ may be covered by a finite number of the sets U_y:

$$f^{-1}(c) \subset W = \bigcup_{i=1,\dots,I} U_{y_i}.$$

Then there exists $\epsilon > 0$ small enough such that $f^{-1}[c - \epsilon, c + \epsilon] \subset W$. Otherwise, there exists sequences $(y_\nu)_\nu \subset \mathbb{R}^n$, $(\epsilon_\nu)_\nu \to 0$ such that $y_\nu \in f^{-1}[c - \epsilon_\nu, c + \epsilon_\nu] \setminus W$, i.e., $f(y_\nu) \in [c - \epsilon_\nu, c + \epsilon_\nu]$ and $y_\nu \notin W$. This implies that $f(y_\nu) \to c$. Because f is proper, there exists a compact set K such that for all ν, $y_\nu \in K$. Up to extracting a subsequence, y_ν converges to $y \in K \setminus W$. By continuity $f(y) = \lim f(y_\nu) = c$, so then $y \in f^{-1}(c) \setminus W$, a contradiction. Therefore, such an ϵ exists and we have

$$f^{-1}[c - \epsilon, c + \epsilon] \subset W = \bigcup_{i=1,\dots,I} U_{y_i}. \tag{4}$$

We can furthermore suppose that $\epsilon < \min_{i=1,\dots,I} \epsilon_{y_i}$. Hence, for $y = y_i$, $\varphi_y^{-1}(f^{-1}[c - \epsilon, c + \epsilon] \cap U_y)$ is a smaller closed solid cylinder $\overline{CL(b_y, \epsilon)}$ included in $V_y = CL(b_y, \epsilon_y)$.

For simplicity, we drop the y in the subscript of U_{y_i} and simply write U_i or V_i instead, where the understanding is that we are referring to y_i.

In fact, the open covering in (4) can be refined into a compact covering. Indeed, there exists a family of closed cylinders $\overline{CL(b_i', \epsilon_i')}$ with $0 < b_i' < b_i$ and $\epsilon < \epsilon_i' < \epsilon$, with $\mu = b_i - b_i' = \epsilon_i - \epsilon_i'$ small enough such that they still cover the interlevel set up to the homeomorphisms φ_i:

$$f^{-1}[c - \epsilon, c + \epsilon] \subset \bigcup_{i=1,\dots,I} K_i, \tag{5}$$

where $K_i = \varphi_i(\overline{CL(b_i', \epsilon_i')}) \subset U_i$ are compact sets. Otherwise, there exist sequences $(\mu_\nu)_\nu \to 0$, $(y_\nu)_\nu \in \mathbb{R}^n$ such that we successively obtain the following:

$$y_\nu \in f^{-1}[c - \epsilon, c + \epsilon] \setminus \bigcup_{i=1,\dots,I} \varphi_i(\overline{CL((b_i')_\nu, (\epsilon_i')_\nu)})$$

$$y_\nu \in \bigcup_{i=1,\dots,I} \varphi_i(CL(b_i, \epsilon_i)) \setminus \varphi_i(\overline{CL((b_i')_\nu, (\epsilon_i')_\nu)})$$

$$y_\nu \in \bigcup_{i=1,\dots,I} \varphi_i(DL_{i,\nu}),$$

where $DL_{i,\nu} = CL(b_i, \epsilon_i) \setminus \overline{CL((b'_i)_\nu, (\epsilon'_i)_\nu)}$. This implies that $\text{dist}(y_\nu, (\cup_i U_i)^C) \to 0$. Because f is proper, up to extracting a converging subsequence, we obtain $y_\nu \to y \in f^{-1}[c - \epsilon, c + \epsilon] \setminus \cup_i U_i$, a contradiction. Therefore (5) holds for small enough μ.

Local deformations. We now define a family of local deformations corresponding to each of the U_i.

By convention, if $x \in \mathbb{R}^n$, we write $x = (\zeta, x_n)$ where $\zeta = (x_1, \ldots, x_{n-1})$.

Consider the following map: For $x \in \mathbb{R}^n$, $t \in [0, 1]$, let

$$
H_i(x, t) = \begin{cases} x & \text{if } x \notin V_i \text{ or } x \in V_i \cap \{x_n \leq -\epsilon\} \\ (\zeta, \max(-\epsilon, x_n - t\, \delta_i(x))) & \text{if } x \in V_i \cap \{x_n > -\epsilon\} \end{cases}
$$

where δ_i is defined similarly to the function introduced by [23]:

$$
\forall\, x \in V_i, \quad \delta_i(x) = \frac{\epsilon_i}{2}\left(1 - \frac{|\zeta|}{b_i}\right)\left(1 - \frac{x_n^2}{\epsilon_i^2}\right) > 0.
$$

Note that $H_i(x, 0) = x$. The quantity $\delta_i(x)$ can be understood as a maximal potential amount of decrease in the last coordinate applied to x. For a given starting point x at $t = 0$, the path formed by the points $H_i(x, \cdot)$ moves down linearly in time along the last dimension; at $t = 1$ the last coordinate has decreased by $\delta_i(x)$, unless the path hits the level $\{x_n = -\epsilon\} \cap V_i$ at an earlier timepoint $t_{\text{hit}} \leq 1$ (which means that $H_i(x, t_{\text{hit}}) = (\xi, -\epsilon)$), in which case it stays there for $t \geq t_{\text{hit}}$.

By compactness, there exists some $\eta > 0$ such that $\forall\, i = 1, \ldots, I, \forall\, x \in \varphi_i^{-1}(K_i)$, $\delta_i(x) \geq \eta$.

Let $\Pi_n(x') = x'_n$ denote the projection onto the last coordinate.

$H_i : \mathbb{R}^n \times [0, 1] \to \mathbb{R}^n$ is continuous and satisfies the following properties:

- $\forall\, x \in \mathbb{R}^n, H_i(x, 0) = x$;
- $\forall\, x \in V_i^C \cup (V_i \cap \{x_n \leq -\epsilon\}), \forall\, t \in [0, 1], H_i(x, t) = x$;
- $\forall\, x \in V_i \cap \{x_n > -\epsilon\}, -\epsilon \leq \Pi_n(H_i(x, 1)) < x_n$;
- $\forall\, x \in \varphi_i^{-1}(K_i), -\epsilon \leq \Pi_n(H_i(x, 1)) \leq \max(-\epsilon,\, x_n - \eta)$.

We can carry this deformation H_i into the original domain through the homeomorphism φ_i. We define:

$$
G_i(y, t) := \begin{cases} \varphi_i \circ H_i(\cdot, t) \circ \varphi_i^{-1} & \text{if } y \in \varphi_i(V_i) \\ y & \text{if } y \notin \varphi_i(V_i). \end{cases}
$$

$G_i : \mathbb{R}^n \times [0, 1] \to \mathbb{R}^n$ is continuous and satisfies the following:

- $\forall\, y \in \mathbb{R}^n, G_i(y, 0) = y$;
- $\forall\, y \in U_i^C \cup (U_i \cap \{f \leq c - \epsilon\}), \forall\, t \in [0, 1], G_i(y, t) = y$;

- $\forall\ y \in U_i \cap \{f > c - \epsilon\},\, c - \epsilon \leq f(G_i(y, 1)) < f(y);$
- $\forall\ y \in K_i,\, c - \epsilon \leq f(G_i(y, 1)) \leq \max(c - \epsilon,\, f(y) - \eta).$

Composition of local deformations. Now consider the map $G : \mathbb{R}^n \times [0, I]$, which results from the successive concatenation of these finitely many local deformations: if $t \in [0, 1]$, we set $G(x, t) = G_1(x, t)$; if $t \in [1, 2]$, we set $G(x, t) = G_2(\cdot, t - 1) \circ G_1(x, 1)$, and so on. If $t \in [I - 1, I]$, we set $G(x, t) = G_I(\cdot,\, t - (I - 1)) \circ G_{I-1}(\cdot, 1) \circ \cdots \circ G_1(x, 1)$. This successive concatenation is denoted by $G = G_I \square \cdots \square G_1$.

G is continuous, and furthermore, using the compact covering $f^{-1}[c - \epsilon, c + \epsilon] \subset \bigcup\limits_{i=1,\ldots,I} K_i$,

- $\forall\ y \in \mathbb{R}^n,\, G(y, 0) = y;$
- $\forall\ y \in f^{-1}(-\infty, c - \epsilon],\forall\ t \in [0, I],\, G(y, t) = y;$
- $\forall\ y \in f^{-1}[c - \epsilon, c + \epsilon],$

$$c - \epsilon \leq f(G(y, I)) \leq \max(c - \epsilon,\, c + \epsilon - \eta)$$

because y belongs to some K_i and then the value of f cannot be increased after applying the other deformations.

The last inequalities mean that $G(f^{-1}[c - \epsilon, c + \epsilon], I) \subset f^{-1}[c - \epsilon, \max(c - \epsilon, c + \epsilon - \eta)]$. In fact, for the k-times self-concatenation of G denoted by $G^{\square k}$, we similarly obtain that $G^{\square k}(f^{-1}[c - \epsilon, c + \epsilon], I) \subset f^{-1}[c - \epsilon, \max(c - \epsilon, c + \epsilon - k\eta)]$. Then, for k large enough, we have

$$G^{\square k}(f^{-1}[c - \epsilon, c + \epsilon], I) \subset f^{-1}(c - \epsilon).$$

On the other hand, since the values of f cannot increase over time through deformation, we have

$$\forall\ t \in [0, k\,I], \quad G^{\square k}(f^{-1}(-\infty, c + \epsilon], t) \subset f^{-1}(-\infty, c + \epsilon].$$

Therefore, $G^{\square k} : \mathbb{R}^n \times [0, k\,I] \to \mathbb{R}^n$, when restricted to $f^{-1}(-\infty, c + \epsilon]$ on the source and target spaces, provides a strong deformation retraction of the space sending $f^{-1}(-\infty, c + \epsilon]$ onto $f^{-1}(-\infty, c - \epsilon]$:

- $\forall\ y \in \mathbb{R}^n,\, G^{\square k}(y, 0) = y;$
- $\forall\ y \in f^{-1}(-\infty, c - \epsilon],\forall\ t \in [0, I],\, G^{\square k}(y, t) = y;$
- $\forall\ y \in f^{-1}(-\infty, c + \epsilon],\, G^{\square k}(y, k\,I) \in f^{-1}(-\infty, c - \epsilon].$

Homotopy equivalence of $f^{-1}(-\infty, a]$ and $f^{-1}(-\infty, b]$. We return to the original problem and see that for any value $c \in [a, b]$ and sufficiently small $\epsilon > 0$, the sublevel set $f^{-1}(-\infty, c - \epsilon]$ is a deformation retract of $f^{-1}(-\infty, c + \epsilon]$.

The compact interval $[a, b]$ may be covered by a finite number of such intervals associated to $c_1, \ldots, c_N$. We may suppose that the value a belongs to the last interval $[c_N - \epsilon_{c_N}, c_N + \epsilon_{c_N}]$ only. Then we obtain the desired strong deformation retraction, by successively applying the strong deformation retractions associated to $c_1, \ldots, c_{N-1}$, and then adapting the definition of the local deformations for the last value c_N, for instance:

$$H_i(x, t) = \begin{cases} x & \text{if } x \notin V_i \text{ or } x \in V_i \cap \{x_n \leq -(c_N - a)\} \\ (\zeta, \max(-(c_N - a), x_n - t \, \delta_i(x))) & \text{if } x \in V_i \cap \{x_n > -(c_N - a)\}. \end{cases}$$

This completes the proof. □

3.1.2. The handle attachment lemma

We now derive the counterpart of the handle attachment lemma of smooth Morse functions ((M2) of Theorem 2) for topological Morse functions. Having derived the isotopy lemma for topological Morse functions, we can adopt the proof in [92] for the smooth case for our purposes here.

Theorem 5 *(Handle Attachment Lemma, topological version). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous proper function. Suppose that there exists $c \in \mathbb{R}$ and $\epsilon > 0$ such that all points in $f^{-1}[c - \epsilon, c + \epsilon]$ are topological regular points for f, except a single point q that is a non-degenerate critical point with index λ and value $d(q) = c$. Then $f^{-1}(-\infty, c + \epsilon]$ has the homotopy type of $f^{-1}(-\infty, c - \epsilon]$ with a λ-cell attached:*

$$f^{-1}(-\infty, c + \epsilon] \simeq f^{-1}(-\infty, c - \epsilon] \cup e^\lambda.$$

Proof. Let $\phi : (N_1, 0) \to (N_2, q)$ be a homeomorphism such that

$$f \overset{\phi}{\sim} c - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2.$$

Furthermore, assume that $N_1 = B(0, r)$ is a small open ball around 0. Denote the quadratic part by $h(x) := -\sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2$.

Notice that by applying the Isotopy Lemma of Theorem 4 as needed, we can assume that ϵ is small enough so that the ellipsoid $\mathcal{E} = \{x = (x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1^2 + \ldots + x_\lambda^2 + 2(x_{\lambda+1}^2 + \ldots + x_n^2) \leq 2\epsilon\}$ is strictly contained inside of $B(0, r)$, while maintaining the statement of the Handle Attachment Lemma in the same form as given above.

Auxiliary function. We construct an auxiliary function F with respect to the set $U = \phi(B(0, r))$, following the prescription of [92] for the smooth case. Let $\mu : \mathbb{R} \to \mathbb{R}$ be a smooth bump function such that

$$\mu(0) > \epsilon$$
$$\mu(s) = 0 \qquad \text{for } s \geq 2\,\epsilon,$$

$$\mu'(s) \in (0,1) \qquad \forall \, s \in \mathbb{R}.$$

Let $F : \mathbb{R}^n \to \mathbb{R}$ be defined as follows:

$$F(y) = \begin{cases} f(y) & \text{if } y \notin U, \\ f(y) - \mu(\tilde{h}(\phi^{-1}(y))) & \text{if } y \in U, \end{cases} \tag{6}$$

where $\tilde{h}(x) = x_1^2 + \ldots + x_\lambda^2 + 2\,(x_{\lambda+1}^2 + \ldots + x_n^2)$. For $x \in B(0,r)$, (6) also rewrites as

$$F(\phi(x)) = c + h(x) - \mu(\tilde{h}(x)) = c - \sum_{i \leq \lambda} x_i^2 + \sum_{i \geq \lambda+1} x_i^2 - \mu \left(\sum_{i \leq \lambda} x_i^2 + 2 \sum_{i \geq \lambda+1} x_i^2 \right).$$

Because μ vanishes beyond the ellipsoid $\mathcal{E}$, which is equal to $\{x \in \mathbb{R}^n \mid \tilde{h}(x) \leq 2\,\epsilon\}$, the auxiliary and original functions coincide outside $\phi(\mathcal{E}^o)$: $\forall \, y \notin \phi(\mathcal{E}^o)$, $F(y) = f(y)$.

Retractions over regular points $F^{-1}[c - \epsilon, c + \epsilon]$. We can verify that the constructed function F is continuous and proper. We show that F only admits topological regular points in the interlevel set $F^{-1}[c - \epsilon, c + \epsilon]$.

Let $y \in F^{-1}[c - \epsilon, c + \epsilon]$. If $y \in F^{-1}[c - \epsilon, c + \epsilon] \setminus \phi(\mathcal{E})$, or equivalently, $y \in f^{-1}[c - \epsilon, c + \epsilon] \setminus \phi(\mathcal{E})$, we know that F and f coincide on a neighborhood of y. By assumption on f, there exists $\varphi : (M_1, 0) \to (M_2, y)$ such that $f \circ \varphi(x) = f(y) + x_n$, i.e., $F \circ \varphi(x) = F(y) + x_n$ by taking M_2 a small enough neighborhood around y such that $F = f$ on it.

On the other hand, if $y \in F^{-1}[c - \epsilon, c + \epsilon] \cap \phi(\mathcal{E})$, then $w = \phi^{-1}(y) \in (F \circ \phi)^{-1}[c - \epsilon, c + \epsilon] \cap \mathcal{E} \subset B(0, r)$. $F \circ \phi$ is smooth on $B(0, r)$, so based on a local integral flow defined by the non-vanishing gradient vector field around w, there exists $\theta : (M_0, 0) \to (M_1, w)$, where we can suppose $M_1 \subset B(0, r)$, such that $F \circ \phi \circ \theta(x) = F(w) + a_n$. Then by composition, F rewrites in regular form up to the homeomorphism $\phi \circ \theta : (M_0, 0) \to (\phi(M_1), y)$.

By Theorem 4 applied to F, the sublevel sets $F^{-1}(-\infty, c - \epsilon]$ and $F^{-1}(-\infty, c + \epsilon]$ are homotopy equivalent.

Handle of auxiliary function retracts into λ-cell. As in [92], we have that $\{F \leq c + \epsilon\} = \{f \leq c + \epsilon\}$. We also have that $\{F \leq c - \epsilon\} = \{f \leq c - \epsilon\} \cup L$, where $L = \{F \leq c - \epsilon\} \setminus \{f \leq c - \epsilon\} = \phi(L')$ is the image via ϕ of a handle L' in the "smooth" domain: $L' = \{h - \mu \circ \tilde{h} \leq -\epsilon\} \setminus \{h \leq -\epsilon\})$ is homeomorphic to $(D^\lambda \times D^{n-\lambda}, \partial D^\lambda \times D^{n-\lambda})$ (D^s refers to the unit disk in dimension s), is contained in the ellipsoid $\mathcal{E}$, and is attached to $\{h \leq -\epsilon\}$.

In the smooth case, [92] showed that the handle can be retracted into a λ-cell, $(e^\lambda)'$ attached to $\{h \leq -\epsilon\}$; more precisely, that $\{F \circ \phi \leq c - \epsilon\} = \{f \circ \phi \leq c - \epsilon\} \cup L' \cong \{f \circ \phi \leq c - \epsilon\} \cup (e^\lambda)'$. By composing the homeomorphism ϕ, we obtain that there is a λ-cell e^λ such that the following spaces are homotopy equivalent:

$$\{f \leq c + \epsilon\} = \{F \leq c + \epsilon\} \simeq \{F \leq c - \epsilon\} \simeq \{f \leq c - \epsilon\} \cup e^\lambda.$$

This completes the proof. □

Remark 3. As in [92] (see Theorem 2, the proof of Theorem 5 may be adapted to show that if all points in $f^{-1}[c - \epsilon, c + \epsilon]$ are regular, except for a finite number of critical points $q_1, \ldots, q_N$ sharing the same critical value c and with indices $\lambda_1, \ldots, \lambda_N$, then

$$f^{-1}(-\infty, c + \epsilon] \simeq f^{-1}(-\infty, c - \epsilon] \cup e^{\lambda_1} \cup \ldots \cup e^{\lambda_N}.$$

Topological Morse functions and persistence modules. Having established the Morse handle attachment and isotopy lemmas for topological Morse functions, we can make the usual inferences on the persistence module of the sublevel set filtration of f, and establish the same observations we had for smooth Morse functions in Corollary 1 for topological Morse functions.

Corollary 3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a proper topological Morse function with finitely many topological critical points. Then:*

- *(i) The persistence module of the sublevel set filtration f is pointwise finite dimensional (p.f.d.);*
- *(ii) For all k, the barcode $\mathrm{Bar}_k(f)$ is a finite multiset of intervals $\{[b, d) \subset \bar{\mathbb{R}}\}$;*
- *(iii) A critical point with index λ either corresponds to the birth of an interval in $\mathrm{Bar}_\lambda(f)$ with homology dimension λ, or the death of an interval in $\mathrm{Bar}_{\lambda-1}(f)$ with homology dimension $\lambda - 1$.*

3.2. Generalization to distance functions

We can now combine the results of the previous sections on Min-type functions and topological Morse functions to eventually deduce the fundamental Morse lemmas for signed distance functions.

In Section 3.2.1, we introduce the setting of general Euclidean distance functions (signed and unsigned); we discuss their properties and in particular their (non-)smoothness, and concentrate on the case of signed distance functions to smooth surfaces. Then in Section 3.2.2, we specify some geometric conditions for Euclidean distance functions to admit Min-type non-degenerate critical points, that in turn are also non-degenerate in the topological sense (Definition 10). In Section 3.2.3, we extend Morse's isotopy and handle attachment lemmas to the signed distance, based on the topological Morse theorems of Section 3.1 and the Min-type results of Section 2.2.

Distance functions, signed or unsigned, appear in many different fields, alongside the medial axis, which is the set of points whose distance-to-boundary is realized by at least two closest points. In topological data analysis, open sets were shown to be homotopy equivalent to their medial axis, by considering an extended gradient flow of the distance [85,30]. In the finite setting, the medial axis is simply the Voronoi diagram of the point cloud [85,8]. In the field of PDEs, the distance is viewed as a viscosity solution of the Eikonal equation $|\nabla d| = 1$, and the homotopy equivalence was shown in

the more general Riemannian setting using a generalized gradient flow [5]. The medial axis and distance function were also studied in non-smooth analysis and singularity theory [31,13]. The distance field (signed or not) and medial axis are involved in multiple fields of applications, ranging from computer graphics to shape analysis [81,87,105,100]. The signed version is also related to phase-field representations, that have numerous applications [107].

3.2.1. Distance function to subsets of Euclidean space

We first review some fundamental concepts and facts about distance functions to subsets of Euclidean space. For $A \subset \mathbb{R}^n$, the distance function $\text{dist}(\cdot, A) : \mathbb{R}^n \to [0, \infty)$ is given by

$$\text{dist}(q, A) = \inf_{y \in A} \|q - y\|. \tag{7}$$

If A is compact, then for any q the infimum over $y \in A$ is attained by some point in A. If $p \in A$ realizes the distance from q to A, i.e., if $\text{dist}(q, A) = \|q - p\|$, then p is a *contact point* or closest point of q in A. We let $\Gamma(q)$ denote the set of contact points:

$$\Gamma(q) = \{p \in A \mid \text{dist}(q, p) = \text{dist}(q, A)\}.$$

The *medial axis* of A is the set of points in $\mathbb{R}^n$ that admit more than one contact point in A [48]

$$\mathcal{M}_A = \{q \in \mathbb{R}^n \mid |\Gamma(q)| \geq 2\}.$$

The closure of the medial axis in $\mathbb{R}^n$ is called the *cut locus* $\overline{\mathcal{M}_A}$ of A. Off the medial axis, we let $\xi : \mathbb{R}^n \setminus \mathcal{M}_A \to A$ denote the projection map that sends a point in Euclidean space to its unique contact point in A.

As the distance function is 1-Lipschitz, Rademacher's theorem implies it is differentiable almost everywhere. The following theorem by [48] describes where the distance function is continuously differentiable and gives an expression of its derivatives.

Theorem 6. (Theorem 4.8(3-5), [48]). *For A a non-empty, closed subset of $\mathbb{R}^n$, the nearest neighbor map $\xi : \mathbb{R}^n \setminus \mathcal{M}_A \to A$ is continuous, and the distance function $\text{dist}(\cdot, A)$ is continuously differentiable on $\mathbb{R}^n \setminus (A \cup \overline{\mathcal{M}_A})$. Where $\text{dist}(\cdot, A)$ is differentiable and $q \notin A$, we have*

$$\nabla \text{dist}(q, A) = \frac{q - \xi(q)}{\|q - \xi(q)\|}. \tag{8}$$

In particular, where $\text{dist}(\cdot, A)$ is differentiable, the distance function satisfies the Eikonal equation $\|\nabla \text{dist}(\cdot, A)\| = 1$. In [85], a discontinuous vector field called the *extended gradient* $\nabla \text{dist}(q, A)$ (with a slight abuse of notation) is defined on all of $\mathbb{R}^n$,

such that it vanishes on A, coincides with the gradient of the distance function on $\mathbb{R}^n \setminus (\overline{\mathcal{M}_A} \cup A)$. The extended gradient is given explicitly as

$$\nabla \text{dist}(q, A) = \frac{q - c(q)}{\text{dist}(q, A)} \tag{9}$$

for points on the cut locus $\overline{\mathcal{M}_A}$; $c(q)$ denotes the center of the unique smallest closed ball enclosing $\Gamma(q)$. [85] shows that if A is the boundary of a bounded open set, a discrete *Euler scheme* integrating the extended gradient produces a continuous semi-flow in the limit of step size going to zero; along such a flow, the distance function is non-decreasing. [85] also shows that the points where the extended gradient vanishes obstruct the deformation retraction of sublevel sets onto another along the aforementioned semi-flow.

Signed distance functions to surfaces. If a subset A is the boundary of a bounded open subset, then we can modify the distance function to A to encompass this extra information. Let $\Omega \subset \mathbb{R}^n$ be a *non-empty bounded open subset* whose boundary $\mathcal{S} = \partial \Omega$ is C^k with $k \geq 2$. Given Ω, we partition the ambient Euclidean space as

$$\mathbb{R}^n = \Omega^- \sqcup \mathcal{S} \sqcup \Omega^+ \tag{10}$$

where $\Omega^- = \Omega$ and $\Omega^+ = (\Omega^c)^o$. We let $\text{sgn} : \mathbb{R}^n \to \{\pm 1\}$ be the function that labels $\text{sgn}(q) = -1$ if $q \in \Omega^-$ and $\text{sgn}(q) = +1$ if $q \in \Omega^+$.

The *signed distance function* $d : \mathbb{R}^n \to \mathbb{R}$ associated to an open subset Ω with boundary $\mathcal{S}$ is given by $d(q) = \text{sgn}(q) \cdot \text{dist}(q, \mathcal{S})$. Equivalently,

$$d = \text{dist}(\cdot, \Omega) - \text{dist}(\cdot, \Omega^c) = \text{dist}(\cdot, \Omega^-) - \text{dist}(\cdot, \Omega^+). \tag{11}$$

Note that the *pure distance function* $\text{dist}(\cdot, \mathcal{S}) : \mathbb{R}^n \to \mathbb{R}$ (i.e., unsigned distance function) is related to the one-sided distances by addition:

$$\text{dist}(\cdot, \mathcal{S}) = \text{dist}(\cdot, \Omega^-) + \text{dist}(\cdot, \Omega^+).$$

The partition of the ambient Euclidean space given by eq. (10) induces a partition of the medial axis into two disjoint subsets, the *inner medial axis* $\mathcal{M}_{\mathcal{S}}^- = \Omega_- \cap \mathcal{M}_{\mathcal{S}}$, and the *outer medial axis* $\mathcal{M}_{\mathcal{S}}^+ = \Omega_+ \cap \mathcal{M}_{\mathcal{S}}$. The inner and outer cut loci are similarly defined. The medial axis is illustrated in Fig. 1.

On $\mathbb{R}^n \setminus (\mathcal{S} \cup \overline{\mathcal{M}_{\mathcal{S}}})$, where $\text{dist}(\cdot, \mathcal{S})$ is differentiable, the gradient of d is

$$\nabla d(q) = \text{sgn}(q) \nabla \text{dist}(q, \mathcal{S}) = \text{sgn}(q) \frac{q - \xi(q)}{\|q - \xi(q)\|}. \tag{12}$$

On a tubular neighborhood of $\mathcal{S}$, the pure and signed distance functions inherit the smoothness of the surface: if $\mathcal{S}$ is C^k for $k \geq 2$, there exists a tubular neighborhood $\mathcal{N}_\mu = \{x \in \mathbb{R}^n \mid \text{dist}(x, \mathcal{S}) \leq \mu\}$ of $\mathcal{S}$, with $\mu > 0$, on which the signed distance d is C^k
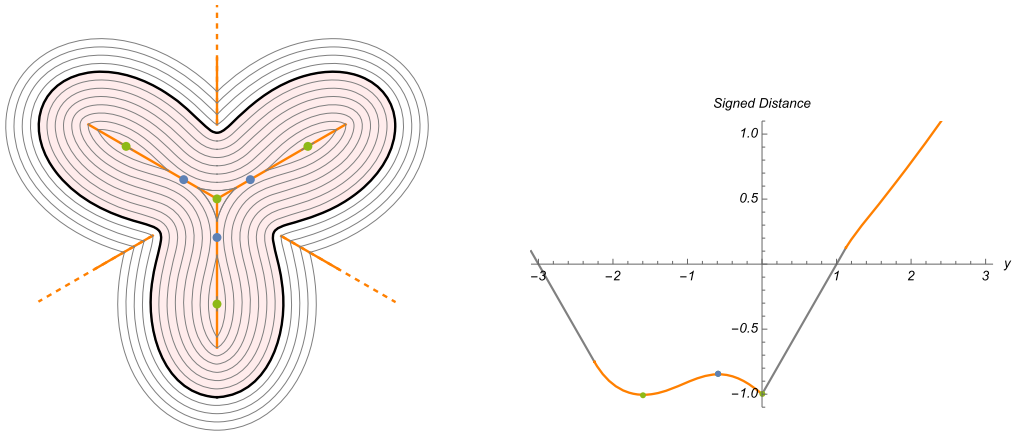
Fig. 1. Signed distance field for a petal-shaped curve. Left: level sets of the field (gray) generated by the curve (black), its cut locus (orange), and its critical points (green for local minima, blue for saddle points). The curve is parameterized by $r(\theta) = 2 + \sin 3\theta$ in polar coordinates. Right: profile of the signed distance, restricted to the vertical axis of the left subfigure. Dots and orange portions indicate the presence of the critical points and the cut locus, with blue dots for minima and green dots for saddle points. The signed distance field is not smooth in $\mathbb{R}^n$, in particular at points belonging to the cut locus. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

[42, Theorem 8.2]. In particular, the gradient of eq. (12) extends smoothly across the surface $\mathcal{S}$, and d satisfies the Eikonal equation $\|\nabla d\| = 1$ on $\mathbb{R}^n \setminus \overline{\mathcal{M}_{\mathcal{S}}}$. In contrast, the pure distance $\mathrm{dist}(\cdot, \mathcal{S})$ is C^k only in $\mathcal{N}_\mu \setminus \mathcal{S}$, with a discontinuous gradient across $\mathcal{S}$ [62, Lemma 14.16], [78].

We can similarly define an extended gradient for d, building on the extended gradient for the pure distance function. For $q \in \mathbb{R}^n \setminus \mathcal{S}$, the extended gradient ∇d (via a slight abuse of notation) of the signed distance function is given by

$$\nabla d(q) = \mathrm{sgn}(q) \cdot \nabla \mathrm{dist}(q, \mathcal{S}) = \frac{q - c(q)}{d(q)}. \tag{13}$$

These vector fields are illustrated in Fig. 2. Having associated an extended vector field with the signed distance function, we can now use it to define critical points of the signed distance function.

Definition 12 *(Critical points of distance functions, signed and unsigned).* A point $q \in \mathbb{R}^n$ is a *critical point of the signed distance function d* if its extended gradient vanishes at q:

$$\nabla d(q) = 0.$$

Notice that $q \in \mathcal{M}$ necessarily, since ∇d is of unit norm outside the medial axis.

Similarly, a point $q \in \mathbb{R}^n$ is a *critical point of the pure distance function* $\mathrm{dist}(\cdot, \mathcal{S})$ if its extended gradient $\nabla \mathrm{dist}(\cdot, \mathcal{S})(q) = 0$. In particular, any point on $\mathcal{S}$ is a critical point of $\mathrm{dist}(\cdot, \mathcal{S})$. On $\mathbb{R}^n \setminus \mathcal{S}$, $\mathrm{dist}(\cdot, \mathcal{S})$ and d share the same critical points.
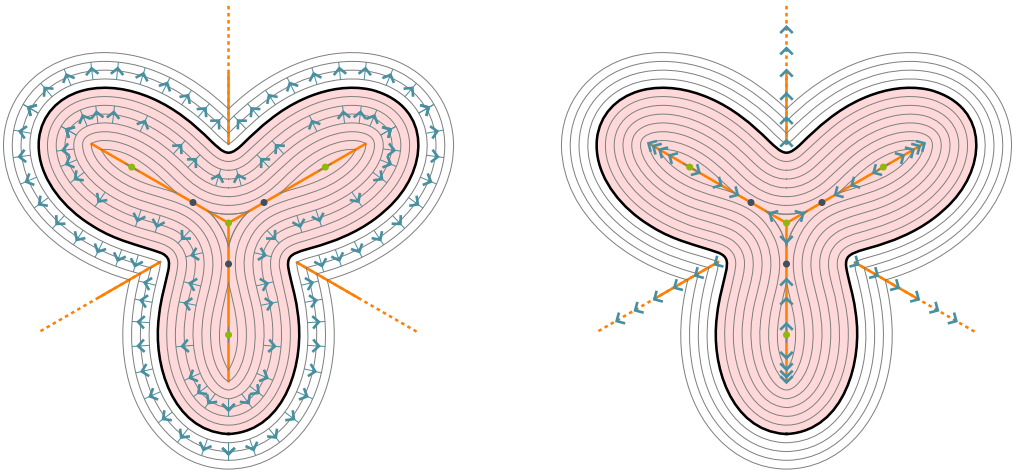
Fig. 2. Extended gradient field of the signed distance function. Left: the gradient field in $\mathbb{R}^n \setminus \overline{\mathcal{M}_{\mathcal{S}}}$, defined in the classical sense, satisfies the Eikonal equation $|\nabla d| = 1$. Right: the extended gradient field on the cut locus $\overline{\mathcal{M}_{\mathcal{S}}}$. Critical points are represented as dots, where by definition the extended gradient field vanishes, with blue for minima and green for saddle points.

We can distinguish *inner and outer critical points* that lie on $\mathcal{M}_{\mathcal{S}}^-$ and $\mathcal{M}_{\mathcal{S}}^+$, so that $\mathrm{Crit}(\mathrm{dist}(\cdot, \mathcal{S})) = \mathrm{Crit}^- \sqcup \mathcal{S} \sqcup \mathrm{Crit}^+$, $\mathrm{Crit}(\mathrm{dist}(\cdot, \Omega^-)) = \mathcal{S} \sqcup \mathrm{Crit}^+$, $\mathrm{Crit}(\mathrm{dist}(\cdot, \Omega^+)) = \mathcal{S} \sqcup \mathrm{Crit}^-$, and $\mathrm{Crit}(d) = \mathrm{Crit}^- \sqcup \mathrm{Crit}^+$.

Definition 12 can in fact be reformulated into several equivalent forms for $\mathrm{dist}(\cdot, \Omega)$ in the Euclidean setting [85,29], and for d_p in the Riemannian setting [65,31].

Lemma 7. *For $q \notin \mathcal{S}$, the following are equivalent and characterize critical points for both d and $\mathrm{dist}(\cdot, \mathcal{S})$:*

1. $\nabla d(q) = 0$;
2. $\nabla \mathrm{dist}(\cdot, \mathcal{S})(q) = 0$;
3. $q = c(q)$ is the center of the unique smallest closed ball containing $\Gamma(q)$;
4. q belongs to the closed convex hull of its contact points $\Gamma(q)$;
5. there is no open half-space containing $\Gamma(q) - q$;
6. for any vector $v \in \mathbb{R}^n$, there exists $p \in \Gamma(q)$ such that $v \cdot (p - q) \geq 0$.

Proof. Conditions 1. and 2. are equivalent, due to eq. (13). Conditions 2. and 3. are equivalent by definition of $c(q)$. Conditions 2. and 4. are known to be equivalent [29, Lemma 2.2]. Conditions 4. and 5. are equivalent due to the hyperplane separation theorem. Finally, conditions 5. and 6. are equivalent: there is an open half-space containing $\Gamma(q) - q = \{p - q\}_{p \in \Gamma(q)}$ if and only if we can find a vector v (perpendicular to the hyperplane and pointing in the other direction) such that $\forall \, p \in \Gamma(q), \, v \cdot (p - q) < 0$. $\quad \square$

Condition 4. was proposed as a definition of critical point by [49] in $\mathbb{R}^n$. Conditions 5. and 6. are adapted from [65,31], where q is critical if for any tangent vector $v \in T_q M$,

there exists a geodesic joining q to p whose speed vector $\dot{\gamma}(0)$ forms an angle less or equal to $\pi/2$ with v.

Remark 4. It can also be shown that the points where the extended gradient of the (signed) distance function vanishes on the medial axis are precisely the critical points of the Clarke subgradient of d [32]. As mentioned previously in Remark 2, the set of critical points of the extended gradient is thus a superset of the points that obstruct the deformation of sublevel sets of (signed) distance functions onto one another.

3.2.2. Min-type non-degenerate critical points of distance functions

We now give sufficient conditions where a distance function can be locally modeled as a C^k-Min-type function (Lemma 9). Using those conditions, we derive in conditions where a critical point of the distance function is a Min-type non-degenerate critical point (Definition 6) in Proposition 1 and Theorem 7, and consequently, a topological non-degenerate critical point endowed with a Morse index (Definition 10). We restrict to the setting where a point in $\mathbb{R}^n \setminus \mathcal{S}$ can only have finitely many contact points on the surface. Conditions on the surface $\mathcal{S}$, the bounded open set Ω that the surface bounds, and the signed distance function d are defined in Section 3.2.1.

We first establish conditions where the distance function $\operatorname{dist}(\cdot, \mathcal{S})$ can be described as locally Min-type at a point q, then consider additional constraints that allow us to find an efficient representation (Definition 2) of $\operatorname{dist}(\cdot, \mathcal{S})$ at q. Recall an efficient representation of a Min-type function f is a local expression

$$f = \min\{\alpha_1, \ldots, \alpha_m\}$$

where m is the minimal number of functions for any expression of f as a minimum over a collection of functions. We proceed with the following extension of Lemma 3.4 in [13], which relates $\operatorname{dist}(\cdot, \mathcal{S})$ at q to local neighborhoods of contact points $\Gamma(q)$ in $\mathcal{S}$.

Lemma 8 (*Existence of contact pieces*). *Let $q \in \mathbb{R}^n \setminus \mathcal{S}$ and suppose that $m = |\Gamma(q)| < \infty$. Denote the contact points of q by $p_1, \ldots, p_m$. Then there exists $r > \operatorname{dist}(q, \mathcal{S})$ such that $\overline{B}(q, r) \cap \mathcal{S}$ contains pairwise-disjoint connected closed subsets $S_i \subset \mathcal{S}$, and a single p_i belongs to the interior of S_i:*

$$\overline{B}(q, r) \cap \mathcal{S} \supset \bigsqcup_{i=1}^{m} S_i \quad with \quad p_i \in \overset{\circ}{S}_i. \tag{14}$$

Moreover, by setting $\alpha_i = \operatorname{dist}(\cdot, S_i)$, there exists a neighborhood N of q on which

$$\forall \, x \in N, \quad \operatorname{dist}(x, \mathcal{S}) = \min_{i=1,\ldots,m} \alpha_i(x). \tag{15}$$

The sets $\{S_i\}$ are said to form a family of contact pieces *of q.*
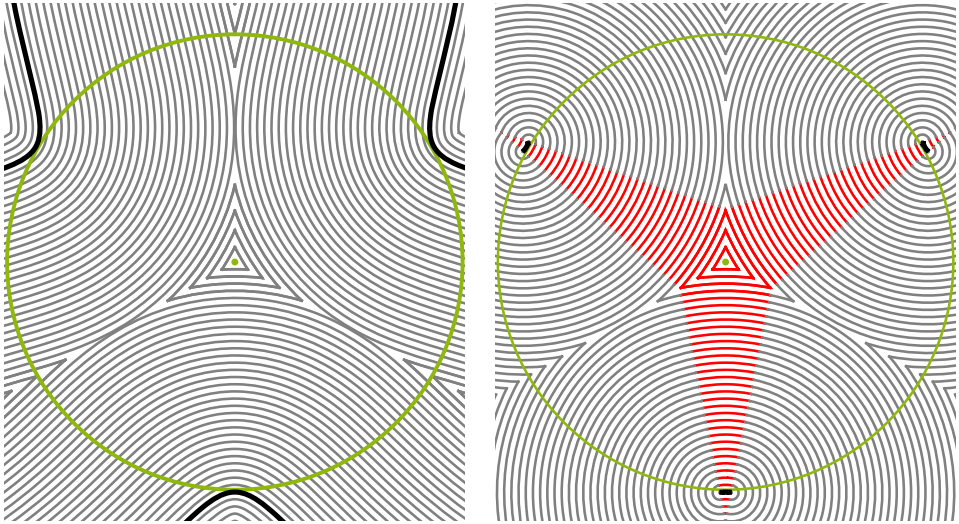
Fig. 3. Contact pieces and induced level sets. This plot compares the level sets of $\mathrm{dist}(\cdot, \mathcal{S})$ (left) v.s. the level sets of $\min\{\alpha_1, \dots, \alpha_m\}$ (right). Red-colored portions of the isolines of $\min\{\alpha_1, \dots, \alpha_m\}$ (right) are those which coincide with the isolines generated by $\mathrm{dist}(\cdot, \mathcal{S})$.

Proof. First, we show that r can be chosen such that the connected components of $\overline{B}(q, r) \cap \mathcal{S}$ contain at most one point p_i each.

Consider disjoint closed balls $\overline{B_i}$ centered at each p_i. There exists $r > \mathrm{dist}(q, \mathcal{S})$ such that $\overline{B}(q, r) \cap \mathcal{S} \subset \sqcup_i(\overline{B_i} \cap \mathcal{S})$. Otherwise, there is a sequence $(y_\nu)_{\nu \in \mathbb{N}^*}$ such that $y_\nu \in \overline{B}(q, d(q, \mathcal{S}) + 1/\nu) \cap \mathcal{S}$ and staying in $\mathcal{S} \setminus \sqcup_i(\overline{B_i} \cap \mathcal{S})$. Up to extracting a converging subsequence, (y_ν) converges to some $p \in \mathcal{S}$ with $\mathrm{dist}(q, p) = \mathrm{dist}(q, \mathcal{S})$, while p does not belong to $\sqcup_i(B_i \cap \mathcal{S})$. This means that p is a closest point of q distinct from $p_1, \dots, p_m$, which is a contradiction.

While fixing r, denote by S_i the connected component of $\overline{B}(x, r) \cap \mathcal{S}$ containing p_i, so that the inclusion (14) is satisfied. Note that necessarily, $p_i \in (S_i)^\circ$ (for the induced topology of $\mathcal{S}$).

Now, we show that there exists a ball $B(q, \epsilon)$ around q such that the equality (15) holds. First, notice that for any point x, $\mathrm{dist}(x, \mathcal{S}) \leq \min_i \mathrm{dist}(x, S_i)$. Suppose for contradiction that there is no ϵ such that $\mathrm{dist}(x, \mathcal{S}) \geq \min_i \mathrm{dist}(x, S_i)$ on $B(q, \epsilon)$. Then there would exist a sequence $(x_\nu)_{\nu \in \mathbb{N}^*}$, with $x_\nu \in B(q, 1/\nu)$, such that $\forall \nu, d(x_\nu, \mathcal{S}) < \min_i d(x_\nu, S_i)$. Up to extraction, we find contact points $y_\nu \in \mathcal{S} \setminus \sqcup_i S_i$ of x_ν that converge to some $y \in \overline{\mathcal{S} \setminus \sqcup_i S_i}$. Then y also realizes the distance $d(q, \mathcal{S})$, so y is a contact point of q, for instance p_1. But then $y \in (\mathcal{S} \setminus (S_1)^\circ) \cap (S_1)^\circ = \emptyset$, a contradiction. $\square$

This lemma states that, in the neighborhood of q, the distance field coincides with the minimum of the distance fields generated by isolated contact pieces around $\Gamma(q)$: points near q only "see" the pieces, as illustrated in Fig. 3.
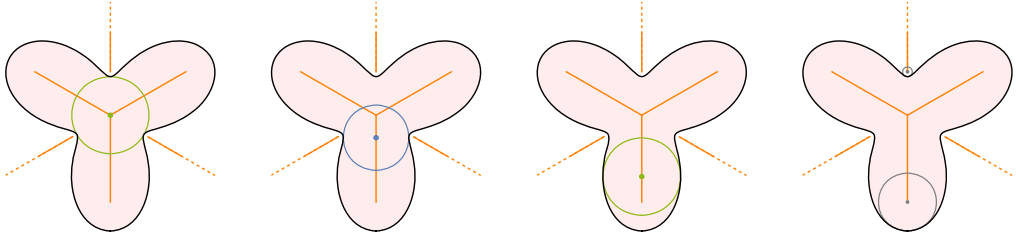
Fig. 4. Contact spheres and the strict ball condition. From left to right: contact spheres with 3, 2, 3 and 1 contact points, respectively. The first three subfigures show critical points: a local minimum, a saddle point, a local minimum, respectively, whose contact spheres all satisfy the strict ball condition. The final subfigure shows two regular points belonging to $\overline{\mathcal{M}_{\mathcal{S}}} \setminus \mathcal{M}_{\mathcal{S}}$, whose contact sphere violates the strict ball condition.

While $|\Gamma(q)| < \infty$ implies $\mathrm{dist}(\cdot, \mathcal{S})$ is a Min-type function at q, whether it is a C^k-smooth Min-type function depends on the regularity of α_i. The smoothness of α_i will be determined by the curvature of $\mathcal{S}$ at the points p_i.

The extrinsic curvature of $\mathcal{S}$ is described by the shape operator: Given the Gauss map $\mathbf{n} : \mathcal{S} \to \mathbb{S}^{n-1}$ (which points inwards into Ω^- by our convention), the *shape operator* $d_p\mathbf{n} : T_p\mathcal{S} \to T_p\mathcal{S}$ is diagonalizable in the *principal directions*:

$$d_p\mathbf{n}(e_i) = -\kappa_i\, e_i,$$

with eigenvalues being the opposite of the *principal curvatures* of $\mathcal{S}$ at p. We order them as $\kappa_{\max} = \kappa_1 \geq \kappa_2 \geq \ldots \geq \kappa_{n-1} = \kappa_{\min}$.

The factor determining whether α_i is smooth at q depends on the curvature at p and the distance between p and q. We give the precise conditions below and an illustration in Fig. 4.

Definition 13 *(Ball conditions).* Given a point $q \in \mathbb{R}^n \setminus \mathcal{S}$ and $p \in \Gamma(q)$, we say that q satisfies the *loose ball condition* at p when

$$\begin{cases} \kappa_{\max}(p) & \leq \frac{1}{\mathrm{dist}(q,\mathcal{S})} \text{ if } q \in \Omega^-, \\ \kappa_{\min}(p) & \geq -\frac{1}{\mathrm{dist}(q,\mathcal{S})} \text{ if } q \in \Omega^+. \end{cases} \tag{16}$$

When the inequalities are strict, we say that q satisfies the *strict ball condition* at p:

$$\begin{cases} \kappa_{\max}(p) & < \frac{1}{\mathrm{dist}(q,\mathcal{S})} \text{ if } q \in \Omega^-, \\ \kappa_{\min}(p) & > -\frac{1}{\mathrm{dist}(q,\mathcal{S})} \text{ if } q \in \Omega^+. \end{cases} \tag{17}$$

The lemma below says that the strict ball condition ensures that α_i is differentiable at q.

Lemma 9 *(Distance as a C^k-Min-type function). Consider $q \in \mathbb{R}^n \setminus \mathcal{S}$ and suppose that $m = |\Gamma(q)| < \infty$. Let $\{S_i\}$, and $\{\alpha_i\}$ be as in Lemma 8. Then if q satisfies the strict ball*

condition (17) *at any contact point* p_i, *the functions* α_i *are* C^k *in a neighborhood of* q. *Therefore,* dist$(\cdot, \mathcal{S})$ *is a* C^k-*Min-type function at* q, *and furthermore the representation* dist$(\cdot, \mathcal{S}) = \min\{\alpha_1, \ldots, \alpha_m\}$ *is efficient at* q.

Proof. We suppose without loss of generality that $q \in \Omega^-$. Note that if $m \geq 2$, then $q \in \mathcal{M}_{\mathcal{S}}$. In any case, $q \notin \mathcal{M}_{S_i}$ because q admits a unique closest point in each piece, by construction of the contact pieces, but this does not say whether q belongs to $\overline{\mathcal{M}_{S_i}}$ or not. In fact, the strict ball condition (17) implies that

$$\forall\, i = 1, \ldots, m, \quad q \notin \overline{\mathcal{M}_{S_i}}. \tag{18}$$

Indeed, suppose for contradiction that for some i, $q \in \overline{\mathcal{M}_{S_i}}$. Then there is a sequence of points (x_ν) converging to q that have at least two contact points on S_i. But it can be shown that there must be a point $y_\nu \in S_i$ whose curvature is greater or equal to the curvature of the contact ball of x_ν: $\kappa_{\max}(y_\nu) \geq \frac{1}{\text{dist}(x_\nu, S_i)}$, such that the sequence (y_ν) converges to p_i (using arguments similar to [97]). However, at the limit this contradicts the condition.

Thus, there is a neighborhood N of q that does not meet the closed set $\overline{\mathcal{M}_{S_i}}$. Thus, $\alpha_i = \text{dist}(\cdot, S_i)$ is differentiable on N. Any point $x \in N$ admits a unique projected point $\Pi_i(x)$ on S_i and we have $\nabla\alpha_i(x) = \frac{x - \Pi_i(x)}{d(x, \mathcal{S})} = \mathbf{n}(\Pi_i(x))$ (see Section 3.2.1).

Next, to show that α_i is C^k in a neighborhood of q, we adapt the proof of [62, Lemma 14.16].

Choose a principal coordinate system $(u, \varphi(u))$ of $\mathcal{S}$ around $p_i = (u_0, \varphi(u_0))$. Recall that $\mathcal{S}$ is locally the graph of the C^k function $\varphi : T_{p_i}\mathcal{S} \cap U \to \mathbb{R}$ where U is an open neighborhood of p_i in $\mathbb{R}^n$, and $\nabla\varphi(u_0) = 0$, and that the axes may be chosen such that the first $n-1$ coordinates are aligned with the principal directions. In this coordinate system the Hessian matrix $\text{Hess}\,\varphi(u_0) = \text{diag}\,(\kappa_1, \kappa_2, \ldots, \kappa_{n-1})$ gives the principal curvatures.

At a point $y(u) = (u, \varphi(u)) \in \mathcal{S}$ with $u \in T_{p_i}\mathcal{S} \cap U$, the inwards normal $\mathbf{n}(y(u))$ is a C^{k-1} function w.r.t. u, since

$$\mathbf{n}(y(u)) = (-\partial_1\varphi(u), \ldots, -\partial_{n-1}\varphi(u), 1) / \sqrt{1 + |\nabla\varphi(u)|^2}.$$

With respect to the principal coordinate system, we have

$$\partial_b(\mathbf{n}_a \circ y)(u_0) = -\kappa_a\, \delta_{a,b} \quad \text{for } a, b = 1, \ldots, n-1.$$

Consider the function $\Psi : (T_{p_i}\mathcal{S} \cap U) \times \mathbb{R} \to \mathbb{R}^n$ defined by

$$\Psi(u, t) = y(u) + t\,\mathbf{n}(y(u)), \quad y = (u, \varphi(u)). \tag{19}$$

Setting $r = \text{dist}(q, \mathcal{S})$, we have $\Psi(u_0, r) = p_i + r\,\mathbf{n}(p_i) = q$. We know that Ψ is C^{k-1} w.r.t. (u, t) because $\mathbf{n}$ is C^{k-1} w.r.t. u. The Jacobian of Ψ at (u_0, r) is equal to

$$D\Psi(u_0, r) = \text{diag}\,(1 - \kappa_1\, r, \dots, 1 - \kappa_{n-1}\, r, 1)$$

and is invertible by the strict ball condition (17):

$$\det(D\Psi(u_0, r)) = (1 - \kappa_{\max}\, r) \dots (1 - \kappa_{\min}\, r) > 0.$$

By the inverse mapping theorem, Ψ is invertible between neighborhoods of (u_0, r) and q and Ψ^{-1} is also C^{k-1}. Thus, around q, the functions $x \mapsto u(x)$ and $x \mapsto y(u(x)) = (u(x), \varphi(u(x)))$ are C^{k-1}. So $\nabla\alpha_i(x) = \mathbf{n}(\Pi_i(x)) = \mathbf{n}(y(u(x)))$ is also C^{k-1}. We conclude that there exists a neighborhood of q on which α_i is C^k.

We now show that $\{\alpha_1, \dots, \alpha_m\}$ is efficient. For any point x on the open ray (q, p_i) connecting q and p_i, the only closest point on $S_1 \sqcup \dots \sqcup S_m$ is p_i. This yields that the germ of (q, p_i) is a subset of $\{\forall\, j \neq i, \alpha_i < \alpha_j\} \subset A_i^\circ$. Hence, $A_i^\circ \neq \emptyset$ as a germ of set.

Let $\text{dist}(\cdot, \mathcal{S}) = \min\{\beta_1, \dots, \beta_l\}$ be another representation. We write $A_i = \cup_j (A_i \cap B_j)$ and since $A_i^\circ \neq \emptyset$, by the Baire category theorem, there exists j such that $(A_i \cap B_j)^\circ = A_i^\circ \cap B_j^\circ \neq \emptyset$, as germs of sets, up to extracting the index j. We can thus build a sequence (x_ν) converging to q such that $\nabla\alpha_i(x_\nu) = \nabla\beta_j(x_\nu)$. At the limit, $\nabla\alpha_i(q) = \nabla\beta_j(q)$ by smoothness. Since the gradients $\{\nabla\alpha_i(q)\}$ are pairwise distinct, the indices i are injectively associated to the indices j, so $m \leq l$. This shows that $\{\alpha_1, \dots, \alpha_m\}$ is an efficient representation. $\square$

Remark 5. A point q such that $|\Gamma(q)| < \infty$ already satisfies a loose ball condition (16) (see [97, Proposition 6.1]).

Lemma 9 establishes the conditions where the distance function is locally a C^k-Min-type function, equipped with an explicit efficient representation. Thus, combining Lemma 9 and Definition 6, we can now give conditions where a point is a non-degenerate Min-type critical point of pure and unsigned distance functions.

Proposition 1. *(Min-type non-degenerate critical point of the distance function) Consider $q \in \mathbb{R}^n \setminus \mathcal{S}$, where*

(N1) $\Gamma(q) = \{p_1, \dots, p_m\}$ is finite; and
(N2) The strict ball condition (eq. (17)) holds at any contact point p_i.

Then q is a Min-type non-degenerate critical point of the pure distance function $\text{dist}(\cdot, \mathcal{S})$ (and thus also of the signed distance function d), if and only if the following hold:

(N3a) $\Gamma(q)$ are in general position in $\mathbb{R}^n$ (efficient-LIG representation);
(N3b) q is strictly in the interior of the convex hull of $\Gamma(q)$; and
(N3c) The restriction $\text{dist}(\cdot, \mathcal{S})_{|G(q)}$ is a Morse function in a neighborhood of q, where $G(q)$ is the germ of set $\{\alpha_1 = \dots = \alpha_m\}$ (Definition 4), and $\{\alpha_i\}$ is the family of distances to contact pieces of q.
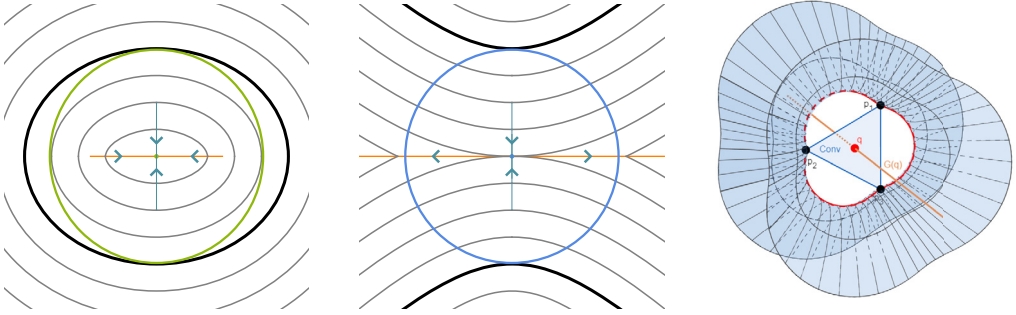
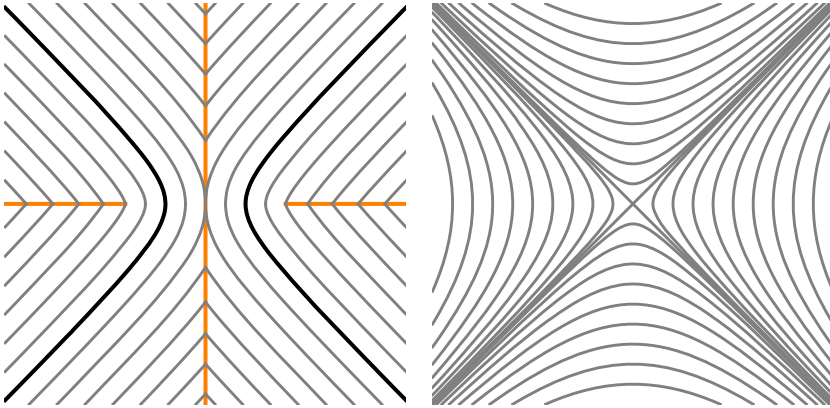**Fig. 5. Non-degenerate critical points and index.** The index can be interpreted as the sum of two terms $\lambda = (m-1) + \lambda'$ measuring how many directions where q is a local maximum, one term for the dimension of the contact subspace spanned by $\Gamma(q)$, and one term for the restriction $\mathrm{dist}(\cdot, \mathcal{S})_{|G(q)}$. Left: q admits 2 contact points in $\mathbb{R}^2$, is a local maximum of $\mathrm{dist}(\cdot, \mathcal{S})$, with index $\lambda = 1 + 1 = 2$. Middle: q admits 2 contact points in $\mathbb{R}^2$, is a saddle point of $\mathrm{dist}(\cdot, \mathcal{S})$, with index $\lambda = 1 + 0 = 1$. Right: q admits 3 contacts points in $\mathbb{R}^3$, is a critical point of $\mathrm{dist}(\cdot, \mathcal{S})$, with index $\lambda = 2 + 0 = 2$. Blue arrows are aligned with the gradient field $\nabla \mathrm{dist}(\cdot, \mathcal{S})$.
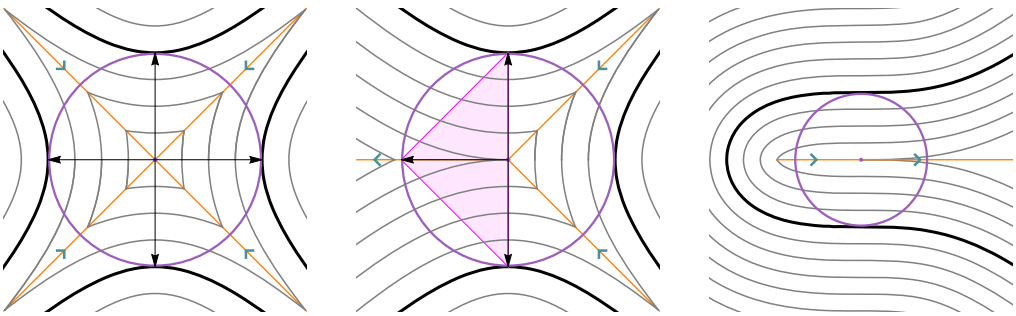
Fig. 6. Degenerate critical points in $\mathbb{R}^2$. Three examples where conditions (N3a), (N3b), and (N3c) of Proposition 1 are violated, respectively. Left: q admits 4 contact points that cannot form a 3-dimensional subspace in $\mathbb{R}^2$. Middle: q is located on the boundary of $\mathrm{Conv}(\Gamma(q))$, but not strictly inside. Right: the distance function restricted on $G(q)$ (which coincides with the medial axis here) is not Morse at q, as its Hessian is degenerate. In the three plots, q is at the center of the contact sphere; the surface is represented in black, the medial axis in orange, and the level sets in gray. Black arrows are aligned with $\nabla \alpha_i(q)$, blue arrows are aligned with $\nabla \mathrm{dist}(\cdot, \mathcal{S})$.

Note that (N3b) is a stronger condition than the one presented in Definition 12 that defines the critical points of $\mathrm{dist}(\cdot, \mathcal{S})$ and d via extended gradient vector fields. From here on, a Min-type non-degenerate critical of either a pure or signed distance function will refer to a point satisfying the conditions listed in Proposition 1. We give some examples of non-degenerate critical points in Fig. 5 and contrast them with degenerate critical points in Fig. 6.

Applying Theorem 3, we can thus obtain the Morse-like normal form for distance functions at a Min-type non-degenerate critical point q. In other words, the pure and signed distance functions at q are locally homeomorphic to the normal form of a non-degenerate critical point of a smooth function. This idea is illustrated in Fig. 7.

Fig. 7. Level sets around a non-degenerate critical point, for a distance field (left) or a smooth function (right). In these examples defined in $\mathbb{R}^2$, the index is $\lambda = 1$ (corresponding to $m = 2$ contact points and $\lambda' = 0$ on the medial axis $G(q)$ on the left). In both cases, as soon as the level crosses the critical value at q, the sublevel set encounters a topological change determined by a λ-handle attachment.

Theorem 7 *(Normal form for distance functions). Let $f = \mathrm{dist}(\cdot, \mathcal{S})$ or $f = d$ denote either the pure or signed distance function. If $q \in \mathbb{R}^n \setminus \mathcal{S}$ satisfies the conditions listed in Proposition 1, then q is a topological non-degenerate critical point (Definition 10): i.e., there exists an almost smooth homeomorphism $\phi : U \to V$ with $\phi(0) = q$, and U and $V = \phi(U)$ open neighborhoods of 0 and q in $\mathbb{R}^n$, such that for $x \in U$,*

$$f \circ \phi (x) = f(q) - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2,$$

where $\lambda = \mathrm{index}(q; f)$. Furthermore, the index of q for the pure distance function is given by

$$\mathrm{index}(q; \mathrm{dist}(\cdot, \mathcal{S})) = (m - 1) + \mathrm{index}(q; \mathrm{dist}(\cdot, \mathcal{S})_{|G(q)}); \tag{20}$$

and the index of q for the signed distance function d is given by

$$\mathrm{index}(q; d) = \begin{cases} n - \mathrm{index}(q; \mathrm{dist}(\cdot, \mathcal{S})) & \text{if } d(q) < 0, \\ \mathrm{index}(q; \mathrm{dist}(\cdot, \mathcal{S})) & \text{if } d(q) > 0. \end{cases} \tag{21}$$

As a consequence of the normal form, there can only be finitely many non-degenerate critical points for a surface $\mathcal{S}$.

Corollary 4. *The set of non-degenerate Min-type critical points of the pure or signed distance to the compact surface $\mathcal{S}$ (i.e., those that satisfy the conditions given in Proposition 1) is finite.*

Proof. From Theorem 3, the non-degenerate critical points described in Proposition 1 are isolated. Because any such critical point is contained in the convex hull of a set of points in $\mathcal{S}$, all critical points are contained in the convex hull of $\mathcal{S}$ (note that some critical points may belong to the outer set Ω^+). As the critical points are isolated and contained in a compact subset, there can only be finitely many of them. □

Remark 6. We can compare the non-degeneracy conditions and the index from Proposition 1 to those of the distance function $d_{\mathcal{P}}$ to a point a cloud $\mathcal{P}$ [16], which is always C^k-Min-type in $\mathbb{R}^n \setminus \mathcal{P}$. A point $q \in \mathbb{R}^n \setminus \mathcal{P}$ is (non-degenerate) critical with index $m-1$ if it admits m contact points $p_1, \ldots, p_m \in \mathcal{P}$ that are in general position around q, and q is strictly in the interior of the convex hull spanned by the contact points. These conditions match with conditions (N3a) and (N3b) of Proposition 1. Condition (N3c) is always satisfied: $G(q)$ in that case is the cell of Voronoi diagram (medial axis) to which q belongs. Moreover, the restriction admits a local minimum at q and the index reduces to $m-1$, the first term of eq. (20). This observation is also true of single-point distance functions d_p defined on negatively curved Riemannian manifolds M [55,73].

3.2.3. Isotopy and handle attachment lemmas

We now show that Morse's isotopy and handle attachment lemmas hold for signed distance functions. We combine the results of Section 3.1 on topological Morse functions to the existence of a normal form for d around non-degenerate points given by the Min-type theory of Section 2.2. Recall that Ω, $\mathcal{S}$, and d are defined as in Section 3.2.1.

For a value $a \in \mathbb{R}$, we define the a-sublevel set

$$V^a = d^{-1}(-\infty, a] = \{x \in \mathbb{R}^n \mid d(x) \leq a\}.$$

For two real values $a < b$, we define the $[a, b]$-interlevel set

$$V_a^b = d^{-1}[a, b] = \{x \in \mathbb{R}^n \mid a \leq d(x) \leq b\}.$$

As d is the signed distance function to a compact subset of Euclidean space, d is proper and the sublevel sets are compact.

Theorem 8 (*Isotopy lemma for signed distance functions*). *Let $a < b$ be two real values. Suppose that V_a^b contains no critical point of d and that this set is compact. Then V^a is a deformation retract of V^b, giving that V^a and V^b are homotopy-equivalent.*

Proof. d is continuous and proper. Moreover, any point $y \in V_a^b$ is in fact a topological regular point (Definition 9), due to results given by [31] and [64]. There is a gradient-like vector field $W : \mathbb{R}^n \setminus \text{Crit}(d) \to T\mathbb{R}^n$ such that for any point $y \in \mathbb{R}^n \setminus \text{Crit}(d) \to T\mathbb{R}^n$, $W(y)$ defines an open half-space strictly containing all the contacts points: $\forall\, p \in \Gamma(y)$, $W(y) \cdot (p - y) < 0$. In a neighborhood U_y of y and for $\tilde{\epsilon} > 0$ small enough, this vector field induces a homeomorphism $V_a^b \cap U_y \simeq (U_y \cap \{d = a\}) \times [a, b]$ where

$a = d(y) - \tilde{\epsilon}$ and $b = d(y) + \tilde{\epsilon}$, so that the values of d are locally represented as the last coordinate (belonging to $[a, b]$) of some homeomorphic system of coordinates. We conclude by applying Theorem 4. $\square$

Remark 7. The same result can be obtained by applying [64, Proposition 1.8] multiple times; they also show that all levels $d^{-1}(t)$ are homeomorphic for $t \in [a, b]$ and V_a^b is homeomorphic to $d^{-1}(a) \times [a, b]$.

Theorem 9 *(Handle attachment lemma for signed distance functions). Let $q \in \mathbb{R}^n \setminus \mathcal{S}$ be a non-degenerate critical point of d with index λ and value $d(q) = c$ (i.e., one that satisfies the conditions of Proposition 1). Suppose furthermore that, for some $\epsilon > 0$, the interlevel set $V_{c-\epsilon}^{c+\epsilon}$ contains no other critical point than q and that it is a compact set.*
Then $V^{c+\epsilon}$ has the homotopy type of $V^{c-\epsilon}$ with a λ-cell attached:

$$V^{c+\epsilon} \simeq V^{c-\epsilon} \cup e^{\lambda}.$$

Proof. d is a continuous and proper function. By Theorem 7, d admits a normal form $d \sim \text{cst} - \sum_{i=1}^{\lambda} x_i^2 + \sum_{i=\lambda+1}^{n} x_i^2$ at q, so q is also non-degenerate in the topological sense of Definition 10. On the other hand, any other point $y \in V_{c-\epsilon}^{c+\epsilon} \setminus \{q\}$ is a topological regular point in the sense of Definition 9 (see the proof of Theorem 8).

By applying Theorem 5 we get

$$V^{c+\epsilon} \simeq V^{c-\epsilon} \cup e^{\lambda}. \quad \square$$

Remark 8. If the critical points of d are non-degenerate, then there are only finitely many of them in any interlevel set $V_{c-\epsilon}^{c+\epsilon}$. This is due to non-degenerate critical points being isolated and the compactness of the interlevel set $V_{c-\epsilon}^{c+\epsilon}$, which follows from $\mathcal{S}$ being compact. As such, for the set of critical points $q_1, \ldots, q_N$ that lie in $d^{-1}(c)$, there is a sufficiently small ϵ such that they are the only critical points in $V_{c-\epsilon}^{c+\epsilon}$. We can use Remark 3 to deduce

$$V^{c+\epsilon} \simeq V^{c-\epsilon} \cup e^{\lambda_1} \cup \ldots \cup e^{\lambda_N}$$

where λ_i denote the indices of q_i.

3.3. Signed distance functions are Morse for generic surfaces

We have shown in Section 3.2.3 that Morse theory can be extended to the signed distance function d if critical points of d are C^k-Min-type non-degenerate; that is, if they satisfy the geometric conditions (N1)-(N3c) outlined in Proposition 1. For surfaces in $\mathbb{R}^3$, these conditions restrict the surface from forming spherical caps and cylindrical necks around q. We show that not only are these conditions satisfied for some surfaces in $\mathbb{R}^3$, but they are in fact satisfied for *generic* embeddings of a surface into $\mathbb{R}^3$. Heuristically,

this means that the conditions are true for nearly every way of embedding the manifold as a surface in $\mathbb{R}^3$.

Definition 14 *(Generic property and residual sets).* A property (P) is said to hold for *generic* elements of a topological space E, or *generically*, if $\{e \in E \mid P(e) \text{ is true}\}$ is a *residual set* in E, i.e., is the countable intersection of dense open subsets of E.

If E is a complete metric space, a residual set is nonempty; in fact, by the Baire category theorem, a residual set is dense in E. Moreover, the intersection of countably many residual sets is also residual, so that if a countable family of properties are individually generic, then they are also simultaneously generic.

In this section, we prove that the signed distance function is a topological Morse function with finitely many critical points for generic embeddings of surfaces in Theorem 13. Our proof relies on *transversality theory* [66]. Building up to the proof, we describe key concepts in transversality theory, and summarize key results in the literature which describe the structure of the cut locus for generic embeddings. This allows us to derive Theorem 13.

Before we begin, we make the class of surfaces that we are considering and the parameter space of manifold embeddings explicit. In this section, M refers to a two-dimensional C^k-smooth ($k \geq 3$) manifold, that is orientable and closed (i.e., compact without boundary). We denote $\mathrm{Emb}^k(M, \mathbb{R}^3)$ to be the space of embeddings of M into $\mathbb{R}^3$ endowed with the $C^k(M, \mathbb{R}^3)$ topology, defined in [69]. Note that for any embedding $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$, the embedded surface $\mathcal{S}_\iota = \iota(M)$ is closed and orientable.

3.3.1. Genericity of transversality

Transversality describes the fact that intersections of spaces are not tangential.

Definition 15 *(Transversality).* Two submanifolds M and N in a manifold Y *intersect transversely*, denoted by $M \pitchfork N$, if at each point of intersection $a \in M \cap N$, $T_a M + T_a N = T_a Y$. Similarly, let $f : X \to Y$ be a smooth map between smooth manifolds and denote its differential map at $x \in X$ by $D_x f : T_x X \to T_{f(x)} Y$. f is *transversal* to a submanifold W of Y and written as $f \pitchfork W$ if for any $x \in X$ such that $f(x) \in W$,

$$D_x f(T_x X) + T_{f(x)} W = T_{f(x)} Y.$$

In particular, if we have an empty intersection $\mathrm{Im}\, f \cap W = \emptyset$, then we also have $f \pitchfork W$. Note that if the intersection is not void and contains $f(x)$, transversality implies:

$$\mathrm{rank}\,(D_x f) + \dim W \geq \dim Y.$$

It is important to note that *transversality is a generic property*. For example, two lines in the plane generically intersect at one point and cross each other transversely. In three dimensions, two lines in $\mathbb{R}^3$ generically never intersect, yet they are nevertheless transversal to

each other. If two lines do intersect in $\mathbb{R}^3$, they can always be perturbed around their intersection to avoid each other by arbitrarily small perturbations.

We begin by stating the *parametric transversality theorem*, which exists in various forms for finite-dimensional [69] and infinite-dimensional manifolds [1].

Theorem 10 *(Parametric transversality theorem [1]). Let E, X, Y be C^s manifolds, $W \subset Y$ a C^s submanifold (not necessarily closed), and $F : E \times X \to Y$ such that the following conditions are satisfied:*

- *X has finite dimension $\dim_X$ and W has finite codimension codim_W in Y;*
- *E and X are second countable (i.e., their topology has a countable base);*
- *F is C^s;*
- *$s > \max\{0, \dim_X - \mathrm{codim}_W\}$;*
- *$F \pitchfork W$.*

Then the set

$$\{e \in E \mid F(e, \cdot) \pitchfork W\}$$

is residual (hence dense) in E.

In particular, E, W and Y may be infinite-dimensional. If W and Y are also finite-dimensional, then s must be strictly greater than $\max\{0, \dim X + \dim W - \dim Y\}$.

Using the parametric transversality theorem, we adopt the following proof strategy to show that the geometric conditions for the signed distance function to be topological Morse with finitely many critical points are generically satisfied (Theorem 13): We will express each condition as a transversal intersection $F(e, \cdot) \pitchfork W$ between a map $F : E \times X \to Y$ and a submanifold $W \subset Y$ where E is set to be the space of embeddings $\mathrm{Emb}^k(M, \mathbb{R}^3)$ of a manifold M into $\mathbb{R}^3$. By carefully constructing the map so that the conditions of Theorem 10 are satisfied, we can show that the set of embeddings on which the geometric conditions are satisfied is residual. To set up the proof of Theorem 13, we first describe the structure of the cut locus of generic surfaces.

3.3.2. Generic structure of the cut locus of surfaces

We aim to show that, for generic embeddings, the critical points of the signed distance function d satisfy the conditions listed in Proposition 1 to be non-degenerate Min-type critical points. These conditions describe geometric constraints between a critical point on the medial axis and its contact points on the surface. Our first step to show that these conditions are satisfied for a generic set of embeddings is to consider the possible configurations of the contact set for general points on the cut locus that need not be critical points of d. This subject is well studied in literature and we recall important results on the finiteness of the contact set and the generic local shape of the medial

axis [113,88,58,27,40]. These results are summarized in Theorem 12, which describes a classification of points on generic cut loci into five possible contact set configurations.

Using this description of the geometric configurations of contact sets, we observe that some of the conditions listed in Proposition 1 are already satisfied by some if not all types of points on the cut loci, and restrict the verification of the remaining conditions to a case-by-case analysis over a small number of contact set configurations. The following theorem by [113] shows that (N1) and (N3a) are in fact generically satisfied by all points on the cut locus.

Theorem 11 *(Finiteness of the contact set [113]). For generic embeddings $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$ (where $k \leq 2$), a point $q \in \mathbb{R} \setminus \mathcal{S}$ where $\mathcal{S} = \iota(M)$ only has a finitely many contact points with $|\Gamma(q)| = 1, 2, 3$ or 4, and $\Gamma(q)$ are in general position.*

To study other constraints such as the strict ball condition (N2), we need a finer description of the contact set $\Gamma(q)$ involving the *singularity types* of points in $\Gamma(q)$. In the following brief summary, we follow the classification of [88] and [22] and refer to [27] for a detailed discussion. For a point $q \in \mathbb{R}^3$ with finitely many contact points, and $p \in \Gamma(q)$, the squared distance function to q restricted to $\mathcal{S}$ can be expressed on the local chart at p as

$$g(x_1, x_2) = r^2 + x_1^2(1 - r\kappa_1) + x_2^2(1 - r\kappa_2) + \text{higher order terms}, \qquad (22)$$

where (x_1, x_2) are local coordinates along the principal directions of curvatures at $p = (0, 0)$, and $r = \|p - q\|$. A contact point p is an A_1 *singularity* of q if there is a diffeomorphism on the chart at p, such that eq. (22) can be written as

$$g = r^2 \pm x^2 \pm y^2.$$

In other words, p is a non-degenerate critical point of g. Note that p is an A_1 singularity of q only if the strict ball condition (eq. (17)) is true. p is an A_3 *singularity* of q if there is a diffeomorphism on the chart at p, such that eq. (22) can be written as

$$g = r^2 \pm x^2 \pm y^4.$$

Note that p is then a degenerate critical point of g and $r = 1/\kappa$ for one of the principal curvatures κ, as one of the leading quadratic terms of eq. (22) vanishes. Note that p is an A_3 singularity of q only if the strict ball condition (eq. (17)) is violated.

A complete list of singularities beyond types A_1 and A_3 can be found in [7], but as we shall see in Theorem 12, points on the cut locus only possess contact points of types A_1 and A_3 for generic surfaces. We can thus classify points on the cut locus by enumerating the number of A_1 and A_3 singularities in its contact sets; $q \in \overline{\mathcal{M}}$ is said to be of $A_1^j A_3^k$ *type* if $\Gamma(q)$ contains j and k many A_1 and A_3 singularities, respectively.

In fact, the following lemma shows that Theorem 11 implies the set of A_1^j are $(4-j)$-dimensional submanifolds. The lemma also provides some intuition on the meaning of $G(q)$ (Definition 4).

Lemma 10. *Consider a generic surface $\mathcal{S}$ satisfying the properties specified in Theorem 11. Let q be an A_1^j point for $j = 1, \ldots, 4$. Then:*

1. *There is an efficient-LIG representation of the distance function on a neighborhood U of q*

$$\operatorname{dist}(\cdot, \mathcal{S}) = \min\{\alpha_1, \ldots, \alpha_j\},$$

   where $\{\alpha_i = \operatorname{dist}(\cdot, S_i)\}$ are the distance functions to the contact pieces S_i of q as defined in Lemma 8;
2. *There is a neighborhood $V_1 \subseteq U$ of q in $\mathbb{R}^3$, such that any point $q' \in V_1$ has at most j many contact points, all of which are A_1 singularities; and*
3. *Let $\sigma = \{x \in U \mid \alpha_1(x) = \cdots = \alpha_j(x) = \operatorname{dist}(x, \mathcal{S})\} \ni q$. Then there is a neighborhood V_2 of q such that $V_2 \cap \sigma$ is a $(4-j)$-dimensional manifold.*

Hence, there is a neighborhood $V = V_1 \cap V_2$ of q, such that the set of A_1^j type points in V is the $(4-j)$-dimensional manifold $V \cap \sigma$. In other words, the set of A_1^j points is equivalent to the $(4-j)$ dimensional manifold $G(q)$ (the germ of σ at q as defined in Definition 4) as germs of sets at q.

Proof.

1. Since an A_1^j type point only has finitely many critical points and satisfy the strict ball condition (eq. (17)) by definition, Lemma 9 implies $\operatorname{dist}(\cdot, \mathcal{S})$ is C^k-Min-type with an efficient representation $\operatorname{dist}(\cdot, \mathcal{S}) = \min\{\alpha_1, \ldots, \alpha_j\}$ on a neighborhood U of q, where α_i are distance functions to disjoint subsets S_i of $\mathcal{S}$, each a neighborhood of $p_i \in \Gamma(q)$ in $\mathcal{S}$ as given by Lemma 8. The representation is efficient due to $\Gamma(q)$ being in general position as given in Theorem 11.
2. Consider $q' \in U$. Since we have on U an efficient-LIG representation of the distance function, q' has $j' \leq j$ contact pieces $S_1, \ldots, S_{j'}$, which are also the contact pieces of q specified in the efficient-LIG representation; furthermore, since we have assumed the conditions specified in Theorem 11, the contact set is finite $\Gamma(q') = \{p'_1, \ldots, p'_j\}$ and $p'_i = \Gamma(q') \cap S_i$. We now investigate the singularity types of p'_i. Recall an isolated point $p'_i \in \Gamma(q')$ is an A_1 singularity of q' iff the principal curvatures at $p'_i \in \mathcal{S}$ are less than $r' = \|q' - p'_i\|$. We note that on U, the distance to the pieces $\alpha_i = \operatorname{dist}(\cdot, S_i)$ and the nearest neighbor projection map $\xi_i : U \to S_i$ onto the contact piece S_i are continuous. We also observe that the principal curvatures $\kappa_\ell : \mathcal{S} \to \mathbb{R}$ for $\ell = 1, 2$ are also continuous. We can thus construct a continuous map $\eta_i : U \to \mathbb{R}$ given by

$\eta_i(x) = 1 - \alpha_i(x)\kappa_{\max}(\xi_i(x))$, where $\kappa_{\max} := \max\{\kappa_1, \kappa_2\}$. Note that $p'_i \in \Gamma(q')$ is an A_1 singularity of q'_i iff $\eta_i(q'_i) > 0$. Since q is A_1^j type, we have $\eta_i(q) > 0$ for all i. By continuity, there is a sufficiently small neighborhood W_i of q for each map η_i, such that $\eta_i > 0$ on W_i and the contact point p'_i of $q' \in W_i$ in S_i is an A_1 singularity. We can then take $V_1 = \cap_{i=1}^j W_i$ to be the neighborhood of q on which any $q' \in V_1$ has at most j many contact points, all of which are A_1 singularities.

3. Since the representation is efficient-LIG, Lemma 5 implies there is a neighborhood $V_2 \subset U$ such that $V_2 \cap \sigma$ is a $(4 - j)$-dimensional manifold.

It follows that for $V = V_1 \cap V_2$, the submanifold $V \cap \sigma$ are A_1^j type points; furthermore, since A_1^j points on U must belong to σ by definition, $V \cap \sigma$ is precisely the set of A_1^j points in the neighborhood V of q. $\quad\square$

Lemma 10 shows that there are submanifolds in generic cut loci consisting of A_1^j points. This is further developed by [88], who showed that generic cut loci can be decomposed into manifolds of the same $A_1^j A_3^k$ type. Specifically, the decomposition is a *Whitney stratification*: the manifolds in the decomposition, called *strata*, fit together following local topological constraints that ensure the stratified space is well-behaved. The full definition of Whitney stratification can be found in [63]; for our purposes, the theorem below gives an explicit geometric picture of how the strata fit together for generic cut loci of surfaces.

Theorem 12 (*Generic shape of the cut locus [88,58,27,40]*). *The cut locus $\overline{\mathcal{M}}$ of a generic embedding ι of a smooth surface into $\mathbb{R}^3$ is a two-dimensional Whitney-stratified set whose strata belong to one of five classes, distinguished by the number of contact points $q \in \overline{\mathcal{M}}$ has on $\mathcal{S}$ and the types of singularities of $\mathrm{dist}(\cdot, q)^2|_{\mathcal{S}}$ at the contact points. Specifically, a stratum of $\overline{\mathcal{M}}$ can either be:*

- *A two-dimensional stratum (or sheet) of A_1^2 points, each of which having two contact points on $\mathcal{S}$ of type A_1;*
- *A one-dimensional stratum (or curve) of A_1^3 points, each of which has three contact points on $\mathcal{S}$ of type A_1. They are curves which lie at the intersection of three A_1^2 sheets;*
- *A zero-dimensional stratum consisting of an A_1^4 point which has four contact points of type A_1. The A_1^4 strata are isolated and lie at the intersection of six A_1^2 sheets and four A_1^3 curves along the boundaries of the A_1^2 sheets;*
- *A one-dimensional stratum of A_3 edge points, each of which has one A_3 contact point on $\mathcal{S}$. Thus, they do not belong to the medial axis $\mathcal{M}$ itself. They bound A_1^2 sheets in $\mathbb{R}^3$;*
- *A zero-dimensional stratum consisting of an $A_1 A_3$ 'fin creation' point, which has two contact points on $\mathcal{S}$ of types A_1 and A_3 respectively. A fin creation point is a common*

end point of two one-dimensional strata of types A_1^3 and A_3 respectively, and the two A_1^3 and A_3 curves bound a common A_1^2 sheet. Fin creation points are isolated.

3.3.3. Distance critical points are generically non-degenerate

We are now in a position to prove our main result: generic signed distance functions are topological Morse functions with finitely many critical points that are all non-degenerate (Theorem 13). Our proof proceeds by showing that each of the conditions listed in Proposition 1 holds on a residual set of embeddings. If this is the case, then the set of embeddings on which all conditions hold is an intersection of residual sets, which is itself residual. As we have shown in the previous section, some of the conditions are already true as generic properties of points on the cut locus. Having validated that (N1) and (N3a) hold for all points on generic cut loci in Theorem 11, we verify in Proposition 4 that the strict ball condition (N2) is generically satisfied by showing that only A_1^j points may be critical, but not $A_1 A_3$ points.

Further, in Proposition 2 and Proposition 3, we also show that (N3b) and (N3c) are generic conditions on A_1^2, A_1^3, and A_1^4 type critical points.

Our proofs rely on the parametric transversality theorem (Theorem 10), which was also used to derive the generic description of the cut locus in Theorem 12. We make repeated use of the following transversality argument: We first show that, for a given embedding $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$, a property listed in Proposition 1 is true if some explicitly constructed map $f^\iota : X \to Y$ is transversal to a submanifold $W \subset Y$. To prove that $f^\iota \pitchfork W$ for generic ι, we consider the function f^ι as a member of a family of maps $F : \mathrm{Emb}^k(M, \mathbb{R}^3) \times X \to Y$ parameterized over the space of embeddings $\mathrm{Emb}^k(M, \mathbb{R}^3)$, such that $f^\iota = F(\iota, \cdot)$. By concluding that F is a submersion (and thus transversal to any arbitrary submanifold of Y, in particular W), Theorem 10 then implies $f^\iota \pitchfork W$ for generic ι. We note that the sufficient conditions for some properties listed in Proposition 1 correspond to having a void intersection $\mathrm{Im}\, f^\iota \cap W = \emptyset$. Showing this property for generic ι can be simply achieved by finding a dimensional insufficiency of the form $\mathrm{rank}\, D f^\iota + \dim W \leq \dim X + \dim W < \dim Y$, as transversality of F will then imply a void intersection between $\mathrm{Im}\, f^\iota$ and W.

We can then conclude that the set of sufficient conditions listed in Proposition 1 for a critical point to be non-degenerate is satisfied on a residual subset of the space of embeddings $\mathrm{Emb}^k(M, \mathbb{R}^3)$.

In the construction of the maps F, we choose manifolds X and Y such that the other assumptions of the parametric transversality theorem (Theorem 10) are satisfied. We choose finite-dimensional spaces for X, Y, and W so that either $\dim X + \dim W - \dim Y = 0$ or -1. $E = \mathrm{Emb}^k(M, \mathbb{R}^3)$ and X will be second countable, F will be C^s with $s \geq k - 2 > \max\{0, \dim X + \dim W - \dim Y\} = 0$. Our constructions of F will also be sufficiently smooth submersions for the situation at hand; this is shown in Lemma 11 (Appendix).

We set out the proof in several steps. The proof mainly deals with the pure distance $\mathrm{dist}(\cdot, \mathcal{S})$, which is sufficient to derive the results for the signed distance. Note that critical

points of d directly correspond to those of $\text{dist}(\cdot, \mathcal{S})$ that are not on $\mathcal{S}$ (Definition 12), so in what follows we consider critical points $q \notin \mathcal{S}$. We make a few notational remarks before we embark on the proofs.

Remark 9. To keep concise notations, we may drop the index from the notation of the objects associated to an embedding ι, such as from $\mathcal{S}_\iota$ and Ω_ι.

Remark 10 *(Notations for the Gauss map).* We set up some shorthand notation for denoting the Gauss map. By the Jordan–Brouwer separation theorem for hypersurfaces [86,91], $\mathcal{S}_\iota$ divides $\mathbb{R}^3 \backslash \mathcal{S}_\iota$ into an "inside" region, denoted by $\Omega_\iota = \Omega_\iota^-$, which is a bounded open set with possibly multiple connected components, and an "outside" region $\mathbb{R}^3 \backslash \overline{\Omega_\iota} = \Omega_\iota^+$. For $\iota \in \text{Emb}^k(M, \mathbb{R}^3)$, let $\mathbf{n} = \mathbf{n}_\iota : \mathcal{S} \to \mathbb{S}^2$ denote the Gauss map, i.e., the globally-defined unit vector field normal to the surface that points inwards (in the direction of Ω^- if $\mathcal{S} = \partial\Omega^-$). Note that $\mathbf{n}_\iota$ is orthogonal to the image of $D_{m_i}\iota : T_{m_i}M \to T_{\iota(m_i)}\mathbb{R}^3$, which is the tangent space of the surface $T_{\iota(m_i)}\mathcal{S} = \text{Im}\, D_m\iota$. Since the embedding ι is implicit in the notation $\mathbf{n}$, we drop the index in $\mathbf{n}_\iota$. We abuse notation and also use $\mathbf{n}$ to denote the Gauss map when we compose $\mathbf{n} : \mathcal{S} \to \mathbb{S}^2$ with the projection map of the tubular neighborhood $\text{Tub}(\mathcal{S})$ onto $\mathcal{S}$, or the inclusion of $\mathbb{S}^2$ into $\mathbb{R}^3$. Hence $\mathbf{n}$ is a shorthand for four different maps, whose domain can either be $\mathcal{S}$ or $\text{Tub}(\mathcal{S})$, and whose codomain is either $\mathbb{S}^2$ or $\mathbb{R}^3$. We will specify the source and target spaces explicitly to avoid ambiguity when we reference $\mathbf{n}$ in the subsequent proofs.

Proposition 2. *Consider a critical point q of the distance function of type A_1^j. Then it is a Min-type critical point (Definition 5) of $\text{dist}(\cdot, \mathcal{S})$. Furthermore, there is a residual subset of embeddings $\iota \in \text{Emb}^k(M, \mathbb{R}^3)$ such that A_1^j critical points q of the distance function are non-degenerate critical points of $\text{dist}(\cdot, \mathcal{S})_{|G(q)}$, the smooth function that is the restriction of $\text{dist}(\cdot, \mathcal{S})$ to the germ $G(q)$ of the submanifold of A_1^j points (as given in Lemma 10). Thus, A_1^j type distance critical points satisfy (N3c).*

Proof. Recall for q an A_1^j type, we have an efficient-LIG representation $\text{dist}(\cdot, \mathcal{S}) = \min\{\alpha_1, \ldots, \alpha_j\}$, and the germ of the stratum $G(q)$ is a $(4 - j)$ dimensional submanifold (Lemma 10). Recall Lemma 7 that q is a distance critical point iff q is in the convex hull of its contact points: i.e., there is a set of non-negative coefficients t_i that sum to one, such that $q = \sum_{i=1}^{j} t_i p_i$. Since $\nabla\alpha_i|_q = (q - p_i)/r$ where $r = \|p_1 - q\| = \cdots = \|p_j - q\|$, we can rewrite this condition as $\sum_{i=1}^{j} t_i \nabla\alpha_i|_q = 0$. In other words, the origin is contained in the convex hull of $\{\nabla\alpha_i|_q\}_{i=1,\ldots,j}$. Thus, q is a Min-type critical point of $\text{dist}(\cdot, \mathcal{S})$ (Definition 5); furthermore, q is a critical point of $\text{dist}(\cdot, \mathcal{S})_{|G_f(q)}$ (Lemma 6).

We now show that the Riemannian Hessian of the pure distance restricted to the submanifold $G(q)$ is non-degenerate at q. Without loss of generality, we assume that $q \in \Omega_\iota^-$ is "inside". Note that the explicit expression of the Riemannian Hessian of distance functions restricted on $G(q)$ is given in Lemma 12 and Lemma 13 of the Appendix, derived using the shape operator of offset surfaces. We proceed by a case by case analysis for $j = 2, 3, 4$.

A_1^2 **critical points**. We start by establishing the proof for a point q with two A_1 contact points. In what follows, $M^{(i)}$ refers to the set of i-tuples of distinct points in M.

Let $X = M^{(2)} \times \mathbb{R}_{>0}$ and $Y = (\mathbb{S}^2)^2 \times (\mathbb{R}^3)^2$. Consider the submanifold $W = \{y = (y_1, \ldots, y_4) \in Y \mid y_1 + y_2 = 0$ and $y_3 - y_4 = 0\}$ of dimension 5 in Y. Define

$$
F : \begin{array}{ccc}
\operatorname{Emb}^k(M, \mathbb{R}^3) \times X & \to & Y \\
(\iota, m_1, m_2, r) & \mapsto & (n_1, n_2, p_1 + r\, n_1, p_2 + r\, n_2)
\end{array}
\tag{23}
$$

where $p_i = \iota(m_i)$ and $n_i = \mathbf{n}(p_i)$. In the expression $p_i + r\, n_i$, n_i is viewed in $\mathbb{R}^3$.

Let us explain the geometric meaning of $f^\iota(x) \in W$, where $f^\iota = F(\iota, \cdot)$: Given some embedding ι, if q is of type A_1^2 with contact points $\{p_1, p_2\}$, then setting $m_i = \iota^{-1}(p_i)$, $r = \operatorname{dist}(q, \mathcal{S})$ and $x = (m_1, m_2, r) = x(q)$, we have $n_1 + n_2 = 0$ and $p_1 + r\, n_1 = p_2 + r\, n_2 = q$. This implies that $f^\iota(x(q)) \in W$. However, note that the relationship $f^\iota(x) \in W$ also describes points $x \in X$ that do not necessarily correspond to A_1^2 critical points, as it does not specify, for instance, whether the p_i should be the closest points on the surface.

We investigate what it means for f^ι to be transversal to W. For a fixed embedding ι, we compute the differential of $F(\iota, \cdot)$ with respect to $x = (m_1, m_2, r)$, obtaining

$$
D_x F(\iota, \cdot) = \begin{array}{c}
\begin{array}{ccc}
T_{m_1} M \qquad\quad & T_{m_2} M \qquad & \mathbb{R}
\end{array} \\
\left[\begin{array}{ccc}
d_{p_1} \mathbf{n} \circ d_{m_1} \iota & 0 & 0 \\
0 & d_{p_2} \mathbf{n} \circ d_{m_2} \iota & 0 \\
(\operatorname{Id} + r\, d_{p_1} \mathbf{n}) \circ d_{m_1} \iota & 0 & n_1 \\
0 & (\operatorname{Id} + r\, d_{p_2} \mathbf{n}) \circ d_{m_2} \iota & n_2
\end{array}\right]
\begin{array}{c}
T_{n_1} \mathbb{S}^2 \\
T_{n_2} \mathbb{S}^2 \\
\mathbb{R}^3 \\
\mathbb{R}^3
\end{array}
\end{array} .
$$

Then $f^\iota \pitchfork W$ if and only if for $x \in X$ such that $f^\iota(x) \in W$, the following map is of rank $\operatorname{codim} W = 5$:

$$
\begin{array}{c}
\begin{array}{ccc}
T_{m_1} M \qquad\quad & T_{m_2} M \qquad & \mathbb{R}
\end{array} \\
\left[\begin{array}{ccc}
d_{p_1} \mathbf{n} \circ d_{m_1} \iota & d_{p_2} \mathbf{n} \circ d_{m_2} \iota & 0 \\
(\operatorname{Id} + r\, d_{p_1} \mathbf{n}) \circ d_{m_1} \iota & -(\operatorname{Id} + r\, d_{p_2} \mathbf{n}) \circ d_{m_2} \iota & n_1 - n_2
\end{array}\right]
\begin{array}{c}
\mathbb{R}^3 \\
\mathbb{R}^3
\end{array}
\end{array} .
$$

Up to the isomorphism

$$
\begin{array}{c}
\begin{array}{ccc}
T_{m_1} M & T_{m_2} M & \mathbb{R}
\end{array} \\
\left[\begin{array}{ccc}
d_{m_1} \iota & 0 & 0 \\
0 & d_{m_2} \iota & 0 \\
0 & 0 & 1
\end{array}\right]
\begin{array}{c}
T_{p_1} \mathcal{S} \\
T_{p_2} \mathcal{S} \\
\mathbb{R}
\end{array}
\end{array} ,
$$

we can consider this map instead

$$\begin{array}{ccc} T_{p_1}\mathcal{S} & T_{p_2}\mathcal{S} & \mathbb{R} \\ \left[\begin{array}{ccc} d_{p_1}\mathbf{n} & d_{p_2}\mathbf{n} & 0 \\ (\mathrm{Id} + r\,d_{p_1}\mathbf{n}) & -(\mathrm{Id} + r\,d_{p_2}\mathbf{n}) & n_1 - n_2 \end{array}\right] & \begin{array}{c} \mathbb{R}^3 \\ \mathbb{R}^3 \end{array} \end{array}.$$

For x such that $f^\iota(x) \in W$, we have in particular that $n_2 = -n_1$ so that $n_1 - n_2 = 2n_1$, and the tangent spaces $T_{p_1}\mathcal{S} \simeq T_{p_2}\mathcal{S}$ are parallel so they may be identified as a single space denoted $T_{p_i}\mathcal{S}$. Still abusing notation, we may consider $d_{p_1}\mathbf{n}$ and $d_{p_2}\mathbf{n}$ as maps $T_{p_i}\mathcal{S} \to T_{p_i}\mathcal{S}$.

The previous rank condition is then equivalent to the following ones:

$$f^\iota \pitchfork W \Leftrightarrow \quad \forall\, x \in X \text{ s.t. } f^\iota(x) \in W,$$

$$\begin{array}{ccc} T_{p_i}\mathcal{S} & T_{p_i}\mathcal{S} & \mathbb{R} \\ \left[\begin{array}{ccc} d_{p_1}\mathbf{n} & d_{p_2}\mathbf{n} & 0 \\ \mathrm{Id} + r\,d_{p_1}\mathbf{n} & -(\mathrm{Id} + r\,d_{p_2}\mathbf{n}) & n_1 - n_2 \end{array}\right] & \begin{array}{c} T_{p_i}\mathcal{S} \\ \mathbb{R}^3 \end{array} \end{array} \text{ is of rank 5}$$

$$\Leftrightarrow \quad \forall\, x \in X \text{ s.t. } f^\iota(x) \in W,$$

$$\begin{array}{ccc} T_{p_i}\mathcal{S} & T_{p_i}\mathcal{S} & \mathbb{R} \\ \left[\begin{array}{ccc} d_{p_1}\mathbf{n} & d_{p_2}\mathbf{n} & 0 \\ \mathrm{Id} + r\,d_{p_1}\mathbf{n} & -(\mathrm{Id} + r\,d_{p_2}\mathbf{n}) & 0 \\ 0 & 0 & 2 \end{array}\right] & \begin{array}{c} T_{p_i}\mathcal{S} \\ T_{p_i}\mathcal{S} \\ \mathbb{R}\,n_1 \end{array} \end{array} \text{ is of rank 5}$$

$$\Leftrightarrow \quad \forall\, x \in X \text{ s.t. } f^\iota(x) \in W,$$

$$\begin{array}{cc} T_{p_i}\mathcal{S} & T_{p_i}\mathcal{S} \\ \left[\begin{array}{cc} d_{p_1}\mathbf{n} & d_{p_2}\mathbf{n} \\ \mathrm{Id} + r\,d_{p_1}\mathbf{n} & -(\mathrm{Id} + r\,d_{p_2}\mathbf{n}) \end{array}\right] & \begin{array}{c} T_{p_i}\mathcal{S} \\ T_{p_i}\mathcal{S} \end{array} \end{array} \text{ is of rank 4.}$$

The last condition involves a block matrix whose determinant can be expressed simply, thanks to a formula due to [106]:

$$AC = CA \quad \Rightarrow \quad \det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - CB).$$

Since $d_{p_1}\mathbf{n}$ commutes with $(\mathrm{Id} + r\,d_{p_1}\mathbf{n})$, we have

$$f^\iota \pitchfork W \Leftrightarrow \forall\, x \in X \text{ s.t. } f^\iota(x) \in W,$$

$$\det\left(d_{p_1}\mathbf{n} \circ (\mathrm{Id} + r\,d_{p_2}\mathbf{n}) + (\mathrm{Id} + r\,d_{p_1}\mathbf{n}) \circ d_{p_2}\mathbf{n}\right) \neq 0.$$

By composing with the isomorphisms $(\mathrm{Id} + r\, d_{p_1}\mathbf{n})^{-1}$ and $(\mathrm{Id} + r\, d_{p_2}\mathbf{n})^{-1}$ on the left and the right sides, this means

$$f^\iota \pitchfork W \Leftrightarrow \forall\, x \in X \text{ s.t. } f^\iota(x) \in W,$$
$$\det\left((\mathrm{Id} + r\, d_{p_1}\mathbf{n})^{-1} \circ d_{p_1}\mathbf{n} + d_{p_2}\mathbf{n} \circ (\mathrm{Id} + r\, d_{p_2}\mathbf{n})^{-1}\right) \neq 0.$$

Finally, $d_{p_1}\mathbf{n}$ also commutes with $(\mathrm{Id} + r\, d_{p_1}\mathbf{n})^{-1}$, and so

$$f^\iota \pitchfork W \Leftrightarrow \forall\, x \in X \text{ s.t. } f^\iota(x) \in W,$$
$$\det\left(d_{p_1}\mathbf{n} \circ (\mathrm{Id} + r\, d_{p_1}\mathbf{n})^{-1} + d_{p_2}\mathbf{n} \circ (\mathrm{Id} + r\, d_{p_2}\mathbf{n})^{-1}\right) \neq 0.$$

Notice that the last inequality gives non-degeneracy of the Riemannian Hessian from Lemma 13. So, for generic embeddings ι, we obtain

$$f^\iota \pitchfork W \Rightarrow \mathrm{Hess}^{\mathrm{Riem}} g(q) \text{ is non-degenerate for any critical point } q \text{ of type } A_1^2.$$

It remains to verify that $f^\iota \pitchfork W$ for generic embeddings $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$. But this is true by the transversality argument stated above, because F is a submersion (Lemma 11).

A_1^3 **critical points.** Likewise, for a point q of type A_1^3, we define some well-chosen spaces and maps.

Let $X = M^{(3)} \times \mathbb{R}_{>0}$ and $Y = (\mathbb{S}^2)^3 \times (\mathbb{R}^3)^3$. For two vectors $u, v \in \mathbb{R}^3$, we denote the usual scalar product by $u \cdot v$ and the vector product by $u \times v$. Consider the submanifold

$$W = \{y = (y_1, \ldots, y_6) \in Y \mid y_3 \cdot (y_1 \times y_2) = 0 \text{ and } y_4 = y_5 = y_6\}$$

which is of dimension in 8 and codimension 7 in Y. Note that the triple product is equal to $y_3 \cdot (y_1 \times y_2) = y_1 \cdot (y_2 \times y_3) = y_2 \cdot (y_3 \times y_1)$, and that it vanishes if and only if one of the three vectors belongs the subspace spanned by the two others (which is of dimension 1 or 2). There is a map ℓ such that $W = \{\ell = 0\} \cap Y \subset E$, with $E = (\mathbb{R}^3)^6$ and $\ell(e_1, \ldots, e_6) = (e_3 \cdot (e_1 \times e_2), e_4 - e_5, e_4 - e_6)$. The tangent space to $\{\ell = 0\}$ at a point in E is given by

$$\mathrm{Ker}\, \nabla\ell = \{(\epsilon_1, \ldots, \epsilon_6) \in E \mid (e_2 \times e_3) \cdot \epsilon_1 + (e_3 \times e_1) \cdot \epsilon_2 + (e_1 \times e_2) \cdot \epsilon_3 = 0\}$$
$$\cap \{(\epsilon_1, \ldots, \epsilon_6) \in E \mid \epsilon_4 - \epsilon_5 = 0 \text{ and } \epsilon_4 - \epsilon_6 = 0\}.$$

Now, we introduce the following map:

$$F : \begin{matrix} Emb \times X & \to & Y \\ (\iota, m_1, m_2, m_3, r) & \mapsto & (n_1, n_2, n_3, p_1 + rn_1, p_2 + rn_2, p_3 + rn_3) \end{matrix}. \tag{24}$$

Again, we study what it means for $f^\iota = F(\iota, \cdot)$ to be transversal to W.

Given an embedding ι, the differential at $x = (m_1, m_2, m_3, r)$ is equal to

$$D_x F(\iota, x) =$$

$$
\begin{array}{cccc}
T_{m_1} M & T_{m_2} M & T_{m_3} M & \mathbb{R} \\
\left[\begin{array}{cccc}
d_{p_1} \mathbf{n} \circ d_{m_1} \iota & 0 & 0 & 0 \\
0 & d_{p_2} \mathbf{n} \circ d_{m_2} \iota & 0 & 0 \\
0 & 0 & d_{p_3} \mathbf{n} \circ d_{m_3} \iota & 0 \\
(\mathrm{Id} + r\, d_{p_1} \mathbf{n}) \circ d_{m_1} \iota & 0 & 0 & n_1 \\
0 & (\mathrm{Id} + r\, d_{p_2} \mathbf{n}) \circ d_{m_2} \iota & 0 & n_2 \\
0 & 0 & (\mathrm{Id} + r\, d_{p_3} \mathbf{n}) \circ d_{m_3} \iota & n_3
\end{array}\right]
& \begin{array}{c}
T_{n_1} \mathbb{S}^2 \\
T_{n_2} \mathbb{S}^2 \\
T_{n_3} \mathbb{S}^2 \\
\mathbb{R}^3 \\
\mathbb{R}^3 \\
\mathbb{R}^3
\end{array}
\end{array}.
$$

Then, as before, we study its projection on $\operatorname{Ker}(\nabla \ell)^{\perp} = \operatorname{Im}((\nabla \ell)^T)$ in E. We obtain that $f^\iota \pitchfork W$ if and only if $\forall\, x \in X$ s.t. $f^\iota(x) \in W$,

$$
\begin{array}{cccc}
T_{p_1} \mathcal{S} & T_{p_2} \mathcal{S} & T_{p_3} \mathcal{S} & \mathbb{R} \\
\left[\begin{array}{cccc}
(n_2 \times n_3) \cdot d_{p_1} \mathbf{n} & (n_3 \times n_1) \cdot d_{p_2} \mathbf{n} & (n_1 \times n_2) \cdot d_{p_3} \mathbf{n} & 0 \\
\mathrm{Id} + r\, d_{p_1} \mathbf{n} & -(\mathrm{Id} + r\, d_{p_2} \mathbf{n}) & 0 & n_1 - n_2 \\
\mathrm{Id} + r\, d_{p_1} \mathbf{n} & 0 & -(\mathrm{Id} + r\, d_{p_3} \mathbf{n}) & n_1 - n_3
\end{array}\right]
& \begin{array}{c}
\mathbb{R} \\
\mathbb{R}^3 \\
\mathbb{R}^3
\end{array}
\end{array}
\quad \text{is of rank 7.}
$$

In particular, if q is of type A_1^3 with contact points $\{p_1, p_2, p_3\}$, we set $m_i = \iota^{-1}(p_i)$, $r = \operatorname{dist}(q, \mathcal{S})$ and $x = (m_1, m_2, m_3, r) = x(q)$. By the geometric relations, we have $n_3 \cdot (n_1 \times n_2) = 0$ (because they span the same two-dimensional subspace) and $p_1 + r\, n_1 = p_2 + r\, n_2 = p_3 + r\, n_3 = q$. This implies that $f^\iota(x(q)) \in W$. Then, because the contact points are of A_1 type and satisfy the strict ball condition (17), the following map is an isomorphism:

$$
\begin{array}{cccc}
T_{p_1} \mathcal{S} & T_{p_2} \mathcal{S} & T_{p_3} \mathcal{S} & \mathbb{R} \\
\left[\begin{array}{cccc}
\mathrm{Id} + r\, d_{p_1} \mathbf{n} & 0 & 0 & 0 \\
0 & \mathrm{Id} + r\, d_{p_2} \mathbf{n} & 0 & 0 \\
0 & 0 & \mathrm{Id} + r\, d_{p_3} \mathbf{n} & 0 \\
0 & 0 & 0 & 1
\end{array}\right]
& \begin{array}{c}
T_{p_1} \mathcal{S} \\
T_{p_2} \mathcal{S} \\
T_{p_3} \mathcal{S} \\
\mathbb{R}
\end{array}
\end{array}.
$$

By applying the inverse of this isomorphism to the previous map on the right side, we obtain that, if $f^\iota \pitchfork W$, then at the point $x(q)$,

$$
\begin{array}{cccc}
T_{p_1}\mathcal{S} & T_{p_2}\mathcal{S} & T_{p_3}\mathcal{S} & \mathbb{R}
\end{array}
$$

$$
\begin{bmatrix}
t_{2,3}\cdot u_1 & t_{3,1}\cdot u_2 & t_{1,2}\cdot u_3 & 0 \\
\mathrm{Id} & -\mathrm{Id} & 0 & n_1 - n_2 \\
\mathrm{Id} & 0 & -\mathrm{Id} & n_1 - n_3
\end{bmatrix}
\begin{array}{l}
\mathbb{R} \\
\mathbb{R}^3 \\
\mathbb{R}^3
\end{array}
\quad \text{is of rank 7, i.e., is full-rank,}
$$

where $t_{i,j} = n_i \times n_j$ and $u_i = d_{p_i}\mathbf{n}\circ(\mathrm{Id} + r\, d_{p_i}\mathbf{n})^{-1}$.

Now, we use the block matrix formula

$$
D \text{ invertible} \quad \Rightarrow \quad \det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A - BD^{-1}C)\det(D)
$$

applied to

$$
\begin{array}{cc}
T_{p_1}\mathcal{S} & T_{p_2}\mathcal{S}
\end{array} \qquad\qquad\qquad
\begin{array}{cc}
T_{p_3}\mathcal{S} & \mathbb{R}
\end{array}
$$

$$
A = \begin{bmatrix}
t_{2,3}\cdot u_1 & t_{3,1}\cdot u_2 \\
\mathrm{Id} & -\mathrm{Id}
\end{bmatrix}
\begin{array}{l}\mathbb{R}\\ \mathbb{R}^3\end{array}
\qquad
B = \begin{bmatrix}
t_{1,2}\cdot u_3 & 0 \\
0 & n_1 - n_2
\end{bmatrix}
\begin{array}{l}\mathbb{R}\\ \mathbb{R}^3\end{array}
$$

$$
\begin{array}{cc}
T_{p_1}\mathcal{S} & T_{p_2}\mathcal{S}
\end{array} \qquad\qquad\qquad\quad
\begin{array}{cc}
T_{p_3}\mathcal{S} & \mathbb{R}
\end{array}
$$

$$
C = \begin{bmatrix}
\mathrm{Id} & 0
\end{bmatrix}\ \mathbb{R}^3
\qquad
D = \begin{bmatrix}
-\mathrm{Id} & n_1 - n_3
\end{bmatrix}\ \mathbb{R}^3 \ .
$$

D is indeed invertible because $n_1 - n_3$ has a non-zero component along n_3 because $n_1 \neq n_3$ (otherwise $p_1 = p_3$).

We consider generic embeddings in the sense of Theorem 12, which allows us to work with the generic cut locus. Let us introduce t, a unit vector directing $T_q G(q)$ and known to be perpendicular to the subspace spanned by $\{n_1, n_2, n_3\}$. In the basis $(t, t\times n_3, n_3)$, we have then

$$
\begin{array}{ccc}
t & t\times n_3 & \mathbb{R}
\end{array}
$$

$$
D = \begin{bmatrix}
-1 & 0 & 0 \\
0 & -1 & t\cdot t_{3,1} \\
0 & 0 & n_1\cdot n_3 - 1
\end{bmatrix}
\begin{array}{l}
t \\
t\times n_3 \\
n_3
\end{array}
$$

so that

$$
\begin{array}{ccc}
t & t\times n_3 & n_3
\end{array}
$$

$$
D^{-1} = \begin{bmatrix}
-1 & 0 & 0 \\
0 & -1 & \dfrac{t\cdot t_{3,1}}{n_1\cdot n_3 - 1} \\
0 & 0 & \dfrac{1}{n_1\cdot n_3 - 1}
\end{bmatrix}
\begin{array}{l}
t \\
t\times n_3 \\
\mathbb{R}
\end{array} \ .
$$

After some computations with compatible choices of bases, we can explicit the expressions of B, C, and A as well. For instance, the product space $T_{p_1}\mathcal{S} \times T_{p_2}\mathcal{S}$ is spanned by the basis loosely denoted by $((t, t \times n_1), (t, t \times n_2))$. We obtain that

$$
A - BD^{-1}C =
\begin{array}{cccc}
t & t \times n_1 & t & t \times n_2 \\
\end{array}
\begin{bmatrix}
(a) & (b) & (c) & (d) \\
1 & 0 & -1 & 0 \\
0 & (e) & 0 & -1 \\
0 & (f) & 0 & 0
\end{bmatrix}
\begin{array}{l}
\mathbb{R} \\
t \\
t \times n_2 \\
n_2
\end{array}
$$

where

$$(a) = t_{2,3} \cdot u_1(t) - t_{2,1} \cdot u_3(t)$$

$$(b) = t_{2,3} \cdot u_1(t \times n_1) - t_{2,1} \cdot u_3(t \times n_3) \left(n_1 \cdot n_3 + \frac{(t \cdot t_{3,1})^2}{n_1 \cdot n_3 - 1} \right)$$

$$(c) = t_{3,1} \cdot u_2(t)$$

$$(d) = t_{3,1} \cdot u_2(t \times n_2)$$

$$(e) = n_1 \cdot n_2 - \frac{(t \cdot t_{2,1})(t \cdot t_{1,3})}{n_1 \cdot n_3 - 1}$$

$$(f) = t \cdot t_{1,2} - \frac{(t \cdot t_{1,3})(n_1 \cdot n_2 - 1)}{n_1 \cdot n_3 - 1}.$$

Therefore, for generic embeddings, being transversal to W means

$$f^\iota \pitchfork W \text{ at } x(q)$$

$$\Rightarrow \quad \det(A - BD^{-1}C) \neq 0$$

$$\Rightarrow \quad (f) \begin{vmatrix} (a) & (c) & (d) \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{vmatrix} \neq 0$$

$$\Rightarrow \quad (f)\,((a) + (c))\neq 0$$

$$\Rightarrow \quad \frac{t \cdot t_{1,2}}{n_1 \cdot n_2 - 1} \neq \frac{t \cdot t_{1,3}}{n_1 \cdot n_3 - 1} \text{ and } t_{2,3} \cdot u_1(t) + t_{3,1} \cdot u_2(t) + t_{1,2} \cdot u_3(t) \neq 0$$

Note that, if $\widehat{(n_1, n_2)}$ denotes the algebraic angle formed by n_1, n_2 in the plane orthogonal to $G(q)$ and oriented by t, we have $t \cdot t_{1,2} = \sin \widehat{(n_1, n_2)}$ and $n_1 \cdot n_2 = \cos \widehat{(n_1, n_2)}$. Hence the first condition involves some function $\frac{\sin \theta}{\cos \theta - 1}$ and just means that the angles $\widehat{(n_1, n_2)}$ and $\widehat{(n_1, n_3)}$ should be different, i.e., that $n_2 \neq n_3$, which is true. In the second condition, we recognize the result of Lemma 13, so that, for generic embeddings,

$f^\iota \pitchfork W \Rightarrow \mathrm{Hess}^{\mathrm{Riem}} g(q)$ is non-degenerate for any critical point q of type A_1^3.

Here again, F is a submersion by Lemma 11, thus $f^\iota \pitchfork W$ for generic embeddings $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$, by the transversality argument stated above.

A_1^4 **critical points**. There is nothing to check here, because the stratum of medial axis $G(q)$ reduces to a single point $\{q\}$ itself. $\quad\square$

Proposition 3. *There is a residual subset of embeddings $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$ for which a distance critical point q of type A_1^2, A_1^3, or A_1^4 lies strictly in the interior of the convex hull of $\Gamma(q)$ (i.e., it satisfies (N3b)).*

Proof. A critical point q of type A_1^2 always lies midway between the two points p_1 and p_2 in $\Gamma(q)$, so that the origin is always strictly inside of the convex hull spanned by $\{\nabla \alpha_1(q), \nabla \alpha_2(q)\} = \{n_1, n_2\}$.

For a critical point q of type A_1^3, we re-introduce the same spaces and maps as above, but with a slight modification to W. Let $X = M^{(3)} \times \mathbb{R}_{>0}$, $Y = (\mathbb{S}^2)^3 \times (\mathbb{R}^3)^3$, $W = \{y = (y_1, \ldots, y_6) \in Y \mid y_3 \cdot (y_1 \times y_2) = 0 \text{ and } y_4 = y_5 = y_6\}$ and

$$F: \begin{array}{ccc} Emb \times X & \to & Y \\ (\iota, m_1, m_2, m_3, r) & \mapsto & (n_1, n_2, n_3, p_1 + rn_1, p_2 + rn_2, p_3 + rn_3) \end{array},$$

as before. Consider the subspace of W defined by

$$W_{1,2} = \{w \in W \mid w_1 + w_2 = 0\},$$

which is of dimension in 7 and codimension 8 in Y. If $f^\iota(x(q))$ belongs to $W_{1,2}$, this corresponds geometrically to a right-angled triangle formed by n_1, n_2, n_3, where the right angle sits on n_3 while $n_1 + n_2 = 0$.

Actually, we already know that F is a submersion (Lemma 11), hence for generic ι, $f^\iota \pitchfork W_{1,2}$. But rank $D_x f^\iota \leq \dim X = 7$, so that rank $D_x f^\iota + \dim W \leq 7 + 7 < 15 = \dim Y$, which means that generically $\mathrm{Im} f^\iota$ does not intersect $W_{1,2}$. Likewise, generically, $\mathrm{Im} f^\iota$ does not intersect any of $W_{1,2}$, $W_{1,3}$ or $W_{2,3}$. In geometric terms, we get that for generic embeddings, there is no A_1^3 critical point lying on the boundary of the convex hull.

Similarly, for the A_1^4 case, by introducing the relevant map F and spaces $W_{i,j,k}$, we obtain a generic void intersection between images and spaces and conclude that for generic embeddings, there is no A_1^4 critical point lying on the boundary of the convex hull, i.e., belonging to a plane spanned by only 3 of the 4 contact points. $\quad\square$

Now, we move on to prove that for generic embeddings, critical points satisfy the strict ball condition (17) at any of their contact points. In particular, this excludes A_3 contact singularities (see [27]). We must therefore show that critical points only admit

A_1 contact points and no A_3 contact points, or equivalently, that there are no critical points of type $A_1 A_3$, but only of types A_1^j, $j = 2, 3, 4$.

Proposition 4. *There is a residual subset of embeddings $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$ that not only satisfies the properties set out in Theorem 12, but also the following: the critical points of $d = d_\iota$ are of types A_1^2, A_1^3, or A_1^4 only, but not $A_1 A_3$ (hence distance critical points of $\mathcal{S} = \iota(M)$ satisfy the strict ball condition (N2)).*

Proof. We make a slight modification to the spaces and maps introduced previously in the proof of Proposition 2 for A_1^2 critical points. Let $X = M^{(2)} \times \mathbb{R}_{>0}$ and $Y = E \times \mathbb{S}^2 \times \mathbb{R}^3 \times \mathbb{R}^3$, where we define a fiber bundle $(E, \mathbb{S}^2, \pi, \mathrm{Sym}((\mathbb{R}y_1)^\perp))$ whose total space is E, base space is $\mathbb{S}^2$, fiber is $\mathrm{Sym}((\mathbb{R}y_1)^\perp)$, and $\pi : E \to \mathbb{S}^2$. Here, for a base point $y_1 \in \mathbb{S}^2$, the notation $\mathrm{Sym}((\mathbb{R}y_1)^\perp)$ refers to the space of symmetrical endomorphisms of the two-dimensional subspace orthogonal to y_1. Define the map

$$F : \begin{array}{ccc} \mathrm{Emb}^k(M, \mathbb{R}^3) \times X & \to & Y \\ (\iota, m_1, m_2, r) & \mapsto & (n_1, \mathrm{Id} + r\, d_{p_1} \mathbf{n}, n_2, p_1 + r\, n_1, p_2 + r\, n_2) \end{array} \tag{25}$$

where $p_i = \iota(m_i)$ and $n_i = \mathbf{n}(p_i)$. Consider the submanifold $W = \{y = ((y_1, \varphi), y_2, \ldots, y_4) \in Y \mid y_1 + y_2 = 0 \text{ and } y_3 - y_4 = 0 \text{ and } \mathrm{rank}\, \varphi = 1\}$ which is of dimension 7 in Y.

Geometrically, if q is a critical point with one A_3 contact point p_1 and one A_1 contact point p_2, then the map $\mathrm{Id} + r\, d_{p_1} \mathbf{n} \in \mathrm{Sym}((\mathbb{R}n_1)^\perp)$ is of rank 1 because the largest principal curvature does not satisfy the strict ball condition, i.e., we have $\kappa_{\max}(p_1) = \frac{1}{\mathrm{dist}(q, \mathcal{S})}$, contrarily to the smallest principal curvature (we may suppose with loss of generality that $q \in \Omega^-$). Thus, for such q, we get $f^\iota(x(q)) \in W$.

But, with the same arguments as before, generically $f^\iota \pitchfork W$, but since $\mathrm{rank}\, D_x f^\iota + \dim W \leq \dim X + \dim W = 5 + 7 < 13 = \dim Y$, this means that the intersection between $\mathrm{Im}\, f^\iota$ and W is generically void, which shows what we need. $\quad\square$

Assembling Theorem 11, Theorem 13, Proposition 2, Proposition 3, Proposition 4, and Corollary 4, we arrive at the main result of this section.

Theorem 13 *(Signed distances are generically Morse with finitely many critical points). There is a residual subset of embeddings $\iota \in \mathrm{Emb}^k(M, \mathbb{R}^3)$, such that the signed distance function d to $\mathcal{S} = \iota(M)$ is a topological Morse function with finitely many critical points. In particular, for such embeddings, the critical points of d are non-degenerate Min-type critical points, satisfying the conditions (N1)-(N3c) in Proposition 1.*

Proof. To show that d is a topological Morse function, we prove that the critical points of d are topologically non-degenerate critical points; in particular, they are non-degenerate Min-type critical points. We do so by verifying conditions (N1)-(N3c) in Proposition 1 are individually true on a residual subset of embeddings; thus there is a residual subset of embeddings where all the conditions (N1)-(N3c) are true. The finiteness of contact points

(N1) and LIG Item (N3a) conditions shown to be generic in Theorem 11. Proposition 4 shows that any critical point on a generic set of embeddings is of A_1^j types and satisfies the strict ball condition (N2); and Proposition 3 shows that the convex hull condition (N3b) is also generic. Finally, as the critical points are non-degenerate, there are only finitely many such critical points Corollary 4. $\quad\square$

4. Signed distance persistent homology (SDPH)

Given the generalized Morse theorems for generic shapes established in the previous Section 3, we may now rigorously define SDPH. Our goal is to study the shape of an object, which requires for it to be rigorously and interpretably quantified. The distance and signed distance fields associated with the object, as well as persistent homology, are tools well-suited to this task. SDPH computes persistent homology with respect to d (previously introduced in Section 2.1).

Topology, in general, can be thought of as a pure mathematical study of shape and topological approaches have been used to study porous materials and vascular networks. For instance, porous materials were characterized by Vietoris–Rips persistent homology [83,98], by cubical filtrations derived from the image itself [102], and by a max-flow study in network models [6]. Discrete Morse theory has been used in conjunction with deep learning to accurately segment thin vascular structures [71,72]. Recently, [108] showcased how persistent homology, computed on a radial filtration or an alpha-complex filtration, is able to characterize tumor vascular networks.

The signed distance field is particularly useful since it naturally delineates the boundary of a shape, which happens when the value of the signed distance function is zero. In the context of persistent homology, filtering the space by the sublevel sets of the signed distance rather than the unsigned one allows features located inside the shape to be distinguished from those outside. The SDPH methodology appeared (under various names) in the discrete cubical setting first and was used to study porous materials with complex morphologies [43,44] and the interactions between fluid flow and microstructure [68,93].

In the works by [43,44,68,93], the SDPH method was described in a purely discrete setting, which is suitable for carrying numerical computations. [68,93] proposed an empirical interpretation of the SDPH diagrams, but without mathematical justification. To mathematically study the SDPH diagrams arising in the discrete settings, several types of discrete Morse theories may be useful [10,50] (see [84,15,11,103]). Here, we are interested in the behavior of SDPH in the smooth case, i.e., in the setting of smooth surfaces as opposed to cubical complexes in existing literature.

We begin by recalling our setting (see Section 3.2.1): Let $\Omega = \Omega^-$ be a bounded open set of $\mathbb{R}^3$ with C^k-smooth boundary $\mathcal{S}$. We write $\mathbb{R}^3 = \Omega^- \sqcup \mathcal{S} \sqcup \Omega^+$. We furthermore assume that $\mathcal{S}$ is a generic embedded surface in the sense of Theorem 13 (Section 3.3): in particular, we assume $k \geq 3$. Recall that the associated signed distance field d is defined by $d = \mathrm{dist}(\cdot, \Omega^-) - \mathrm{dist}(\cdot, \Omega^+)$.

Consider the *sublevel set filtration* of d, namely, the nested family of sublevel sets $X_r = \{x \in \mathbb{R}^3 \mid d(x) \le r\}_{r \in \mathbb{R}}$,

$$X_\bullet: \quad X_a \subseteq X_b, \quad \forall\, a \le b \in \mathbb{R}$$

endowed with the inclusion maps. Applying the homology functor with coefficients in $\mathbb{Z}_2$ to the filtration $X_\bullet$ gives the *persistence module*

$$\mathrm{PH}(X_\bullet): \quad \mathrm{H}(X_a) \to \mathrm{H}(X_b), \quad a \le b.$$

In $\mathbb{R}^3$, we are interested in the persistent homologies in dimensions $k = 0, 1, 2$.

$\mathrm{PH}(X_\bullet)$ is in fact determined by a finite sequence:

$$\mathrm{PH}(X_\bullet): \mathrm{H}(X_{r_1}) \to \ldots \to \mathrm{H}(X_{r_N})$$

where $r_1, \ldots, r_N$ are the critical values of d. Indeed, by combining Theorem 13 (signed distances are generically Morse with finitely many critical points) with Theorem 8 (isotopy lemma) and Theorem 9 (handle attachment lemma), we obtain that the persistence module is of finite type in the sense of [29, Section 3.8].

Corollary 5. *For a generically embedded surface $\mathcal{S}$ (as defined in Theorem 13), the corresponding signed distance d is a proper, lower-bounded, topological Morse function (Definition 11) that admits finitely many critical points. By Corollary 3, the persistence module $\mathrm{PH}(X_\bullet)$ is pointwise finite-dimensional; it can be decomposed into a finite direct sum of birth–death intervals in closed–open form $[b, d) \subset \bar{\mathbb{R}}$; and a critical point with index λ corresponds to either a birth in homology dimension λ, or a death in homology dimension $\lambda - 1$.*

Recall from Section 2.1 that the collections of birth–death pairs in each homology dimension are called the *persistence diagrams*. Here, the birth and death times correspond to the critical values $r_1, \ldots, r_N$.

If the critical level $d^{-1}(r_i)$ contains a single critical point with index λ, then crossing that level consists of attaching a λ-cell, which either creates a new homology generator in dimension λ (birth), or kills a homology generator in dimension $\lambda - 1$ (death). Similarly, if the critical level contains multiple critical points, each of them contributes to a handle. Therefore, any birth–death point in the persistence diagrams corresponds to a pair of creator–destroyer critical points. Conversely, any critical point is associated to a birth–death point of the diagram, either as a creator or as a destroyer, except one critical point that creates the homological component in PH_0 that has infinite lifetime and is paired with no other critical point.

This allows for a rigorous mathematical definition of the SDPH methodology.

Definition 16. Given a generic input binary shape Ω with boundary $\mathcal{S}$, the *signed distance persistent homology* (SDPH) methodology consists of building the sublevel set filtration $X_\bullet$ and computing the persistence diagrams of $\mathrm{PH}(X_\bullet)$.

4.1. Six types of critical points

Based on the results of Section 3, we provide now a geometric interpretation of the non-degenerate critical points of d and investigate the role of their index. This will enable us to describe SDPH diagrams precisely, since they consist of pairings of critical points across consecutive homology dimensions. Recall that the index is also the maximal dimension of a submanifold around q such that the restriction of the function has a maximum at q.

Non-degenerate critical points may be classified into 6 different types, denoted by λ^{sgn}, where $\lambda = 0, 1, 2, 3$ is the index of the critical point q (Proposition 1), and sgn $= \pm 1$ is the sign of $d(q)$. By definition, if μ refers to the index defined relative to the pure distance function (Proposition 1), we have that $\lambda = 3 - \mu$ if $q \in \Omega^-$ and $\lambda = \mu$ if $q \in \Omega^+$. Locally, the shape of the surface around a critical point $q \in \Omega^+$ of type λ^+ is the same as for a critical point $q \in \Omega^-$ of type $(3 - \lambda)^-$, as can be seen by exchanging the roles of Ω^- and Ω^+ locally.

For each type, one or several subtypes may be distinguished further, knowing that $\mu = (m - 1) + \mu'$, where $m = 2, 3, 4$ is the number of contact points and μ' is the Morse index of $\mathrm{dist}(\cdot, \mathcal{S})$ restricted to the stratum of medial axis $G(q)$ (Proposition 1). For example, $\mu = 3$ leads to three subtypes corresponding to $3 = 1 + 2$, $2 + 1$, $3 + 0$. Because the index roughly corresponds to the number of directions in which the function decreases from the point, the subtypes encode how many of these directions belong to the subspace spanned by the contact points versus the stratum of medial axis. For instance, points of type 2^+ (subtype $2 + 0$) have two decreasing directions in the contact subspace and none along the medial axis, which locally corresponds to a "bottleneck" shape.

Finally, there are no critical points of type 0^+ or 3^-. Indeed, by definition $\mu = (m - 1) + \mu' \geq 1$ since $m \geq 2$. Intuitively, this is because the signed distance admits no local minimum in Ω^+ (and equivalently, no local maximum in Ω^-).

In summary, we have 6 different types and a total of 12 subtypes, reported in Table 1 and depicted in Figs. 8 and 9. The plots were generated with Geogebra [70].
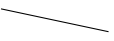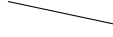
## 4.2. Seven types of birth–death persistence pairs

In SDPH, the critical points are paired into different types of birth–death persistence intervals, which we explain now using the classification of critical points from the previous section. Crossing a critical value with index λ is topologically equivalent to attaching a λ-cell, which either creates a new homological component in dimension λ, or kills a previously-existing component in dimension $\lambda - 1$. Therefore, we obtain the diagram shown in Fig. 10, where we can distinguish seven types of persistence pairings.

Table 1

Classification of non-degenerate critical points of the signed distance function in dimension 3. There are 6 types and 12 subtypes. The index $\lambda = \mathrm{index}(q; d)$ is related to the index $\mu = \mathrm{index}(q; \mathrm{dist}(\cdot, \mathcal{S}))$ through $\lambda = 3 - \mu$ if $q \in \Omega^-$ and $\lambda = \mu$ if $q \in \Omega^+$ (see Proposition 1). Geometrically, the local shape of the surface around points of type λ^+ is the same as around those of type $(3 - \lambda)^-$.

	$\lambda = 0$	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$
$d < 0$	**type** 0^- 3 subtypes $\mu = 1+2,\ 2+1,\ 3+0$	**type** 1^- 2 subtypes $\mu = 1+1,\ 2+0$	**type** 2^- 1 subtype $\mu = 1+0$	
$d > 0$		**type** 1^+ 1 subtype $\mu = 1+0$	**type** 2^+ 2 subtypes $\mu = 1+1,\ 2+0$	**type** 3^+ 3 subtypes $\mu = 1+2,\ 2+1,\ 3+0$
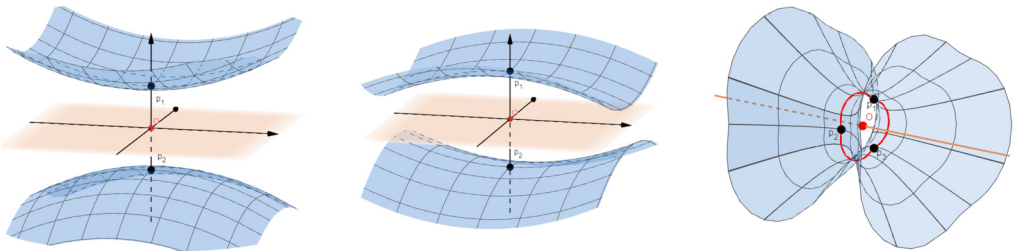
Fig. 8. Local shape of non-degenerate critical points of type 1^+ or 2^- (subtype $1+0$), type 2^+ or 1^- (subtype $1+1$) and type 2^+ or 1^- (subtype $2+0$) respectively. The red dot is q, black dots are $p_1, \ldots, p_m \in \Gamma(q)$, the orange plane or line is $G(q)$. The type of point λ^+ or $(3-\lambda)^-$ depends on whether we consider $q \in \Omega^+$ or $q \in \Omega^-$.
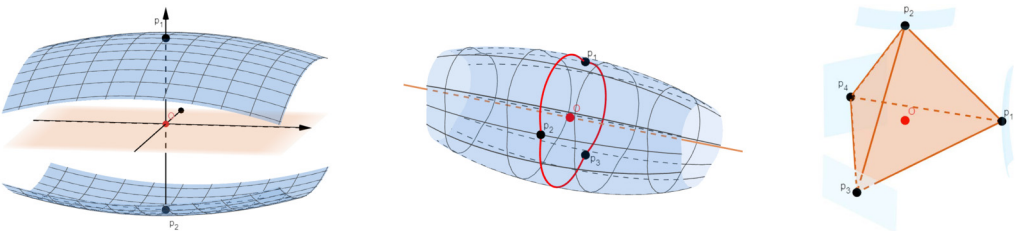

Fig. 9. Local shape of non-degenerate critical points of type 3^+ or 0^-, with subtypes $1+2$, $2+1$ and $3+0$ respectively. The red dot is q, black dots are $p_1, \ldots, p_m \in \Gamma(q)$, the orange plane or line is $G(q)$. For subtype $3+0$, $G(q) = q$. The type of point 3^+ or 0^- depends on whether we consider $q \in \Omega^+$ or $q \in \Omega^-$.
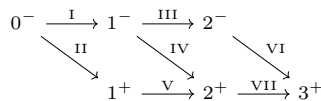
Fig. 10. 7 different types of birth–death persistence pairs in the SDPH diagrams, based on the classification of non-degenerate critical points given in Table 1.

Pairings of types I and II contribute to PH_0, types III, IV, V contribute to PH_1, and types VI and VII contribute to PH_2. Recall that no critical points belong to $\mathcal{S}$, hence there is no zero critical value, and no birth or death point belonging to a main axis $b = 0$ and $d = 0$ of the diagrams. Furthermore, the points on the persistence
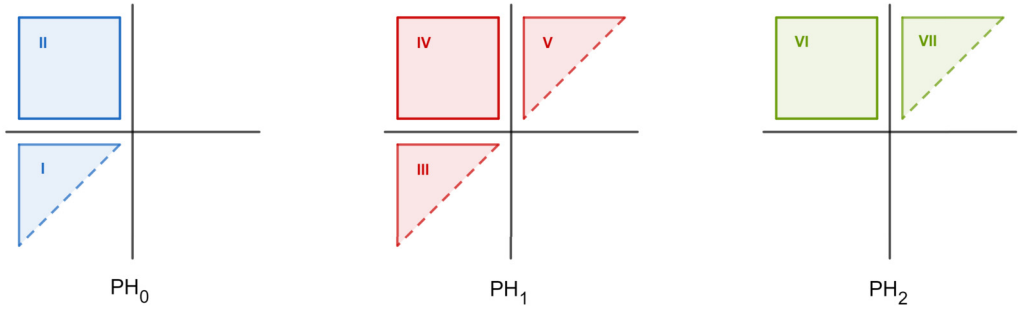
Fig. 11. Pairing types and quadrants in the SDPH persistence diagrams. birth–death points (not represented) are located in the quadrants accordingly to their type (Fig. 10).

diagrams are bounded away from the axes by the positive reach of surface, as the critical values of the distance functions must be greater than the reach [97]. Using the sign of the critical values, we can break down the types into further relationships with respect to the quadrants of the persistence diagrams. Before doing so, we observe that unlike usual persistence diagrams (where filtrations are parameterized by some $\epsilon > 0$, as for Vietoris–Rips filtrations) that contain birth–death points in the NE quadrants only of each PH_k, SDPH diagrams may contain points in the other quadrants as well since the signed distance has negative critical values. Pairs of type I and II are located in PH_0 SW and NW respectively; III, IV, V pairs are in PH_1 SW, NW, and NE respectively; and VI, VII pairs are in PH_1 NW, NE respectively. In particular, the absence of critical points of type 0^+ or 3^- implies that no persistence pairs are located in the quadrants PH_0 NE and PH_2 NW. Because death times are strictly larger than birth times, all birth–death points lie strictly above the diagonal. Additionally, there is a point with infinite death time in PH_0. Note that the number of components and the genus of $\mathcal{S}$ correspond to the number of birth–death points of type II and IV, respectively. These properties are summarized in Fig. 11 (generated with Geogebra [70]).

4.3. Numerical implementation

To compute SDPH diagrams, we follow Algorithm 1.

Algorithm 1 SDPH: quantify shape texture of binary shape Ω.

1: **procedure** SDPH(Ω)
 output: persistence diagrams PH_0, PH_1, PH_2

2: $d \leftarrow \text{dist}(\cdot, \Omega) - \text{dist}(\cdot, \Omega^c)$ ▷ build signed distance field
3: $K_\bullet \leftarrow$ cubical complex filtered by sublevels of d ▷ T-constructed
4: $PH_0, PH_1, PH_2 \leftarrow \text{Dgm}(K_\bullet)$ ▷ compute persistence diagrams

In practice, a 3D binary shape Ω is encoded as a 3D binary digital image (i.e., discrete array) with black and white values. We first used the `scikit-fmm`[1] Python package to convert Ω into a discrete signed distance field d in the form of a 3D array. Next, we considered a finite *cubical complex* filtered by the sublevel sets of d [75,112]:

$$K_\bullet : K_1 \to \ldots \to K_N = K.$$

Cubical complexes are the natural representation of digital images. We then computed the persistence diagrams (with coefficients in $\mathbb{Z}_2$)

$$\mathrm{PH}(K_\bullet) : \mathrm{H}(K_1) \to \ldots \to \mathrm{H}(K_n)$$

in dimensions $0, 1, 2$ with the `giotto-tda` Python package [109].

Giotto-tda uses the GUDHI package [110,45] as a back-end, and hence computes a T-constructed filtration [14]. Informally, this corresponds to increasing a level and considering each time the groups of voxels in d having value less than this level. By adding groups of voxels at a time, components, loops, and voids are formed or destroyed. In vascular patterns, for instance, loops outlining cycles in the vessels start to form at negative values of d, before being filled out at positive values, which corresponds to pairings of type IV in PH_1 NW (Fig. 11).

Note that the `CubicalRipser` [76] package may also be used to keep track of the critical voxels giving rise to or terminating homological components. Hence, a spatial analysis is also possible.

Remark 11. As seen in the Introduction, [44], [68] and [93] studied SDPH directly in the setting of cubical complexes. A natural question to explore as future work is to relate the discrete Morse theories [10,50,11,84,15,103] to our setting. In particular, it remains to be understood how critical points in the discrete settings relate to the original critical points, as the sampling resolution goes to infinity. Furthermore, how the gradient flow in the continuous case relates to the gradient flow induced by a point-cloud distance field $d_\mathcal{P}$ may be studied. The latter, which was studied by [61,59,60], is useful in applications such as surface reconstruction based on surface sampling.

5. Quantifying texture in porous shapes

In this section, we showcase several applications of SDPH for quantifying texture in porous shapes. Such shapes may exhibit high morphological complexity and irregularity, including tubular, membranous and spherical elements, or mixtures of them, forming pores, cavities, channels, and so on.

---

[1] Available at https://github.com/scikit-fmm/scikit-fmm.

We computed SDPH diagrams for three case studies; two entail synthetic data and the third entails real data. For the two case studies pertaining to synthetic data, the first involved simulated data from realizations of Gaussian random fields (specifically, level surfaces of GRFs), while the second involved simulated shapes generated using curvatubes [107]; the third case study involves real samples of bone marrow vessels segmented from 3D confocal images (data courtesy of Antoniana Batsivari and Dominique Bonnet, The Francis Crick Institute, London). The choice of these three case studies is to investigate the behavior where the randomness of the underlying data generating process derives from a classical probability distribution, from a phase-field approximation, and from real-world noise (such as measurement, errors), respectively. The data from each of the case studies are described in Sections 5.1, 5.2, and 5.3 respectively. Their SDPH diagrams are compared in Section 5.4, where a general interpretation of SDPH diagrams is given. The case studies involving synthetic data suggest in Section 5.5 that SDPH diagrams may be stable with respect to the texture of the input shape, which in turn is related to the variability of its curvature.

Note on the output figures. The same layout is used in Figs. 12, 13, 14, 15, and 16. The first column shows a 3D view of the surface bounding the binary shape, plotted with *ParaView* [9]. The second column shows a 2D section taken in the middle of the volume, after the shape has been smoothly closed near the boundaries of the simulation domain. A scalar field (random field, phase-field, or binary segmentation, depending on the data) is represented in there, whose zero level set gives the input surface. Finally, the three last columns display the SDPH diagrams PH_0, PH_1, and PH_2 in the form of scatter plots, where birth–death points are colored by density (Gaussian kernel with $\sigma = 0.5$). Birth–death points (birth, death) with persistence death $-$ birth less than 0.5 were thresholded out.

5.1. Gaussian random fields

We first computed SDPH diagrams on the zero level set of a *Gaussian random field* (GRF) [3,53]. GRFs may be used to model the random morphology of porous materials, such as rocks.

A GRF on $\mathbb{R}^d$ is a real-valued stochastic process $f(t)$ indexed over points $t \in \mathbb{R}^d$, such that any finite dimensional distribution of f indexed over any arbitrary finite set of points $T = \{t_1, \ldots, t_N\} \subset \mathbb{R}^d$ is a multivariate normal distribution:

$$f(t_1), \ldots, f(t_N) \sim N(\mu_T, \Sigma_T).$$

A classical GRF model is the zero-mean *squared exponential covariance* model (a.k.a. *"Gaussian covariance model"*), where $\mu = 0$ and $\Sigma(s, t) = \mathrm{Cov}\left(f(s), f(t)\right) = e^{\frac{\|s - t\|^2}{2\lambda^2}}$. The covariance between $f(s)$ and $f(t)$ is controlled by a lengthscale parameter $\lambda > 0$: for a given spatial separation $\|s - t\|$, a larger λ implies stronger correlation between $f(s)$ and $f(t)$, hence lower frequency features in the level sets at this scale.

Table 2
Parameters of the Gaussian covariance model used to generate the random fields of Fig. 12.

GRF model	ℓ_1	ℓ_2	ℓ_3	other
F_1	8	8	8	
F_2	8	$\frac{\ell_2}{\ell_1} = 0.7$	$\frac{\ell_3}{\ell_1} = 0.85$	
F_3	5	5	5	
F_4				$f_4 = f_3 - 0.5$
F_5				$f_5 = f_1 + f_3$

For GRFs with Gaussian covariance model, any realization f (viewed as a function $\mathbb{R}^d \to \mathbb{R}$) admits a smooth version $\tilde{f}$ such that $\mathbb{P}(f(t) = \tilde{f}(t)) = 1$ for all t. Indeed, the continuity and differentiability of GRF samples are dictated by conditions on the covariance function [3, Theorem 3.4.1], which in this case are satisfied by the Gaussian covariance model.

We used *GSTools* [96] to generate 5 random fields $f_1, \ldots, f_5$ in a $100 \times 100 \times 100$ domain size, using the parameter values presented in Table 2. The covariance kernel here is defined by $\mathrm{Cov}\,(f(s),\,f(t)) = C(s - t) = \sigma^2\,\rho_0(\|\mathrm{A}\,(s - t)\|)$, where $\rho_0(r) = e^{-\frac{\pi}{4}r^2}$, $\mathrm{A} = \mathrm{diag}(\frac{1}{\ell_1}, \frac{1}{\ell_2}, \frac{1}{\ell_3}) \times \mathrm{R}$ and R is a rotation matrix. The lengthscales ℓ_1, ℓ_2, ℓ_3 are given in terms of voxel size units. The matrix A encodes the isotropy of the process: differences in the lengthscales represent the anisotropy w.r.t. the orthonormal frame rotated by R. In all simulations, we took $\sigma^2 = 1$, since changing the variance does not statistically affect the behavior of the zero level set.

The SDPH diagrams were computed with input shape $\Omega = \{f \geq 0\}$, the region of non-negative values, after smoothly closing the shape near the boundaries of the simulation domain. The results are displayed in Fig. 12.

5.2. Curvatubes

The *curvatubes* model [107] was constructed to randomly simulate a large variety of organic-looking porous shapes. Here, the output of curvatubes is precisely the synthetic realization of the shapes and textures that we aim to study with our SDPH framework and its process to produce random realizations is different in nature to the GRF case study. The curvatubes framework assumes that a porous shape optimizes some curvature functional involving a second-degree polynomial of the principal curvatures of the surface,

$$\mathbf{F}(\mathcal{S}) = \int_{\mathcal{S}} \left(a_{2,0}\ \kappa_1^2 + a_{1,1}\ \kappa_1\kappa_2 + a_{0,2}\ \kappa_2^2 + a_{1,0}\ \kappa_1 + a_{0,1}\ \kappa_2 + a_{0,0} \right)\ dA,$$

with a constraint of constant volume enclosed. Effectively, this geometric problem is approximated by a phase-field problem that is parameterized by a transition scale ϵ and a mass m encoding the volume constraint. The optimization starts with random noise and
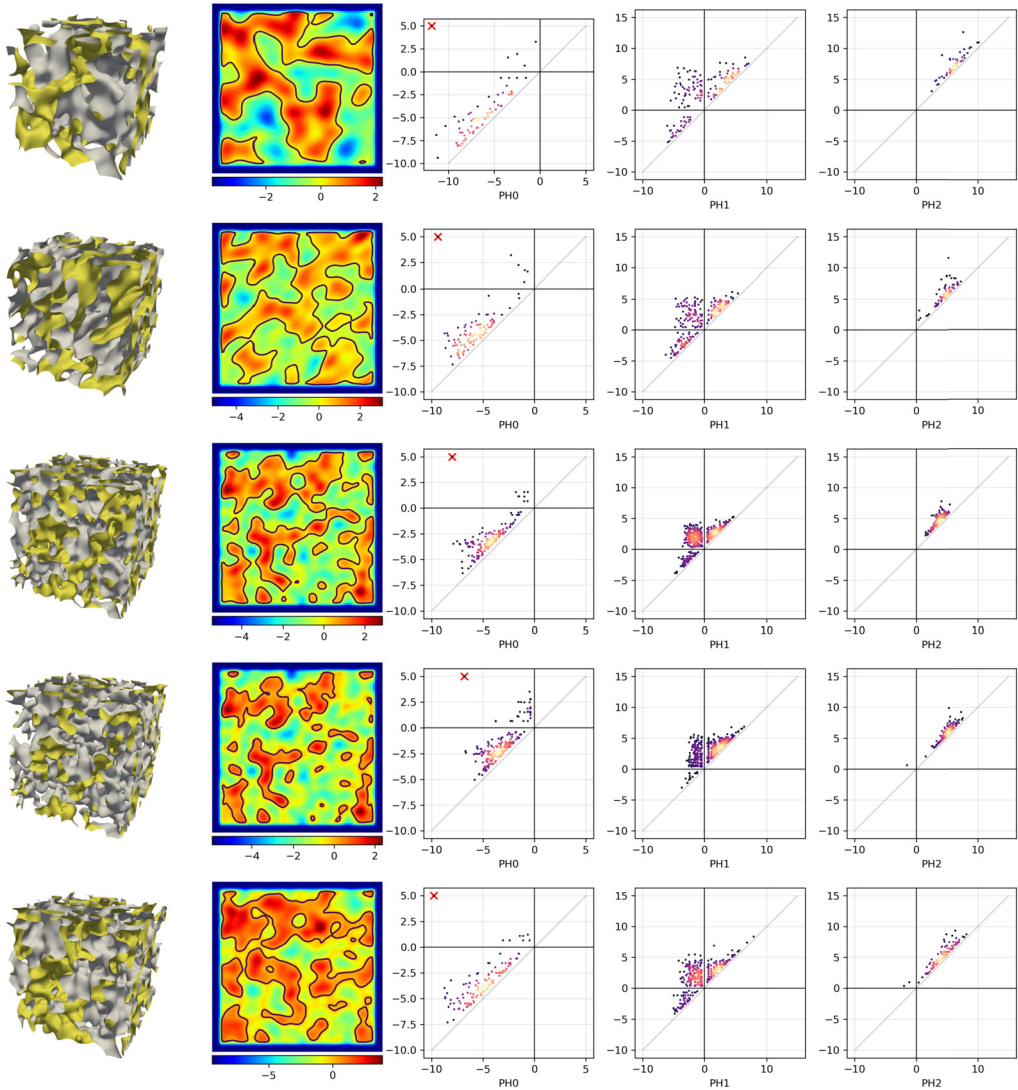
Fig. 12. SDPH diagrams of the zero level set of 5 Gaussian random fields $F_1, \ldots, F_5$ generated with the parameters shown in Table 2.

converges towards a local optimum. The output surface is then implicitly represented as the zero level set of the phase-field.

We used curvatubes to generate 10 different shapes $(S_1, \ldots, S_{10})$ in a $100 \times 100 \times 100$ domain. The parameter values are presented in Tables 3 and 4, where ϵ is defined relative to the side of the domain supposed to be 1 unit. Their SDPH diagrams were then computed after smoothly closing the shapes near the boundaries of the simulation domain. The results are shown in Figs. 13 and 14.

Table 3
Parameters of the curvatubes model used for Fig. 13.

Shape	Coefficients	m	ϵ
S_1	$(1, 2, 6, -40, -40, 400)$	-0.6	0.02
S_2	$(1, -1.284, 9.626, -77.283, 39.681, -633.849)$	-0.421	0.02
S_3	$(1, 2.8, 2, -10, -10, 25)$	-0.25	0.02
S_4	$(1, 3.225, 10.87, -146.309, 143.487, -2920.302)$	-0.499	0.02
S_5	$(1, -0.238, 11.988, -175.909, -27.167, 2117.037)$	-0.648	0.02

Table 4
Parameters of the curvatubes model used for Fig. 14.

Shape	Coefficients	m	ϵ
S_6	$(1, 2.034, 11.166, 14.553, 28.829, -565.092)$	-0.356	0.02
S_7	$(1, 0.63, 4.399, 132.459, 195.066, -2378.53)$	-0.364	0.02
S_8	$(1, 1, 1, 0, 0, 0)$	-0.4	0.02
S_9	$(1, 0.396, 1.095, -28.64, 190.906, 2062.082)$	-0.598	0.02
S_{10}	$(1, 4.185, 2.053, 19.375, 29.607, 120.265)$	-0.534	0.02

5.3. Vascular data

As our final case study, we considered 3 samples B_1, B_2, B_3 of bone marrow vessels ("BM shapes") taken from a healthy control mouse. These samples were selected at different anatomical locations of the same femur bone, and segmented on 3D confocal images. The selected vascular crops were $100 \times 200 \times 200$ in size, where a voxel side represents 2 µm in physical length. Due to noise in the signal, and the complexity of the task, the segmentation may not be fully accurate.

5.4. Interpreting SDPH diagrams

From the classification of critical points and pairings described in Sections 4.1 and 4.2, we may derive a *general interpretation* of SDPH diagrams, supported by a comparison of the diagrams in Figs. 12, 13, 14, and 15.

SDPH may be interpreted as follows:

1. PH_0 SW measures variations of thicknesses in tubular and membranous structures.
2. PH_0 NW counts the number of components and measures their characteristic size and separation.
3. PH_1 SW detects the presence of dimples (similar to the hollowed shape of red blood cells).
4. PH_1 NW counts the number of loops and measures their characteristic thickness and size.
5. PH_1 NE measures variations in the proximity of the structures, creating at a distance curvy and non-convex loops.
6. PH_2 NW indicates the presence of small voids trapped inside the shape, which do not exist in tubular shapes.

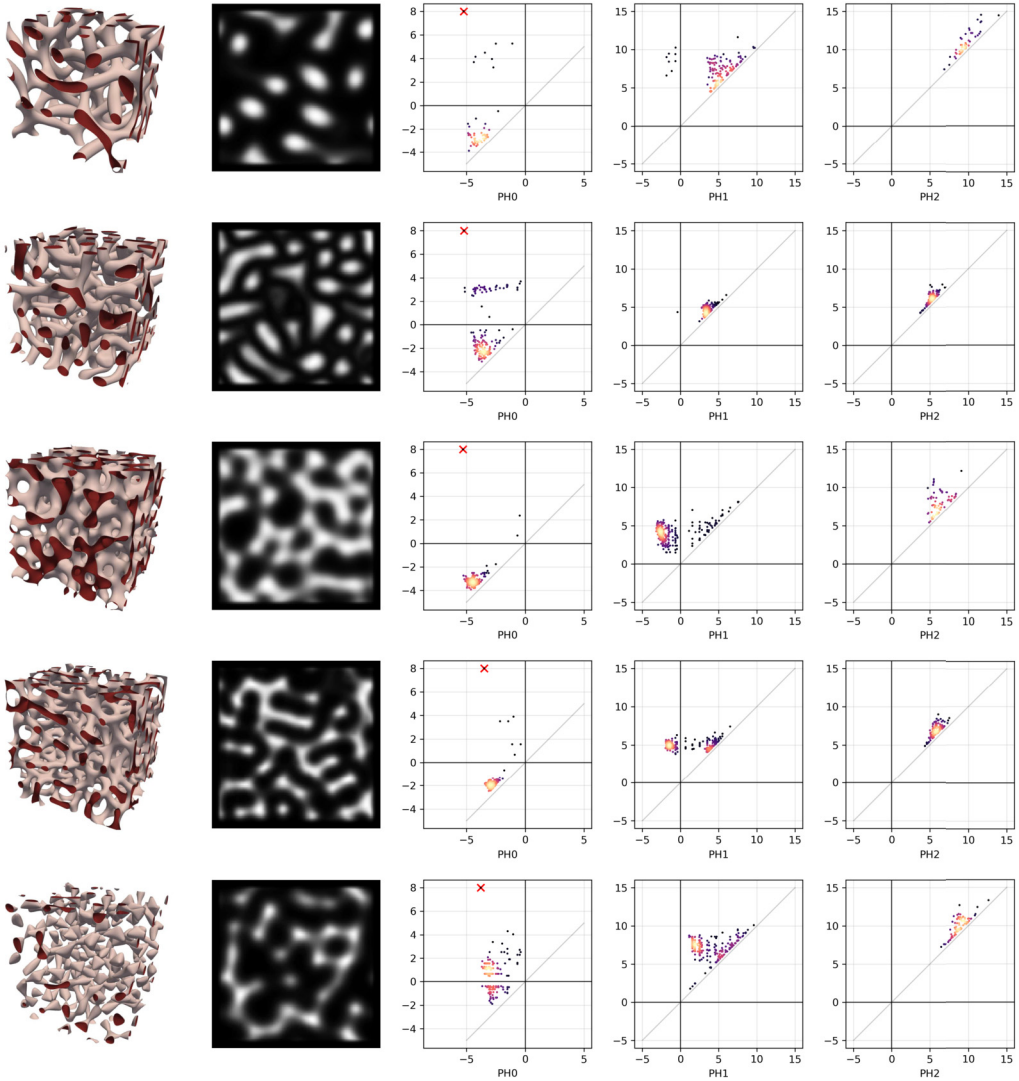**Fig. 13.** SDPH diagrams of 5 tubular shapes $S_1, \ldots, S_5$ generated with curvatubes using the parameters shown in Table 3.

7. PH$_2$ NE measures the characteristic sizes of interspaces separating plain structures and the density of tubular branching, so that larger sizes mean sparser structures.

Each of these properties may be observed in the following examples. For instance,

1. Thicknesses are much more variable in the biological shapes B_i than the synthetic shapes S_i.
2. Both of S_2 and S_5 have a large number of disconnected components, but in S_2 they are typically separated by the same distance.
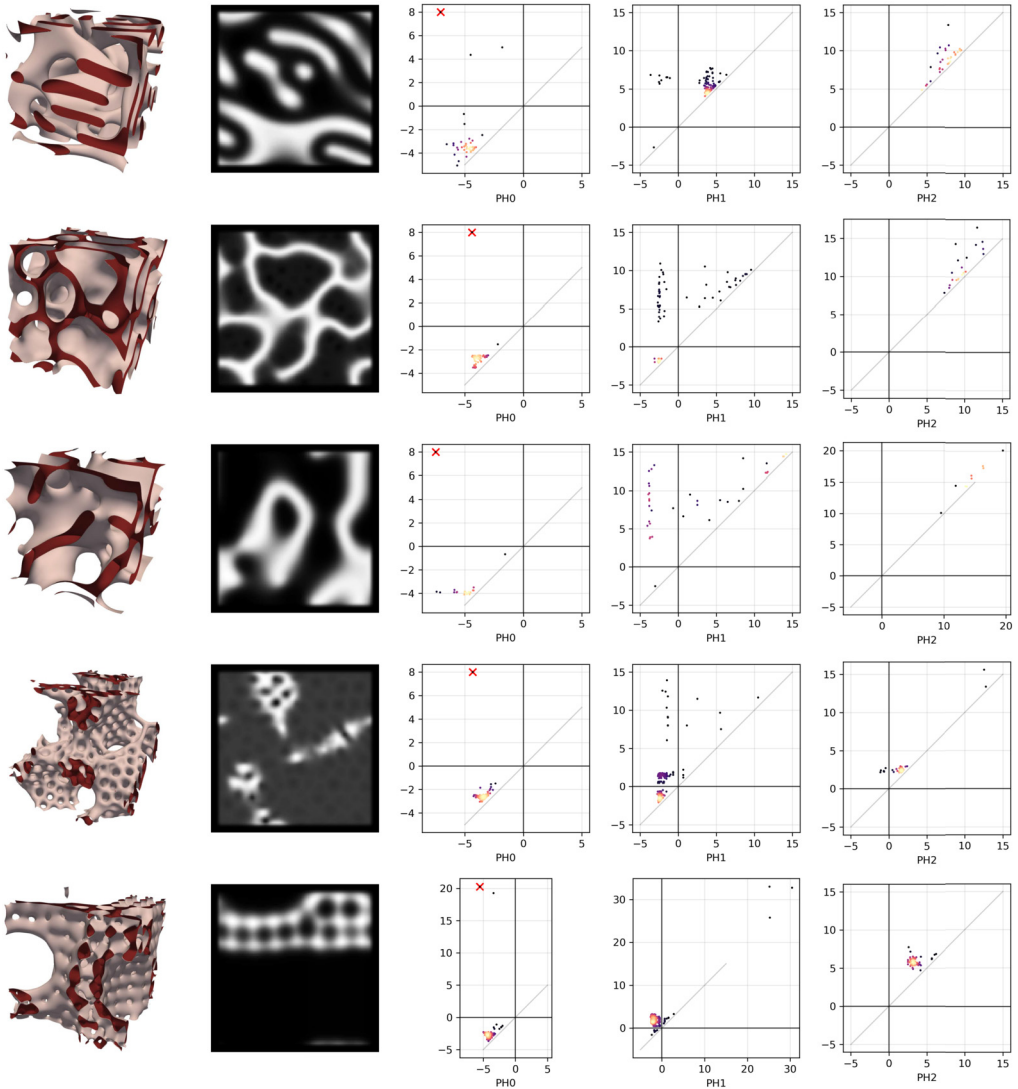
Fig. 14. SDPH diagrams of 5 membranous and porous shapes $S_6, \ldots, S_{10}$ generated with curvatubes using the parameters shown in Table 4.

3. Dimples are mostly found in the GRF shapes F_i and the BM shapes B_i, but not in the tubular shapes $S_1, \ldots, S_4$.

4. There is an increasing number of loops forming in F_3 than F_1, as the length scale parameter to generate the GRFs is decreased.

5. When the shape is thinned, by considering different textures, as for S_3, S_4, S_5, or by considering different level sets, as for $F_3 = \{f_3 = 0\}$ and $F_4 = \{f_3 = 0.5\}$, homological loops tend to disappear from the initial structure, in favor of those that are formed at a distance in a non-convex curvy way.
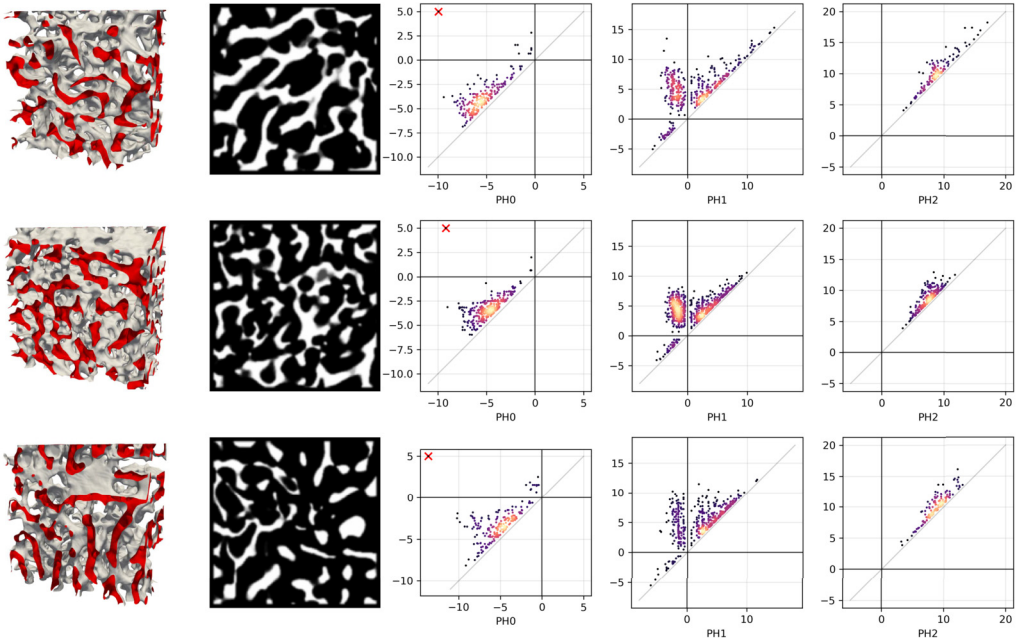
Fig. 15. SDPH diagrams of 3 bone marrow vascular samples B_1, B_2, B_3 in a healthy mouse (data courtesy of Antoniana Batsivari and Dominique Bonnet, The Francis Crick Institute, London).

6. Only rarely small voids are spotted inside the shape, except for shapes S_9, F_4 and F_5.
7. Finally, while homological voids formed at a distance are mostly of the same size in the densely windowed shapes S_9 and S_{10}, they are found at much larger and variable sizes in the sparse vascular BM shapes B_1 and B_3.

Finally, note that *multi-scale texture* can be quantified by SDPH, as can be seen in the diagram of S_9, where most loops form at similar critical sizes, but a few of them appear at variable larger scales; or for F_5, which by construction mixes features taken from F_1 and F_3 at different scales. Generally, mixtures of different textures in space or in scale are apparent in the SDPH diagrams. Further observations can be made: In all quadrants, a concentration (or spread) of the scatter plot indicates *(in)homogeneities*, not only in the distribution of the critical sizes, but particularly in terms of the morphological patterns described above for each respective quadrant. For instance, the shapes S_7, S_8, S_9 form cycles of irregular sizes, in contrast to the shapes S_3 and S_4; B_3 forms irregular-sized loops compared to B_2. Moreover, birth–death points close to the main axes arise in *noisy* structures whose topological features have close to zero critical sizes, whereas *smooth* structures induce a clearer gap from the axes. This effect is apparent here: curvatubes shapes are smooth and lead to clearly distinct diagrams, compared to GRF shapes which are topologically noisy and whose diagrams are more difficult to interpret, which may be expected due to the difference in the nature of their randomness.

5.5. Texture, curvatures, and stability of SDPH

Texture in shapes may be characterized as spatially repeated patterns of the surface, subject to some randomness, but not easily discriminated by visual perception [74,101, 80]. There is hence some quantity, left unchanged by some type of geometric randomness, that represents texture. In a previous work, we suggested that one possible texture descriptor is the behavior of the curvatures on the surface, or "curvature diagram" [107]. Given the same choice of generation parameters **a**, but with random initializations in curvatubes, we illustrated that the output curvature diagram was regularly stable with respect to **a**.

Here, it is natural to ask whether SDPH diagrams also quantify texture and exhibit a similar stability behavior. We generated five synthetic shapes with the same choice of curvatubes parameters **a** but random initializations of the optimization flow. The results are shown in Fig. 16, where we used a $150 \times 150 \times 150$ simulation grid, with mass $m = -0.5$, $\epsilon = 0.02$ (one pixel being of size 0.01) and coefficients $\mathbf{a} = (6, 2.5, 6, -230, 20, 2350)$. These values correspond to the non-reduced curvature polynomial $p(\kappa_1, \kappa_2) = (H-15)^2 + .5\,K + 5\,(\kappa_1 - 20)^2 + 5\,(\kappa_2 + 5)^2$ where $H = \kappa_1 + \kappa_2$ and $K = \kappa_1\,\kappa_2$.

We found that, for a given choice of texture, the SDPH method produced very similar-looking output diagrams, especially in contrast to those produced for other generation parameters, shown in Figs. 13 and 14. Empirically, we observed the same consistent behavior for other choices of shape textures (data not shown here, but is available on our GitHub repository). In other words, despite the shapes being very different geometrically, the distribution of their topological patterns as measured by SDPH appears to be very similar.

The well-known Stability Theorem [33] from classical persistent homology theory states that two functions which are close in infinity-norm lead to persistence diagrams which are close in bottleneck distance (see Section 2.1, Theorem 1); namely, that similar shapes lead to similar persistence diagrams. This guarantees stability when the input shapes are geometrically perturbed by noise (which results in a small infinity-norm between them), a property consequently satisfied by SDPH diagrams. Here, our observations suggest that *"SDPH diagrams are stable with respect to texture"*, which may give rise to a new kind of stability result that extends beyond the existing stability results in classical persistent homology. The novelty of this kind of stability study lies in a more geometric and probabilistic direction, as well as the fact that there is no existing stability result entailing SDPH, to the best of our knowledge, since SDPH has been comparatively less studied than classical persistence. Finally, it will be interesting to explore as future work how curvature diagrams relate to SDPH diagrams.

6. Discussion

Morse theory is about the study of critical points of a function, and how they characterize geometric and topological information from the underlying space. While classical
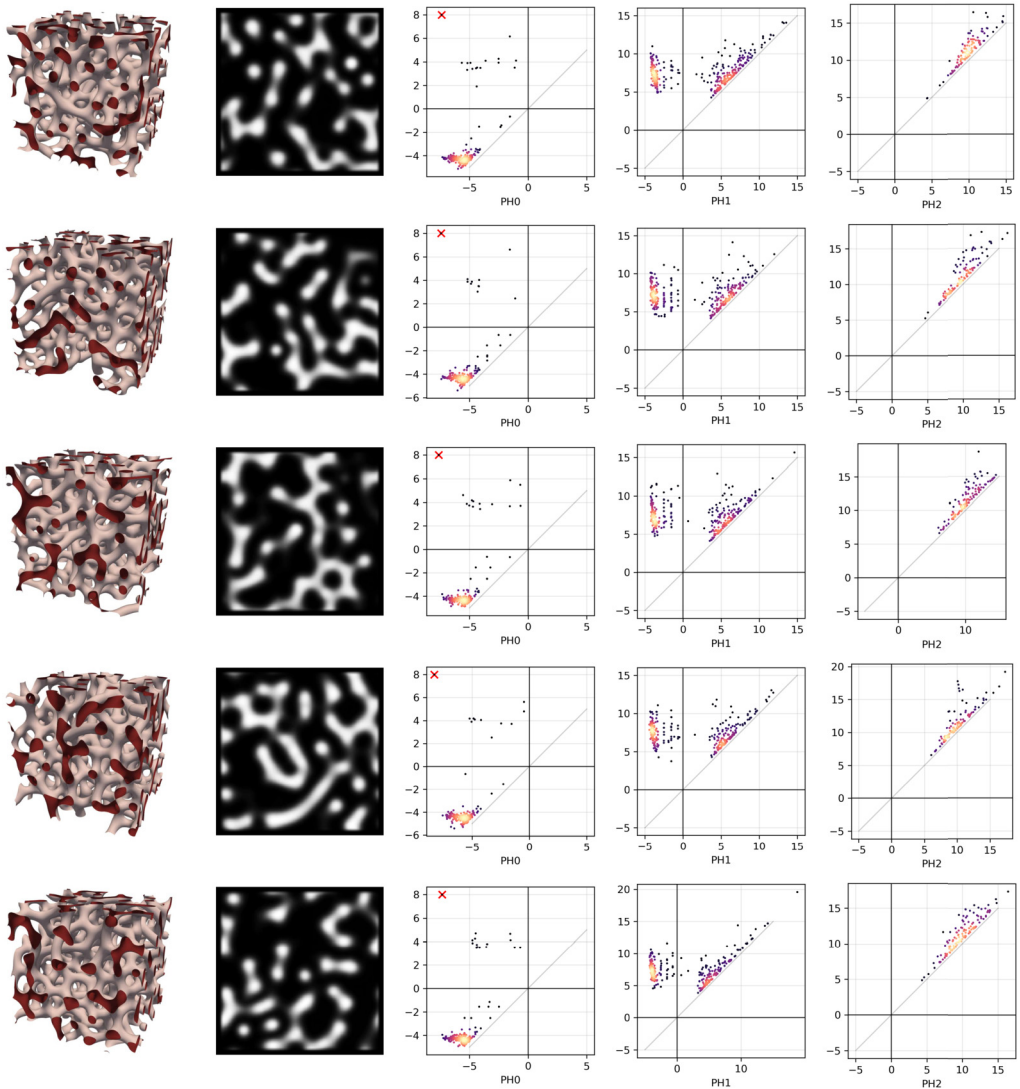
Fig. 16. Five random shapes sharing a similar texture and similar SDPH diagrams. They were generated with cur-vatubes using the same parameters but different initializations.

Morse theory deals with smooth functions, this paper generalizes it to a broad class of (signed and unsigned) Euclidean distance functions.

We have successfully established the two basic Morse lemmas (isotopy lemma and handle attachment lemma), namely, that the topological changes encountered by the sublevel sets of Euclidean distance functions (signed or not) with non-degenerate critical points are exactly dictated by the latter ones. To do this, we reframed Euclidean distances as Min-type functions (signed distances are then locally Min-type outside the shape and Max-type inside), which in turn may be recast as topological Morse functions. As a

further contribution, we have also provided complete proofs extending the Morse lemmas to topological Morse functions.

As a focal crux of our study, we have identified the correct geometric conditions corresponding to the Min-type non-degeneracy at a critical point, and shown that, in $\mathbb{R}^3$, for generic embeddings of the shape, the signed distance field only admits a finite number of critical points that are all non-degenerate. In other words, generic signed distances are topological Morse functions with finitely many critical points. This has important implications, since it guarantees that essentially any shape may be filtered by the sublevel sets of its signed distance field and that the persistent homology of this filtration may be computed, namely the "signed distance persistent homology" (SDPH).

We have demonstrated how to use SDPH as a quantifier of *texture* in shapes, based on the patterns of birth–death pairs in persistence diagrams. We have proposed a rigorous geometric interpretation of the critical points involved in the diagrams, alongside a classification of persistence pairs, which leads to a practical guide for interpreting SDPH diagrams in real applications. Furthermore, SDPH seems to attribute similar diagrams to shapes with similar texture.

Case studies have been carried out on both simulated and real data, demonstrating that our proposed framework may be readily applied to several fields, including shape analysis, imaging, computer graphics, materials science, biology, environmental science and geospatial science to rigorously and interpretably quantify shapes and textures.

Software and data availability

The code to compute SDPH, generate the synthetic data, and the real data are all publicly available on the GitHub repository https://github.com/annasongmaths/SDPH. Details of the numerical implementation are as above in Section 4.3.

Appendix A. Technical lemmas for the genericity theorem

In this Appendix, we present three important components that were used to prove Theorem 13. As before, we make use of simplified notations for the Gauss map (see Remark 10).

Lemma 11 *(The functions F are submersions). The functions F introduced in the proof of Theorem 13 are sufficiently smooth submersions. More precisely,*

- *Let $X = M^{(2)} \times \mathbb{R}_{>0}$ and $Y = (\mathbb{S}^2)^2 \times (\mathbb{R}^3)^2$. Define*

$$
F : \begin{array}{rc} \mathrm{Emb}^k(M, \mathbb{R}^3) \times X \to & Y \\ (\iota, m_1, m_2, r) \mapsto & (n_1, n_2, p_1 + r\, n_1, p_2 + r\, n_2) \end{array} \tag{A.1}
$$

  where $p_i = \iota(m_i)$ and $n_i = \mathbf{n}(p_i)$. In the expression $p_i + r\, n_i$, n_i is viewed in $\mathbb{R}^3$.
  Then F is a C^{k-1} submersion.
- *Likewise, define a function $F(\iota, m_1, m_2, m_3, r) = (n_1, n_2, n_3, p_1 + r\, n_1, p_2 + r\, n_2, p_3 + r\, n_3)$. Then F is a C^{k-1} submersion.*
- *Let $X = M^{(2)} \times \mathbb{R}_{>0}$ and $Y = \Xi \times \mathbb{S}^2 \times \mathbb{R}^3 \times \mathbb{R}^3$. Here, $(\Xi, \mathbb{S}^2, \pi, \mathrm{Sym}((\mathbb{R}\, n)^\perp))$ is a fiber bundle of base space $\mathbb{S}^2$ and fiber space $\mathrm{Sym}((\mathbb{R}\, n)^\perp)$, the space of symmetric endomorphisms defined on the subspace of $\mathbb{R}^3$ orthogonal to $n \in \mathbb{S}^2$. Define the map*

$$
F : \begin{array}{rc} \mathrm{Emb}^k(M, \mathbb{R}^3) \times X \to & Y \\ (\iota, m_1, m_2, r) \mapsto & (n_1, \mathrm{Id} + r\, d_{p_1}\mathbf{n}, n_2, p_1 + r\, n_1, p_2 + r\, n_2) \end{array} \tag{A.2}
$$

  where $p_i = \iota(m_i)$ and $n_i = \mathbf{n}(p_i)$.
  Then F is a C^{k-2} submersion.

Proof. We show first that the following simplified function is a submersion; in fact, $D_\iota F$ alone will already be surjective. Consider the fiber bundle Ξ having base space $\mathbb{S}^2$ and fiber space $\mathrm{Sym}((\mathbb{R}\, n)^\perp)$.

$$
F : \begin{array}{rc} \mathrm{Emb}^k(M, \mathbb{R}^3) \times M \to & \mathbb{R}^3 \times \Xi \\ (\iota, m) \mapsto & (p, n, dn), \end{array} \tag{A.3}
$$

where $p = \iota(m)$, $n = \mathbf{n}(p)$, $dn = d_p\mathbf{n}$.

First, F is a C^{k-2} map. Let us fix ι_0 and m_0. Now, given some arbitrary element (v, N, A) tangential to $\mathbb{R}^3 \times \Xi$ at $(p_0, n_0, dn_0) = F(\iota_0, m_0)$, where $v \in T_{p_0}\mathbb{R}^3 \simeq \mathbb{R}^3$, $N \in T_{n_0}\mathbb{S}^2$, $A \in T_{dn_0}\mathrm{Sym}((\mathbb{R}\, n_0)^\perp) \simeq \mathrm{Sym}((\mathbb{R}\, n_0)^\perp)$, the aim is to find some tangential element $V \in T_\iota \mathrm{Emb}^k(M, \mathbb{R}^3) \simeq C^k(M, \mathbb{R}^3)$ such that $D_\iota F(\iota_0, m_0)[V] = (v, N, A)$, by local perturbations of the embedding. We can decompose this into three independent problems, by finding preimages of $(v, 0, 0)$, $(0, N, 0)$, and $(0, 0, A)$.

Set $p(t) = p_0 + t\,v$: we have $p(0) = p_0$ and $p'(0) = v$. We want to find a path $\iota(t)$ such that $\iota(0) = \iota_0$, $\iota'(t) = V$, and satisfying $F(\iota(t), m_0) = (p(t), 0, 0)$. Let U be some neighborhood of m_0 in M and $K \subset U$ a compact subset. There exists a bump function $g \in C^k(M, \mathbb{R})$ such that $g \equiv 1$ in K and $g \equiv 0$ outside U. Apply a small translation in the direction of $v \in \mathbb{R}^3$ on ι in the neighborhood of m_0, by introducing $\iota(t) = \iota_0 + t\,g\,v$, which for small t belongs to $\mathrm{Emb}^k(M, \mathbb{R}^3)$ since it is open in $C^k(M, \mathbb{R}^3)$. Then $F(\iota(t), m_0) = (p(t), 0, 0)$, and the first problem is solved by taking $V = g\,v \in C^k(M, \mathbb{R}^3)$.

There exists a path $n(t)$ in the sphere such that $n(0) = n_0$ and $n'(0) = N$, which defines a rotation in $\mathbb{R}^3$, denoted by R_t, whose axis is $(n_0, N)^\perp$ and angle $t\,\|N\|$. Applying a small such rotation in the neighborhood of m_0, we define $\iota(t) = (1 - g)\,\iota_0 + g\,R_t(\iota_0)$, which brings the normal n_0 to $n(t)$ while fixing p_0 and dn_0 (g is the same as in the previous paragraph). Thus, $F(\iota(t), m_0) = (0, n(t), 0)$, which solves the second problem.

For the third problem, we can work with principal coordinate systems. There exists a C^k function $h_0 : T_{p_0}\mathcal{S} \cap \mathcal{U} \to \mathbb{R}$ such that $\mathcal{S}$ is locally the graph of h_0, where $\mathcal{U}$ is an open neighborhood of p, and the embedding is described by $\iota_0(m) = p = (a, b, h_0(a, b)) \in \mathcal{S}$ where $a, b \in T_{p_0}\mathcal{S} \cap \mathcal{U}$. We may assume that $\partial_a h = \partial_b h = 0$ at $(0,0)$, so that the normal is $(0, 0, 1)$, and that the principal directions at p_0 are $(1, 0, 0)$ and $(0, 1, 0)$, with $\partial_{a,b}h = \partial_{b,a}h = 0$ and $\partial^2_{a,a}h = \kappa_1$, $\partial^2_{b,b}h = \kappa_2$. Define a bump function (still denoted by g) relative to the neighborhood $T_{p_0}\mathcal{S} \cap \mathcal{U}$, which is in correspondence to some neighborhood U of m_0 in M.

Now, the tangential element A is represented as a 2×2 symmetric matrix in the chosen basis of $T_{p_0}\mathcal{S}$ and has the form $A = R_\theta \operatorname{diag}(\lambda, \mu)\,R_\theta^{-1}$ with $\lambda, \mu \in \mathbb{R}$. Define $\phi : T_{p_0}\mathcal{S} \to \mathbb{R}$ such that $\phi(a, b) = \frac{\lambda}{2}\,a^2 + \frac{\mu}{2}\,b^2$, whose graph is a surface with curvatures λ and μ at 0. Set $h_t = h_0 + t\,g\,\phi \circ R_\theta$, and finally $\iota_t(m) = (a(m), b(m), h_t(a(m), b(m)))$ if $m \in U$, $\iota_t(m) = \iota_0(m)$ elsewhere on M. Then one can show that the perturbed surface fixes p_0 and n_0, but that its shape operator is $dn_t = dn_0 + t\,A$, which is what we wanted.

To conclude, one can see that the proofs for the various forms of F are similar. If m_1 and m_2 are distinct points in M, one can choose non-intersecting neighborhoods and respective bump functions to apply local perturbations to the embedding. $\quad\square$

Lemma 12 *(Shape operator of offset level surfaces). Let q be a point in $\mathbb{R}^3 \setminus \mathcal{S}$ and p one contact point with normal n and at distance r such that q satisfies the strict ball condition* (17) *at p. We have $q = p + r\,n$. Let α be the distance function to an associated contact piece. Then the shape operator at q of the offset surface $\{\alpha = r\}$ is*

$$d_q\mathbf{n}_{\alpha=r} = d_p\mathbf{n}\,(\mathrm{Id} + r\,d_p\mathbf{n})^{-1}. \tag{A.4}$$

Proof. Suppose first that p is not an umbilical point, i.e., $\mathcal{S}$ admits two distinct principal curvatures there. A remarkable property of the principal directions $e_1(\rho), e_2(\rho)$ of the offset surfaces $\{\alpha = \rho\}$ for $\rho \in [0, r]$ is that the orthonormal frame (n, e_1, e_2) is unchanged when traveling in the normal direction from p to q [90]. On the other hand, the principal curvatures obey the following differential equation in the normal direction:

$$D_n(\kappa_1) = \kappa_1^2,$$

$$D_n(\kappa_2) = \kappa_2^2.$$

If $\kappa_i(0) = 0$ at the contact point p, then $\kappa_i(\rho) \cong 0$ for any offset, and eq. (A.4) holds. If $\kappa_i(0) \neq 0$, by assumption $\kappa_i(0) \neq r$ at p, then $\kappa_i(\rho) = \frac{1}{\frac{1}{\kappa_i(0)} - \rho}$ for $\rho \in [0, r]$. In particular, in the basis (n, e_1, e_2),

$$d_q\mathbf{n}_{\alpha=r} = \mathrm{Hess}\,\alpha(q) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\kappa_1(r) & 0 \\ 0 & 0 & -\kappa_2(r) \end{pmatrix} = d_p\mathbf{n}\,(\mathrm{Id} + r\,d_p\mathbf{n})^{-1}.$$

Even if p is umbilical, all offset surfaces $\{\alpha = \rho\}$ are umbilical at $p + \rho n$ and we can choose consistent frames so that eq. (A.4) still holds. $\square$

Lemma 13 *(Riemannian Hessian of the restricted distance function). Let ι be a generic embedding in the sense of Theorem 12. Let q be a critical point of the pure distance $\mathrm{dist}(\cdot, \mathcal{S})$, with value r, contact points $\Gamma(q) = \{p_1, \ldots, p_m\}$ and normals $n_1, \ldots, n_m$ at these contacts. Denote by $\{\alpha_1, \ldots, \alpha_m\}$ the distance functions to some associated contact pieces, and by $g = \mathrm{dist}(\cdot, \mathcal{S})_{|G(q)}$ the restriction of this function to $G(q)$ (Definition 4).*

- *If q is of type A_1^2, then on the A_1^2 sheet stratum $G(q)$ orthogonal to $n_1 = -n_2$,*

$$\mathrm{Hess}^{\mathrm{Riem}} g(q) = \frac{1}{2}(d_q\mathbf{n}_{\alpha_1=r} + d_q\mathbf{n}_{\alpha_2=r}) \tag{A.5}$$

$$= \frac{1}{2}\left(d_{p_1}\mathbf{n}\,(\mathrm{Id} + r\,d_{p_1}\mathbf{n})^{-1} + d_{p_2}\mathbf{n}\,(\mathrm{Id} + r\,d_{p_2}\mathbf{n})^{-1}\right), \tag{A.6}$$

  where $d_q\mathbf{n}_{\alpha_1=r}$ is the shape operator of the submanifolds $\{\alpha_i = r\}$.
- *If q is of type A_1^3, then the A_1^3 curve stratum $G(q)$ is directed by some unit vector t orthogonal to the plane spanned by n_1, n_2, n_3. Then,*

$$\mathrm{Hess}^{\mathrm{Riem}} g(q) \text{ is non-degenerate } \Leftrightarrow t_{2,3} \cdot u_1(t) + t_{3,1} \cdot u_2(t) + t_{1,2} \cdot u_3(t) \neq 0$$

  where $t_{i,j} = n_i \times n_j$ and $u_i = d_{p_i}\mathbf{n} \circ (\mathrm{Id} + r\,d_{p_i}\mathbf{n})^{-1} = d_q\mathbf{n}_{\alpha_i=r}$.

Proof. We apply an extrinsic formula for the Riemannian Hessian of a smooth function defined on a submanifold $\mathcal{M}$ of $\mathbb{R}^3$, given by [2]. If $f : \mathcal{M} \to \mathbb{R}$ is defined on the submanifold and $\bar{f} : \mathbb{R}^3 \to \mathbb{R}$ extends f to a neighborhood of $\mathcal{M}$, then for $p \in \mathcal{M}$ and $v \in T_p\mathcal{M}$,

$$\mathrm{Hess}^{\mathrm{Riem}} f(p)[v] = \Pi_p \mathrm{Hess}\,\bar{f}(p)[v] + W_p(v, \Pi_p^\perp \nabla \bar{f}(p)),$$

where the Weingarten map W_p of the submanifold $\mathcal{M}$ at p is the operator that associates to $v \in T_p\mathcal{M}$ and $\mathbf{N}_0 \in T_p^\perp \mathcal{M}$ the value $W_p(v, \mathbf{N}_0) = -\Pi_p D_v \mathbf{N}$ for any local extension $\mathbf{N}$

of $\mathbf{N}_0$ to a normal vector field on $\mathcal{M}$. Here, $D_v\mathbf{N}$ designates the directional derivative of $\mathbf{N}$ along the direction v, Π_p is the projection onto the tangent space $T_p\mathcal{M}$, and $\Pi_p^\perp$ the projection onto the normal space $T_p^\perp\mathcal{M}$. Hess $\bar{f}(p)$ and $\nabla\bar{f}(p)$ are the usual Hessian and gradient of a scalar function defined in $\mathbb{R}^3$.

Reformulating the previous equality, we obtain

$$\text{Hess}^{\text{Riem}}f(p)[v] = \Pi_p\,\text{Hess}\bar{f}(p)[v] - \Pi_p D_v(\Pi_{(\cdot)}^\perp\nabla\bar{f}(\cdot)).$$

We can now express the Riemannian Hessian at q of the function $\text{dist}(\cdot,\mathcal{S})_{|G(q)}$ defined on the (germ of) submanifold $G(q) = \{\alpha_1 = \ldots = \alpha_m = \text{dist}(\cdot,\mathcal{S})\}$.

Here, the extension to consider is any function α_i, for which we take advantage of well-known properties of distance functions. In particular, $\text{Hess}\,\alpha_i(q)[n_i] = 0$ and $\nabla\alpha_i = n_i \in T_q^\perp G(q)$. Observe that $\text{Hess}\,\alpha_i(q)$ is just the usual shape operator $d_q\mathbf{n}_{\alpha_i=r} : T_q\{\alpha_i = r\} \to T_q\{\alpha_i = r\}$ of the surface defined by $\{\alpha_i = r\}$, with orientation given by n_i at q. Also, let $\mathbf{n}_G^i$ denote a local extension of $\nabla\alpha_i$ to a normal vector field on the stratum $G(q)$.

In particular, if $G(q)$ is a A_1^2 sheet, $\mathbf{n}_G^i$ is the Gauss map with the same orientation as n_i at q, so that $d_q\mathbf{n}_G^i : T_q G(q) \to T_q G(q)$ will just be the usual shape operator, but this time of the submanifold $G(q)$. The operators $d_q\mathbf{n}_{\alpha_i=r}$ and $d_q\mathbf{n}_G^i$ are defined on the same tangent space: $T_q\{\alpha_i = r\} = T_q G(q)$. If $G(q)$ is a A_1^3 curve, then its tangent direction t is contained in any $T_q\{\alpha_i = r\} \supsetneq T_q G(q) = \mathbb{R}\,t$.

q **is of type** A_1^2**.** In this case, we get

$$\text{Hess}^{\text{Riem}}g(q)[v] = \text{Hess}\,\alpha_1(q)[v] - D_v(\nabla\alpha_{1|G(q)}))$$
$$= d_q\mathbf{n}_{\alpha_i=r}[v] - d_q\mathbf{n}_G^i[v].$$

Since $n_1 = -n_2$, we can choose $\mathbf{n}_G^1 = -\mathbf{n}_G^2$, to obtain

$$\text{Hess}^{\text{Riem}}g(q) = d_q\mathbf{n}_{\alpha_1=r} - d_q\mathbf{n}_G^1$$
$$= d_q\mathbf{n}_{\alpha_2=r} - d_q\mathbf{n}_G^2$$
$$= \frac{d_q\mathbf{n}_{\alpha_1=r} + d_q\mathbf{n}_{\alpha_2=r}}{2}.$$

Using that $d_q\mathbf{n}_{\alpha_1=r} = d_{p_i}\mathbf{n}\,(\text{Id} + r\,d_{p_i}\mathbf{n})^{-1}$ thanks to Lemma 12, we establish the first case.

q **is of type** A_1^3**.** Now, the expression involves a projection on $\mathbb{R}\,t$. We use the concise notation $u_i(t) = d_q\mathbf{n}_{\alpha_i=r}[t]$. Applying the formula above, we get

$$h = \text{Hess}^{\text{Riem}}g(q)[t] \cdot t = t \cdot (u_1(t) - d_q\mathbf{n}_G^1[t])$$
$$= t \cdot (u_2(t) - d_q\mathbf{n}_G^2[t])$$
$$= t \cdot (u_3(t) - d_q\mathbf{n}_G^3[t]) \in \mathbb{R}.$$

Two of the vectors n_i, say n_1 and n_2, form a basis of the plane spanned by n_1, n_2, n_3. In that basis, $n_3 = \lambda\, n_1 + \mu\, n_2$. Importantly, we can choose the local extensions in such a way that $\mathbf{n}_G^3 = \lambda\, \mathbf{n}_G^1 + \mu\, \mathbf{n}_G^2$, hence $d_q \mathbf{n}_G^3[t] = \lambda\, d_q \mathbf{n}_G^1[t] + \mu\, d_q \mathbf{n}_G^2[t]$.

After summing the three lines with the weights $-\lambda, -\mu, 1$ respectively, the second terms of each line cancel out and we obtain

$$(1 - \lambda - \mu)\, h = t \cdot (u_3(t) - \lambda\, u_1(t) - \mu\, u_2(t)).$$

By some trigonometric computations, one can see that

$$\lambda = \frac{\sin \theta_{2,3}}{\sin \theta_{2,1}} \quad \mu = \frac{\sin \theta_{1,3}}{\sin \theta_{1,2}}$$

where $\theta_{i,j} = \widehat{(n_i, n_j)}$ is the algebraic angle in the plane orthogonal to t, with sign convention given by t. This way, $t_{i,j} = n_i \times n_j = \sin \theta_{i,j}\, t$. Also, $1 \neq \lambda + \mu$ because $n_3 \neq n_1$ and $n_3 \neq n_2$. Therefore, we are led to the desired inequality:

$$h \neq 0 \Leftrightarrow t \cdot \left(u_3(t) - \frac{\sin \theta_{2,3}}{\sin \theta_{2,1}}\, u_1(t) - \frac{\sin \theta_{1,3}}{\sin \theta_{1,2}}\, u_2(t) \right) \neq 0$$

$$\Leftrightarrow t \cdot (\sin \theta_{1,2}\, u_3(t) + \sin \theta_{2,3}\, u_1(t) + \sin \theta_{3,1}\, u_2(t)) \neq 0$$

$$\Leftrightarrow t_{1,2} \cdot u_1(t) + t_{2,3} \cdot u_2(t) + t_{3,1} \cdot u_3(t) \neq 0. \quad \square$$

References

[1] R. Abraham, J. Robbin, Transversal Mappings and Flows, WA Benjamin, New York, 1967.
[2] P.A. Absil, R. Mahony, J. Trumpf, An extrinsic look at the Riemannian Hessian, in: F. Nielsen, F. Barbaresco (Eds.), Geometric Science of Information, in: Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2013, pp. 361–368.
[3] R.J. Adler, The Geometry of Random Fields, Classics in Applied Mathematics, Society for Industrial and Applied Mathematics, 2010, https://epubs.siam.org/doi/book/10.1137/1.9780898718980.
[4] A. Agrachev, D. Pallaschke, S. Scholtes, On Morse theory for piecewise smooth functions, J. Dyn. Control Syst. 3 (4) (1997) 449–469, https://doi.org/10.1007/BF02463278.
[5] P. Albano, P. Cannarsa, K.T. Nguyen, C. Sinestrari, Singular gradient flow of the distance function and homotopy equivalence, Math. Ann. 356 (1) (2013) 23–43, https://doi.org/10.1007/s00208-012-0835-8.
[6] R.T. Armstrong, Z. Lanetc, P. Mostaghimi, A. Zhuravljov, A. Herring, V. Robins, Correspondence of max-flow to the absolute permeability of porous systems, Phys. Rev. Fluids 6 (5) (2021) 054003, https://doi.org/10.1103/PhysRevFluids.6.054003, https://link.aps.org/doi/10.1103/PhysRevFluids.6.054003.
[7] V.I. Arnol'd, Normal forms of functions in neighbourhoods of degenerate critical points, Russ. Math. Surv. 29 (2) (1974) 10.
[8] D. Attali, J.-D. Boissonnat, H. Edelsbrunner, Stability and computation of medial axes - a state-of-the-art report, in: T. Möller, B. Hamann, R.D. Russell (Eds.), Mathematical Foundations of Scientific Visualization, Computer Graphics, and Massive Data Exploration, Mathematics and Visualization, Springer, Berlin, Heidelberg, 2009, pp. 109–125, https://doi.org/10.1007/b106657_6.
[9] U. Ayachit, The ParaView Guide: A Parallel Visualization Application, Kitware, Inc., 2015.
[10] T. Banchoff, Critical points and curvature for embedded polyhedra, J. Differ. Geom. 1 (3–4) (Jan. 1967), https://doi.org/10.4310/jdg/1214428092, https://projecteuclid.org/journals/journal-of-differential-geometry/volume-1/issue-3-4/Critical-points-and-curvature-for-embedded-polyhedra/10.4310/jdg/1214428092.full.

[11] U. Bauer, Persistence in discrete Morse theory, Ph.D. thesis, Georg-August-Universität Göttingen, Jul. 2011, https://ediss.uni-goettingen.de/handle/11858/00-1735-0000-0006-B3E6-C.

[12] U. Bauer, C. Lange, M. Wardetzky, Optimal topological simplification of discrete functions on surfaces, Discrete Comput. Geom. 47 (2012) 347–377, https://doi.org/10.1007/s00454-011-9350-z.

[13] L. Birbrair, M.P. Denkowski, Medial axis and singularities, J. Geom. Anal. 27 (3) (2017) 2339–2380, https://doi.org/10.1007/s12220-017-9763-x, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6294241/.

[14] B. Bleile, A. Garin, T. Heiss, K. Maggs, V. Robins, The persistent homology of dual digital image constructions, in: E. Gasparovic, V. Robins, K. Turner (Eds.), Research in Computational Topology 2, in: Association for Women in Mathematics Series, Springer International Publishing, Cham, 2022, pp. 1–26, https://doi.org/10.1007/978-3-030-95519-9_1.

[15] E.D. Bloch, Polyhedral representation of discrete Morse functions, Discrete Math. 313 (12) (2013) 1342–1348, https://doi.org/10.1016/j.disc.2013.02.020, https://linkinghub.elsevier.com/retrieve/pii/S0012365X13001052.

[16] O. Bobrowski, R.J. Adler, Distance functions, critical points, and the topology of random Čech complexes, Homol. Homotopy Appl. 16 (2) (2014) 311–344, https://doi.org/10.4310/HHA.2014.v16.n2.a18, http://www.intlpress.com/site/pub/pages/journals/items/hha/content/vols/0016/0002/a018/.

[17] O. Bobrowski, G. Oliveira, Random Čech complexes on Riemannian manifolds, Random Struct. Algorithms 54 (3) (2019) 373–412, https://doi.org/10.1002/rsa.20800Citations.

[18] O. Bobrowski, S. Weinberger, On the vanishing of homology in random Čech complexes, Random Struct. Algorithms 51 (1) (2017) 14–51, https://doi.org/10.1002/rsa.20697.

[19] K. Borsuk, On the imbedding of systems of compacta in simplicial complexes, Fundam. Math. 35 (1) (1948) 217–234, https://eudml.org/doc/213158.

[20] R. Bott, Morse theory indomitable, Publ. Math. IHÉS 68 (1988) 99–114.

[21] N. Bourbaki, General Topology: Chapters 1–4, vol. 18, Springer Science & Business Media, 2013.

[22] J.W. Bruce, J.W. Bruce, P. Giblin, Curves and Singularities: A Geometrical Introduction to Singularity Theory, Cambridge University Press, 1992.

[23] J. Cantwell, Topological non-degenerate functions, Tohoku Math. J. 20 (2) (1968) 120–125, https://doi.org/10.2748/tmj/1178243171, https://projecteuclid.org/journals/tohoku-mathematical-journal/volume-20/issue-2/Topological-non-degenerate-functions/10.2748/tmj/1178243171.full.

[24] Y. Cao, A. Monod, Approximating persistent homology for large datasets, arXiv:2204.09155 [cs, math, stat], May 2022, https://doi.org/10.48550/arXiv.2204.09155.

[25] G. Carlsson, Topology and data, Bull. Am. Math. Soc. 46 (2) (2009) 255–308, https://doi.org/10.1090/S0273-0979-09-01249-X, https://www.ams.org/bull/2009-46-02/S0273-0979-09-01249-X/.

[26] F. Cazals, F. Chazal, T. Lewiner, Molecular shape analysis based upon the Morse-Smale complex and the Connolly function, in: Proceedings of the Nineteenth Annual Symposium on Computational Geometry, 2003, pp. 351–360.

[27] F. Cazals, M. Pouget, Differential topology and geometry of smooth embedded surfaces: selected topics, Int. J. Comput. Geom. Appl. 15 (05) (2005) 511–536, https://doi.org/10.1142/S0218195905001816, https://www.worldscientific.com/doi/abs/10.1142/S0218195905001816.

[28] F. Chazal, L.J. Guibas, S.Y. Oudot, P. Skraba, Persistence-based clustering in Riemannian manifolds, J. ACM 60 (6) (2013) 1–38, https://doi.org/10.1145/2535927.

[29] F. Chazal, A. Lieutier, Stability and computation of topological invariants of solids in Rⁿ, Discrete Comput. Geom. 37 (4) (2007) 601–617, https://doi.org/10.1007/s00454-007-1309-8.

[30] F. Chazal, R. Soufflet, Stability and finiteness properties of medial axis and skeleton, J. Dyn. Control Syst. 10 (2) (2004) 149–170, https://doi.org/10.1023/B:JODS.0000024119.38784.ff.

[31] J. Cheeger, Critical points of distance functions and applications to geometry, in: P. de Bartolomeis, F. Tricerri (Eds.), Geometric Topology: Recent Developments, in: Lecture Notes in Mathematics, vol. 1504, Springer, Berlin Heidelberg, Berlin, Heidelberg, 1991, pp. 1–38, http://link.springer.com/10.1007/BFb0094288.

[32] F.H. Clarke, Optimization and Nonsmooth Analysis, SIAM, 1990.

[33] D. Cohen-Steiner, H. Edelsbrunner, J. Harer, Stability of persistence diagrams, Discrete Comput. Geom. 37 (1) (2007) 103–120, https://doi.org/10.1007/s00454-006-1276-5.

[34] A. Commaret, Generalized Morse theory for tubular neighborhoods, arXiv preprint, arXiv:2401.04034, 2024.

[35] J.-N. Corvellec, On the second deformation lemma, Topol. Methods Nonlinear Anal. 17 (1) (2001) 55–66, https://apcz.umk.pl/TMNA/article/view/TMNA.2001.004.

[36] J.-N. Corvellec, M. Degiovanni, M. Marzocchi, Deformation properties for continuous functionals and critical point theory, Topol. Methods Nonlinear Anal. (1993) 151–171, https://apcz.umk.pl/TMNA/article/view/TMNA.1993.012.

[37] L. Crawford, A. Monod, A.X. Chen, S. Mukherjee, R. Rabadán, Predicting clinical outcomes in glioblastoma: an application of topological and functional data analysis, J. Am. Stat. Assoc. 115 (531) (2020) 1139–1150, https://doi.org/10.1080/01621459.2019.1671198.

[38] W. Crawley-Boevey, Decomposition of pointwise finite-dimensional persistence modules, J. Algebra Appl. 14 (05) (2015) 1550066, https://doi.org/10.1142/S0219498815500668.

[39] J. Curry, R. Ghrist, V. Nanda, Discrete Morse theory for computing cellular sheaf cohomology, Found. Comput. Math. 16 (2016) 875–897, https://doi.org/10.1007/s10208-015-9266-8.

[40] J. Damon, The global medial structure of regions in R3, Geom. Topol. 10 (4) (2006) 2385–2429, https://doi.org/10.2140/gt.2006.10.2385, http://www.msp.org/gt/2006/10-4/p11.xhtml.

[41] H.-L. de Kergorlay, U. Tillmann, O. Vipond, Random Čech complexes on manifolds with boundary, Random Struct. Algorithms 61 (2) (2022) 309–352, https://doi.org/10.1002/rsa.21062.

[42] M.C. Delfour, J.P. Zolésio, Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization, second edition, Society for Industrial and Applied Mathematics, 2011, http://epubs.siam.org/doi/book/10.1137/1.9780898719826.

[43] O. Delgado-Friedrichs, V. Robins, A. Sheppard, Morse theory and persistent homology for topological analysis of 3D images of complex materials, in: 2014 IEEE International Conference on Image Processing (ICIP), 2014, pp. 4872–4876, https://doi.org/10.1109/ICIP.2014.7025987.

[44] O. Delgado-Friedrichs, V. Robins, A. Sheppard, Skeletonization and partitioning of digital images using discrete Morse theory, IEEE Trans. Pattern Anal. Mach. Intell. 37 (3) (2015) 654–666, https://doi.org/10.1109/TPAMI.2014.2346172.

[45] P. Dlotko, Cubical complex (2021). URL, https://gudhi.inria.fr/doc/3.4.1/group___cubical__complex.html.

[46] H. Edelsbrunner, J. Harer, Persistent homology—a survey, in: J.E. Goodman, J. Pach, R. Pollack (Eds.), Contemporary Mathematics, vol. 453, American Mathematical Society, Providence, Rhode Island, 2008, pp. 257–282, https://doi.org/10.1090/conm/453/08802, http://www.ams.org/conm/453/.

[47] H. Edelsbrunner, J. Harer, A. Zomorodian, Hierarchical Morse complexes for piecewise linear 2-manifolds, in: Proceedings of the Seventeenth Annual Symposium on Computational Geometry, 2001, pp. 70–79.

[48] H. Federer, Curvature measures, Trans. Am. Math. Soc. 93 (3) (1959) 418–491.

[49] S. Ferry, When eps-boundaries are manifolds, Fundam. Math. 90 (3) (1976) 199–210, https://eudml.org/doc/214901.

[50] R. Forman, Morse theory for cell complexes, Adv. Math. 134 (1) (1998) 90–145, https://doi.org/10.1006/aima.1997.1650, https://linkinghub.elsevier.com/retrieve/pii/S0001870897916509.

[51] J.H. Fu, Curvature measures and generalized Morse theory, J. Differ. Geom. 30 (3) (1989) 619–642.

[52] P. Gabriel, Unzerlegbare darstellungen I, Manuscr. Math. 6 (1972) 71–103, https://doi.org/10.1007/BF01298413.

[53] C. Gaetan, X. Guyon, Spatial Statistics and Modeling, Springer Series in Statistics, Springer New York, New York, NY, 2010, http://link.springer.com/10.1007/978-0-387-92257-7.

[54] V. Gershkovich, Singularity theory for Riemannian distance functions on non-positively curved surfaces, in: Geometry from the Pacific Rim: Proceedings of the Pacific Rim Geometry Conference Held at National University of Singapore, Republic of Singapore, December 12–17, 1994, De Gruyter, 1997, p. 117.

[55] V. Gershkovich, H. Rubinstein, Morse theory for Min-type functions, Asian J. Math. 1 (4) (1997) 696–715, https://doi.org/10.4310/AJM.1997.v1.n4.a3, http://www.intlpress.com/site/pub/pages/journals/items/ajm/content/vols/0001/0004/a003/.

[56] R. Ghrist, Barcodes: the persistent topology of data, Bull. Am. Math. Soc. 45 (1) (2008) 61–75, https://doi.org/10.1090/S0273-0979-07-01191-3, https://www.ams.org/bull/2008-45-01/S0273-0979-07-01191-3/.

[57] R.W. Ghrist, Elementary Applied Topology, Createspace, Seattle, 2014.

[58] P. Giblin, Symmetry sets and medial axes in two and three dimensions, in: R. Cipolla, R. Martin (Eds.), The Mathematics of Surfaces IX, Springer, London, 2000, pp. 306–321.

[59] J. Giesen, M. John, The flow complex: a data structure for geometric modeling, Comput. Geom. 39 (3) (2008) 178–190, https://doi.org/10.1016/j.comgeo.2007.01.002, https://www.sciencedirect.com/science/article/pii/S0925772107000247.

[60] J. Giesen, L. Kuehne, A parallel algorithm for computing the flow complex, in: Proceedings of the Twenty-Ninth Annual Symposium on Computational Geometry, SoCG'13, Association for Computing Machinery, New York, NY, USA, 2013, pp. 57–66, https://doi.org/10.1145/2462356.2462384.

[61] J. Giesen, E.A. Ramos, B. Sadri, Medial axis approximation and unstable flow complex, in: Proceedings of the Twenty-Second Annual Symposium on Computational Geometry, SCG'06, Association for Computing Machinery, New York, NY, USA, 2006, pp. 327–336, https://doi.org/10.1145/1137856.1137905.

[62] D. Gilbarg, N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, Classics in Mathematics, Springer, Berlin, Heidelberg, 1977.

[63] M. Goresky, R. MacPherson, Stratified Morse theory, in: Stratified Morse Theory, Springer Berlin Heidelberg, Berlin, Heidelberg, 1988, pp. 3–22, http://link.springer.com/10.1007/978-3-642-71714-7_1.

[64] K. Grove, Critical point theory for distance functions, Proc. Symp. Pure Math. 54 (3) (1993) 357–385.

[65] K. Grove, K. Shiohama, A generalized sphere theorem, Ann. Math. 106 (1) (1977) 201–211, https://doi.org/10.2307/1971164, https://www.jstor.org/stable/1971164.

[66] V. Guillemin, A. Pollack, Differential Topology, AMS Chelsea Publishing, 2010, https://books.google.co.uk/books?id=FdRhAQAAQBAJ.

[67] S. Harker, K. Mischaikow, M. Mrozek, V. Nanda, Discrete Morse theoretic algorithms for computing homology of complexes and maps, Found. Comput. Math. 14 (2014) 151–184, https://doi.org/10.1007/s10208-013-9145-0.

[68] A.L. Herring, V. Robins, A.P. Sheppard, Topological persistence for relating microstructure and capillary fluid trapping in sandstones, Water Resour. Res. 55 (1) (2019) 555–573, https://doi.org/10.1029/2018WR022780, https://onlinelibrary.wiley.com/doi/10.1029/2018WR022780.

[69] M.W. Hirsch, Differential Topology, Springer New York, New York, NY, 1976, http://archive.org/details/springer_10.1007-978-1-4684-9449-5.

[70] M. Hohenwarter, M. Borcherds, G. Ancsin, B. Bencze, M. Blossier, a. Delobelle, C. Denizet, J. Éliás, a. Fekete, L. Gál, Z. Konečný, Z. Kovács, S. Lizelfelner, B. Parisse, G. Sturr, GeoGebra 4.4, Dec. 2013.

[71] X. Hu, F. Li, D. Samaras, C. Chen, Topology-preserving deep image segmentation, Adv. Neural Inf. Process. Syst. 32 (2019), https://proceedings.neurips.cc/paper/2019/file/2d95666e2649fcfc6e3af75e09f5adb9-Paper.pdf.

[72] X. Hu, D. Samaras, C. Chen, Learning probabilistic topological representations using discrete Morse theory, arXiv:2206.01742 [cs, eess], Oct. 2022, http://arxiv.org/abs/2206.01742.

[73] J.-I. Itoh, T. Sakai, Cut loci and distance functions, Math. J. Okayama Univ. 49 (1) (2007) 65–92.

[74] B. Julesz, Visual pattern discrimination, IRE Trans. Inf. Theory 8 (2) (1962) 84–92, https://doi.org/10.1109/TIT.1962.1057698.

[75] T. Kaczynski, K. Mischaikow, M. Mrozek, Cubical homology, in: T. Kaczynski, K. Mischaikow, M. Mrozek (Eds.), Computational Homology, in: Applied Mathematical Sciences, Springer, New York, NY, 2004, pp. 39–92, https://doi.org/10.1007/0-387-21597-2_2.

[76] S. Kaji, T. Sudo, K. Ahara, Cubical Ripser: software for computing persistent homology of image and volume data, Tech. Rep., arXiv:2005.12692 [cs, math], Jun. 2020, https://doi.org/10.48550/arXiv.2005.12692.

[77] R.C. Kirby, L.C. Siebenmann, Foundational Essays on Topological Manifolds, Smoothings, and Triangulations (AM-88), Princeton University Press, 1977, https://www.jstor.org/stable/j.ctt1b9s024.

[78] S.G. Krantz, H.R. Parks, Distance to Ck hypersurfaces, J. Differ. Equ. 40 (1) (1981) 116–120, https://doi.org/10.1016/0022-0396(81)90013-9, https://www.sciencedirect.com/science/article/pii/0022039681900139.

[79] D. Landis, M. Morse, Tractions in critical point theory, Rocky Mt. J. Math. 5 (3) (1975) 379–399, https://www.jstor.org/stable/44236445.

[80] M. Landy, N. Graham, Visual perception of texture, in: L. Chalupa, J. Werner (Eds.), The Visual Neurosciences, MIT Press, Cambridge, MA, 2004, p. 1106.

[81] D.T. Lee, Medial axis transformation of a planar shape, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-4 (4) (1982) 363–369, https://doi.org/10.1109/TPAMI.1982.4767267.

[82] J.M. Lee, Introduction to Smooth Manifolds, 2nd edition, Graduate Texts in Mathematics, vol. 218, Springer, New York, London, 2013, oCLC: ocn800646950.

[83] Y. Lee, S.D. Barthel, P. Dłotko, S.M. Moosavi, K. Hess, B. Smit, Quantifying similarity of pore-geometry in nanoporous materials, Nat. Commun. 8 (1) (2017) 15396, https://doi.org/10.1038/ncomms15396, https://www.nature.com/articles/ncomms15396.

[84] T. Lewiner, Critical sets in discrete Morse theories: relating Forman and piecewise-linear approaches, Comput. Aided Geom. Des. 30 (6) (2013) 609–621, https://doi.org/10.1016/j.cagd.2012.03.012, https://linkinghub.elsevier.com/retrieve/pii/S0167839612000386.

[85] A. Lieutier, Any open bounded subset of Rn has the same homotopy type than its medial axis, in: Proceedings of the Eighth ACM Symposium on Solid Modeling and Applications, SM'03, Association for Computing Machinery, New York, NY, USA, 2003, pp. 65–75, https://doi.org/10.1145/781606.781620.

[86] E.L. Lima, The Jordan-Brouwer separation theorem for smooth hypersurfaces, Am. Math. Mon. 95 (1) (1988) 39–42, https://doi.org/10.2307/2323445, https://www.jstor.org/stable/2323445.

[87] W.B. Lindquist, S.-M. Lee, D.A. Coker, K.W. Jones, P. Spanne, Medial axis analysis of void structure in three-dimensional tomographic images of porous media, J. Geophys. Res., Solid Earth 101 (B4) (1996) 8297–8310, https://doi.org/10.1029/95JB03039, eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/95JB03039.

[88] J.N. Mather, Distance from a submanifold in Euclidean space, Proc. Symp. Pure Math. 40 (1983) 199–216.

[89] V.I. Matov, Singularities of the maximum function on a manifold with boundary, J. Sov. Math. 33 (4) (1986) 1103–1127, https://doi.org/10.1007/BF01086112.

[90] D. Mayost, Applications of the Signed Distance Function to Surface Geometry, Thesis, University of Toronto, 2014, https://tspace.library.utoronto.ca/handle/1807/80184.

[91] P. McGrath, On the smooth Jordan Brouwer separation theorem, Am. Math. Mon. 123 (3) (2016) 292–295, https://doi.org/10.4169/amer.math.monthly.123.3.292, https://www.jstor.org/stable/10.4169/amer.math.monthly.123.3.292.

[92] J.W. Milnor, Morse Theory, Princeton University Press, 1963.

[93] C. Moon, S.A. Mitchell, J.E. Heath, M. Andrew, Statistical inference over persistent homology predicts fluid flow in porous media, Water Resour. Res. 55 (11) (2019) 9592–9603, https://doi.org/10.1029/2019WR025171, https://onlinelibrary.wiley.com/doi/10.1029/2019WR025171.

[94] M. Morse, *F*-deformations and *F*-tractions, Proc. Natl. Acad. Sci. USA 70 (6) (1973) 1634–1635, https://doi.org/10.1073/pnas.70.6.1634, https://pnas.org/doi/full/10.1073/pnas.70.6.1634.

[95] M. Morse, Topologically non-degenerate functions on a compact n-manifold M, J. Anal. Math. 7 (1) (1959) 189–208, https://doi.org/10.1007/BF02787685.

[96] S. Müller, L. Schüler, A. Zech, F. Heße, GSTools v1.3: a toolbox for geostatistical modelling in Python, Geosci. Model Dev. 15 (7) (2022) 3161–3182, https://doi.org/10.5194/gmd-15-3161-2022, https://gmd.copernicus.org/articles/15/3161/2022/.

[97] P. Niyogi, S. Smale, S. Weinberger, Finding the homology of submanifolds with high confidence from random samples, Discrete Comput. Geom. 39 (1) (2008) 419–441, https://doi.org/10.1007/s00454-008-9053-2.

[98] I. Obayashi, T. Nakamura, Y. Hiraoka, Persistent homology analysis for materials research and persistent homology software: HomCloud, J. Phys. Soc. Jpn. 91 (9) (2022) 091013, https://doi.org/10.7566/JPSJ.91.091013, https://journals.jps.jp/doi/10.7566/JPSJ.91.091013.

[99] S. Osher, R. Fedkiw, Signed distance functions, in: J.E. Marsden, S.S. Antman, L. Sirovich (Eds.), Level Set Methods and Dynamic Implicit Surfaces, in: Applied Mathematical Sciences, vol. 153, Springer, New York, New York, NY, 2003, pp. 17–22, http://link.springer.com/10.1007/0-387-22746-6_2.

[100] J.J. Park, P. Florence, J. Straub, R. Newcombe, S. Lovegrove, DeepSDF: learning continuous signed distance functions for shape representation, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 165–174, https://doi.org/10.1109/CVPR.2019.00025.

[101] J. Portilla, E.P. Simoncelli, A parametric texture model based on joint statistics of complex wavelet coefficients, Int. J. Comput. Vis. 40 (1) (2000) 49–70, https://doi.org/10.1023/A:1026553619983.

[102] V. Robins, P.J. Wood, A.P. Sheppard, Theory and algorithms for constructing discrete Morse complexes from grayscale digital images, IEEE Trans. Pattern Anal. Mach. Intell. 33 (8) (2011) 1646–1658, https://doi.org/10.1109/TPAMI.2011.95.

[103] E. Saucan, Discrete Morse theory, persistent homology and Forman-Ricci curvature, arXiv:2003.03844 [math], https://doi.org/10.48550/arXiv.2003.03844, http://arxiv.org/abs/2003.03844, Mar. 2020.

[104] L.C. Siebenmann, Deformation of homeomorphisms on stratified sets, Comment. Math. Helv. 47 (1) (1972) 123–163, https://doi.org/10.1007/BF02566793.

[105] C. Sigg, R. Peikert, M. Gross, Signed distance transform using graphics hardware, in: IEEE Visualization 2003 (VIS 2003), 2003, pp. 83–90.

[106] J.R. Silvester, Determinants of block matrices, Math. Gaz. 84 (501) (2000) 460–467, https://doi.org/10.2307/3620776, https://www.jstor.org/stable/3620776.

[107] A. Song, Generation of tubular and membranous shape textures with curvature functionals, J. Math. Imaging Vis. 64 (1) (2022) 17–40, https://doi.org/10.1007/s10851-021-01049-9, https://link.springer.com/article/10.1007/s10851-021-01049-9.

[108] B.J. Stolz, J. Kaeppler, B. Markelc, F. Braun, F. Lipsmeier, R.J. Muschel, H.M. Byrne, H.A. Harrington, Multiscale topology characterizes dynamic tumor vascular networks, Sci. Adv. 8 (23) (2022) eabm2456, https://doi.org/10.1126/sciadv.abm2456, https://www.science.org/doi/10.1126/sciadv.abm2456.

[109] G. Tauzin, U. Lupo, L. Tunstall, J.B. Pérez, M. Caorsi, A.M. Medina-Mardones, A. Dassatti, K. Hess, giotto-tda: a topological data analysis toolkit for machine learning and data exploration, J. Mach. Learn. Res. 22 (1) (2022) 39:1834–39:1839.

[110] The GUDHI Project, GUDHI User and Reference Manual, GUDHI Editorial Board, 2015, http://gudhi.gforge.inria.fr/doc/latest/.

[111] J. Tierny, V. Pascucci, Generalized topological simplification of scalar fields on surfaces, IEEE Trans. Vis. Comput. Graph. 18 (12) (2012) 2005–2013, https://doi.org/10.1109/TVCG.2012.228.

[112] H. Wagner, C. Chen, E. Vuçini, Efficient computation of persistent homology for cubical data, in: R. Peikert, H. Hauser, H. Carr, R. Fuchs (Eds.), Topological Methods in Data Analysis and Visualization II, in: Mathematics and Visualization, Springer, Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 91–106, http://link.springer.com/10.1007/978-3-642-23175-9_7.

[113] Y. Yomdin, On the local structure of a generic central set, Compos. Math. (1981) 15.

[114] A. Zomorodian, G. Carlsson, Computing persistent homology, in: Proceedings of the Twentieth Annual Symposium on Computational Geometry, 2004, pp. 347–356.