Skeletonization and partitioning of digital images using discrete Morse theory

Olaf Delgado-Friedrichs, Vanessa Robins, Adrian Sheppard

Abstract—We show how discrete Morse theory provides a rigorous and unifying foundation for defining skeletons and partitions of greyscale digital images. We model a greyscale image as a cubical complex with a real-valued function defined on its vertices (the voxel values). This function is extended to a discrete gradient vector field using the algorithm presented in Robins, Wood, Sheppard TPAMI 33:1646 (2011). In the current paper we define basins (the building blocks of a partition) and segments of the skeleton using the stable and unstable sets associated with critical cells. The natural connection between Morse theory and homology allows us to prove the topological validity of these constructions; for example, that the skeleton is homotopic to the initial object. We simplify the basins and skeletons via Morse-theoretic cancellation of critical cells in the discrete gradient vector field using a strategy informed by persistent homology. Simple working Python code for our algorithms for efficient vector field traversal is included. Example data are taken from micro-CT images of porous materials, an application area where accurate topological models of pore connectivity are vital for fluid-flow modelling.

Index Terms—curve skeleton, surface skeleton, medial axis transform, watershed transform, discrete Morse theory, persistent homology.

1 Introduction

Skeletonization and partitioning operations are well established approaches to summarising shapes in digital image analysis. The skeleton of a shape is a lower-dimensional object that retains essential aspects of the shape's topology and geometry, typically via a centred curve but it may also incorporate medial surface patches [1]. Partitioning, or shape decomposition, refers to the identification of simple regions that divide the shape into meaningful parts [2], [3]. Many researchers have observed the interrelation between the two constructions; intuitively the skeleton describes how the regions of the partition are joined together. Despite this correspondence, the two operations are usually treated as completely separate computations, although some recent work has started to combine the two [4]. In this paper we define skeletonization and partitioning operations for digital images based on a single construction from discrete Morse theory.

Morse theory connects the topology of the sub-level sets of a real-valued function to its critical points [5]. With digital images the real-valued function, f, might be the intensity values of a greyscale image or a distance function derived from a binarized image and the sub-level set consists of all voxels x where $f(x) \leq c$. The relevant definitions are given in Section 2. The Morse partition defined in this paper is based on stable basins surrounding local minima and is closely related to the watershed transform. The Morse skeleton is built from paths between critical points and generalises constructions such as the contour or component tree (also called the Reeb graph), and the medial axis. Both are defined here via a discrete vector field derived from a digital image using the algorithm described in [6]. As in that paper, the machinery of discrete Morse theory [7] allows us to prove the topological validity of our constructions. These constructions are described in Section 3.

Any practical algorithm for image analysis must incorporate a method for removing the artefacts of noise; in greyscale images this manifests as extraneous local extrema. A further advantage of Morse theory is that it has a built-in technique for cancelling pairs of critical points. We exploit this and the theory of persistent homology to simplify the Morse skeleton and partition simultaneously in a way that preserves the most important structural features of an image. This is described in Section 4.

We have implemented efficient serial and parallel algorithms for constructing, traversing and simplifying the discrete gradient vector field in C++ and provide working Python code for some of the essential routines in Section 5. The code is applied to some reference 3D images, and some micro-CT images of porous rocks.

1.1 Related work

There is a vast literature on the subject of skeletonization for shape analysis spread across a number of application domains, and a comprehensive review is beyond the scope of this paper. A broad overview may be found in the book by Siddiqi and Pizer [1], and a comparative survey of various algorithms in [8]. Typically, a skeleton is a lower-dimensional object that is homotopic to the

• *All authors are at the Department of Applied Mathematics, Research School of Physics and Engineering, The Australian National University, Canberra ACT 0200.*
E-mail: vanessa.robins@anu.edu.au

initial shape, centred with respect to the shape's boundary; some definitions have an associated function that reconstructs the original shape. The basic concept has its roots in the 1967 paper by Blum where the skeleton or medial axis is imagined as the set of corner points generated as a front propagates from the boundary of a 2d shape [9]. In current applications the initial data may be binary or greyscale, 2D, 3D or 4D digital images, surface meshes, or unstructured point clouds [1]. Algorithms for constructing a skeleton from data are based on finding singularities and ridges in distance or potential functions, thinning by boundary propagation, or geometric methods including Voronoi diagrams [8], [10]. For digital images in particular, topology-preserving thinning by removing simple points is a popular technique with the most successful implementations often using distance or potential functions to direct the removal of voxels, see [11] for a recent example. The discrete Morse-theoretic skeleton defined in this paper is closely related to all of these approaches and provides a theoretical framework for unifying thinning and ridge-finding algorithms.

Discrete Morse theory also provides a framework for partitioning, or shape decomposition, that is closely related to the watershed transform [2], [12]. The connection between watersheds and Morse theory has long been recognized, both concepts are traced back to a paper by Maxwell [13]. Applications of Morse theory to shape and image analysis are covered in the 2008 review article [14].

Connections between Morse theory and persistent homology in the context of data and image analysis were first explored in [15], which stimulated a significant stream of research in scientific visualization of two- and three-dimensional data. For example, in [16] a Morse-theoretic approach to segmentation with simplification informed by persistent homology is implemented using simplicial piece-wise linear approximations to grey-scale images. Earlier (non-persistent) homology computations from image data are found in [17]. In the past five years the use of discrete Morse theory in digital image analysis has received increasing attention [6], [18], [19], [20], [21].

Recent applications of persistent homology to characterise three-dimensional structure of interest to physicists include the distribution of galaxies in the universe [22], [23], and porous or granular materials [24], [25], [26].

1.2 Contributions of this paper

This paper contributes a theoretical understanding, based on discrete Morse theory, for skeletonization and watershed transform operations in digital image analysis. It provides a definition of the Morse skeleton as the union of unstable complexes of critical cells and a proof that the Morse skeleton is homotopic to the initial complex. The stable basin of a minimum is defined, and it is shown that the basins of all local minima partition the image. The relationship between adjacencies of critical points in the Morse skeleton and adjacencies of the basins is clarified by introducing the concept of a bridge. The final theoretical contribution is to elucidate connections between simplification of the discrete Morse vector field and the persistent homology of the sequence of level subcomplexes, in a similar vein to work in [19], [27]. Alongside the theory, efficient serial algorithms are presented for the essential tasks of traversing the discrete vector field and determining paths between critical cells; Python and C++ implementations are available from the authors.

2 Prerequisites

2.1 Digital images as cubical complexes

A three-dimensional greyscale digital image can be modelled by a function $g\colon \mathcal{D} \to \mathbb{R}$, where $\mathcal{D} \subset \mathbb{Z}^3$ is typically a rectangular subset of the discrete lattice:

$$\mathcal{D} = \{(i,j,k) \in \mathbb{Z}^3 \mid 0 \le i \le I, 0 \le j \le J, 0 \le k \le K\}.$$

A point $x \in \mathcal{D}$ is called a *voxel* (or *pixel* in two dimensions).

Following Kovalevsky [28], we model digital images by cubical complexes. In our setting, the voxels $(i,j,k) \in \mathcal{D}$ are the vertices (0-cells) of the complex. Higher-dimensional cells are the unit edges (1-cells) between voxels whose coordinates differ by one in a single axis, unit squares (2-cells), and unit cubes (3-cells). We write $\mathcal{K}$ for the collection of all such p-cells built from voxels in $\mathcal{D}$ with $p = 0, 1, 2, 3$. We often denote the dimension of a cell by a superscript: $\alpha^{(p)}$. A cell $\alpha^{(p)} \in \mathcal{K}$ is a *face* of another cell $\beta^{(q)}$ if $p < q$ and the vertices of α are a subset of the vertices of β. If this is the case then we write $\alpha^{(p)} < \beta^{(q)}$. We can also say that β is a *coface* of α. If $\alpha^{(p)} < \beta^{(q)}$ with $q = p+1$, we call α a *facet* of β and β a *cofacet* of α. A set of such cells $\mathcal{S}$ is a *complex* if for any $\alpha \in \mathcal{S}$, all its faces are also in $\mathcal{S}$. The set $\mathcal{K}$ is a complex, and a complex $\mathcal{X} \subset \mathcal{K}$ is called a *subcomplex* of $\mathcal{K}$.

We initially extend the greyscale function g to the full cubical complex $\mathcal{K}$ by taking the maximal value from the vertices of a cell:

$$g(\beta) := \max\{g(\alpha) \mid \alpha^{(0)} < \beta, \alpha \in \mathcal{D}\}.$$

Given any function defined on a complex $f : \mathcal{K} \to \mathbb{R}$ we can study the subcomplexes derived from the level cuts on f,

$$\mathcal{K}_f(c) := \{\alpha \mid \alpha \le \beta \text{ for } f(\beta) \le c\}.$$

This is called the *level subcomplex* at value c with respect to f. For the greyscale image function g extended to the cubical complex as above, these level subcomplexes are simply $\mathcal{K}_g(c) = \{\beta \mid g(\beta) \le c\}$. The connectivity of the cubical level subcomplex with voxels as vertices is equivalent to using the standard 6-neighbourhood of 3D digital topology: voxel (i,j,k) is connected to a voxel on a face diagonal, say $(i+1, j+1, k)$, if and only if at least one of the voxels $(i+1, j, k)$ and $(i, j+1, k)$ is present. However, as will be seen below, extending the image to a complex in this way allows us to apply

a number of sophisticated theoretical tools, particularly discrete Morse theory.

We often need to study the local structure of a complex via the *star* $\mathrm{St}(\alpha)$ of a vertex $\alpha \in \mathcal{D}$:

$$\mathrm{St}(\alpha) = \{\beta \mid \alpha < \beta\},$$

i.e., the set of all cells in $\mathcal{K}$ that contain α as a face. When α is in the interior of the rectangular region $\mathcal{D}$ its star consists of six unit edges, twelve unit squares and eight unit cubes. The star of a vertex is not a subcomplex of $\mathcal{K}$, so it is sometimes useful to consider its closure, the *closed star*.

Now suppose we have a total order on the vertices of a complex: $\alpha_0 \prec \alpha_1 \prec \cdots \prec \alpha_N$, then the *lower star* $\mathrm{L}(\alpha_i)$ is the subset of cells from $\mathrm{St}(\alpha_i)$ that have vertices that precede α_i in the order. We can then define a sequence of subcomplexes of $\mathcal{K}$ by adding the lower star of each vertex according to the given ordering:

$$\mathcal{K}_0 = \alpha_0, \quad \mathcal{K}_i = \mathcal{K}_{i-1} \cup \mathrm{L}(\alpha_i).$$

This is called the *lower star filtration* of $\mathcal{K}$. For a greyscale image with voxels ordered by their intensity values, the level subcomplexes $\mathcal{K}_g(c)$ for some increasing sequence of c-values form a subsequence of the lower star filtration.

2.2 Homology: topology from cell complexes

Mathematically speaking, two objects A and B have the same topology when there is a continuous map $f\colon A \to B$ with a continuous inverse; this is not true of a shape and its skeleton, which are instead usually related by a homotopy. When a shape and its skeleton, are said to "have the same topology" this typically means they have the same connectivity as measured, for example, by their homology groups. The homology groups are algebraic objects that quantify topological structures of different dimensions present in a cell complex. The ranks of these groups are called the *Betti numbers* and for subsets of three-dimensional space, the Betti numbers have a simple interpretation as the number of components (0-dimensional homology), inequivalent loops (1-dimensional homology), and enclosed voids (2-dimensional homology). The algebraic formulation of homology is extremely powerful: it lets us quantify the connectivity of a cell complex in a routine, algorithmic way, and it also reveals a connection between critical points and gradient flow lines of a Morse function on a manifold and the homology of that manifold. A detailed knowledge of homology theory is not needed for the constructions and results in this paper, but it provides context and background to our perspective on digital images. Further details are available in [29].

Persistent homology is defined for a sequence of subcomplexes i.e. a *filtration*. It provides a tool for measuring the significance of topological features, thereby potentially identifying features that are artefacts arising from image noise. In this paper the filtration is derived from level cuts on the digital image as described in Section 2.1. When we talk about topology-preserving simplification of the skeleton or partition, we mean that topological features with persistence greater than a specified level are guaranteed to remain. We will provide some further details in Section 4, and a comprehensive introduction to the subject is given in [30].

2.3 Discrete Morse Theory

In this section, we briefly summarize the definitions and results we need from discrete Morse theory, a combinatorial analog of Morse theory developed by Forman [7], [31]. We work with a cubical complex $\mathcal{K}$, but note that Forman's theory holds for general CW-complexes.

We start by recalling some concepts from Whitehead's simple homotopy theory [32]. Consider cells $\alpha^{(p-1)} < \beta^{(p)}$ such that α has no other cofacets. Then α is called a *free face* of β and (α, β) is called a *free pair*. The subcomplex $\mathcal{K}'$ obtained from $\mathcal{K}$ by removing a free pair is called an *elementary collapse* of $\mathcal{K}$. We say that $\mathcal{K}$ collapses to a subcomplex $\mathcal{K}'$ if there is a sequence of elementary collapses from $\mathcal{K}$ to $\mathcal{K}'$. The inverse of a collapse is an *expansion*. We say that two complexes are *simple homotopy equivalent* if there is a sequence of collapses and expansions from one complex to the other. An elementary collapse is a strong deformation retraction (see [30] for a definition). It follows therefore that if two complexes are simple homotopy equivalent, they are homotopy equivalent. Note that there is no way to remove a single cell from a complex while preserving homotopy; only simultaneous removal of free pairs achieves this.

A function $f\colon \mathcal{K} \to \mathbb{R}$ is a *discrete Morse function* if for each cell $\alpha^{(p)} \in \mathcal{K}$, the following two conditions hold:

$$\#\{\beta^{(p+1)} > \alpha \mid f(\beta) \le f(\alpha)\} \le 1$$
$$\#\{\gamma^{(p-1)} < \alpha \mid f(\gamma) \ge f(\alpha)\} \le 1.$$

Here, $\#$ denotes the number of elements or cardinality of the set. In other words, α can have at most one cofacet on which f takes a value no larger than on α itself and at most one facet on which the value is no smaller. A cell is *critical* if both cardinalities are 0. Lemma 2.5 of [7] shows that at most one cardinality can be 1 for any given cell. In this way, we can define a pairing of non-critical cells in which $\alpha^{(p)} < \beta^{(p+1)}$ form an ordered pair (α, β) if and only if $f(\alpha) \ge f(\beta)$ holds. This pairing connects with the simple homotopy described above because (α, β) are a free pair for the level subcomplex $\mathcal{K}(f(\beta))$.

More generally, any pairing of cells, V, with the property that each cell is a member of at most one pair is called a *discrete vector field*. If a discrete vector field is derived from a discrete Morse function, f, it is a *gradient vector field* V_f. We often treat V as a function by writing $V(\alpha) = \beta$, $V(\beta) = 0$ when $(\alpha^{(p)}, \beta^{(p+1)}) \in V$, and $V(\gamma) = 0$ when γ is a critical cell (and so unpaired). When $V_f(\alpha) = \beta$, we can imagine an arrow pointing

from α into β; these arrows point in the unique decreasing direction for f.

A flow line in a discrete vector field is called a *V-path*, i.e. a sequence of cells

$$\alpha_0^{(p)}, \beta_0^{(p+1)}, \alpha_1^{(p)}, \beta_1^{(p+1)}, \alpha_2^{(p)}, \ldots, \beta_{r-1}^{(p+1)}, \alpha_r^{(p)}$$

where $(\alpha_i, \beta_i) \in V$, $\beta_i > \alpha_{i+1}$ and $\alpha_i \neq \alpha_{i+1}$ for all $i = 0, \ldots, r-1$. We call a V-path *trivial* if $r = 0$ and *closed* if $\alpha_r = \alpha_0$. Forman has shown that if no non-trivial, closed V-path exists for a given discrete vector field V, then V is indeed the gradient vector field of some discrete Morse function f (Theorem 9.3 of [7]).

It is often simpler to construct and work with gradient vector fields directly. Our algorithm for building a discrete Morse function from a greyscale image does so via a gradient vector field, as described in Sec 2.4.

Discrete Morse theory allows us to capture the topological essence of level subcomplexes derived from a digital image in a significantly compressed form thanks to the following results. A complex $\mathcal{K}$ with a Morse function f can be shown to be homotopy equivalent to a CW-complex with exactly one cell of dimension p for each critical cell of $\mathcal{K}$ with dimension p (Corollary 3.5 of [7]). Even though constructing a full geometric representation of the cell complex $\mathcal{M}$ may be difficult in practice, we can reconstruct its homology by examining the V-paths between critical cells, to construct the *Morse chain complex* $\mathcal{M}$; see [7] for details. There is a natural filtration on the Morse chain complex defined by level subcomplexes of the function f restricted to $\mathcal{M}$. The persistent homology of this filtration captures the changes in topology that occur as each critical value is passed.

2.4 Building a gradient vector field from a greyscale image

In [6] we described an algorithm (*ProcessLowerStar*) for extending a function defined on the vertices of a cell complex $\mathcal{K}$ to a discrete Morse function defined on all cells of $\mathcal{K}$. The algorithm takes the lower star of each vertex and grows it via simple-homotopy expansions where possible, adding a critical cell only when necessary. The simple-homotopy expansions define pairs in a gradient vector field compatible with the initial function on the vertices. The only requirement is that the star of each vertex has a total order defined on it (by a globally consistent simulated perturbation such as the linear ramp described in [6] if necessary). As proven in [6] for cubical complexes, the algorithm is correct in the sense that it identifies exactly the type of critical cells needed to characterise the changes in topology of the level subcomplexes when adding each lower star. An example of the action of *ProcessLowerStar* is shown in Fig. 1. Note that the lower stars of any two vertices are disjoint, so this algorithm is immediately able to execute in parallel.

The Morse chain complex is derived by following V-paths in the gradient vector field. A naive algorithm uses breadth-first-search, but this can have exponential worst-case behaviour because a single cell can belong to many V-paths. For example, in cubical complexes $(0,1)$ V-paths can merge, $(2,3)$ V-paths can branch and $(1,2)$ V-paths can both merge and branch. Some approaches to managing this are discussed in Sec. 5.

3 Stable and Unstable Sets and the Morse Skeleton

We now explore how to use the gradient vector field to define discrete analogues of the stable and unstable manifolds of a critical point. The Morse skeleton is then defined using unstable sets of critical cells. We finish the section looking at the stable sets of local minima as a partition of the digital image and investigate adjacencies between the basins of this partition. These definitions and results are the new theoretical work presented in this paper. Note that our stable and unstable sets are defined geometrically by the V-paths of the discrete gradient vector field, rather than the algebraic formulation obtained using Forman's definition of gradient flow [7], [33].

Throughout this section, let $\mathcal{K}$ be a finite, n-dimensional regular CW-complex (such as a cubical complex) with discrete Morse function $f\colon \mathcal{K} \to \mathbb{R}$ and associated gradient vector field V (as constructed by the *ProcessLowerStar* algorithm, for example). We assume without loss of generality that the values of f are unique on the cells of $\mathcal{K}$.

3.1 Stable and unstable sets

Let α be any p-cell in $\mathcal{K}$. We use V-paths that end or start at α to define its stable or unstable sets respectively. Formally, the *stable set* of α, denoted $\mathcal{S}_{\mathcal{K}}(\alpha)$, is the smallest set of p-cells in $\mathcal{K}$ such that $\alpha \in \mathcal{S}_{\mathcal{K}}(\alpha)$ and

$$V(\delta^{(p)}) = \gamma^{(p+1)} > \beta^{(p)} \in \mathcal{S}_{\mathcal{K}}(\alpha) \text{ implies } \delta \in \mathcal{S}_{\mathcal{K}}(\alpha).$$

Similarly, the *unstable set* of α, denoted $\mathcal{U}_{\mathcal{K}}(\alpha)$, is the smallest set of p-cells such that $\alpha \in \mathcal{U}_{\mathcal{K}}(\alpha)$ and

$$V^{-1}(\delta^{(p)}) = \gamma^{(p-1)} < \beta^{(p)} \in \mathcal{U}_{\mathcal{K}}(\alpha) \text{ implies } \delta \in \mathcal{U}_{\mathcal{K}}(\alpha).$$

We will omit the subscript $\mathcal{K}$ when the complex referred to is obvious from the context.

Note that the stable and unstable sets of $\alpha^{(p)}$ contain cells only of dimension p. It is useful therefore, to also define an *unstable complex* of α as the closure of its unstable set.

$$\mathcal{W}_{\mathcal{K}}(\alpha) := \{\gamma \mid \gamma \leq \delta \in \mathcal{U}_{\mathcal{K}}(\alpha)\}.$$

We do not make a dual definition of stable complex as it would most naturally involve the cofaces of elements in $\mathcal{S}(\alpha)$ (rather than faces) and therefore not be a subcomplex of $\mathcal{K}$. However, in Sec. 3.3 we will show that an n-dimensional subcomplex can be defined from the 0-cells that comprise the stable set of a local minimum.

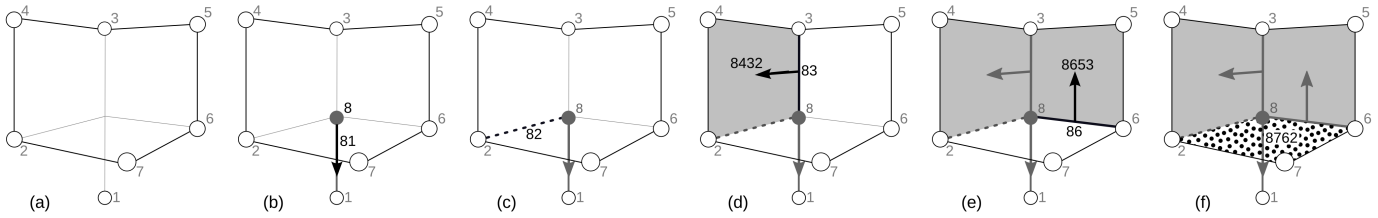

Fig. 1. In this example we build the lower star of vertex 8 in a cubical complex by incrementally adding simple-homotopy pairs when possible. (a) The vertices 1–7 and six dark edges are present in the subcomplex $\mathcal{K}(7)$; the lower star of vertex 8 contains four edges and three square faces. (b) Vertex 8 is paired with the edge 81. (c) Edge 82 is added as a critical cell. (d) Edge 83 is paired with face 8432. (e) Edge 86 is paired with face 8653. (f) Face 8762 is added as a critical cell. This configuration of voxels represents a compound critical point for the (6,26) digital topology; the addition of two critical cells accurately reflects the change in the topology of the level subcomplexes in going from $\mathcal{K}(7)$ to $\mathcal{K}(8)$.

Certain subcomplexes $\mathcal{H} \subset \mathcal{K}$ have the property of being closed under the flow induced by the vector field. We formalize this by saying a subcomplex $\mathcal{H}$ of $\mathcal{K}$ is *hermetic* if for each cell $\alpha \in \mathcal{H}$ either $V(\alpha) = 0$ or $V(\alpha) \in \mathcal{H}$. The most important examples of hermetic subcomplexes are the level subcomplexes $\mathcal{K}_f(c)$, a consequence of the following lemma.

Lemma 1: Let c be a real value and $\beta^{(p)}$ a cell in $\mathcal{K}_f(c)$. If β is critical, then $f(\beta) \leq c$. If β is not critical and $V(\beta) = \gamma^{(p+1)} \neq 0$, then $f(\gamma) \leq c$.

Proof: First, for any cell $\beta^{(p)} \in \mathcal{K}_f(c)$ we know that either $f(\beta) \leq c$ or there is some coface, $\tau > \beta$ with $f(\tau) \leq c$. From Lemma 3.2 of [7], we know that in the latter case, there exists $\tau'^{(p+1)} < \tau$ with $f(\tau') \leq f(\tau)$ and $\beta < \tau'$.

Assume now that β is critical with $f(\beta) > c$. By the above, there is a $\tau^{(p+1)} > \beta$ with $f(\tau) \leq c$. But since β is a critical cell, we have $f(\beta) < f(\tau) \leq c$, a contradiction.

Now let $V(\beta) = \gamma$. Then $f(\gamma) > c$ would imply $f(\beta) \geq f(\gamma) > c$, so as above, there must then be a $\tau^{(p+1)} > \beta$ with $f(\tau) \leq c < f(\beta)$. But then by the definition of a discrete Morse function, it follows that $\tau = \gamma$, in contradiction to the assumption that $f(\gamma) > c$. □

The next result shows that hermetic subcomplexes preserve the stable and unstable sets of their cells.

Lemma 2: For α a cell in a hermetic subcomplex $\mathcal{H}$ of $\mathcal{K}$, we have

$$\mathcal{S}_{\mathcal{H}}(\alpha) = \mathcal{S}_{\mathcal{K}}(\alpha) \cap \mathcal{H}$$

and

$$\mathcal{U}_{\mathcal{H}}(\alpha) = \mathcal{U}_{\mathcal{K}}(\alpha).$$

Proof: The inclusions $\mathcal{S}_{\mathcal{H}}(\alpha) \subseteq \mathcal{S}_{\mathcal{K}}(\alpha) \cap \mathcal{H}$ and $\mathcal{U}_{\mathcal{H}}(\alpha) \subseteq \mathcal{U}_{\mathcal{K}}(\alpha)$ are obvious.

Consider now a cell $\delta^{(p)} \in \mathcal{S}_{\mathcal{K}}(\alpha) \cap \mathcal{H}$ and assume $\delta \notin \mathcal{S}_{\mathcal{H}}(\alpha)$ with $f(\delta)$ minimal for all such cells. By definition, there must exist cells $\beta^{(p)}$ and $\gamma^{(p+1)}$ such that $V(\delta) = \gamma > \beta \in \mathcal{S}_{\mathcal{K}}(\alpha)$. We know that $\gamma \in \mathcal{H}$ since $\mathcal{H}$ is hermetic, which implies $\beta \in \mathcal{H}$ since $\mathcal{H}$ is a subcomplex, and then by the assumed minimality of $f(\delta)$, also $\beta \in \mathcal{S}_{\mathcal{H}}(\alpha)$. But this would imply $\delta \in \mathcal{S}_{\mathcal{H}}(\alpha)$, a contradiction.

Finally, the inclusion $\mathcal{U}_{\mathcal{K}}(\alpha) \subset \mathcal{U}_{\mathcal{H}}(\alpha)$ when $\alpha \in \mathcal{H}$ follows directly from the definitions. Both hermetic subcomplexes and unstable sets include their 'downstream' cells. □

Putting Lemmas 1 and 2 together we have the following results about the stable and unstable sets.

Lemma 3: For $c \in \mathbb{R}$ and α a cell in $\mathcal{K}_f(c)$, we have

$$\mathcal{S}_{\mathcal{K}(c)}(\alpha) = \mathcal{S}_{\mathcal{K}}(\alpha) \cap \mathcal{K}(c)$$

and

$$\mathcal{U}_{\mathcal{K}(c)}(\alpha) = \mathcal{U}_{\mathcal{K}}(\alpha).$$

3.2 The Morse Skeleton

For a complex $\mathcal{K}$ with a discrete Morse function $f\colon \mathcal{K} \to \mathbb{R}$, the *Morse skeleton* (or axis) is now defined to be the subcomplex built from unstable complexes of critical cells,

$$\mathcal{A}_{\mathcal{K}} := \bigcup_{\alpha \text{ critical}} \mathcal{W}_{\mathcal{K}}(\alpha).$$

When $\mathcal{K}$ and f are derived from a greyscale digital image on a rectangular domain, the skeleton is effectively the entire complex; it is when we examine the level subcomplexes that the Morse skeleton gives us an interesting summary of structure.

First, we have that

$$\mathcal{A}_{\mathcal{K}(c)} = \bigcup_{\substack{\alpha \text{ critical} \\ f(\alpha) \leq c}} \mathcal{W}_{\mathcal{K}(c)}(\alpha) = \bigcup_{\substack{\alpha \text{ critical} \\ f(\alpha) \leq c}} \mathcal{W}_{\mathcal{K}}(\alpha).$$

The first equality follows from Lemma 1 and the second from Lemma 3. This means we need only determine the unstable subcomplexes of critical points with respect to the complex $\mathcal{K}$, there is no need to recompute for the level subcomplexes.

Next we show that a complex is homotopy equivalent to its Morse skeleton by a collapse from one to the other that follows the gradient vector field. First some terminology: If (α, β) is a free pair of cells in $\mathcal{K}$ such that $V(\alpha) = \beta$, we call the elementary collapse defined by removing α and β from $\mathcal{K}$ *regular*. We call a series of regular elementary collapses a *regular collapse* and write

$\mathcal{K} \searrow \mathcal{K}'$ if there is a regular collapse from the complex $\mathcal{K}$ to the complex $\mathcal{K}'$. Note that the result of a regular collapse is a hermetic subcomplex.

Theorem 4: There is a regular collapse from a complex $\mathcal{K}$ to its Morse skeleton $\mathcal{A}$.

Proof: Let $\mathcal{X}$ be a subcomplex of $\mathcal{K}$ such that

$$\mathcal{K} \searrow \mathcal{X} \supseteq \mathcal{A},$$

and $\mathcal{X}$ is the smallest possible subcomplex with this property. We claim that $\mathcal{X} = \mathcal{A}$.

Assume otherwise and consider a cell $\beta^{(p)} \in \mathcal{X} \backslash \mathcal{A}$ such that p is maximal and among all such cells, $f(\beta)$ is also maximal. We distinguish three cases: (a) β is critical, (b) $V^{-1}(\beta) = \alpha \neq 0$, and (c) $V(\beta) = \gamma \neq 0$.

Case (a), β is critical. Then $\beta \in \mathcal{A}$ by definition, so we get a contradiction immediately.

Case (b), $V^{-1}(\beta) = \alpha^{(p-1)}$. Let $\beta'^{(p)} > \alpha$ be any other cofacet of α. By the definition of the skeleton, $\beta' \in \mathcal{A}$ would imply $\beta \in \mathcal{A}$, violating the assumption. So no cofacets of α are in $\mathcal{A}$. From the definition of a discrete Morse function we see that $f(\beta') > f(\beta)$, so $\beta' \notin \mathcal{X}$ because we chose $f(\beta)$ to be maximal. Furthermore, because $\mathcal{X}$ is a subcomplex, we must have $\alpha \in \mathcal{X}$. On the other hand, $\alpha \notin \mathcal{A}$ because α is neither critical, nor is $V^{(-1)}(\alpha)$ defined, nor does it have any cofaces that are in $\mathcal{A}$. Thus (α, β) is a free pair with respect to $\mathcal{X}$ with $V(\alpha) = \beta$ that avoids $\mathcal{A}$, but this contradicts the minimality of $\mathcal{X}$.

Case (c), $V(\beta) = \gamma^{(p+1)}$. By assumption, we must have either $\gamma \in \mathcal{A}$ or $\gamma \notin \mathcal{X}$. But the former would imply $\beta \in \mathcal{A}$, which has been excluded. On the other hand, since $\mathcal{X}$ was derived from $\mathcal{K}$ by a series of regular elementary collapses, the one that removed γ from the complex would have to have removed β, as well. □

Applied to digital image analysis, the regular collapse from a level subcomplex to its Morse skeleton takes on the role of a thinning process. The use of discrete Morse theory means that the collapse (or thinning) does not have to be performed explicitly and there are no ambiguities in its construction. Moreover, the choice of threshold for the level subcomplex can be changed without having to recompute anything other than to find which critical cells have values below the new threshold. Theorem 4 guarantees that the skeleton is homotopy equivalent to the original complex; in particular the Morse skeleton has the same number of pieces as the initial object. Geometric properties of the skeleton (such as being centred with respect to the object boundary) depend on the initial function; in general the unstable complexes that form the skeleton follow paths of steepest-descent.

One natural application is to consider the Euclidean distance transform of a binary image, so that the greyscale function g encodes the distance of each voxel from the boundary between foreground (black) and background (white) voxels. By making the distance negative in the foreground and positive in the background, we see that the Morse skeleton at threshold $c = 0$ includes $0, 1$ and 2-dimensional unstable complexes and thus defines a type of medial surface of the foreground.

3.3 Basins and bridges

Next we investigate the role of stable sets in partitioning the vertex set of the complex $\mathcal{K}$ with discrete Morse function f into regions similar to those derived by watershed algorithms.

Lemma 5: Each vertex of $\mathcal{K}$ is in the stable set of exactly one local minimum (i.e. a critical vertex).

Proof: First note that a vertex cannot belong to the stable set of more than one minimum because $(0, 1)$ V-paths can merge but not branch. Now let $\mathcal{E}$ be the set of vertices that are in the stable set of exactly one critical vertex. Assume there is a vertex $\alpha \notin \mathcal{E}$, which we may choose to minimize $f(\alpha)$ among all such vertices. Since α can not be critical, we must have $V(\alpha) = \beta^{(1)} \neq 0$. Let γ be the other vertex of β, then from the definition of a discrete Morse function we have $f(\alpha) \geq f(\beta) > f(\gamma)$. Now if $\gamma \in \mathcal{E}$, we also have $\alpha \in \mathcal{E}$ by the definition of a stable set. If $\gamma \notin \mathcal{E}$ then $f(\alpha)$ was not minimal. Either case contradicts our assumption. □

We have already noted that in general the stable set of a critical cell cannot be extended to a subcomplex in a manner dual to the unstable complex. But for critical 0-cells, i.e. minima, we can define the *basin* $\mathcal{B}_{\mathcal{K}}(\alpha)$ as the maximal subcomplex of $\mathcal{K}$ that has a regular collapse onto α. Uniqueness follows from:

Lemma 6: For any hermetic subcomplex $\mathcal{H}$ of $\mathcal{K}$, there is a unique maximal hermetic subcomplex $\widetilde{\mathcal{H}}$ such that $\widetilde{\mathcal{H}} \searrow \mathcal{H}$.

Proof: Because $\mathcal{H}$ is hermetic and regular expansions add cells in V-pairs, any complex $\widetilde{\mathcal{H}}$ that collapses to $\mathcal{H}$ must also be hermetic.

Now let $\widetilde{\mathcal{H}}_1 \neq \widetilde{\mathcal{H}}_2$ both be maximal subcomplexes of $\mathcal{K}$ that collapse regularly onto $\mathcal{H}$. Without restriction, assume that $\widetilde{\mathcal{H}}_2$ contains cells not in $\widetilde{\mathcal{H}}_1$. Let $\alpha^{(p)}$ be such a cell with first p and then $f(\alpha)$ minimal. Since $\widetilde{\mathcal{H}}_2 \searrow \mathcal{H}$, we must have $0 \neq \beta := V(\alpha) \in \widetilde{\mathcal{H}}_2 \backslash \widetilde{\mathcal{H}}_1$. Then $f(\alpha') < f(\alpha)$ for any other facet α' of β, implying $\alpha' \in \widetilde{\mathcal{H}}_1$. But then $\widetilde{\mathcal{H}}_1 \cup \{\alpha, \beta\} \searrow \widetilde{\mathcal{H}}_1$ is a regular collapse, a contradiction to the maximality of $\widetilde{\mathcal{H}}_1$. □

It also follows easily from the definition that basins are simply-connected, pairwise disjoint and free of critical cells other than their defining minima. Basins also contain the stable sets of their minima: $\mathcal{S}_{\mathcal{K}}(\alpha) \subset \mathcal{B}_{\mathcal{K}}(\alpha)$. And similar to Lemma 2 for stable sets we have the following for basins.

Lemma 7: For α a critical 0-cell in a hermetic subcomplex $\mathcal{H}$ of $\mathcal{K}$,

$$\mathcal{B}_{\mathcal{H}}(\alpha) = \mathcal{B}_{\mathcal{K}}(\alpha) \cap \mathcal{H}.$$

Although each vertex in $\mathcal{K}$ must belong to the basin of some minimum, the same is not true of the k-cells,

$k \geq 1$. We call a 1-cell that is not contained in any basin a *bridge*. If $\beta^{(1)}$ is a bridge, then we must have $V^{-1}(\beta) = 0$, so either β is critical (a 1-saddle), or $V(\beta) \neq 0$. Then the following result implies that every 1-cell from $\mathcal{K}$ is contained either in the basin of a minimum or in the stable set of one or more 1-saddles.

Lemma 8: Every bridge is in the stable set of at least one critical bridge.

Proof: First we show that if a bridge β is such that $V(\beta) = \gamma \neq 0$ then γ has another facet $\beta' < \gamma$, $\beta' \neq \beta$ such that β' is also a bridge. We will obtain a contradiction by assuming that all facets $\sigma < \gamma$, $\sigma \neq \beta$ belong to some basin. Note that if a 1-cell belongs to a basin $\mathcal{B}(\alpha)$ then both its vertices must belong to the same basin, since basins are sub-complexes. This implies that the facets σ all belong to the same basin. But then (β, γ) form a free pair with respect to $\mathcal{B}(\alpha)$, implying that $\mathcal{B}(\alpha)$ was not maximal.

Now let $\mathcal{V}$ be the set of all bridges that belong to stable sets of 1-saddle bridges, and pick a bridge $\beta \notin \mathcal{V}$ with $f(\beta)$ minimal. In particular, β cannot itself be a saddle and so $V(\beta) = \gamma \neq 0$. At least one other facet, $\beta' < \gamma$ must be a bridge by the above argument. Clearly, $f(\beta') < f(\beta)$, so $\beta' \in \mathcal{V}$ and then also $\beta \in \mathcal{V}$, contradicting our assumption. □

Now observe that when β is a bridge between the basins of two minima $\mathcal{B}(\alpha)$ and $\mathcal{B}(\gamma)$, the V-paths that start at each vertex of β form a path between α and γ that we call the *canonical path* defined by β. Note that $\alpha = \gamma$ is possible and then the canonical path will be homotopic to a circle (we only have homotopy here, not homeomorphism because (0,1) V-paths can merge). If β is a critical 1-cell, this path is the unstable complex of β and forms part of the Morse skeleton of $\mathcal{K}$, so we have established that when two minima are in the unstable complex of a 1-saddle, their basins are adjacent at that 1-saddle. Notice that the converse is not true: two basins may be adjacent when there is no critical 1-cell bridging their minima. The canonical path of a non-critical bridge captures this additional connectivity between minima.

A further characterisation of these canonical paths is given by the following.

Theorem 9: For any bridge, β, there are critical bridges $\beta_1, \ldots, \beta_k$ such that the canonical path through β can be deformed, keeping its end-point minima fixed, to the concatenation of canonical paths through the β_i.

Proof: We first show that any path through the 1-skeleton of $\mathcal{K}$ is homotopic with fixed end points to one which does not cross any non-saddle bridges. Let φ be a counterexample which minimizes the largest f-value of any non-saddle bridge crossed. Let β be the bridge in question. We can then find cells γ and β' as in the proof of Lemma 8 above, and push the path across γ to obtain a new path φ' which now crosses the bridge β' with $f(\beta') < f(\beta)$. In fact all non-saddle bridges crossed by φ' must have f-values smaller than $f(\beta)$. This is true for those already crossed by φ by the assumed uniqueness of f-values on cells of $\mathcal{K}$ and for those new to φ' by the same argument that showed $f(\beta') < f(\beta)$. But if φ was a counterexample to the statement, so is φ' and then φ could not have minimised the largest f-value for a non-saddle bridge crossed.

Finally, since basins are simply-connected, any path within a particular basin is homotopic to any other path with the same end points. So the original canonical path from α to α' through β, which is deformed to some path from α to α' that passes only through critical bridges, is broken into segments interior to a single basin. Each of these segments is homotopic to a path that is built from canonical paths of the relevant critical bridges. □

4 Simplification

The initial watershed partition of a real image (even one with small-amplitude noise) usually creates more basins than necessary to effectively summarize the structures present; naturally this issue also occurs when we construct the Morse basins of a real image. There are many strategies presented in the literature for merging watershed basins that are close in some sense [34], [35], [36], [37], [38]. As we saw in the previous section, adjacent basins may or may not be connected via a 1-saddle, and identifying this is crucial to merging basins in a way that preserves important topological information (as measured by persistent homology). Our Morse theoretic approach uses built-in techniques for cancelling critical points and the natural connection with homology means we can simplify the topology of the level sets in a controlled way. As well as merging basins, Morse theory incorporates higher-order topological changes such as filling in extraneous loops and voids using the same techniques.

Discrete Morse theory provides a very simple method for cancelling two critical cells, $\alpha^{(p)}$ and $\beta^{(p+1)}$, by reversing the pairings along a V-path between them. Forman [7] shows that when there is exactly one V-path from the boundary of β to α, the modified vector field is a gradient vector field (if there is more than one V-path, the reversal introduces loops). The question now is which critical cells should be cancelled, and this is where we require the information from persistent homology. Morse cancellation of critical points and the connection with persistent homology has been studied in detail for functions on triangulations of 2-manifolds [15], [19], [39]. For three-dimensional complexes there are known obstructions to simplifying some configurations of persistence pairs as discussed in [19], [27], [40], [41].

We emphasise that there are two cell-complex representations of our data. The lower-level description is the cubical complex $\mathcal{K}$ with a discrete Morse function f, and the higher-level description is the Morse chain complex $\mathcal{M}$. The latter is the basis for the persistent homology calculations, but it is the cubical complex where we require a simplification operation to reduce the number of critical cells by adjusting the discrete Morse function

and thereby merging basins, for example. The Morse chain complex is an algebraic description of the object. There is a boundary operator $\partial_{\mathcal{M}}$ that records incidences between critical cells that are connected via V-paths. The V-path incidences are counted with orientations and with respect to a coefficient group; we use $\mathbb{Z}_2$, but other choices are possible, see [7] for details. This algebraic structure means the Morse chain complex ignores cell incidences that cancel out, for example, when there are two V-paths between a pair of critical cells.

4.1 Persistent homology

Persistent homology captures the changes in topology of a sequence of nested subcomplexes, i.e. a filtration. The filtration is determined by the order of insertion of cells from the cell complex $\mathcal{K}$. We write $\alpha \prec \beta$ if α is inserted before β. To ensure each step in the filtration is a subcomplex we must have faces inserted before their cofaces: $\alpha \prec \beta$ if $\alpha < \beta$. The persistence of features is measured via a time-of-insertion function $t : \mathcal{K} \to \mathbb{R}$ so that $t(\alpha) \leq t(\beta)$ whenever $\alpha \prec \beta$. For example, when the filtration is determined by level cuts of a function f we might have

$$t(\alpha) = \min\{c \mid \alpha \in \mathcal{K}_f(c)\}.$$

Now examine what happens when a single p-cell α is added to a subcomplex $\mathcal{K}_i$. The boundary of α must already be in $\mathcal{K}_i$, so α either creates a new p-cycle or fills in a $(p-1)$-cycle. The idea behind persistent homology is that these events are paired and define the lifetime of a topological feature. Specifically, suppose that z is a p-cycle created by adding the cell α at time $t(\alpha)$ and that z' is a p-cycle that is homologous to z with $z' = \partial\beta$, where β is a $(p+1)$-cell added at time $t(\beta) \geq t(\alpha)$. The *persistence* of z, and of its homology class $[z]$, is equal to $t(\beta) - t(\alpha)$. The cell α is called the *creator* and β the *destroyer* of $[z]$. Some cycles never become boundaries and these are given infinite persistence.

4.2 Close pair cancellation

Ideally, we would like to adjust the discrete Morse function on the cubical complex so as to remove all low-persistence features up to a given threshold, τ. A naive approach to simplification might be to first determine all these persistence pairs and then cancel out the ones with persistence below the threshold τ. There are at least two potential problems with this idea. Firstly, determining all creator-destroyer pairs is relatively expensive, and secondly, destroyers are not necessarily incident to creators, so it is not obvious how such a cancellation would be performed in practice.

An alternative approach not requiring the persistence pairing, might be to identify pairs of critical cells $(\alpha^{(p)}, \beta^{(p+1)})$ in the Morse chain complex, $\mathcal{M}$, where $\alpha \in \partial_{\mathcal{M}}(\beta)$ and $t(\beta) - t(\alpha) \leq \tau$ (t is the time-of-insertion function for the filtration on $\mathcal{M}$). In this homological setting, the cancellation of adjacent cells corresponds to adding the boundary of β to the boundary of each remaining coboundary cell of α and then removing α and β from $\mathcal{M}$.

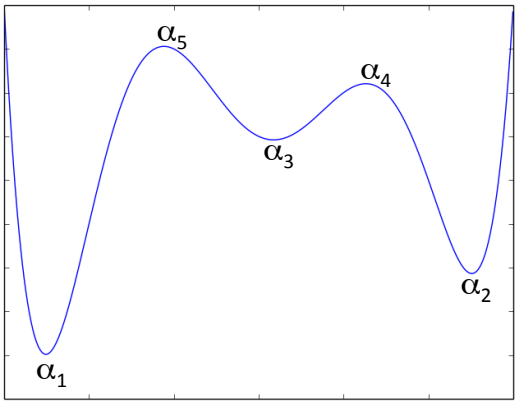

Fig. 2. A 1D function with three minima $\alpha_1, \alpha_2, \alpha_3$ and two maxima α_4, α_5. The Morse complex of this function is analysed in the text. Persistent homology pairs α_4 with α_3 and α_5 with α_2.

Unfortunately, cancellation of pairs as described above does not guarantee that features with persistence higher than the chosen threshold are always preserved. To illustrate this, consider the following 1-complex derived from the Morse complex of the function shown in Fig. 2:

$$(\alpha_1) - \alpha_5 - (\alpha_3) - \alpha_4 - (\alpha_2).$$

Here, $\alpha_1, \ldots, \alpha_5$ represent cells in the complex and enclosing parentheses denote 0-cells, whereas enclosing dashes denote 1-cells. The filtration order is given by the subscripts, i.e. $t(\alpha_i) = i$, so the persistent homology is as follows. The cells α_1, α_2 and α_3 create 0-cycles (i.e., connected components), the cell α_4 destroys the cycle created by α_3, and the cell α_5 destroys the cycle created by α_2. A correct persistence cancellation with threshold $\tau = 2$ starts by cancelling out α_4 against α_3, leading to this complex:

$$(\alpha_1) - \alpha_5 - (\alpha_2).$$

The persistence of the cycle created by α_2 is $5 - 2 = 3$ and thus above the threshold, so this is already the final result of the process.

By contrast, when cancelling pairs arbitrarily just as long as the difference between their t-values is no larger than the threshold, we might first cancel out α_5 against α_3, leading to

$$(\alpha_1) - \alpha_4 - (\alpha_2)$$

and then α_4 against α_2, yielding

$$(\alpha_1).$$

This is clearly not the correct result, as we have now eliminated the cycle of persistence 3 generated by α_2.

This brings us to the concept of a *close pair* of critical cells. Suppose $\mathcal{M}$ is a filtration with time-of-insertion

function t, then (α, β) form a close pair when $\alpha^{(p)}$ is the last boundary cell of $\beta^{(p+1)}$ to be added in the filtration, and β is the first coboundary cell of α to be inserted. In the example above, (α_3, α_4) form a close pair and (α_3, α_5) do not. The following two results hold:

Theorem 10: If $\mathcal{M}$ contains a cycle with persistence τ, then it contains a close pair $\alpha \in \partial_{\mathcal{M}}(\beta)$ with $t(\beta) - t(\alpha) \leq \tau$.

Proof: Let us begin by assuming that the cycle in question is created by a cell γ and destroyed by a cell ϵ. If ϵ is not a coboundary cell of γ, there must be another coboundary cell, say ϵ', such that $\gamma \prec \epsilon' \prec \epsilon$ and thus $t(\epsilon') - t(\gamma) \leq t(\epsilon) - t(\gamma) = \tau$. Indeed, every coboundary cell of γ must be inserted after γ, and if no coboundary cells of γ are part of the complex by the time ϵ is inserted, there is no way for ϵ to destroy the cycle created by γ. This proves that $\mathcal{M}$ must contain a pair with t-value difference no larger than τ.

Now consider a cell β and $\alpha \in \partial_{\mathcal{M}}(\beta)$ such that $t(\beta) - t(\alpha) \leq \tau$ and the number of cells inserted between α and β is minimal. Such a pair exists by the above, and because $\mathcal{M}$ is finite. Assume there is a coboundary cell β' of α such that $\beta' \prec \beta$. But then $t(\beta') \geq t(\beta)$ and thus $t(\beta') - t(\alpha) \leq t(\beta) - t(\alpha) \leq \tau$, and moreover, the number of cells inserted between α and β' would be smaller than the number of cells inserted between α and β, contradicting our assumption. In the same manner, one shows that there can be no boundary cell α' of β with $\alpha \prec \alpha'$, and consequently α and β must already form a close pair. □

Lemma 11: If $\alpha \in \partial_{\mathcal{M}}(\beta)$ form a close pair in $\mathcal{M}$, then α must create a homological cycle which is destroyed by β.

Proof: Indeed, since α is the last boundary cell of β inserted, and no coboundary cell of α can be present at the time α is added, α must create a cycle that is homologous to the boundary of β in the cubical complex. Until β is added, there can be no chain containing α in its boundary, so that cycle must still exist, and thus is indeed destroyed by β. □

Now if $\mathcal{M}$ contains any cycle with persistence no larger than a threshold τ, Theorem 10 tells us that there exists a close pair with t-value difference also no larger than τ, and Lemma 11 establishes that such a pair must indeed form a creator-destroyer pair for a homological cycle. Thus by performing a cancellation on the close pair, we obtain a smaller complex with the same τ-persistent homology. If the original cycle was not eliminated in this round, the induction principle shows that it eventually will be. In conclusion, by repeatedly cancelling out close pairs up to a given persistence threshold, we obtain a smallest simplified complex with the same τ-persistent homology as the original $\mathcal{M}$.

As a special case, setting $p = \infty$ will successively eliminate all features of $\mathcal{M}$ with finite persistence and in effect generate a complete sequence of homological creator-destroyer pairs. In this sense, simplification of a complex via close pair cancellation in the purely homological setting can be viewed as a generalisation of classical persistence computation. We emphasise that this simplification happens on the *algebraic Morse chain complex*, not on the gradient vector field of the original cubical complex.

Note that the above results show that close pairs are persistence pairs, but clearly not all persistence pairs are close. Results similar to Theorem 10 and Lemma 11 were first discussed for 2-dimensional simplicial complexes in [15], and for general cell complexes in [42]. More recently, an analogous approach has been used in persistent homology algorithms developed in [43].

In terms of the gradient vector field on the initial (cubical) complex $\mathcal{K}$, a *cancellable close pair* $(\alpha^{(p)}, \beta^{(p+1)})$ is a critical cell pair such that there is a single V-path from the boundary of β to α, with the condition that any other critical cell $\gamma^{(p)}$ that is reachable via a V-path from the boundary of β has $f(\gamma) \leq f(\alpha)$ and similarly for any critical cell $\delta^{(p+1)}$ for which there is a V-path terminating at α has $f(\delta) \geq f(\beta)$. A cancellable close pair in the cubical complex must be a close pair in the Morse chain complex, but the converse does not always hold, because of the possible *algebraic* cancellation of paths between critical cells, e.g. three V-paths adding mod-2 to a single incidence in the chain complex.

Reversing the V-path between a cancellable close pair (α, β) with persistence $t(\beta) - t(\alpha) < \tau$ results in a new vector field with the same h-persistent homology as the original, for $h \geq \tau$. However, the recursive reversal of V-paths between cancellable close pairs with $t(\beta) - t(\alpha) \leq \tau$ may not remove *all* homological features with persistence less than τ.

5 Implementation

We assume that a discrete gradient vector field, V, is already defined on a given cell complex, $\mathcal{K}$. For a digital image, with a function f defined on its voxels, this could be the output from the ProcessLowerStar algorithm given in [6]. To build a Morse skeleton and partition for f and V requires us to compute the unstable complexes and stable sets of critical cells, and also to simplify the vector field by cancelling close pairs of critical cells. Each of these operations requires us to trace V-paths through the discrete vector field, particularly between critical cells. We present Python implementations for the essential routines required to do this efficiently below. The full code, including some examples of its usage, is available from the authors. For practical use, we have also implemented several versions of these algorithms in C++, including one that supports distributed memory parallelism.

The algorithms shown below are implemented using Python generators. This means that the routine performing the algorithm and the code from which it is called

execute concurrently. The `yield` keyword is used to pass results successively to the caller, which consumes them one by one as they are produced, typically within a *for* loop. We chose this form for its resemblance to the *output* statement often found in pseudo-code.

We use the following parameters to represent the gradient vector field and underlying cell complex:

V A function representing a discrete vector field:

$$\mathrm{V}(x) = \begin{cases} x & \text{if } x \text{ is critical,} \\ y & \text{if } x \text{ is paired with a cofacet } y, \\ \emptyset & \text{otherwise.} \end{cases}$$

coV The reverse vector field in the same form.
I A function that lists a cell's facets.
coI A function that lists a cell's cofacets.

5.1 Vector field traversal and the Morse chain complex

For any pair of critical cells $\gamma^{(i+1)}$, $\delta^{(i)}$ the Morse chain complex counts the number of V-paths that start in a facet of γ and end at δ. In practice, we typically do not require the exact counts. Instead, it is often sufficient to know whether there is an even or odd number of paths (in order to compute homology over $\mathbb{Z}_2$) and whether there is no path, exactly one, or several (in order to determine whether Morse cancellation can be applied). But this still leaves common graph traversal algorithms such as breadth-first search (BFS) unsuitable for our purposes without modifications.

The method proposed in [6] was to use a modified BFS that may traverse the same edge multiple times, which is very easy to implement and in practice often very fast, but unfortunately suffers from an exponential worst case behaviour. This is because $(1, 2)$ V-paths in 3D complexes can branch and re-merge, and each such branch/re-merge event leads to (at least) a two-fold increase in the total number of paths. A better behaved algorithm was proposed in [20], based on the following idea: in a traditional BFS, a vertex is added to the queue as soon as it is first reached via an incoming edge. If instead one waits until it is reached for the last time, information such as the total number of paths leading to this vertex can be accumulated before being passed on further.

Depending on the value of a parameter (*coordinated*), our algorithm *traverseFlow* performs either an unmodified BFS or the modified version that delays following the outgoing edges of a node in the graph until all its incoming edges have been seen. We refer to the latter version as a *coordinated breadth-first search*.

With *coordinated* false, we have a straightforward BFS. A queue, here implemented as a list, holds cells to be processed. The set *seen* contains all cells that have been added to the queue at some point, and effectively ensures that each cell is only processed once. Both the queue and the set are initialised with the start cell s. The outer loop pulls the first cell a from the queue, then the inner loop looks at each cofacet b of a in turn, first determining the facet c it is paired with, if any. If c is defined and different from a, the cell triple is returned, and c is added to the queue unless it had been added earlier or is a critical cell, the latter being indicated by $b = c$.

If *coordinated* is true, a standard BFS first determines how often in total each cell appears as c. This number is stored in the map *nrIn*. A new map k is then used, during a second BFS pass, to count cell appearances again. Cells are prevented from being added to the queue as long as their *nrIn* and k values still differ.

```
def traverseFlow(s, V, I, coordinated = False):
    """
    Produces all cell triples of the form (a, b, c),
    where a is in the unstable set of the cell s, b
    is an element of I(a), and c = V(b) != a. A cell
    other than s only appears as a after appearing
    as c at least once. If coordinated is True, its
    appearances as a strictly follow those as c.
    """

    if coordinated:
        nrIn = {}
        k = {}
        for (a, b, c) in traverseFlow(s, V, I):
            nrIn[c] = nrIn.get(c, 0) + 1

    queue = [s]
    seen = {s}

    while len(queue) > 0:
        a = queue.pop(0)
        for b in I(a):
            c = V(b)
            if c != a and c is not None:
                yield (a, b, c)
                if c not in seen and c != b:
                    if coordinated:
                        k[c] = k.get(c, 0) + 1
                        if k[c] != nrIn[c]:
                            continue
                    queue.append(c)
                    seen.add(c)
```

The algorithm *traverseFlow* is used in a variety of circumstances. Note that the third item c in the sequence of cell triples *traverseFlow* produces ranges over the unstable set of s (with the exception of s itself), plus the critical cells with dimension one less than s that can be reached by V-paths starting at facets of s (since we have $c = V(b) = b$ when b is critical). Results from Section 3 show that the Morse skeleton of a level subcomplex is simply a union of unstable complexes for critical cells belonging to the subcomplex, so this is all the computation required. We can also traverse the flow in the opposite direction by passing in coV instead of V and coI instead of I, and so compute stable sets and thus the basins of local minima.

To obtain the homological Morse chain complex we need to compute boundary operators as follows. For each cell c that can be reached by V-paths from a facet of a given critical cell s, we count the total number of such paths, which is obviously the sum of path counts for all cells from which c can be reached in a single step. Updating these counts in the order prescribed by

a coordinated traversal ensures that incoming counts are always correct. An exponential number of V-paths could potentially lead to excessively large counts, so we only store three discrete values (1, 2 for an even number, and 3 for all larger odd numbers) as indicated in our implementation *morseBoundary* below. To account for the possibility of the same critical cell being reached independently from multiple directions, we keep track of these in the set *boundary* and only output them together with their path counts after the traversal is finished.

```
def morseBoundary(s, V, I):
    """
    Produces all pairs (c, k) where c is a critical
    cell that can be reached from a facet of s via
    V-paths, with k=1 if there is a single such path
    and k>1 indicating multiple paths, with k=2 for
    an even and k=3 for an odd number.
    """

    counts = { s: 1 }
    boundary = set()

    for (a, b, c) in traverseFlow(s, V, I, True):
        n = counts[a] + counts.get(c, 0)
        counts[c] = n if n <= 3 else n % 2 + 2
        if b == c:
            boundary.add(c)

    for c in boundary:
        yield (c, counts[c])
```

5.2 Simplification

To implement the simplification of a gradient vector field via Morse cancellation of close pairs, as discussed in Section 4, we first determine all cancellable close pairs of given dimensions, which is straightforward given the list of critical cells and the *morseBoundary* algorithm. We can introduce further conditions at this point, such as an upper limit on the persistence of a close pair.

In general, the algorithm *connections* given below, constructs the union of all V-paths between a given pair (s, t) of critical cells by first marking all cells in the stable set of t via a backwards traversal, and then restricting the results of a forward traversal accordingly.

```
def connections(s, t, V, coV, I, coI):
    """
    Produces all triples of the form (a, b, c),
    where a is in the unstable set of the cell s, b
    is an element of I(a) which is also in the
    stable set of t, and c = V(b) != a. A cell other
    than s only appears as a after appearing as c at
    least once.
    """

    active = { t }
    for (a, b, c) in traverseFlow(t, coV, coI):
        active.add(c)

    for (a, b, c) in traverseFlow(s, V, I):
        if b in active:
            yield (a, b, c)
```

For a cancellable close pair (s, t), the algorithm *connections* yields triples (a, b, c) along the single V-path from s to t. The gradient vector field is simplified by pairing the cell b with its cofacet a, so reversing the gradient flow along that V-path. Because of the restriction to *cancellable* close pairs, all these path reversals can be performed concurrently without influencing each other. This is due to two properties of cancellable close pairs. First, given two cancellable close pairs (s_1, t_1) and (s_2, t_2) the V-path from s_1 to t_1 cannot intersect that from s_2 to t_2. Second, close pairs are persistence pairs, and each critical k-cell can either create a k-cycle or destroy a $(k-1)$-cycle, but not both. In a more general setting, path reversals between critical cells in a 3D discrete gradient vector field can cause considerable changes in the adjacencies between nearby critical cells; this has been explored in detail in [27].

The new gradient vector field obtained after a simplification pass as described above may contain new cancellable close pairs. We therefore repeat the process until no further cancellable close pairs that fulfill any additional restrictions are found.

6 Examples

We present some examples to illustrate our methods. Table 3 shows execution times and memory consumption for a number of three-dimensional data sets. The first eight, from [44], are included for comparison with [20], [45]. The remaining four were derived from binary images [46] by applying a signed Euclidean distance transform, as discussed in Section 3. All computations were performed on a single core of an Intel Core 2 Duo T8100 laptop CPU running at 2.1GHz. We used an optimised C++ implementation of the algorithms described in Section 5.

The long execution times for the simplification of the gradient vector field are due to the fact that our current implementation recomputes the full Morse chain complex after each pass. This could be improved by modifying the algorithm to make incremental updates. We also observe that computing the chain complex took significantly longer after simplification for some images, whereas for others it became faster. This is likely due to competition between a reduction in the number of critical cells and an associated increase in size of the stable and unstable sets associated with the critical cells. It is clear from these results that the computational cost of these methods is low enough that they can easily be applied to multi-gigavoxel images given a parallel implementation and the multi-core computing hardware available in 2013.

Figure 4 shows our methods applied to a 2d dataset. The input data was obtained by extracting a 510x510 slice from the Mt Gambier limestone image at [46] and applying a signed Euclidean distance transform on the resulting binary 2d image. We show the resulting features after simplifying the gradient vector field with a threshold of 1 pixel unit. This level of simplification is sufficient to capture both large- and small-scale details in a natural way. We observe that some gradient paths tend to follow axis directions leading to partition boundaries

Name	Dimensions	# Critical Cells		Memory (MB)				Time (sec)			
				vector	simpli-	chain complex		vector	simpli-	chain complex	
		before	after	field	fication	before	after	field	fication	before	after
Silicium	98x34x34	1109	813	1.4	1.8	1.3	1.2	0.28	0.61	0.25	0.26
Fuel	64x64x64	667	135	2.5	2.9	1.5	2.0	0.67	1.36	0.13	0.30
Neghip	64x64x64	5709	489	2.5	3.2	1.8	2.0	0.68	2.00	0.26	0.36
Hydrogen	128x128x128	24257	17	16.9	20.7	10.8	10.1	5.58	23.83	5.33	0.80
Engine	256x256x128	1035127	105573	66.0	25.4	11.5	44.3	24.60	709.07	32.65	76.02
Aneurysm	256x256x256	75485	44657	131.6	145.9	71.7	79.3	44.41	2099.63	68.76	442.58
Bonsai	256x256x256	344277	130033	131.6	193.9	95.4	85.5	45.56	6682.86	162.99	950.81
Foot	256x256x256	1658617	1210291	131.6	425.5	194.2	163.8	46.75	2718.83	107.57	415.07
Limestone	510x510x510	1708665	19133	1036.8	1343.6	649.6	521.1	345.73	1991.57	197.15	68.99
Sandstone	510x510x510	2320429	61041	1036.8	1437.3	688.1	523.6	358.62	1643.09	150.67	69.69
Sand Pack	510x510x510	2942231	41611	1036.8	1525.4	725.3	521.9	352.94	1586.24	126.46	57.25
Sphere Pack	510x510x510	984625	3755	1036.8	1191.8	584.7	519.4	323.91	1192.34	48.00	40.16

Fig. 3. Execution times and memory consumption for some 3d images. Numbers of critical cells before and after simplification with persistence threshold 1 are listed. Memory consumption and execution times are given for the gradient vector field computation, simplification, and Morse chain complex computation before and after simplification.

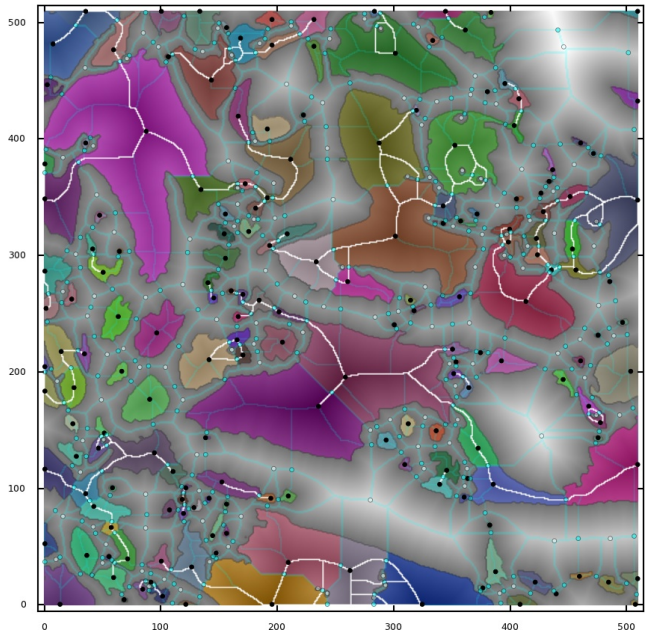

Fig. 4. Morse analysis of a 2d Euclidean distance image (Mt. Gambier limestone). The solid phase (background) is shown in gray, and the void space (foreground) is divided into colored pores formed by its intersections with the basins. Critical cells are indicated by round dots, black for minima, cyan for saddles, and white for maxima. The Morse skeleton is shown in white, and other V-paths between critical cells in cyan.

that do not match their intuitive best location. This is an inherent problem with discretised versions of Morse theory, and a technique to correct the geometric aberration using randomised variations in the gradient pairings is presented in [47].

In Figure 5, we present a visualisation (using [48]) of a silica sphere pack, also from [46], together with the Morse skeleton of its pore-space, again after simplifying with a threshold of 1 voxel unit. Here, we extracted a 254x254x254 section from the center of the original image after applying the distance transform. The Morse skeleton contains 1-dimensional and 2-dimensional elements; the latter formed by unstable complexes of critical 2-cells. The presence of persistent critical 2-cells — each resulting in a single 2-dimensional element in the skeleton — suggests that this pore geometry would be poorly represented by a 1D line skeleton. Extracting networks from binary digital images of porous materials is commonly performed as input for pore-network models that are used to simulate the flow of two or three immiscible fluids through the pore space [49], [50]. The vertices of the pore network carry weights such as the volume of the associated pore, and the edges of the network have weights that determine how easy it is for fluid to move from one pore to the next. Without information about the critical 2-cells as seen in Fig. 5, such models must currently either replace each surface patch with a single line [49], [50], or abandon homotopy equivalence by puncturing the patch and using two or more paths through the pore [51]. A better solution would be to extend the pore-network model to include the 2-dimensional patches and an improved physical model of fluid flow in these critical 2-cell pores. While surface skeletons can be derived from digital images [52], such skeletons are not naturally combinatorial and their surface may branch in a complex manner rendering them difficult to use for modelling applications. A combinatorial skeleton that does contains 2-cells, was defined in [53] for the complement of objects built from a union of balls. While potentially applicable to 3D digital images of porous materials, the method in [53] is too computationally intensive for routine 3D application. Our Morse skeleton is both geometrically realised as a digital object, has topologically accurate combinatorial structure, and is dual in a rigorous sense to the Morse partition. The persistence and geometric extent of the 2-dimensional elements associated with critical 2-cells in the pore space will be used to provide better physical modelling of fluid displacements via improved pore-network models.

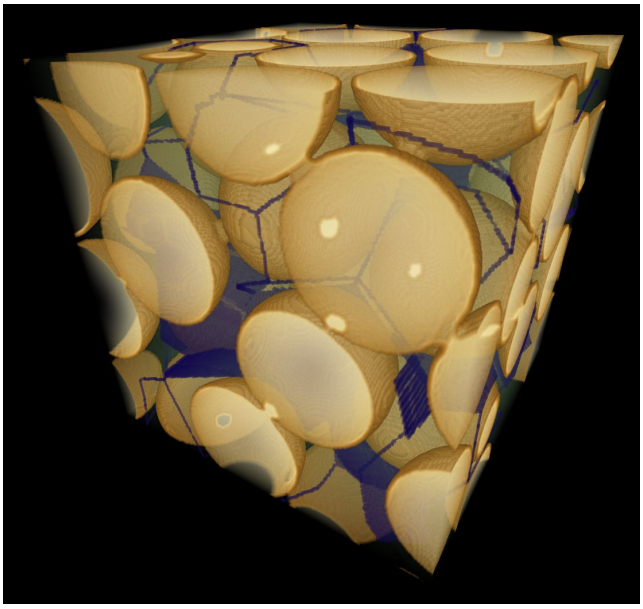

Fig. 5. A sphere pack with the Morse skeleton of its pore space shown in blue. The blue patches are unstable complexes of persistent critical 2-cells.

7 Conclusion

This paper has shown how the mathematical formalism of discrete Morse theory permits rigorous definitions of two popular operations in digital image analysis, namely skeletonization and partitioning. Our techniques allow us to compute these simultaneously with a single algorithm and provide topologically combinatorial and compatible structures while maintaining a geometric representation at the voxel level. We address the effects of noise and discretisation by the iterative removal of *cancellable close pairs* of critical points. These pairs are a subset of those found by a persistent homology analysis, but we do not need to perform the persistence computations to identify them. Our definition of skeleton includes both curve and surface elements that are related in a combinatorial complex. The surface elements provide important information about the structure of the pore-space of granular and porous materials that we intend to incorporate into future models of fluid displacements in these geometries.

Acknowledgments

This research was supported in part by ARC Discovery Project DP110102888.

References

[1] K. Siddiqi and S. Pizer, *Medial Representations*. Berlin: Springer, 2008.
[2] S. Beucher and F. Meyer, "The morphological approach to segmentation: the watershed transformation," *Optical Engineering* Marcell-Dekker Inc., New York, vol. 34, p. 433, 1992.
[3] L.J. Latecki and R. Lakämper, "Convexity rule for shape decomposition based on discrete contour evolution," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 441–454, Mar. 1999.
[4] J.-M. Lien, J. Keyser, and N.M. Amato, "Simultaneous shape decomposition and skeletonization," in *Proceedings of the 2006 ACM symposium on Solid and physical modeling*, 2006, p. 219–228.
[5] Y. Matsumoto, *An Introduction to Morse Theory*. Providence, RI: AMS Bookstore, 2002.
[6] V. Robins, P. J. Wood, and A. P. Sheppard, "Theory and algorithms for constructing discrete Morse complexes from grayscale digital images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, p. 1646–1658, 2011.
[7] R. Forman, "Morse theory for cell complexes," *Advances in Mathematics*, vol. 134, p. 90–145, 1998.
[8] N. C. Cornea, D. Silver, and P. Min, "Curve-skeleton properties, applications and algorithms," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 3, p. 530–548, 2007.
[9] H. Blum, "A transformation for extracting descriptors of shape," in *Models for the Perception of Speech and Visual Forms*. MIT Press, 1967, pp. 362–380.
[10] S. Biasotti, D. Attali, J.-D. Boissonnat, et al., "Skeletal structures," in *Shape Analysis and Structuring*, ser. Mathematics and Visualization, L. Floriani and M. Spagnuolo, Eds. Springer Berlin, 2008, pp. 145–183.
[11] C. Arcelli, G. Sanniti di Baja, and L. Serino, "Distance-driven skeletonization in voxel images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, p. 709–720, 2011.
[12] S. Beucher and C. Lantuejoul, "Use of watersheds in contour detection," *International Workshop on Image Processing: Real-time Edge and Motion Detection/Estimation*, Rennes, France. Sep. 1979.
[13] J. C. Maxwell, "On hills and dales," *The Philosophical Magazine*, vol. 40, no. 269, p. 421–427, 1870.
[14] S. Biasotti, L. De Floriani, B. Falcidieno, et al., "Describing shapes by geometrical-topological properties of real functions," *ACM Comput. Surv.*, vol. 40, no. 4, p. 1–87, 2008.
[15] H. Edelsbrunner, J. Harer, and A. Zomorodian, "Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds," *Discrete and Computational Geometry*, vol. 30, p. 98–107, 2003.
[16] H. Edelsbrunner and J. Harer, "The persistent Morse complex segmentation of a 3-manifold," in *Modelling the Physiological Human*, ser. Lecture Notes in Computer Science. Berlin: Springer-Verlag, 2009, no. 5903, pp. 36–50.
[17] M. Allili, K. Mischaikow, and A. Tannenbaum, "Cubical homology and the topological classification of 2D and 3D imagery," vol. 2. IEEE, 2001, pp. 173–176.
[18] A. Gyulassy, P. T. Bremer, B. Hamann, and V. Pascucci, "A practical approach to Morse-Smale complex computation: Scalability and generality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, p. 1619–1626, 2008.
[19] U. Bauer, C. Lange, and M. Wardetzky, "Optimal topological simplification of discrete functions on surfaces," *Discrete Comput. Geom.*, vol. 47, no. 2, p. 347–377, Mar. 2012.
[20] S. Guenther, J. Reininghaus, H. Wagner, and I. Hotz, "Efficient computation of 3D Morse-Smale complexes and persistent homology using discrete Morse theory," *The Visual Computer*, vol. 28, no. 10, p. 959–969, 2012.
[21] K. Mischaikow and V. Nanda, "Morse theory for filtrations and efficient computation of persistent homology," *Discrete Comput. Geom.*, vol. 50, no. 2, pp. 330–353, Sep. 2013.
[22] R. v. d. Weygaert, G. Vegter, H. Edelsbrunner, et al., "Alpha, Betti and the megaparsec universe: On the topology of the cosmic web," in *Transactions on Computational Science XIV*, ser. Lecture Notes in Computer Science, M. L. Gavrilova, C. J. K. Tan, and M. A. Mostafavi, Eds. Springer Berlin, Jan. 2011, no. 6970, pp. 60–101.
[23] T. Sousbie, C. Pichon, and H. Kawahara, "The persistent cosmic web and its filamentary structure: II. illustrations," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 1, pp. 384–403, Jun. 2011.
[24] A. Gyulassy, M. Duchaineau, V. Natarajan, V. Pascucci, E. Bringa, A. Higginbotham, and B. Hamann, "Topologically clean distance fields," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, p. 1432–1439, 2007.
[25] L. Kondic, A. Goullet, C. S. O'Hern, M. Kramar, K. Mischaikow, and R. P. Behringer, "Topology of force networks in compressed granular media," *EPL*, vol. 97, no. 5, p. 54001, Mar. 2012.
[26] O. Delgado-Friedrichs, V. Robins, and A.P. Sheppard, "Morse theory and persistent homology for topological analysis of 3d images of complex materials," *IEEE International Conference on Image Processing*, Paris, France, Oct. 2014.

[27] D. Guenther, J. Reininghaus, H.-P. Seidel, and T. Weinkauf, "Notes on the simplification of the Morse-Smale complex," in *Topological Methods in Data Analysis and Visualization III*, ser. Mathematics and Visualization, P.-T. Bremer, I. Hotz, V. Pascucci, and R. Peikert, Eds. Springer, Berlin. Jan. 2014, pp. 135–150.
[28] V. A. Kovalevsky, "Finite topology as applied to image analysis," *Comput. Vision Graph. Image Process.*, vol. 46, no. 2, p. 141–161, 1989.
[29] V. Robins, "Algebraic Topology," in *Mathematical Tools for Physicists*. 2nd Edition, Berlin: Wiley, 2014, pp. 211–238.
[30] H. Edelsbrunner and J. Harer, *Computational Topology: An introduction*. Providence, Rhode Island: American Mathematical Society, 2010.
[31] R. Forman, "A user's guide to discrete Morse theory," *Seminaire Lotharingien de Combinatoire*, vol. 48, 2002.
[32] J. H. C. Whitehead, "Simplicial spaces, nuclei, and m-groups," *Proc. London Math. Soc.*, vol. 45, no. 2, pp. 243–327, 1939.
[33] T. Lewiner, "Geometric discrete Morse complexes," Ph.D. dissertation, Department of Mathematics, PUC-Rio, 2005. [Online]. Available: http://www.matmidia.mat.puc-rio.br/tomlew/pdfs/phd_thesis_puc.pdf
[34] F. Meyer and S. Beucher, "Morphological segmentation," Journal of Visual Communication and Image Representation, vol. 1, no. 1, pp. 2146, 1990.
[35] S. Beucher, "Watershed, hierarchical segmentation and waterfall algorithm," in *Mathematical Morphology and Its Applications to Image Processing*, ser. Computational Imaging and Vision, J. Serra and P. Soille, Eds. Springer Netherlands, Jan. 1994, no. 2, pp. 69–76.
[36] M. Couprie and G. Bertrand, "Topological gray-scale watershed transformation," *Optical Science, Engineering and Instrumentation '97*, pp. 136–146, Oct. 1997.
[37] K. Haris, S. Efstratiadis, and N. Maglaveras, "Watershed-based image segmentation with fast region merging," in *International Conference on Image Processing*, 1998, pp. 338–342 vol.3.
[38] A. Bleau and L. Leon, "Watershed-based segmentation and region merging," *Computer Vision and Image Understanding*, vol. 77, no. 3, pp. 317–370, Mar. 2000.
[39] T. Lewiner, "Critical sets in discrete Morse theories: Relating Forman and piecewise-linear approaches," *Computer Aided Geometric Design*, vol. 30, pp. 609–621, 2013.
[40] N. Shivashankar and V. Natarajan, "Parallel computation of 3D Morse-Smale complexes," *Computer Graphics Forum*, vol. 31, pp. 965–974, Jun. 2012.
[41] A. Gyulassy, V. Natarajan, V. Pascucci, and B. Hamann, "Efficient computation of Morse-Smale complexes for three-dimensional scalar functions," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, p. 1440–1447, 2007.
[42] H. King, K. Knudson, and N. Mramor, "Genering discrete Morse functions from point data," *Experimental Mathematics*, vol. 14, pp. 435–444, 2005.
[43] U. Bauer, M. Kerber, and J. Reininghaus, "Clear and compress: Computing persistent homology in chunks," in *Topological Methods in Data Analysis and Visualization III*, ser. Mathematics and Visualization, P.-T. Bremer, I. Hotz, V. Pascucci, and R. Peikert, Eds. Springer, Jan. 2014, pp. 103–117.
[44] M. Meissner, "Volvis: Voxel data repository," 2001. [Online]. Available: http://volvis.org
[45] H. Wagner, C. Chen, and E. Vucini, "Efficient computation of persistent homology for cubical data," in *Topological Methods in Data Analysis and Visualization II*, ser. Mathematics and Visualization, R. Peikert, H. Hauser, H. Carr, and R. Fuchs, Eds. Springer Berlin, Jan. 2012, pp. 91–106.
[46] A.P. Sheppard, "The network generation comparison forum," 2005. [Online]. Available: http://xct.anu.edu.au/network_comparison.
[47] A. Gyulassy, P. Bremer, and V. Pascucci, "Computing Morse-Smale complexes with accurate geometry," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2014–2022, Dec. 2012.
[48] D. Whitehouse, "Voluminous - the web based volume renderer," Dec. 2013. [Online]. Available: http://voluminous.nci.org.au
[49] M. Prodanovic, W. Lindquist, and R. Seright, "3D image-based characterization of fluid displacement in a Berea core," *Advances in Water Resources*, vol. 30, no. 2, pp. 214–226, Feb. 2007.
[50] S. Bakke and P. Oren, "3-d pore-scale modelling of sandstones and flow simulations in the pore networks," *SPE Journal*, vol. 2, no. 2, pp. 136–149, 1997.
[51] D. Silin and T. Patzek, "Pore space morphology analysis using maximal inscribed spheres," *Physica A: Statistical Mechanics and its Applications*, vol. 371, no. 2, pp. 336–360, 2006.
[52] G. Borgefors, I. Nyström, and G. S. d. Baja, "Skeletonizing volume objects part II: from surface to curve skeleton," in *Advances in Pattern Recognition*, ser. Lecture Notes in Computer Science, A. Amin, D. Dori, P. Pudil, and H. Freeman, Eds. Springer Berlin, 1998, pp. 220–229.
[53] R. Glantz and M. Hilpert, "Tight dual models of pore spaces," *Advances in Water Resources*, vol. 31, no. 5, pp. 787–806, 2008.

Olaf Delgado-Friedrichs Olaf Delgado-Friedrichs received his M.Sc., PhD and Dr.habil. degrees in Mathematics at the University of Bielefeld in 1990, 1994 and 2002. He has held research positions at the University of Bielefeld, the Arizona State University, the University of Tübingen, and the Australian National University. Since 2012, he works as a contractor based in Melbourne, Australia. His research interests are in discrete and computational mathematics and its applications.

Vanessa Robins Vanessa Robins completed her B.Sc.(Hons) degree in Mathematics at the Australian National University in 1994 and received her Ph.D. degree in Applied Mathematics from the University of Colorado, Boulder in 2000. She has been a research fellow in the Department of Applied Mathematics at the Australian National University since 2000. Her research interests are in computational topology and geometry, and the enumeration and classification of periodic nets. She works on mathematical foundations, algorithms and scientific applications in both these fields.

Adrian Sheppard Adrian Sheppard completed his B.Sc (hons) degree in Mathematical Physics at The University of Adelaide in 1992 and his PhD degree in Physics at the Australian National University in 1996. He has held research positions at the Université Libre de Bruxelles, the University Of New South Wales and the Australian National University, where he has worked since 2000. His research interests include the characterisation of shape, fluid flow in porous media, parallel algorithms for image processing and tomographic imaging.