ERC Advanced Grant 2017
Research Proposal [Part B2]

Alpha Shape Theory Extended

Alpha

Section a: State-of-the-art and objectives

The theory we propose to fully develop started some 25 years ago with the inception of alpha shapes, which later lead to the wrap algorithm for surface reconstruction and to persistent homology for topological data analysis. We show that combining these concepts and extending them further opens a range of new opportunities. Focusing on low-dimensional applications, we state concrete objectives whose solutions will be milestones in this development.

A. ALPHA SHAPES

The motivating question for the inception of alpha shapes [19, 23, 25] was the generalization of the convex hull of a finite set of points to something like its shape by Jarvis [33]. The answer depends on the scale the observer prefers, so α was introduced as a real parameter that controls the amount of detail the shape reveals, but we will write r for this radius and use alpha merely as a name.

We now give a formal introduction to the concept in three steps, writing $\mathbb{R}^d$ for the d-dimensional Euclidean space and $X \subseteq \mathbb{R}^d$ for a finite set of points in this space. For each $x \in X$, we collect the points $a \in \mathbb{R}^d$ for which x is the closest point in X to form the *Voronoi domain*: $\mathrm{dom}(x) = \{a \in \mathbb{R}^d \mid \|a - x\| \leq \|a - y\|, \forall y \in X\}$. It is the intersection of finitely many closed half-spaces and therefore a convex polyhedron. The *Voronoi tessellation* of X is the collection of Voronoi domains, one for each point $x \in X$; see the left panel in Figure 1. The domains cover the entire $\mathbb{R}^d$, and their common intersections are convex polyhedra of dimension $d - 1$ or less. Assuming the points in X are in general position, the tessellation is *primitive*, by which we mean that every non-empty common intersection of i Voronoi domains is a convex polyhedron of dimension $d + 1 - i$. In particular, this means that every $d + 2$ or more domains have an empty common intersection.

Considering a non-empty common intersection of Voronoi domains, and letting $Y \subseteq X$ contain the points whose domains share this intersection, we call its convex hull, $\mathrm{conv}\, Y$, a *Delaunay cell*. By construction, there exists a sphere such that the points of Y lie on and all points of $X \setminus Y$ lie outside the sphere. To express both of these properties, we call this an *empty circumscribed sphere* of Y. Assuming general position, all Delaunay cells are simplices of dimension from 0 to d. For this reason, the collection

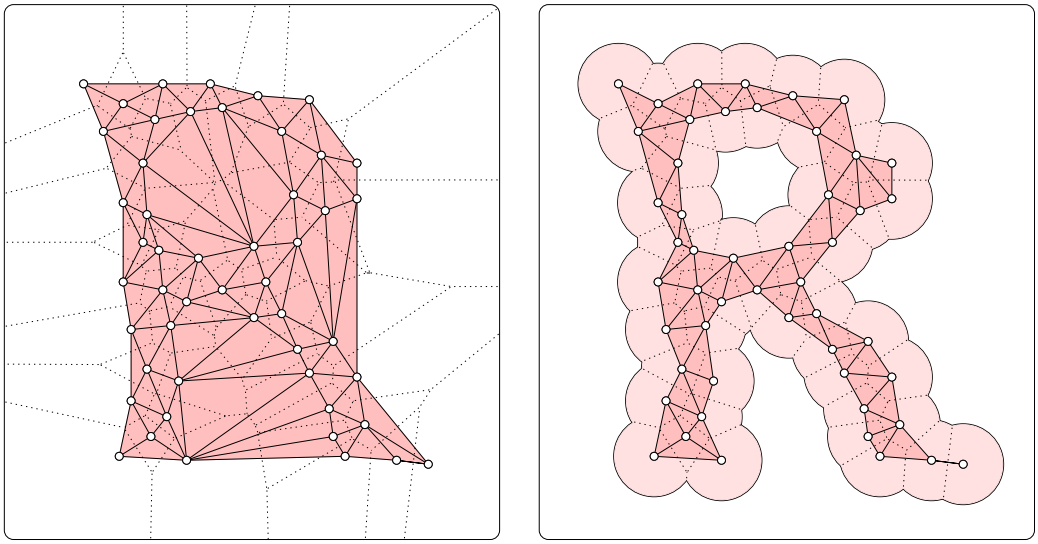

Figure 1: *Left:* a Voronoi tessellation in the plane (dotted) and its dual Delaunay mosaic (solid and shaded) superimposed. Only the Voronoi edges inside the rectangular window and their dual Delaunay edges are shown. *Right:* the Voronoi decomposition of the union of balls centered at the data points and the dual alpha complex superimposed.

of Delaunay cells is often referred to as the *Delaunay triangulation* of X. Notwithstanding, we will adopt the term from stochastic geometry and refer to this collection as the *Delaunay mosaic* [38]. It will however be convenient to assume general position so that the Delaunay mosaic is a simplicial complex in $\mathbb{R}^d$.

The union of cells in the Delaunay mosaic is the convex hull of X, which is topologically trivial. Interesting structures arise when we clip each Voronoi domain to within a neighborhood of the generating point. To do this in a uniform fashion, we specify a radius $r \geq 0$, write $B_r(x)$ for the closed ball with radius r centered at $x \in X$, and intersect the Voronoi domain of x with this ball; see the right panel in Figure 1. The dual *alpha complex* consists of all Delaunay cells, $\operatorname{conv} Y$, for which the corresponding clipped Voronoi domains have a non-empty common intersection, denoted $\operatorname{Alpha}_r(X)$. Assuming general position, $\operatorname{Alpha}_r(X)$ is isomorphic to the *nerve* of the clipped Voronoi domains, which is the system of subcollections of domains with non-empty intersections. By the Nerve Theorem of algebraic topology [7, 35], the nerve has the same homotopy type as the union of clipped domains. It follows that the alpha complex for radius r has the same homotopy type as the union of balls of radius r, and is therefore a topologically faithful proxy of the latter space. The *alpha shape* for radius r is the underlying space of the alpha complex, which formally is the subset of $\mathbb{R}^d$ covered by the cells in $\operatorname{Alpha}_r(X)$ together with the topology inherited from $\mathbb{R}^d$.

B. WRAP COMPLEXES IN TERMS OF DISCRETE MORSE THEORY

A discrete counterpart of the classic Morse theory for generic smooth functions [36] was introduced in 1998 by Robin Forman [30]. We describe a slightly generalized version of this discrete theory [31] that is a better fit for our geometric constructs. We begin with a general formulation and will return to the geometric setting afterwards.

Let K be a simplicial complex. A *generalized discrete vector field* is a partition of K into intervals, in which an *interval* is determined by simplices $P, R \in K$, with $P \subseteq R$, and consists of all $Q \in K$ with $P \subseteq Q \subseteq R$. We write $[P, R]$ for the interval and call P its *lower* and R its *upper bound*. If $P = R$ then $[P, R]$ consists of a single simplex and we refer to it as a *singular interval*. The partition into intervals is *acyclic* if there exists a function $f\colon K \to \mathbb{R}$ such that for every pair of simplices, $P \subseteq Q$,

we have $f(P) \leq f(Q)$ and $f(P) = f(Q)$ iff P and Q belong to a common interval. Such an f is called a *generalized discrete Morse function*, and the partition of K is its *generalized discrete gradient*. The *critical simplices* of f are the ones contained in singular intervals. This terminology is motivated by the incremental construction of K, one interval at a time in the ordering of their values. Adding a critical simplex changes the homotopy type, while adding the simplices in a non-singular interval preserves the homotopy type. It is common to refer to the reverse of the latter operation as a *collapse*: it removes the simplices in a non-singular interval while making sure that the complex remains a complex. To get a feeling why these notions might be useful to reason about the simplicial complex, let $K_r = f^{-1}(-\infty, r]$ and $K_s = f^{-1}(-\infty, s]$ be *sublevel sets* of the function for $r \leq s$, and note that $K_r \subseteq K_s$ are both subcomplexes of K. One of the theorems in [30] asserts that if $(r, s]$ does not contain the value of any critical simplex, then we can get K_r in a sequence of collapses from K_s, which we denote as $K_s \searrow K_r$. A collapse is a special deformation retraction and therefore preserves the homotopy type. Hence $K_s \searrow K_r$ implies that the two complexes have the same homotopy type.

It will be useful to have a more constructive characterization of when a generalized discrete vector field is acyclic. Consider the directed graph whose nodes are the intervals in the partition, and there is an arc from a first node, $[P, R]$, to a second node, $[P', R']$, if there are simplices $P \subseteq Q \subseteq R$ and $P' \subseteq Q' \subseteq R'$ with $Q \subseteq Q'$. Then the partition into intervals is acyclic iff the thus constructed graph has no directed cycle. The left panel of Figure 2 illustrates that not all partitions into intervals are acyclic.

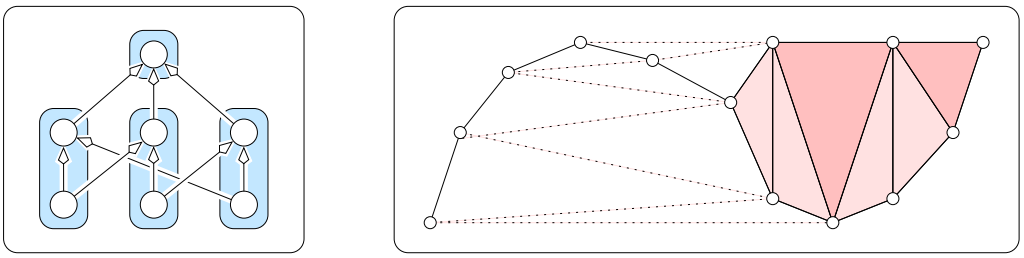

Figure 2: *Left:* the Hasse diagram of a triangle (top), its three edges (middle), and its three vertices (bottom). The partition pairs each edge with one of its vertices. The directed cycle in the graph of intervals is proof that the partition is not acyclic. *Right:* The Delaunay mosaic of a few points in the plane, whose critical triangles are dark shaded. We get the wrap complex for $r = \infty$ by collapsing the white triangles together with the dotted edges.

To return to the geometric setting, assume $X \subseteq \mathbb{R}^d$ is in general position and consider the Delaunay mosaic, $K = \mathrm{Del}(X)$. Let $f\colon K \to \mathbb{R}$ be the *radius function* defined by mapping $Q \in K$ to the radius of the smallest empty circumscribed sphere. It can be seen that f is a generalized discrete Morse function [3], and we get the alpha complexes as sublevel sets: $\mathrm{Alpha}_r(X) = f^{-1}[0, r]$. To define the wrap complex, we recall the directed graph constructed from the Hasse diagram, which for the radius function is necessarily acyclic. For each critical simplex Q of f, we write $\downarrow Q$ for the *lower set*, which is the union of the intervals that lie on directed paths ending at Q. The *wrap complex* for radius r is the union of the lower sets of all critical simplices with $f(Q) \leq r$, denoted as $\mathrm{Wrap}_r(X)$. It is immediate that $\mathrm{Wrap}_r(X) \subseteq \mathrm{Alpha}_r(X)$. The right panel of Figure 2 illustrates that the wrap complex can be obtained from the alpha complex for the same radius by a sequence of collapses. In other words, $\mathrm{Alpha}_r(X) \searrow \mathrm{Wrap}_r(X)$, which implies that the two complexes also have the same homotopy type.

The above introduction of the wrap complex downplays its difference to the alpha complex, on which its primary use for surface reconstruction is based; see [18] for specific and [17] for general background. Indeed, we may think of the wrap algorithm as an adaptive version of the alpha algorithm: whether or not a simplex belongs to $\mathrm{Wrap}_r(X)$ does not only depend on the radius but also on the environment in the Delaunay mosaic, and the relevant environment is usually but not necessarily local. We will see later that this difference offers significant advantages in the reconstruction of shapes and of holes in shapes.

C. PERSISTENT HOMOLOGY IN A NUTSHELL

Call a function $f\colon K \to \mathbb{R}$ on a simplicial complex *monotonic* if $P \subseteq Q$ implies $f(P) \leq f(Q)$, and note that every sublevel set, $f^{-1}(-\infty, r]$, is a subcomplex of K. For example, generalized discrete Morse functions are monotonic. Indexing the sublevel sets in sequence, we get a *filtration*, $\emptyset = K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n = K$. For each K_i and each dimension $p \geq 0$, the *p-th homology group* of K_i, denoted $\mathsf{H}_p(K_i)$, represents the p-dimensional cycles of K_i, and here we refer to standard literature in algebraic topology [32, 37] because we do not have the space to formally introduce this group. In informal discussions, we sometimes prefer the more intuitive language of *holes*, tacitly assuming equivalent formal statements in terms of cycles and homology groups. Other than on the complex, $\mathsf{H}_p(K_i)$ also depends on the coefficients used to produce formal sums of simplices. We assume field coefficients throughout, for which the group has the particularly simple structure of a vector space. To further simplify, we take the direct sum of the homology groups over all dimensions and write $H_i = \bigoplus_{p\geq 0} \mathsf{H}_p(K_i)$. Since $K_i \subseteq K_{i+1}$, there is an induced linear map between the homology groups, $h_i^{i+1}\colon H_i \to H_{i+1}$. The sequence of homology groups, $H_0 \to H_1 \to \ldots \to H_n$, together with the linear maps $h_i^j\colon H_i \to H_j$ obtained by composition from the maps between contiguous groups is referred to as a *persistence module*. We call the images of these linear maps the *persistent homology groups* of f. For h_i^j, this image consists of the classes that already exist in H_i and still exist in H_j. There are generalizations to bi-directional sequences and other graphs in the literature [9], but to keep things simple, we stay with uni-directional modules.

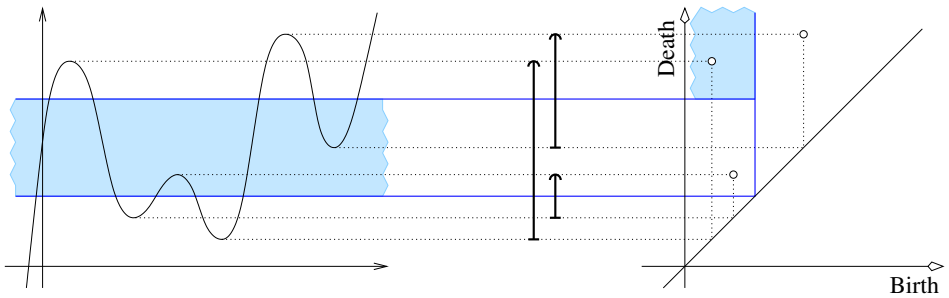

Figure 3: The graph of a 1-dimensional function on the *left* and the corresponding persistence diagram on the *right*. Only points with finite coordinates are shown. The shaded quadrant on the right contains a single point, which represents the component that exists in all sublevel sets defined by lines in the shaded horizontal strip on the left.

Introduced in [24], persistent homology has found many practical applications, facilitated by its elementary combinatorial representation and its fast algorithms. This representation is a collection of intervals, called a *barcode* or, equivalently, a multiset of points in the plane extended by points at infinity, called a *persistence diagram*; see Figure 3. Each interval marks the *birth* and *death* of a homology class as we move from left to right in the filtration. An algebraic interpretation is based on the decomposition of the module into indecomposables, but we will not need this here. Importantly, this decomposition is unique and the intervals characterize the persistence module up to isomorphisms. A most useful property of persistence diagrams is their stability under perturbations. It is formulated in terms of the *bottleneck distance*, which is the L_∞-length of the longest edge in a minimizing perfect matching between two diagrams, where we borrow points from the diagonal if needed. Given two functions, $f, g\colon K \to \mathbb{R}$, the theorem asserts that the bottleneck distance between the diagrams of f and g is bounded from above by $\|f - g\|_\infty$. The first proof of this insight can be found in [13]. Recently, the result has been generalized by relating the persistence diagram directly to the persistence module [4, 10].

The connection between persistent homology and alpha complexes should be clear: letting f be the radius function on the Delaunay mosaic, we compute its persistence diagram and get the ranks of the homology groups of $\mathrm{Alpha}_r(X)$ by counting the points in the upper-left quadrant, $[-\infty, r) \times [r, \infty]$. For

each point in the diagram, we also get information about its *persistence*, which we define as the absolute difference between its horizontal coordinate (the time of birth) and its vertical coordinate (the time of death). We have the same connection between persistence diagrams and the wrap complexes. Indeed, the alpha and wrap complexes for the same radius have the same homotopy type and therefore isomorphic homology groups. It is therefore not difficult to show that the filtration of wrap complexes generates the same persistence diagram as the filtration of alpha complexes.

D. EXTENSIONS

The three main concepts in our theory have far-reaching extensions. Indeed, for every generalized discrete Morse function on a simplicial complex, $g\colon K \to \mathbb{R}$, the sublevel sets, unions of lower sets, and the persistence diagram of g generalize the alpha shapes, wrap complexes, and persistent diagrams of the radius function of a Delaunay mosaic. Besides these automatic extensions, we are interested in Brillouin zones and in vineyards, as we will explain shortly.

Weighted points. In quantitative studies of molecules, it is common to represent atoms by spheres and to distinguish different atom types by mapping them to different size spheres. This motivates the generalization of the Voronoi tessellation to the *weighted Voronoi tessellation* in which every data point has a real weight. The influence of the point $x \in \mathbb{R}^d$ with weight $w_x \in \mathbb{R}$ on its surrounding is quantified by the *power function*, $\pi_x\colon \mathbb{R}^d \to \mathbb{R}$ defined by $\pi_x(a) = \|a - x\|^2 - w_x$. Substituting the power for the squared Euclidean distance, we define the *weighted Voronoi tessellation*, which is similar to the tessellation in the unweighted case and decomposes $\mathbb{R}^d$ into convex polyhedra. This concept is nearly as old as Voronoi tessellations and known by a variety of different names, including *power diagrams* and *Laguerre tessellations* [2]. Its dual is what we call the *weighted Delaunay mosaic*, and it has also been studied under a variety of names, including *regular*, *coherent*, and *convex triangulations* [16].

Assuming all weights are positive, we may think of each weighted point as a round ball, namely the sublevel set $\pi_x^{-1}(-\infty, 0]$. To get a filtration, we vary the upper bounds of the intervals, $\pi_x^{-1}(-\infty, w]$, where we write w for the parameter and not r because it behaves similar to the square of the radius. In analogy to the unweighted case, this gives a radius function on the weighted Delaunay mosaic and thus a filtration of *weighted alpha complexes* and of *weighted wrap complexes*.

Dense data. For simplicity, let us return to unweighted points, $X \subseteq \mathbb{R}^d$, although this is not necessary. For $Q \subseteq X$, consider the $a \in \mathbb{R}^d$ for which the points in Q are at least as close as those not in Q: $\mathrm{dom}(Q) = \{a \in \mathbb{R}^d \mid \|a - x\| \leq \|a - y\|, \forall x \in Q, y \in X \setminus Q\}$. Fixing an integer k, the *order-k Voronoi tessellation* of X is the collection of domains for sets $Q \subseteq X$ of cardinality k. Mapping each domain to the average of the k points that define it, we get a geometric realization of the dual, which we refer to as the *order-k Delaunay mosaic* of X. We note that beyond $d = 2$ dimensions, this mosaic is not necessarily simplicial even if we assume X be in general position.

To give more concrete geometric meaning to these concepts, fix a radius $r \geq 0$ and consider the *k-fold cover*, which consists of all points $a \in \mathbb{R}^d$ that are at distance at most r from at least k points in X. Now observe that the order-k Voronoi tessellation decomposes the k-fold cover into convex pieces, and restricting the dual to these pieces gives a subcomplex of the order-k Delaunay mosaic, which we refer to as the *order-k alpha complex* for radius r. Increasing r from 0 to ∞ gives a filtration of subcomplexes and therefore a monotonic radius function on the mosaic. It does not satisfy the requirements of a generalized discrete Morse function to the letter, but there are similarities so that a further generalization seems to be in order. For completeness we point out that the order-k Delaunay mosaic is the order-1 Delaunay mosaic of a set of points with weights, but the radius function defined by the k-fold cover is in general different from the one defined by these weighted points.

Brillouin zones. Letting $X \subseteq \mathbb{R}^d$ be locally finite and k a positive integer, the *k-th Brillouin zone* of $x \in X$ is the set of points $a \in \mathbb{R}^d$ for which x is the k-closest point of X, with ties ignored; see Figure 4. The 1-st Brillouin zone is the Voronoi domain of x, the union of the first two Brillouin zones is the union of domains $\mathrm{dom}(Q)$ in the order-2 Voronoi tessellation for which $x \in Q$, and so on. The Brillouin zones have originally been introduced in crystallography [8] and they were primarily used to study lattices, which are affine images of $\mathbb{Z}^d$ in $\mathbb{R}^d$. For example, Bieberbach proved that for a lattice the volume of any Brillouin zone is the same, namely the determinant of the matrix of basis vectors spanning the lattice [5]. This property generalizes to other periodic measures [21]. It is also interesting to note that for k going to infinity, the k-th Brillouin zone approaches the shape of a round sphere. Indeed, this is true for every lattice but not for more general point sets.

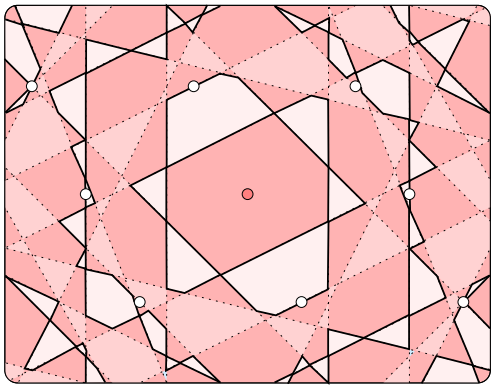

Figure 4: The Brillouin zones centered at the point in the middle are distinguished by alternating shading. To improve visibility, we emphasize the boundaries of the 2-nd, 6-th, and 10-th zones.

Vineyards. To study a time-series of functions, we may consider the corresponding stack of persistence diagrams, which we call a vineyard [14]. Imagine, for example, a slowly folding protein. At any moment in time, $t \in [0, 1]$, the data is a collection of (possibly weighted) points in $\mathbb{R}^3$, each the center of an atom of the protein. Correspondingly, we get the persistence diagram of the radius function on the Delaunay mosaic, which provides a multi-scale description of the shape at time t. Following the folding process from its beginning at $t = 0$ to the end at $t = 1$, we get a 1-parameter family of persistence diagrams. At this juncture, we recall that the bottleneck distance between the persistence diagrams of functions f, g on the same domain is bounded from above by $\|f - g\|_\infty$. In our example, the function at time t is $f_t \colon \mathbb{R}^3 \to \mathbb{R}$ defined by mapping $a \in \mathbb{R}^3$ to $f_t(a) = \min_{x \in X} \|a - x(t)\|$. This function varies continuously with t, which implies that also the points in the persistence diagram vary continuously with t. Stacking up the persistence diagrams of the 1-parameter family, we see that each point moves along a continuous path, which we call a *vine*. The collection of such vines is the *vineyard* of the time-series of distance functions or, more descriptively of the folding protein. We mention that vineyards can also be defined for two or more independent time-parameters. These extensions have not been studied much in the literature, and we are not aware of any published algorithms or software.

E. OBJECTIVES

We are now ready to flesh out Objectives (I) to (IV) mentioned in the Proposal Summary of Part B1. We will do so by decomposing each objective into smaller tasks, and by exploring possible extensions.

Adaptive reconstruction. The wrap complex may be viewed as an adaptive alpha complex, in the sense that simplices are selected not solely based on the radius but also depending on their shape and the sur-

rounding. Similarly, we can ask whether it is possible to have an adaptive version of the wrap complex, in the sense that the holes are not solely determined by the radii of the critical simplices but also by their persistence. The plan of how to approach this challenge will be explained in Section b. One ingredient in our approach is the canonical basis of a persistence module, which is of independent interest.

Obj. (I.1) *Adapt the wrap algorithm for a finite set of points in $\mathbb{R}^d$ so that the complex has all holes of persistence larger than some threshold and none of the holes with smaller persistence.*

Obj. (I.2) *Evaluate the known heuristics for speeding up the conventional persistence algorithm, eg. [11], when applied to computing the canonical basis of a persistence module.*

We expect that the primary applications of the adaptive wrap complex will be to biomolecules in $\mathbb{R}^3$ and to time-series of such molecules. This suggests we combine the adaptive reconstruction with vineyards.

Obj. (I.3) *Customize the vineyard algorithm for kinetic alpha complexes in $\mathbb{R}^3$.*

Obj. (I.4) *Study vineyards for two and more independent time-parameters and develop algorithms to construct them.*

The main challenge in (I.3) will be the implementation of the algorithm in [14] for complexes that change combinatorially with time, as do Delaunay mosaics. Objective (I.4) is interesting because a biomolecule might have more than one degree of substantial flexibility. Mapping each degree to a time-dimension, we can use such an algorithm to search for an interesting path, for example one that propagates a bubble (void) from one end of the molecule to another.

Stochastic analysis. During the last two years, we had some success in the stochastic analysis of radius functions on Delaunay mosaics [26, 27]. For example, we now have integral expressions for the expected number of critical simplices if $X \subseteq \mathbb{R}^d$ is a stationary Poisson point process, but also for intervals of every possible type. While the analytic evaluation of the integral expressions is still a challenge, we have all constants up to dimension $d = 4$ in the unweighted case, and for dimension $d = 2$ in the weighted case. To turn expected numbers of critical simplices into expected Betti numbers, we need to better understand the division into birth and death. We are not hopeful that an analytic solution to this question can be found, but experimental studies provide intuition and are independently useful. The mentioned integral expressions for the intervals imply relations for the expected numbers of simplices in the alpha complexes. Because of the adaptive nature of the wrap algorithm, these relations do not generalize to wrap complexes.

Obj. (II.1) *For each dimension $0 \leq p \leq d$, study the distribution of critical p-simplices that give birth to p-dimensional holes and compare it with the distribution of critical p-simplices that give death to $(p-1)$-dimensional holes.*

Obj. (II.2) *Give bounds for the expected number of simplices in the wrap complexes of a Poisson point process in $\mathbb{R}^d$.*

An answer to (II.1) would give us the expected Betti numbers of the alpha complexes but not how long the corresponding homology classes persist. The latter is likely to be an even harder question, and bounds on the expected maximum persistence have only recently become available [6]. A somewhat different but related question concerns the shape of the cells in weighted Voronoi tessellations. Writing ν_d for the d-dimensional volume of the unit ball in $\mathbb{R}^d$ and σ_d for the $(d-1)$-dimensional volume of the unit sphere in $\mathbb{R}^d$, we quantify the *roundness* of a body $A \subseteq \mathbb{R}^d$ with boundary ∂A by

$$\operatorname{round}(A) = \frac{\sigma_d^{d/d-1}}{\nu_d} \cdot \frac{\operatorname{vol}_d(A)}{\operatorname{vol}_{d-1}(\partial A)^{d/d-1}}, \tag{1}$$

which is a real number in $[0, 1]$ and attains its maximum at 1 for every round ball. Letting X be a Poisson point process in $\mathbb{R}^n$, with $d < n$, we get a d-dimensional weighted Voronoi tessellation by intersecting the (unweighted) Voronoi tessellation of X with a d-plane. We have experimental evidence that for $d = 2$ the roundness of the two-dimensional cells increases as n grows.

Obj. (II.3) *Determine the distribution of roundness of the d-cells in a weighted Voronoi tessellation in $\mathbb{R}^d$ obtained as a slice of the (unweighted) Voronoi tessellation of a Poisson point process in $\mathbb{R}^n$.*

Obj. (II.4) *Determine whether the points that generate non-empty d-cells in the weighted Voronoi tessellation obtained from a Poisson point process in $\mathbb{R}^n$ approach a regular configuration, as n goes to infinity, and if yes, which one.*

Related to the computational evidence about roundedness, we have evidence that in $d = 2$ dimensions the distribution of volumes of the cells gets narrower as n goes to infinity. This is consistent with an affirmative answer to (II.4) in the plane, where we would expect the regular hexagonal grid as limit configuration.

Dense data. Let $X \subseteq \mathbb{R}^d$ be finite and k a positive integer. Write $X_r(k)$ for the k-fold cover for radius r and recall that the order-k Voronoi tessellation of X decomposes $X_r(k)$ into convex pieces. This construction defines a radius function on the order-k Delaunay mosaic and a corresponding filtration of order-k alpha complexes. As proved in [1], there is a set of weighted points whose order-1 Voronoi tessellation is the order-k Voronoi tessellation of X, but interestingly enough the weighted points define a different radius function. Note that we have two filtrations which capture different aspects of the family of covers, one obtained by fixing k and varying r, and the other obtained by fixing r and varying k, which we call the *depth*.

Obj. (III.1) *Implement and optimize the incremental algorithm that computes the order-k from the order-$(k-1)$ Delaunay mosaic. Furthermore, compute the radius function of the order-k Delaunay mosaic and its persistence diagram.*

Obj. (III.2) *Compare the persistence diagram in (III.1) with that obtained for the weighted points whose order-1 Voronoi tessellation mosaic is the order-k Voronoi tessellation of X.*

Obj. (III.3) *Develop an algorithm for computing the persistence diagram of the filtration of covers for fixed radius, $\emptyset \subseteq \ldots \subseteq X_r(k) \subseteq X_r(k-1) \subseteq \ldots \subseteq \mathbb{R}^d$.*

Obj. (III.4) *Study the 2-parameter filtration of covers by comparing the sequences of persistence diagrams, one for each depth value $k \geq 1$, with the vineyard in which the radius is treated as time.*

To approach (III.3), it seems opportune to represent each k-fold cover by the corresponding order-k alpha complex, but there are challenges because the complexes are not nested and it is not obvious how to construct maps between them that induce the proper homomorphisms between the homology groups. Ideas how to overcome this obstacle can be found in Section b.

Long-range order. If $X \subseteq \mathbb{R}^d$ is a lattice, then the radius function of the Delaunay mosaic necessarily has a simple persistence diagram. For example if $X = \mathbb{Z}^3$ in $\mathbb{R}^3$, then the persistence diagram has only three points, each with infinite multiplicity: $(0, 1/2)$ for the gaps, $(1/2, \sqrt{2}/2)$ for the tunnels, and $(\sqrt{2}/2, \sqrt{3}/2)$ for the voids. When we perturb X, these piles of points spread out to form local clusters in the diagram. Before we address sets that are near-lattices, we need to better understand the lattices.

Obj. (IV.1) *Characterize the persistence diagrams of the Delaunay radius functions of the family of lattices in $\mathbb{R}^d$.*

Obj. (IV.2) *Generalize (IV.1) to order-k Delaunay mosaics.*

We can reasonably expect characterizations of the diagrams only for very small values of d and k, but weaker results such as bounds on the number of different points might be possible more generally. Objective (IV.1) is related to classifying Voronoi domains in lattices. It is easy to see that there are only two combinatorially different kinds in $\mathbb{R}^2$, namely the convex hexagon and the rectangle. There are 5 kinds in $\mathbb{R}^3$, which was known already to Fedorov [29], 52 kinds in $\mathbb{R}^4$, as proved by Delaunay [15] and Shtogrin [40], and 103,769 in $\mathbb{R}^5$, as determined by Engel [28]. It may be interesting to extend these classifications to order-k Voronoi tessellations.

An interesting tool to measure long-range order in a configuration is the sequence of Brillouin zones. Recall that each lattice point has a sequences of such zones, which decompose $\mathbb{R}^d$ like the layers of an onion. For a lattice, all Brillouin zones have the same volume, and the shape approaches that of a round sphere as k goes to infinity. We expect that this regularity deteriorates quickly when we perturb the points, but we do not know how quickly.

Obj. (IV.3) *Quantify how fast the Brillouin zones of a lattice approach the shape of a sphere as k goes to infinity.*

Obj. (IV.4) *Compare the Brillouin zones for lattices, near-lattices, and random point sets to get a quantification for how far a set is from being a lattice.*

There are obvious connections between the zones and the tessellations, namely the k-Brillouin zone of a point $x \in X$ is the union of the $\mathrm{dom}(Q)$ minus the union of the $\mathrm{dom}(P)$, in which $\mathrm{card}\, Q = 1 + \mathrm{card}\, P = k$ and $x \in P, Q$. We therefore foresee a cross-fertilizing effect between the tools developed for dense data and for long-range order.

Section b: Methodology

In this section, we describe plans how to approach the objectives listed in E. For some of them, we have specific ideas, and for others, our thoughts are more speculative. As a general principle, we believe in the power of *alternating and combining complementary activities* in our work. We use applications to motivate the mathematical and computational questions, which we pursue with theoretical means. We alternate between proving theorems and running computational experiments. Indeed, solid insights are necessary to see order in experimental results, and exceptions to or extensions of this order gleaned from the experiments can motivate conjectures that have the potential to deepen our insights or open new directions of inquiry. We also believe that algorithm design and software development support each other. Clearly, the design needs to proceed the implementation, but information flowing back can be productive, such as early computational experiments identifying bottlenecks and other weaknesses in the design.

F. ADAPTIVE RECONSTRUCTION

A motivating problem for Objective (I) is the search for *transport pathways* in proteins [12]. For example cell membrane proteins often have precisely one functional channel — which is likely the most persistent tunnel in the alpha shape filtration — but experience shows that there may be no radius for which the alpha shape has the functional channel as its sole tunnel. We propose to use the interval structure of the Delaunay radius function to adapt the wrap algorithm so it automatically reconstructs the shape with the most persistent as sole tunnel. Generalizations to other selections of tunnels but also gaps and voids are possible. Our approach to this problem requires more information about the persistence module than contained in the persistence diagram, namely its *canonical basis*, which we discuss first.

To develop an intuition for this basis, consider a connected graph, fix a spanning tree, and distinguish between *tree edges* and *non-tree edges* of the graph. Every non-tree edge connects two vertices in the spanning tree and thus forms a unique cycle of which this is the only non-tree edge. We get such a cycle for every non-tree edge, and the collection of such cycles forms a basis of the 1-st homology group of the graph. There are other such bases, but given a total order of the edges, there is a unique one, which can be constructed with the greedy algorithm. This works because the system of cycle-free edges forms a matroid. Similarly, the system of cycle-free p-simplices in a simplicial complex form a matroid. We can therefore apply the greedy algorithm to simultaneously compute all canonical bases, one for each dimension $p \geq 0$. To describe this algorithm, we let K be a simplicial complex, we assume a total order of its n simplices, and we write ∂ for the correspondingly ordered boundary matrix. The algorithm uses two matrices: U to store chains and R to store the boundaries of the chains. Initially, U is the identity matrix and $R = \partial U$. Using left-to-right column operations to reduce R, the algorithm writes $low(i)$ for the row index of the lowest non-zero item in the i-th column of R, setting $low(i) = 0$ if this is the zero column. Using modulo-2 arithmetic, every non-zero item is equal to 1.

```
for j = 1 to n do
    while ∃i < j such that R[low(i), j] = 1 do
        add the i-th column to the j-th column, both in R and in U.
```

We call this the *exhaustive reduction algorithm* because it does not stop when the lowest non-zero item in a column has been determined; compare with the standard reduction algorithm [20, chapter VII]. Objective (I.2) asks for more efficient ways to exhaustively reduce the boundary matrix than described above. It is not difficult to see that the produced matrices, R and U, are unique in the sense that they do not depend on the sequence of the column operations. The *canonical basis* can be retrieved from U by collecting the cycles stored in the columns of birth-giving simplices. The remainder of the algorithm is straightforward (to describe but not to prove correct). Suppose the j-th simplex, Q_j, gives death, and observe that the birth-giving simplices that contain Q_j in their canonical cycles are characterized by having a non-zero item in the j-th row of U. To construct the *adaptive wrap complex*, we therefore

STEP 1. retrieve the desired tunnel from the persistence diagram, with death-giving triangle Q_j;

STEP 2. collect the birth-giving triangles that contain Q_j in their corresponding canonical cycles;

STEP 3. output the union of lower sets over all critical simplices whose lower sets contain neither Q_j nor any birth-giving triangles collected in Step 2.

We have implemented a more general version of this algorithm and are encouraged by the positive results. More interesting than reconstructing the functional channel in a static model is finding such a channel in a dynamic model, by which we mean a time-series of static models. In an optimistic scenario, we would get the same channel at all moments in time, and thus a vineyard consisting of a single vine. Rather than scrutinizing less optimistic scenarios, we briefly discuss the algorithmic challenges for computing this vineyard. We imagine the data consists of finitely many points continuously moving in space, $x_i \colon [0,1] \to \mathbb{R}^3$. Computing the Delaunay mosaic for the points $x_i(0)$, and its radius function and persistence diagram, we get the vineyard by maintaining this information as the time increases from $t = 0$ to $t = 1$. An elementary step of the vineyard algorithm [14] is a transposition of two simplices in the total order defined by the radius function. In our particular setting, there is a second kind of elementary step, namely the change of the interval structure, which happens when the points move through a configuration that violates general position, such as 5 points on a common sphere. An efficient solution to Objective (I.3) will require the effective scheduling of these events and data structures that are optimized for the elementary steps.

G. STOCHASTIC GEOMETRY

A first motivation for the stochastic analysis of random data is the creation of a baseline expectation to which we can compare results computed for non-random data. A second motivation is information about the complexes that might be useful in the design or optimization of our algorithms. Since our data is usually a finite set of points, the most natural choice of random data is a stationary Poisson point process, $X \subseteq \mathbb{R}^d$. Considering the radius function on the Delaunay mosaic of X, we have used integral geometry relations combined with discrete geometric reasoning to determine the expected density of intervals of any type and with radius at most some threshold [27]. Objective (II.1) asks to split the distribution of the expected numbers of critical simplices (in singular intervals) into the distributions of the ones that give birth and the others that give death. Observe that this question has an easy solution if we drop the dependence on the radius and just ask for the total numbers. Indeed, the expected number of death-giving critical edges is 1 less than the expected number of vertices, the expected number of birth-giving critical edges is the expected number of critical edges minus the expected number of death-giving critical edges, and so on for critical triangles and higher-dimensional simplices. So we are really after intermingling the two types of critical simplices, but it is unlikely that our tools will suffice to shed significant light on this stochastic question. Indeed, the difference between death- and birth-giving is of global nature, which is bad news for the local tools in probability theory. This suggests we switch from the analytic to an experimental approach, returning to proving theorems after formulating conjectures gleaned from the data.

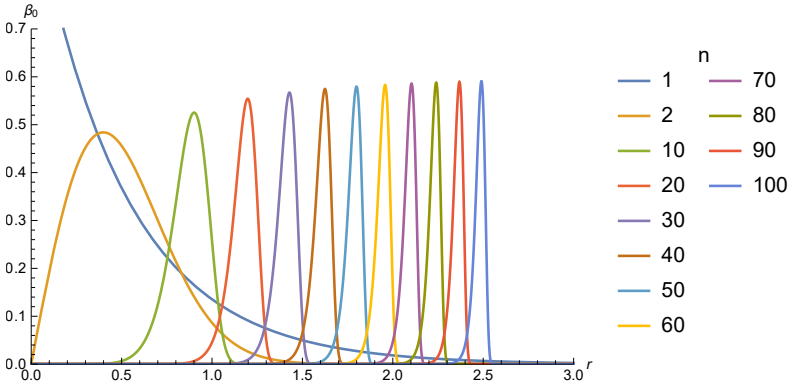

Figure 5: Graphs of the expected number of connected components in a random 1-dimensional weighted alpha complex per unit length. For increasing n, the graphs tend to get more concentrated, suggesting that the critical vertices get progressively more evenly spaced along the line.

Let us change gears and discuss Objective (II.3) on the roundness of top-dimensional cells in d-dimensional weighted Voronoi tessellations, which we generate by taking a d-dimensional slice of the n-dimensional unweighted Voronoi tessellation of a Poisson point process in $\mathbb{R}^n$. We have experimental evidence that for growing n, the cells get rounder. We will collect more data to be able to quantify this conjecture. There is a related conjecture for which we have analytic support in dimension $d = 1$. Considering the distribution of the distances between adjacent vertices of the weighted Delaunay mosaic, we conjecture that this distribution gets progressively more narrow as n goes to infinity. In dimension $d = 1$, the weighted Delaunay mosaic alternates between critical vertices and critical edges, padding the gaps between them with vertex-edge pairs. The critical vertices start components, the pairs grow them, and the critical edges merge them. Since we have analytic expressions depending on the radius for all three [26], we get an analytic expression for the expected number of components of the weighted alpha shape for radius r; see Figure 5. We clearly see that the distributions get progressively more narrow as n grows. A possible explanation is that the critical vertices get progressively more evenly spaced and their weights get

progressively closer to each other. The generalization of this question to $d \geq 2$ dimensions is interesting as its answer would give information on how to approach Objective (II.4).

H. DENSE DATA

There are two motivations for research on this topic, namely ordered or unordered particle systems and the mundane challenge of large data sets in topological data analysis. We approach this subject by revisiting the generalization of Voronoi tessellations to order-k [34, 39]. We find it useful to recast the concept in the dual, capturing all tessellations in a single structure, which we call a *rhomboid tiling*; see Figure 6 for a low-dimensional example. To define it, let $X \subseteq \mathbb{R}^d$ be in general position. For any subset Q of $d+1$ points, consider the unique sphere that passes through all points of Q, and let $E \subseteq X$ be the subset of points strictly enclosed by the sphere. The pair Q, E corresponds to a rhomboid in the tiling, and to describe this rhomboid we map every subset $F \subseteq Q$ to the point $(\sum E + \sum F, \operatorname{card} E + \operatorname{card} F)$ in $\mathbb{R}^{d+1}$. There are 2^{d+1} such points, namely the vertices of the rhomboid spanned by the vectors $(x, 1)$ with $x \in Q$.

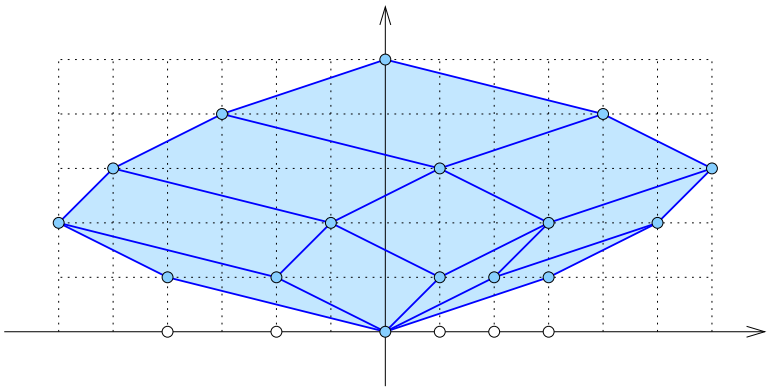

Figure 6: The rhomboid tiling of 5 points on the real line. The horizontal line at height k intersects the tiling in a geometric realization of the order-k Delaunay mosaic of the 5 points.

We skip the argument that this construction really gives a tiling with rhomboids in $\mathbb{R}^{d+1}$. It is not difficult to see that the horizontal d-plane at height k intersects the tiling in a geometric realization of the order-k Delaunay mosaic. The top-dimensional cells in the mosaic have therefore the same combinatorial structure as particular slices of the $(d+1)$-cube. For $d = 2$ we have triangles, for $d = 3$ we have tetrahedra and the octahedron, and in general we have convex hulls of face barycenters of the d-simplex. Perhaps mildly surprising, this shows that in high dimensions the cells in the order-k Delaunay mosaic that are simplicial are the exception rather than the norm. We plan to exploit the rhomboid tiling to approach all four objectives stated for dense data.

1. The tiling clarifies the incremental algorithm as it shows how we can get the cells of the order-k from those of the order-$(k-1)$ Delaunay mosaic. The gaps left between the cells derived from the order-$(k-1)$ mosaic need to be filled with the d-simplices of local order-1 Delaunay mosaics.

2. Given the order-k Delaunay mosaic, it is not difficult to compute the two radius functions — the one determined by the k-fold covers and the other by the weighted points — and to compare their persistence diagrams. Can we say something about the interleaving between their persistence modules?

3. Introducing a radius, r, we get a filtration of order-k alpha complexes whose persistence diagram is the same as that of the filtration of k-fold covers. Using the rhomboid tiling, we characterize the increments in the filtration as a generalization of the interval structure of discrete Morse theory.

4. Fixing r but varying k, we get a filtration of alpha complexes, one for every k. These complexes are not related by inclusions, but we can use the slabs of the rhomboids between heights k and $k-1$ to construct a map from the order-k to the order-$(k-1)$ Delaunay mosaic.

Not all details have been worked out, but the preliminary results show that we are on the right track, suggesting that there are no serious obstacles to completing Objectives (III.1) to (III.3). Objective (III.4) is more open ended. It would be interesting to see whether there are special properties that help avoid some of the algebraic difficulties of multi-parameter persistence.

J. LONG-RANGE ORDER

The motivation for this topic derives from materials whose atoms are globally ordered, or approximately so. Specifically, we are interested in characterizing and quantifying local interruptions of the global order. There are two extreme types: local defects within otherwise perfect order, and slight deviations from the global order everywhere. We are interested in both, and everything in between, and propose to use our geometric tools to shed light on the subject. To start, we will study persistence diagrams of lattices, a topic that has not received much attention yet. Objectives (IV.1) and (IV.2) ask for classifications of the persistence diagrams of lattices — under some appropriate notion of equivalence. This is an interesting exercise for $k = 1$ in $d = 2$ dimensions, and extension to larger values of k may be possible. The case $k = 1$ and $d = 3$ is challenging but also most interesting and perhaps doable; see [22] for the exploration of a 1-parameter sub-family of 3-dimensional lattices. With increasing k and d, the classification gets more difficult and soon impossible. This suggests we adopt goals that are less ambitious than classification, such as proving bounds on the number of different points in the persistence diagrams.

Another geometric tool that captures aspects of global order is the sequence of Brillouin zones. Recall that for a lattice, all Brillouin zones have the same volume, and as k goes to infinity, the Brillouin zone converges to the shape of a round sphere. To approach Objectives (IV.3) and (IV.4), we investigate near-periodic configurations by quantifying how these two properties — constant measure and spherical shape in the limit — deteriorate when the strict periodicity is relaxed. To quantify how fast the Brillouin zones approach the shape of a sphere is interesting even for lattices, where it relates to number theoretic questions. A useful tool in this context is the *radial distance function*, which maps every direction to the distance between the center and the closest point of the k-th Brillouin zone in this direction. We propose to compute the persistence diagram of this function, which for a lattice consists of two points at infinity — representing the non-trivial classes in the 0-th and the $(d-1)$-st homology groups of the sphere — and otherwise of points near the diagonal. We expect that amorphous deviations from the global order, as we see in glass, increase the noise near the diagonal but otherwise maintain this structure of the persistence diagram. For local defects to order, we expect a dependence of the diagram on the location of the center of the Brillouin zones. The method even works for data sets that are very far from globally ordered. While the Brillouin zones will generally not approach a spherical shape, they are star-convex so that the radial function is well defined. Its persistence diagram reflects density differences that lead to the distortions of the zones. In $\mathbb{R}^d$, we have a $(d+1)$-dimensional family of such diagrams — the depth direction and d spatial directions for the center — and it might be interesting to develop summary statistics over this family that characterizes the varying density of the data.

References

[1] F. AURENHAMMER. A new duality result concerning Voronoi diagrams. *Discrete Comput. Geom.* **5** (1990), 243–254.

[2] F. AURENHAMMER. Voronoi diagrams — a study of a fundamental geometric data structure. *ACM Comput. Surveys* **23** (1991), 345–405.

[3] U. BAUER AND H. EDELSBRUNNER. The Morse theory of Čech and Delaunay complexes. *Trans. Amer. Math. Soc.* **369** (2017), 3741–3762.

[4] U. BAUER AND M. LESNICK. Induced matchings of barcodes and the algebraic stability of persistence. In *Proc. 30th Sympos. Comput. Geom., 2014*, 355–364.

[5] L. BIEBERBACH. Über die Inhaltsgleichheit der Brillouinschen Zonen. *Monatsh. Math. Phys.* **48** (1939), 509–515.

[6] O. BOBROWSKI, M. KAHLE AND P. SKRABA. Maximally persistent cycles in random geometric complexes. *Ann. Appl. Probab.*, to appear.

[7] K. BORSUK. On the imbedding of systems of compacta in simplicial complexes. *Fund. Math.* **35** (1948), 217–234.

[8] L. BRILLOUIN. Les électrons dans les métaux et le classement des ondes de de Broglie correspondantes. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* **191** (1930), 292.

[9] G. CARLSSON AND V. DE SILVA. Zigzag persistence. *Found. Comput. Math.* **10** (2010), 367–405.

[10] F. CHAZAL, V. DE SILVA, M. GLISSE, AND S. OUDOT. *The Structure and Stability of Persistence Modules.* SpringerBriefs in Mathematics, Springer, 2116.

[11] C. CHEN AND M. KERBER. Persistent homology computation with a twist. In *Proc. 27th European Workshop Comput. Geom., 2011*.

[12] E. CHOVANCOVA, A. PAVELKA, P. BENES, O. STRNAD, J. BREZOVSKY, B. KOZLIKOVA, A. GORA, V. SUSTR, M. KLVANA, P. MEDEK, L. BIEDERMANNOVA, J. SOCHOR AND J. DAMBORSKY. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLOS Computational Biology* (2012), doi.org/10.1371/journal.pcbi.1002708.

[13] D. COHEN-STEINER, H. EDELSBRUNNER AND J. HARER. Stability of persistence diagrams. *Discrete Comput. Geom.* **37** (2007), 103–120.

[14] D. COHEN-STEINER, H. EDELSBRUNNER, AND D. MOROZOV. Vines and vineyards by updating persistence in linear time. In *Proc. 22nd Sympos. Comput. Geom., 2006*, 119–126.

[15] B.N. DELONE. Sur la partition régulière de l'espace à 4 dimensions. *Bull. Acad. Sci. URSS* (1929), 79–110, ibid. 145–164.

[16] J. DE LOERA, J. RAMBAU AND F. SANTOS. *Triangulations.* Springer, Heidelberg, Germany, 2010.

[17] T.K. DEY. *Curve and Surface Reconstruction.* Cambridge Univ. Press, Cambridge, England, 2006.

[18] H. EDELSBRUNNER. Surface reconstruction by wrapping finite sets in space. In *Discrete and Computational Geometry. The Goodman-Pollack Festschrift*, 379–404, eds.: B. Aronov, S. Basu, J. Pach and M. Sharir, Springer-Verlag, 2003.

[19] H. EDELSBRUNNER. Alpha shapes — a survey. In *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*, eds.: R. van de Weygaert, G. Vegter, J. Ritzerveld and V. Icke, Springer-Verlag, to appear.

[20] H. EDELSBRUNNER AND J. L. HARER. *Computational Topology. An Introduction.* Amer. Math. Soc., Providence, Rhode Island, 2010.

[21] H. EDELSBRUNNER AND M. IGLESIAS-HAM. On the optimality of the FCC lattice for soft sphere packing. Manuscript, IST Austria, Klosterneuburg, Austria, 2017.

[22] H. EDELSBRUNNER AND M. KERBER. Covering and packing with spheres by diagonal distortion in $\mathbb{R}^n$. In *Rainbow of Computer Science*, 20–35, eds.: C. Calude, G. Rozenberg and A. Salomaa, Springer Lecture Notes in Computer Science **6570**, 2011.

[23] H. EDELSBRUNNER, D.G. KIRKPATRICK AND R. SEIDEL. On the shape of a set of points in the plane. *IEEE Trans. Inform. Theory* **IT-29** (1983), 551–559.

[24] H. EDELSBRUNNER, D. LETSCHER AND A. ZOMORODIAN. Topological persistence and simplification. *Discrete Comput. Geom.* **28** (2002), 511–533.

[25] H. EDELSBRUNNER AND E.P. MÜCKE. Three-dimensional alpha shapes. *ACM Trans. Graphics* **13** (1994), 43–72.

[26] H. EDELSBRUNNER AND A. NIKITENKO. Weighted Poisson–Delaunay mosaics. Manuscript, IST Austria, Klosterneuburg, Austria, 2017.

[27] H. EDELSBRUNNER, A. NIKITENKO AND M. REITZNER. Expected sizes of Poisson–Delaunay mosaics and their discrete Morse functions. *Adv. Appl. Probab.* **49** (2017), to appear.

[28] P. ENGEL. The contraction types of parallelohedra in $\mathbb{R}^5$. *Acta Cryst. A* **56** (2000), 491–496.

[29] E.S. FEDOROV. *Nachala ucheniya o fgurakh.* (in Russian). 1885.

[30] R. FORMAN. Morse theory for cell complexes. *Adv. Math.* **134** (1998), 90–145.

[31] R. FREIJ. Equivariant discrete Morse theory. *Discrete Math.* **309** (2009), 3821–3829.

[32] A. HATCHER. *Algebraic Topology.* Cambridge Univ. Press, Cambridge, England, 2002.

[33] R.A. JARVIS. Computing the shape hull of points in the plane. *In* "Comput. Soc. Conf. Pattern Recognition and Image Processing, 1977", 231–241.

[34] D.-T. LEE. On k-nearest neighbor Voronoi diagrams in the plane. *IEEE Trans. Comput.* **C-31** (1982), 478–487.

[35] J. LERAY. Sur la forme des espaces topologiques et sur les points fixes des représentations. *J. Math. Pures Appl.* **24** (1945), 95–167.

[36] J. MILNOR. *Morse Theory.* Princeton Univ. Press, Princeton, New Jersey, 1963.

[37] J.R. MUNKRES. *Elements of Algebraic Topology.* Perseus, Cambridge, Massachusetts, 1984.

[38] R. SCHNEIDER AND W. WEIL. *Stochastic and Integral Geometry.* Springer, Berlin, Germany, 2008.

[39] M.I. SHAMOS AND D. HOEY. Closest-point problems. *In* "Proc. 16th Ann. ACM Sympos. Found. Comput. Sci., 1975", 151–162.

[40] M.I. SHTROGRIN. Regular Dirichlet–Voronoi partitions for the second triclinic group. *Proc. Steklov Inst. Math.* **123** (1973).

Section c: Resources (including project costs)

As explained in the technical sections of this proposal, the project consists of two symbiotic activities: the design and creation of the software and experimentation, and the development of the mathematics needed for the analysis of the software and the interpretation of the results obtained with it. We request financial support for **2 postdocs** and for **2 students** over most of the lifetime of this project. An exception is the first year for which we request support for one postdoc less, since we will need some time to find the right person and to get the different aspects of the project started and coordinated. We will require a mix of expertise in programming, discrete geometry, mesh generation, computational topology, and stochastic geometry. We will make use of our scientific network to identify and attract students and postdocs with these qualifications. At the beginning of the 5-year period, we plan to work with the following two students and one postdoc who are currently part of the PI's group at IST Austria: **Katharina Ölsböck** working on alpha and wrap complexes, **Georg Osang** working on density regimes, as well as **Arseniy Akopyan** who has a strong background in discrete geometry. To assume a successful continuation of the project work, we will hire students and postdocs in critical areas. Currently, we are talking to **Therese Heiss** (beginning student at IST Austria), and we will be talking to finishing students in the areas of computational geometry and computational topology. We request funds to attend **2 conferences** per year per member of the project team. Since this is a cross-disciplinary project, it is essential that the team members attend conferences in computer science and mathematics but also in meetings on molecular structure and on granular systems. In addition, we request funds for **computing time** and **data storage** as well as **printing material** to print shapes with the locally available 3D printers. Finally, we ask for funds to reimburse **invited speakers** and to pay for catering and other services as part of organizing **a major workshop** two years into the project to facilitate the communication with the wider research community.

The time-plan for the 5-year project is more complete near the beginning than near the end. We list the students and postdocs we expect to work during the beginning of the project, as well as the main topics in the four scientific themes. The later years depend on the early results and a more detailed plan will be developed as the project evolves.

	Year 1	Year 2	Year 3	Year 4	Year 5
postdoc I	A. Akopyan	A. Akopyan	TBD	TBD	TBD
postdoc II		TBD	TBD	TBD	TBD
student I	K. Ölsböck	K. Ölsböck	T. Heiss	T. Heiss	TBD
student II	G. Osang	G. Osang	TBD	TBD	TBD
Adapt. Recon.	canonical pers.	optimized softw.	3D pathways	custom. vnyd.	4D pathways
Stoch. Geom.	birth vs. death	expect. wrap	expect. pers.	expect. round.	optimized round.
Dense Data	increm. constr.	pers. in scale	pers. in depth	compare pers.	density vnyd.
Lg.-Rng. Ord.	pers. lattices	Brillouin algor.	stoch. Brillouin	pers. Brillouin	Brillouin vnyd.