

# Parallel Computation of Alpha Complexes for Biomolecules

Talha Bin Masood<sup>1</sup> 

Scientific Visualization Group, Linköping University, Norrköping, Sweden  
talha.bin.masood@liu.se

Tathagata Ray

BITS Pilani, Hyderabad Campus, Hyderabad, India  
rayt@hyderabad.bits-pilani.ac.in

Vijay Natarajan 

Department of Computer Science and Automation, Indian Institute of Science, Bangalore, India  
vijayn@iisc.ac.in

---

## Abstract

The alpha complex, a subset of the Delaunay triangulation, has been extensively used as the underlying representation for biomolecular structures. We propose a GPU-based parallel algorithm for the computation of the alpha complex, which exploits the knowledge of typical spatial distribution and sizes of atoms in a biomolecule. Unlike existing methods, this algorithm does not require prior construction of the Delaunay triangulation. The algorithm computes the alpha complex in two stages. The first stage proceeds in a bottom-up fashion and computes a superset of the edges, triangles, and tetrahedra belonging to the alpha complex. The false positives from this estimation stage are removed in a subsequent pruning stage to obtain the correct alpha complex. Computational experiments on several biomolecules demonstrate the superior performance of the algorithm, up to a factor of 50 when compared to existing methods that are optimized for biomolecules.

**2012 ACM Subject Classification** Theory of computation → Parallel algorithms; Computing methodologies → Shape modeling; Applied computing → Molecular structural biology

**Keywords and phrases** Delaunay triangulation, parallel algorithms, biomolecules, GPU

**Digital Object Identifier** 10.4230/LIPIcs.SoCG.2020.17

**Related Version** The full version of this paper is available at [29], <https://arxiv.org/abs/1908.05944>.

**Supplement Material** Source code available at: [https://bitbucket.org/vgl\\_iisc/parallelac/](https://bitbucket.org/vgl_iisc/parallelac/)

**Funding** *Vijay Natarajan*: This work is partially supported by a Swarnajayanti Fellowship from the Department of Science and Technology, India (DST/SJF/ETA-02/2015-16); a Mindtree Chair research grant; and the Robert Bosch Centre for Cyber Physical Systems, IISc.

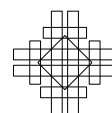
**Acknowledgements** Part of this work was done when the first author was at Indian Institute of Science, Bangalore. The authors would like to thank Sathish Vadhiyar and Nikhil Ranjanikar for helpful discussions and suggestions during the early phase of this work.

## 1 Introduction

The alpha complex of a set of points in  $\mathbb{R}^3$  is a subset of the Delaunay triangulation. A size parameter  $\alpha$  determines the set of simplices (tetrahedra, triangles, edges, and vertices) of the Delaunay triangulation that are included in the alpha complex. It is an elegant representation of the shape of the set of points [16, 18, 14] and has found various applications, particularly

---

<sup>1</sup> Corresponding author



in molecular modeling and molecular graphics. The atoms in a biomolecule are represented by weighted points in  $\mathbb{R}^3$ , and the region occupied by the molecule is represented by the union of balls centered at these points. The geometric shape of a biomolecule determines its function, namely how it interacts with other biomolecules. The alpha complex represents the geometric shape of the molecule very efficiently. It has been widely used for computing and studying geometric features such as cavities and channels [25, 26, 10, 30, 33, 28, 23]. Further, an alpha complex based representation is also crucial for accurate computation of geometric properties like volume and surface area [24, 12, 27].

Advances in imaging technology have resulted in a significant increase in the size of molecular structure data. This necessitates the development of efficient methods for storing, processing, and querying these structures. In this paper, we study the problem of efficient construction of the alpha complex with particular focus on point distributions that are typical of biomolecules. In particular, we present a parallel algorithm for computing the alpha complex and an efficient GPU implementation that outperforms existing methods. In contrast to existing algorithms, our algorithm does not require the explicit construction of the Delaunay triangulation.

## 1.1 Related work

The Delaunay triangulation has been studied within the field of computational geometry for several decades and numerous algorithms have been proposed for its construction [1]. Below, we describe only a few methods that are most relevant to this paper.

A tetrahedron belongs to the Delaunay triangulation of a set of points in  $\mathbb{R}^3$  if and only if it satisfies the *empty circumsphere* property, namely no point is contained within the circumsphere of the tetrahedron. The Bowyer-Watson algorithm [4, 34] and the incremental insertion algorithm by Guibas *et al.* [21] are based on the above characterization of the Delaunay triangulation. In both methods, points are inserted incrementally and the triangulation is locally updated to ensure that the Delaunay property is satisfied. The incremental insertion method followed by bi-stellar flipping works in higher dimensions also [20] and can construct the Delaunay triangulation in  $O(n \log n + n^{\lceil d/2 \rceil})$  time in the worst case, where  $n$  is the number of input points in  $\mathbb{R}^d$ . A second approach for constructing the Delaunay triangulation is based on its equivalence to the convex hull of the points lifted onto a  $(d + 1)$ -dimensional paraboloid [19].

A third divide-and-conquer approach partitions the inputs points into two or generally multiple subsets, constructs the Delaunay triangulation for each partition, and merges the pieces of the triangulation finally. The merge procedure depends on the ability to order the edges incident on a vertex and hence works only in  $\mathbb{R}^2$ . The extension to  $\mathbb{R}^3$  requires that the merge procedure be executed first [6]. The divide-and-conquer strategy directly extends to a parallel algorithm [31, 5]. The DeWall algorithm [6] partitions the input point set into two halves and first constructs the triangulation of points lying within the boundary region of the two partitions. The Delaunay triangulation of the two halves is then constructed in parallel. The process is repeated recursively resulting in increased parallelism. Cao *et al.* [5] have developed a GPU parallel algorithm, *gDel3D*, that constructs the Delaunay triangulation in two stages. In the first stage, points are inserted in parallel followed by flipping to obtain an approximate Delaunay triangulation. In the second stage, a star splaying procedure works locally to convert non-Delaunay tetrahedra into Delaunay tetrahedra. The algorithm can be extended to construct the weighted Delaunay triangulation for points with weights. Cao *et al.* report a speed up of up to a factor of 10 over a sequential implementation for constructing the weighted Delaunay triangulation of 3 million weighted points.

Existing algorithms for constructing the alpha complex [11, 18, 27, 9] often require that the Delaunay triangulation be computed in a first step, with the exception of a recent method that guarantees output sensitive construction under mild assumptions on weights [32] or a possible construction from Čech complexes [2]. Simplices that belong to the alpha complex are identified using a size filtration in a second step. In the case of biomolecules, only small values of the size parameter are of interest and the number of simplices in the alpha complex is a small fraction of those contained in the Delaunay triangulation. Hence, the Delaunay triangulation construction is often the bottleneck in the alpha complex computation. The key difficulty lies in the absence of a direct characterization of simplices that belong to the alpha complex.

## 1.2 Summary of results

We propose an algorithm that avoids the expensive Delaunay triangulation computation and instead directly computes the alpha complex for biomolecules. The key contributions of this paper are summarized below:

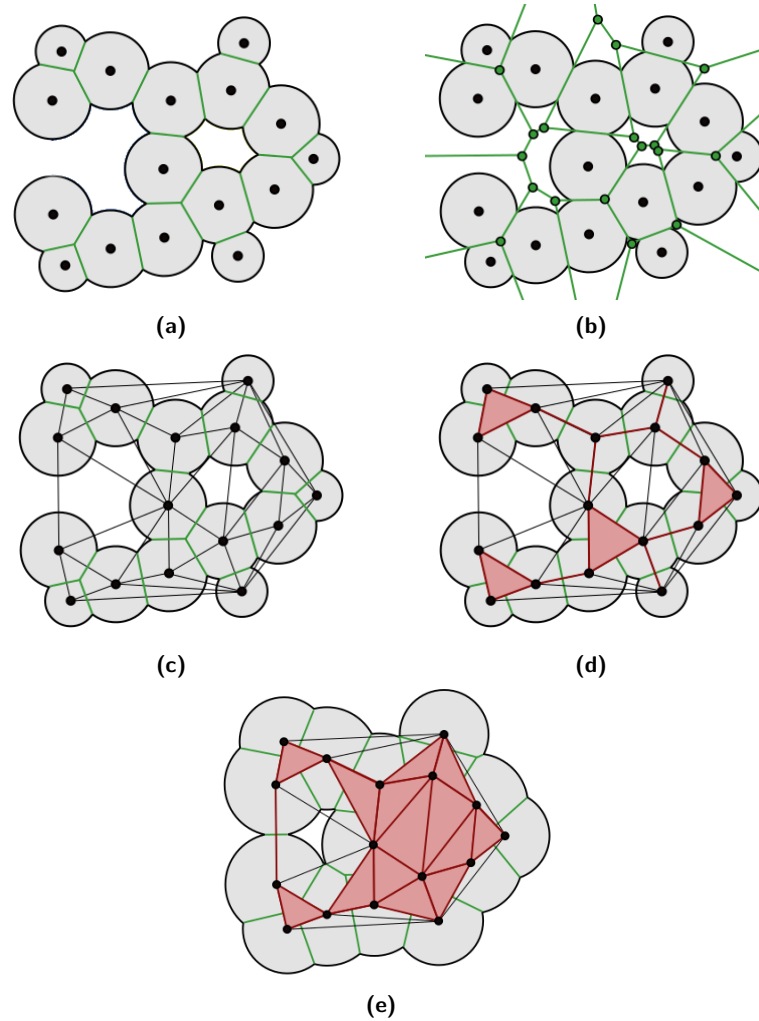
- A new characterization of the alpha complex – a set of conditions necessary and sufficient for a simplex to be a part of the alpha complex.
- A new algorithm for computing the alpha complex of a set of weighted points in  $\mathbb{R}^3$ . The algorithm identifies simplices of the alpha complex in decreasing order of dimension without computing the complete weighted Delaunay triangulation.
- An efficient CUDA-based parallel implementation of this algorithm for biomolecular data that can compute the alpha complex for a 10 million point dataset in approximately 10 seconds.
- A proof of correctness of the algorithm and comprehensive experimental validation to demonstrate that it outperforms existing methods.

While the experimental results presented here focus on biomolecular data, the algorithm is applicable to data from other application domains as well. In particular, the efficient GPU implementation may be used for points that arise in smoothed particle hydrodynamics (SPH) simulations, atomistic simulations in material science, and particle systems that appear in computational fluid dynamics (CFD).

## 2 Background

In this section, we review the necessary background on Delaunay triangulations required to describe the algorithm and also establish a new characterization of the alpha complex that does not require the Delaunay triangulation. For a detailed description of Delaunay triangulations, alpha complexes, and related structures, we refer the reader to various books on the topic [1, 13, 15].

Let  $B = \{b_i\}$  denote a set of balls or weighted points, where  $b_i = (p_i, r_i)$  represents a ball centered at  $p_i$  with radius  $r_i$ . We limit our discussion to balls in  $\mathbb{R}^3$ , so  $p_i = (x_i, y_i, z_i) \in \mathbb{R}^3$ . Further, we assume that the points in  $B$  are in general position, *i.e.*, no two points have the same location, no three points are collinear, no four points are coplanar, and no subset of five points are equidistant from a point in  $\mathbb{R}^3$ . Such configurations are called *degeneracies*. In practice, a degenerate input can be handled via symbolic perturbation [17].



■ **Figure 1** 2D weighted Delaunay triangulation and alpha complex. (a) A set of weighted points  $B$  in  $\mathbb{R}^2$  shown as disks. (b) The weighted Voronoi diagram of  $B$ . Voronoi edges and vertices are highlighted in green. (c) The weighted Delaunay complex is the dual of the weighted Voronoi diagram. (d) The alpha complex  $K_\alpha$  for  $\alpha = 0$  is shown in red. This is the dual of the intersection of the weighted Voronoi diagram and union of balls. (e) The alpha complex shown for an  $\alpha > 0$ . It is the dual of the intersection of the weighted Voronoi diagram and union of balls after growing them to have radius  $\sqrt{r_i^2 + \alpha}$ .

## 2.1 Simplex and simplicial complex

A  $d$ -dimensional *simplex*  $\sigma^d$  is defined as the convex hull of  $d + 1$  affinely independent points. Assuming the centres of balls in  $B$  are in general position, all  $(d + 1)$  sized subsets of  $B$  form a simplex  $\sigma^d = (p_0^\sigma, p_1^\sigma, \dots, p_d^\sigma)$ . For simplicity, we sometimes use  $b_i$  instead of the center  $p_i$  to refer to points incident on a simplex. For example, we may write  $\sigma^d = (b_0^\sigma, b_1^\sigma, \dots, b_d^\sigma)$ .

A non-empty strict subset of  $\sigma^d$  is also a simplex but with dimension smaller than  $d$ . Such a simplex is called a *face* of  $\sigma^d$ . Specifically, a  $(d - 1)$ -dimensional face of  $\sigma^d$  is referred to as a *facet* of  $\sigma^d$ . A set of simplices  $K$  is called a *simplicial complex* if: 1) a simplex  $\sigma \in K$  implies that all faces of  $\sigma$  also belong to  $K$ , and 2) for two simplices  $\sigma_1, \sigma_2 \in K$ , either  $\sigma_1 \cap \sigma_2 \in K$  or  $\sigma_1 \cap \sigma_2 = \emptyset$ .

## 2.2 Power distance and weighted Voronoi diagram

The *power distance*  $\pi(p, b_i)$  between a point  $p \in \mathbb{R}^3$  and a ball  $b_i = (p_i, r_i) \in B$  is defined as

$$\pi(p, b_i) = \|p - p_i\|^2 - r_i^2.$$

The *weighted Voronoi diagram* is an extension of the Voronoi diagram to weighted points. It is a partition of  $\mathbb{R}^3$  based on proximity to input balls  $b_i$  in terms of the power distance. Points  $p \in \mathbb{R}^3$  that are closer to the ball  $b_i$  compared to all other balls  $b_j \in B$  ( $j \neq i$ ) constitute the *Voronoi region* of  $b_i$ . Points equidistant from two balls  $b_i, b_j \in B$  and closer to these two balls compared to other balls constitute a *Voronoi face*. Similarly, points equidistant from three balls and four balls constitute *Voronoi edges* and *Voronoi vertices* of the weighted Voronoi diagram, respectively. Figure 1b shows the weighted Voronoi diagram for a set of 2D weighted points or disks on the plane. Similar to the unweighted case, the Voronoi regions of the weighted Voronoi diagram are convex and linear. However, the weights may lead to a configuration where the Voronoi region of  $b_i$  is disjoint from  $b_i$ . This occurs when  $b_i$  is contained within another ball  $b_j$ . Further, the Voronoi region of  $b_i$  may even be empty.

## 2.3 Weighted Delaunay triangulation

The *weighted Delaunay triangulation* is the dual of the weighted Voronoi diagram, see Figure 1c. It is a simplicial complex consisting of simplices that are dual to the cells of the weighted Voronoi diagram. The following equivalent definition characterizes a simplex  $\sigma^d$  belonging to a Delaunay triangulation  $D$ .

► **Definition 1** (Weighted Delaunay Triangulation). *A simplex  $\sigma^d = (p_0^\sigma, p_1^\sigma, \dots, p_d^\sigma)$ ,  $0 \leq d \leq 3$ , belongs to the weighted Delaunay triangulation  $D$  of  $B$  if and only if there exists a point  $p \in \mathbb{R}^3$  such that*

**DT1:**  $\pi(p, b_0^\sigma) = \pi(p, b_1^\sigma) = \dots = \pi(p, b_d^\sigma)$ , and

**DT2:**  $\pi(p, b_0^\sigma) \leq \pi(p, b_i)$  for  $b_i \in B - \sigma^d$ .

A point  $p$  that satisfies the above two conditions, DT1 and DT2, is called a *witness* for  $\sigma^d$ . We call a point that minimizes the distance  $\pi(p, b_0^\sigma)$  and satisfies both conditions as the *closest witness*, denoted by  $p_{\min}^\sigma$ . This minimum distance  $\pi(p_{\min}^\sigma, b_0^\sigma)$  is called the **Size** of the simplex  $\sigma^d$ . A point that minimizes the distance  $\pi(p, b_0^\sigma)$  and satisfies DT1 is called the *ortho-center*  $p_{\text{ortho}}^\sigma$  of simplex  $\sigma^d$ . The distance  $\pi(p_{\text{ortho}}^\sigma, b_0^\sigma)$  is called the **OrthoSize** of the simplex  $\sigma^d$ . Clearly, the **Size** of a simplex is lower bounded by its **OrthoSize**. Figure 2 shows the two possible scenarios, namely when **OrthoSize** = **Size** and **OrthoSize** < **Size**.

## 2.4 Alpha complex

Given a parameter  $\alpha \in \mathbb{R}$ , we can construct a subset of the weighted Delaunay triangulation by filtering simplices whose **Size** is less than or equal to  $\alpha$ , see Figures 1d and 1e. The resulting subset, called the *alpha complex*, is a subcomplex of the Delaunay complex and is denoted  $K_\alpha$ :

$$K_\alpha = \{\sigma^d \in D \text{ such that } \text{Size}(\sigma^d) \leq \alpha\}.$$

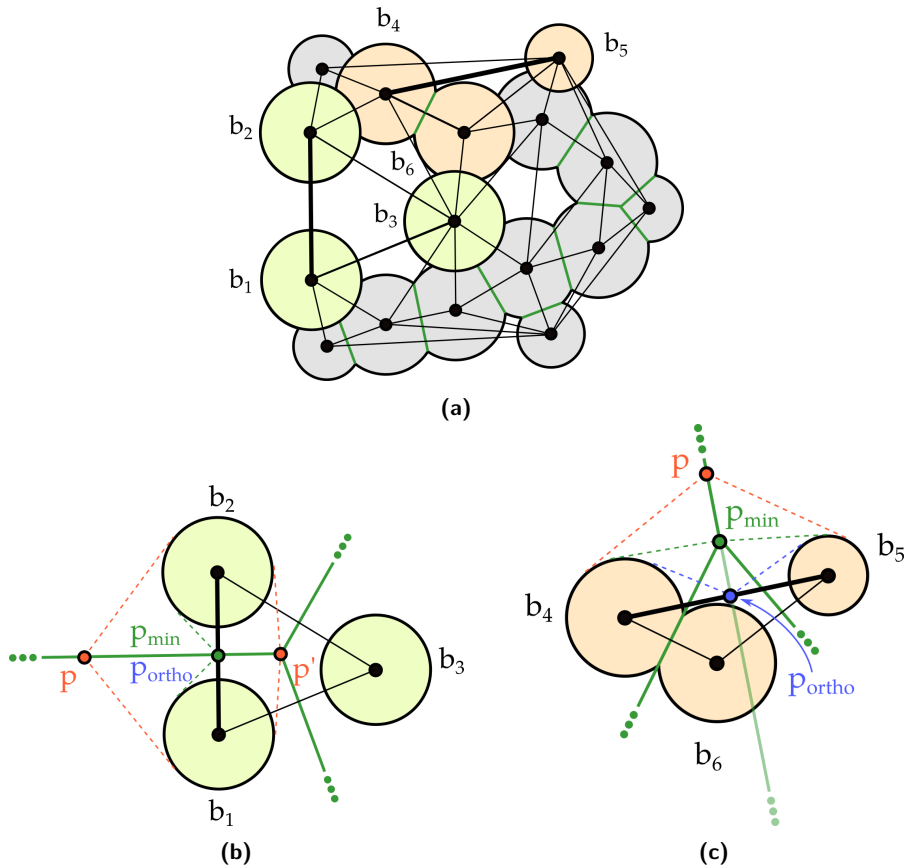
The following equivalent definition characterizes simplices of the alpha complex without explicitly referring to the Delaunay triangulation.

► **Definition 2** (Alpha complex). *A  $d$ -dimensional simplex  $\sigma^d = (p_0^\sigma, p_1^\sigma, \dots, p_d^\sigma)$ ,  $0 \leq d \leq 3$ , belongs to the alpha complex  $K_\alpha$  of  $B$  if and only if there exists a point  $p \in \mathbb{R}^3$  such that the following three conditions are satisfied:*

- AC1:**  $\pi(p, b_0^\sigma) = \pi(p, b_1^\sigma) = \dots = \pi(p, b_d^\sigma)$ ,
- AC2:**  $\pi(p, b_0^\sigma) \leq \pi(p, b_i)$  for  $b_i \in B - \sigma^d$ , and
- AC3:**  $\pi(p, b_0^\sigma) \leq \alpha$  or equivalently, the Size of  $\sigma^d$  is at most  $\alpha$ .

### 3 Algorithm

We now describe an algorithm to compute the alpha complex and prove its correctness. The algorithm utilizes the characterizing conditions introduced above. It first identifies the tetrahedra that belong to the alpha complex, followed by the set of triangles, edges and vertices. Figure 3 illustrates the algorithm as applied to disks on the plane.



■ **Figure 2** Size and OrthoSize of a simplex. (a) A set  $B$  of weighted points. Two edges (bold) belong to the Delaunay triangulation. (b) The Size of edge  $b_1b_2$  is equal to its OrthoSize. Points  $p$ ,  $p'$ ,  $p_{\min}$  and  $p_{\text{ortho}}$  are witnesses. Each one is equidistant from  $b_1$  and  $b_2$  and farther away from other disks in  $B$ . The distance is proportional to the length of the tangent to the disk that represents the weighted point. The next closest disk from these points is  $b_3$ . In this case,  $p_{\min}$  and  $p_{\text{ortho}}$  coincide and hence Size = OrthoSize. (c)  $b_4b_5$  is also a Delaunay edge. The location of a neighboring disk  $b_6$  could lead to a different configuration. The point  $p_{\text{ortho}}$  is closest to  $b_4$  and  $b_5$  among all the points that are equidistant from both. However  $p_{\text{ortho}}$  is closer to  $b_6$  as compared to  $b_4$  and  $b_5$ . The closest point  $p_{\min}$  that satisfies DT1 and DT2 is farther away, hence for  $b_4b_5$  Size is greater than OrthoSize.

### 3.1 Outline

The alpha complex of a point set in  $R^3$  consists of simplices of dimensions 0–3,  $K_\alpha = K_\alpha^0 \cup K_\alpha^1 \cup K_\alpha^2 \cup K_\alpha^3$ , where  $K_\alpha^d \subset K_\alpha$  is the set of  $d$ -dimensional simplices in  $K_\alpha$ . We initialize  $K_\alpha^d = \emptyset$  and construct  $K_\alpha$  in five steps described below:

**Step 1:** For  $0 \leq d \leq 3$ , compute the set of all simplices  $\sigma^d$  such that  $\text{OrthoSize}(\sigma^d) \leq \alpha$ . Let this set be denoted by  $\Sigma_{\text{ortho}} = \Sigma_{\text{ortho}}^0 \cup \Sigma_{\text{ortho}}^1 \cup \Sigma_{\text{ortho}}^2 \cup \Sigma_{\text{ortho}}^3$ .

**Step 2:** For all tetrahedra  $\sigma^3 \in \Sigma_{\text{ortho}}^3$ , check condition AC2 using  $p = p_{\text{ortho}}^\sigma$ . If  $\sigma^3$  satisfies AC2 then insert it into  $K_\alpha^3$ .

**Step 3:** Insert all triangles that are incident on tetrahedra in  $K_\alpha^3$  into  $K_\alpha^2$ . Let  $\Sigma_{\text{free}}^2 = \Sigma_{\text{ortho}}^2 - \text{Facets}(K_\alpha^3)$ , where  $\text{Facets}(K_\alpha^3)$  denotes the set of facets of tetrahedra in  $K_\alpha^3$ . For all triangles  $\sigma^2 \in \Sigma_{\text{free}}^2$ , check condition AC2 using  $p = p_{\text{ortho}}^\sigma$ . If  $\sigma^2$  satisfies AC2 then insert it into  $K_\alpha^2$ .

**Step 4:** Insert all edges incident on triangles in  $K_\alpha^2$  into  $K_\alpha^1$ . Let  $\Sigma_{\text{free}}^1 = \Sigma_{\text{ortho}}^1 - \text{Facets}(K_\alpha^2)$ , where  $\text{Facets}(K_\alpha^2)$  denotes the set of facets of triangles in  $K_\alpha^2$ . For all edges  $\sigma^1 \in \Sigma_{\text{free}}^1$ , check condition AC2 using  $p = p_{\text{ortho}}^\sigma$ . If  $\sigma^1$  satisfies AC2 then insert it into  $K_\alpha^1$ .

**Step 5:** Insert all endpoints of edges in  $K_\alpha^1$  into  $K_\alpha^0$ . Let  $\Sigma_{\text{free}}^0 = \Sigma_{\text{ortho}}^0 - \text{Facets}(K_\alpha^1)$ , where  $\text{Facets}(K_\alpha^1)$  denotes the set of balls incident on edges in  $K_\alpha^1$ . For all balls  $b_i = (p_i, r_i) \in \Sigma_{\text{free}}^0$ , check condition AC2 using  $p = p_i$ . If  $p_i$  satisfies AC2 then insert it into  $K_\alpha^0$ .

Step 1 selects simplices that satisfy AC3. Step 2 recognizes tetrahedra that belong to the alpha complex by checking AC2 using  $p = p_{\text{ortho}}^\sigma$ . Triangle faces of these tetrahedra also belong to  $K_\alpha$ . The other “free” triangles belong to  $K_\alpha^2$  if they satisfy AC2. Step 4 identify edges similarly. First all edge faces of triangles in  $K_\alpha^2$  are inserted followed by those “free” edges that satisfy AC2. Vertices are identified similarly in Step 5.

A notion related to free simplices, called *unattached* simplices, was introduced by Edelsbrunner [11]. However, the characterization of unattached simplices depends on the fact that they belong to the Delaunay complex.

### 3.2 Proof of correctness

We now prove that the algorithm described above correctly computes the alpha complex of the given set of weighted points by proving the following four claims. Each claim states that the set of simplices computed in Steps 2, 3, 4 and 5 are exactly the simplices belonging to the alpha complex. We assume that the input is non-degenerate.

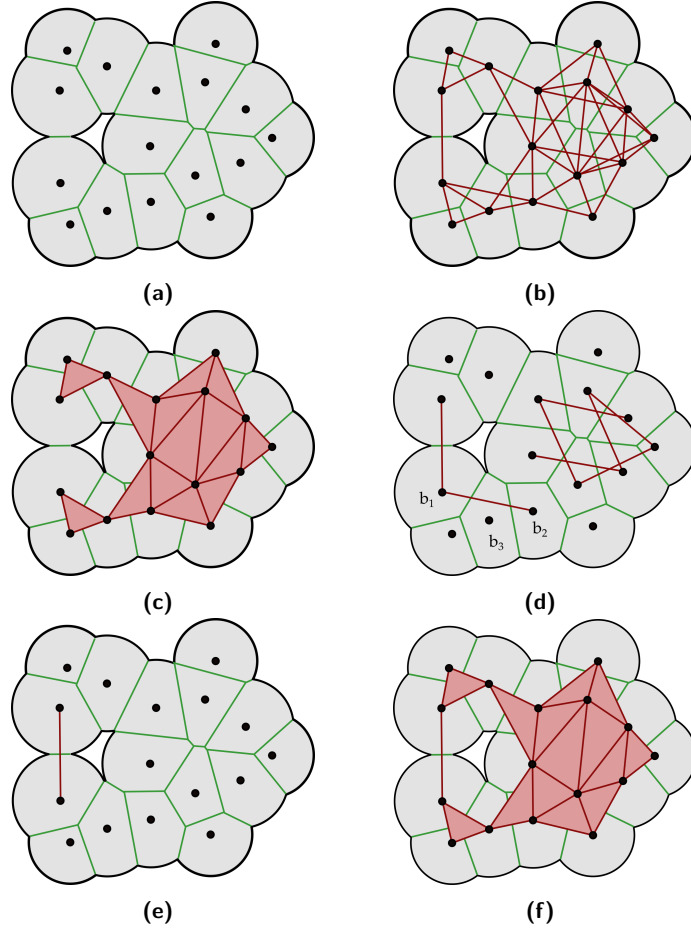
▷ **Claim 3. Step 2 computes  $K_\alpha^3$  correctly.**

**Proof.** For a tetrahedron  $\sigma^3$ ,  $p_{\text{ortho}}^\sigma$  is the only point that satisfies condition AC1. In Step 2 of the proposed algorithm, we check if AC2 holds for  $p_{\text{ortho}}^\sigma$ . If yes, then  $p_{\text{ortho}}^\sigma$  is a witness for  $\sigma^3$ , i.e.,  $p_{\text{ortho}}^\sigma = p_{\text{min}}^\sigma$ . Further, since  $\text{OrthoSize}(\sigma^3) \leq \alpha$  and  $p_{\text{ortho}}^\sigma = p_{\text{min}}^\sigma$ , we have  $\text{Size}(\sigma^3) \leq \alpha$  thereby satisfying AC3. Therefore,  $\sigma^3$  belongs to  $K_\alpha^3$  because it satisfies all three conditions. ◀

We now prove that the algorithm correctly identifies the triangles of the alpha complex.

► **Lemma 4.** *A triangle  $\sigma^2 \in \Sigma_{\text{free}}^2$  belongs to  $K_\alpha^2$  if and only if it satisfies AC2 with  $p = p_{\text{ortho}}^\sigma$ .*

**Proof.** We first prove the backward implication, namely if  $\sigma^2 \in \Sigma_{\text{free}}^2$  satisfies AC2 with  $p = p_{\text{ortho}}^\sigma$ , then  $\sigma^2 \in K_\alpha^2$ . Note that  $p_{\text{ortho}}^\sigma$  satisfies AC1 by definition. Further, it satisfies AC2 by assumption and hence  $\text{Size}(\sigma^2) = \text{OrthoSize}(\sigma^2)$ . We also have  $\text{OrthoSize}(\sigma^2) \leq \alpha$

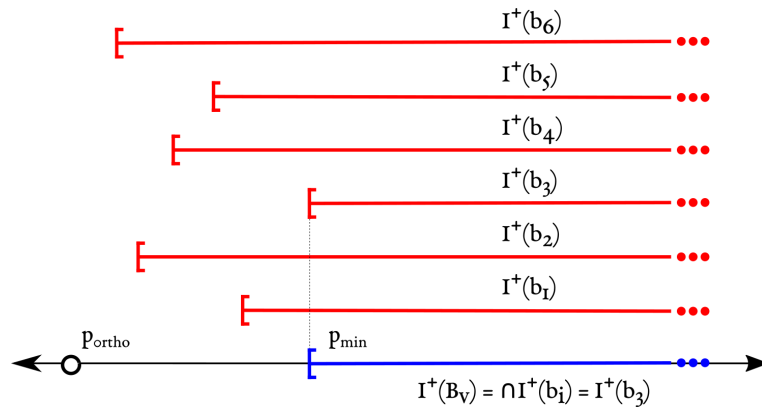


■ **Figure 3** Illustration of the proposed algorithm in 2D. (a) The set of disks  $B$  grown by the parameter  $\alpha$ . (b) First, compute the set of edges  $\Sigma_{\text{ortho}}^1$  whose  $\text{OrthoSize} \leq \alpha$  (red). The triangles  $\Sigma_{\text{ortho}}^2$  that satisfy this condition are also computed but they are not shown here. (c) Next, identify the triangles that satisfy AC2 (red). (d) Collect edges in  $\Sigma_{\text{ortho}}^1$  that are not incident on triangles in  $K_{\alpha}^2$  into  $\Sigma_{\text{free}}^1$ . Check if these edges satisfy AC2 with  $p = p_{\text{ortho}}^{\sigma}$ . For example, the edge  $b_1b_2$  does not satisfy this condition because  $b_3$  is closer to  $p_{\text{ortho}}^{\sigma}$  than  $b_1$  and  $b_2$ . (e) One edge survives the AC2 check and thus belongs to  $K_{\alpha}$ . (f) The alpha complex is obtained as the union of  $K_{\alpha}^2$ ,  $K_{\alpha}^1$  and  $K_{\alpha}^0$ .

because  $\sigma^1 \in \Sigma_{\text{free}}^2 \subseteq \Sigma_{\text{ortho}}^2$ . So,  $\text{Size}(\sigma^2) \leq \alpha$  thereby satisfying AC3. The triangle  $\sigma^2$  with  $p = p_{\text{ortho}}^{\sigma}$  satisfies all three conditions and hence belongs to  $K_{\alpha}^2$ .

We will now prove the forward implication via contradiction. Suppose there exists a triangle  $\sigma^2 \in \Sigma_{\text{free}}^2$  that belongs to  $K_{\alpha}^2$  but does not satisfy AC2 with  $p = p_{\text{ortho}}^{\sigma}$ . In other words, there exists a ball  $b_i \in B - \sigma^2$  for which  $\pi(p_{\text{ortho}}^{\sigma}, b_i) < \pi(p_{\text{ortho}}^{\sigma}, b_0^{\sigma})$ . Let  $B_v$  denote the set of all such balls  $b_i$ . The set of points that are equidistant from the three balls  $(b_0^{\sigma}, b_1^{\sigma}, b_2^{\sigma})$  corresponding to  $\sigma^2$  form a line perpendicular to the plane containing  $\sigma^2$  called the *radical axis*. Each ball  $b_i \in B_v$  partitions the radical axis into two half-intervals based on whether the point on radical axis is closer to  $b_i$  or to  $b_0^{\sigma}$ , see Figure 4. Let  $I^+(b_i)$  denote the half interval consisting of points that are closer to  $b_0^{\sigma}$  compared to  $b_i$ . Let  $I^+(B_v)$  denote the intersection of all such half intervals  $I^+(b_i)$ . We have assumed that  $\sigma^2 \in K_{\alpha}^2$ , so there must exist a closest witness  $p_{\text{min}}^{\sigma}$ , and it lies within  $I^+(B_v)$ . Thus,  $I^+(B_v)$  is non-empty. In fact,  $I^+(B_v) = I^+(b_j)$  for some  $b_j \in B_v$  and  $p_{\text{min}}^{\sigma}$  is exactly the end point of  $I^+(b_j)$ . This implies





■ **Figure 4** The radical axis of a triangle  $\sigma^2$  is drawn such that  $p_{ortho}^\sigma$  is at the origin. A ball  $b_i \in B - \sigma^2$  divides the radical axis into two half intervals. Points in the half interval  $I^+(b_i)$  are closer to  $b_i$  as compared to  $b_0$ , i.e. for all  $p \in I^+(b_i)$ ,  $\pi(p, b_i) < \pi(p, b_0)$ . Consider the set  $B_v$  of balls that are closer to  $p_{ortho}^\sigma$  as compared to  $b_0^\sigma$ . So,  $I^+(b_i)$  does not contain  $p_{ortho}^\sigma$ . The intersection of these intervals, denoted by  $I^+(B_v)$ , is equal to one of the intervals  $I^+(b_i)$ . Here,  $I^+(B_v) = I^+(b_3)$ . The end point of the interval  $I^+(b_3)$  is the closest witness for the tetrahedron  $\sigma^2 \cup b_3$ .

that  $p_{min}^\sigma$  is also a closest witness for the tetrahedron  $\sigma^3 = (b_0^\sigma, b_1^\sigma, b_2^\sigma, b_j)$ . So,  $\sigma^3$  belongs to  $K_\alpha^3$  and its Size is equal to  $\text{Size}(\sigma^2)$ . However, this means that  $\sigma^2 \notin \Sigma_{free}^2$ , a contradiction. So, the forward implication in the lemma is true. ◀

▷ **Claim 5. Step 3 computes  $K_\alpha^2$  correctly.**

**Proof.** If a simplex  $\sigma^3$  belongs to  $K_\alpha$  then naturally all of its faces also belong to  $K_\alpha$ . The algorithm includes such triangles into  $K_\alpha^2$  and remove them from  $\Sigma_{ortho}^2$  to obtain the set of free triangles  $\Sigma_{free}^2$ . It follows directly from Lemma 4 that AC2 is a necessary and sufficient condition for a triangle in  $\Sigma_{free}^2$  to belong to  $K_\alpha^2$ . Hence, Step 3 correctly computes the triangles belonging to  $K_\alpha^2$ . ◀

The above arguments need to be extended to prove that the edges of the alpha complex are also correctly identified.

▶ **Lemma 6.** An edge  $\sigma^1 \in \Sigma_{free}^1$  belongs to  $K_\alpha^1$  if and only if it satisfies the condition AC2 with  $p = p_{ortho}^\sigma$ .

The proof is similar to that of Lemma 4 and appears in the full version [29]. A general result is likely true for  $d$ -dimensional simplices in  $\Sigma_{free}^d$ . However, given the focus on alpha complexes in  $\mathbb{R}^3$ , we prefer to state and prove these results specific to lower dimensions. We also prefer to provide individual proofs for edges and triangles because it simplifies the exposition and could also potentially help in the design of improved data structures to accelerate computation of different steps of the algorithm.

▷ **Claim 7. Step 4 computes  $K_\alpha^1$  correctly.**

**Proof.** All edge faces of triangles in  $K_\alpha^2$  naturally belong to  $K_\alpha^1$ . Step 4 inserts all edges incident on triangles in  $K_\alpha^2$  into  $K_\alpha^1$  as valid edges and removes them from  $\Sigma_{ortho}^1$  to obtain the set of free edges  $\Sigma_{free}^1$ . It follows directly from Lemma 6 that AC2 is a necessary and sufficient condition for an edge  $\sigma^1 \in \Sigma_{free}^1$  to belong to  $K_\alpha^1$ . Therefore, Step 4 correctly computes the edges belonging to  $K_\alpha^1$ . ◀

▷ Claim 8. Step 5 computes  $K_\alpha^0$  correctly.

**Proof.** All vertices incident on  $K_\alpha^1$  naturally belong to  $K_\alpha^0$ . Step 5 inserts all such vertices in  $K_\alpha^0$  as valid vertices and removes them from  $\Sigma_{\text{ortho}}^0$  to obtain the set of free vertices  $\Sigma_{\text{free}}^0$ . Next, the vertices in  $\Sigma_{\text{free}}^0$  for which the center of the ball  $b_i = (p_i, r_i)$  satisfies AC2 are also inserted into  $K_\alpha^0$ . Clearly, these vertices also satisfy AC3 because they belong to  $\Sigma_{\text{ortho}}^0$ . The condition AC1 is not relevant for 0-dimensional simplices. Therefore, these vertices clearly belong to the alpha complex. Similar to Lemmas 4 and 6, it is easy to prove that checking for AC2 for  $p = p_i$  is necessary and sufficient condition to decide whether a vertex in  $\Sigma_{\text{free}}^0$  belongs to the alpha complex. That is, it is possible to show that vertices in alpha complex that have non-empty Voronoi regions but do not satisfy AC2 for  $p = p_i$  would be incident on some edge in  $K_\alpha^1$ , and therefore must have been already detected by Step 4 and hence can not belong to  $\Sigma_{\text{free}}^0$ . Therefore, Step 5 correctly computes the vertices belonging to  $K_\alpha^0$ . ◀

## 4 Parallel algorithm for biomolecules

Although the algorithm as described above is provably correct, a straightforward implementation will be extremely inefficient with a worst-case running time of  $O(n^5)$ , where  $n$  is the number of weighted points in  $B$ . This is because Step 1 requires  $O(n^4)$  time to generate all possible tetrahedra. In later steps, we need  $O(n)$  effort per simplex to check AC2. However, the input corresponds to atoms in a biomolecule. We show how certain properties of biomolecules can be leveraged to develop a fast parallel implementation.

### 4.1 Biomolecular data characteristics

Atoms in a biomolecule are well distributed. The following three properties of biomolecules are most relevant:

- The radius of an atom is bounded. The typical radius of an atom in a protein molecule ranges between 1Å to 2Å [3]. Further, a protein molecule contains upwards of thousand atoms. So, the radius is small compared to the total size of the molecule.
- There is a lower bound on the distance between the centres of two atoms. This is called the *van der Waals contact distance*, beyond which the two atoms start repelling each other. In the case of atoms in protein molecules, this distance is at least 1Å. This property together with the upper bound on atomic radii ensures that no atom is completely contained inside another. This means that the weighted Voronoi regions corresponding to the atoms in a biomolecule can be always be assumed to be non-empty.
- Structural biologists are interested in small values of  $\alpha$ . The two crucial values are 0Å and 1.4Å. The former corresponds to using van der Waals radius and the latter corresponds to the radius of water molecule, which acts as the solvent.

In the light of the above three properties, we can say that the number of simplices of the alpha complex that are incident on a weighted point (atom) is independent of the total number of input atoms and hence bounded by a constant [22].

### 4.2 Acceleration data structure

The algorithm will benefit from an efficient method for accessing points of  $B$  that belong to a local neighborhood of a given weighted point. We store the weighted points in a grid-based data structure. Let  $r_{max}$  denote the radius of the largest atom and assume that the value of the parameter  $\alpha$  is available as input. First, we construct a grid with cells of side length

$\sqrt{r_{max}^2 + \alpha}$  and then bin the input atoms into the grid cells. In our implementation, we do not store the grid explicitly because it may contain several empty cells. Instead, we compute the cell index for each input atom and sort the list of atoms by cell index to ensure that atoms that belong to a particular cell are stored at consecutive locations. The cell index is determined based on a row-major or column-major order. Alternatively, a space-filling curves like the Hilbert curve could also be used to order the cells.

After the atoms are stored in grid cells, the alpha complex is computed in two stages. In the first stage, we employ a bottom-up approach to obtain a conservative estimate of the edges, triangles, and tetrahedra belonging to the alpha complex. The false positives from the first stage are removed in a subsequent pruning stage resulting in the correct alpha complex. We describe these two stages in the following subsections.

### 4.3 Potential simplices

The first stage essentially corresponds to Step 1 of the algorithm described in the previous section. We compute the set  $\Sigma_{ortho}$  of potential simplices for which  $OrthoSize(\sigma^d) \leq \alpha$ . However, for efficiency reasons we process the simplices in the order of increasing dimension. First, we identify edges that satisfy the AC3 condition as described below. Given the size of the grid cell, endpoints of edges that satisfy the condition either lie within the same grid cell or in adjacent cells. So, the grid data structure substantially reduces the time required to compute the list of potential edges  $\Sigma_{ortho}^1$ . Beginning from this set of edges, we construct the set of all possible triangles and retain the triangles whose  $OrthoSize$  is no greater than  $\alpha$ , resulting in the set  $\Sigma_{ortho}^2$ . Finally, we use the triangles in  $\Sigma_{ortho}^2$  to construct the list of tetrahedra that satisfy the  $OrthoSize \leq \alpha$  condition. The above procedure works because the  $OrthoSize$  of a simplex is always greater than or equal to the  $OrthoSize$  of its faces. The set of simplices identified in this stage contains all simplices of the alpha complex. False positives are pruned in the second stage described below.

### 4.4 Pruning

The second stage corresponds to Steps 2-5 of the algorithm and processes the potential simplices in the decreasing order of dimension. This stage checks the characterizing condition AC2 to prune  $\Sigma_{ortho}$  into  $K_\alpha$ . The tetrahedra are processed by checking if any of the input balls are closer to the *ortho-center* than the balls incident on the tetrahedron. If yes, the tetrahedron is pruned away. Else, the tetrahedron is recognized as belonging to the alpha complex and inserted into  $K_\alpha^3$ . Triangles incident on these tetrahedra also belong to the alpha complex and are inserted into  $K_\alpha^2$  after they are removed from the list of potential triangles  $\Sigma_{ortho}^2$ . Next, the triangles in  $\Sigma_{ortho}^2$  are processed by checking if they satisfy AC2. If yes, they are inserted into  $K_\alpha^2$ . Otherwise, they are pruned away. All edges incident on triangles belonging to  $K_\alpha^2$  are inserted into  $K_\alpha^1$  and removed from the set  $\Sigma_{ortho}^1$ . Next, the edges in  $\Sigma_{ortho}^1$  are processed by checking if they satisfy AC2. Edges that satisfy AC2 are inserted into  $K_\alpha^1$  and the others are pruned away. All the vertices in  $\Sigma_{ortho}^0$  are directly inserted into  $K_\alpha^0$  without the AC2 check because for biomolecular data we assume that Voronoi regions of all the atoms are non-empty. The check for condition AC2 for each simplex is again made efficient by the use of the grid data structure. Atoms that may violate AC2 lie within the same cell as that containing the *ortho-center* or within the adjacent cells. Atoms that lie within other cells may be safely ignored.

## 4.5 CUDA implementation

We use the CUDA framework [7] and the `thrust` library [8] within CUDA to develop a parallel implementation of the algorithm that executes on the many cores of the GPU. The grid computation is implemented as a CUDA kernel where all atoms are processed in parallel. The computation of potential simplices and pruning stages are broken down into multiple CUDA kernels and parallelized differently in order to increase efficiency. We now describe the parallelization strategy in brief.

For computing the set of potential edges, the initial enumeration of possible edges incident on an atom is done using the atoms in the corresponding grid cell and its neighbouring cells. This is done per atom in parallel, the thread corresponding to the atom  $i$  being responsible for generating the edges  $ij, j > i$  to ensure no duplicate edges are generated. Subsequently, the AC3 condition is checked for the edges in parallel to finally generate the list of potential edges  $\Sigma_{\text{ortho}}^1$ . For computing potential triangles  $\Sigma_{\text{ortho}}^2$ , the potential edge list is used as a starting point for the initial enumeration of all possible triangles. This step is also parallelized per atom, the thread  $i$  being responsible for generation of triangles of the type  $ijk; j, k > i$  if all three edges  $ij, ik$  and  $jk$  are potential edges. The AC3 condition for the triangle is checked next within a separate kernel and parallelized per triangle to generate the potential triangles  $\Sigma_{\text{ortho}}^2$ . A similar strategy is used for computing the set of potential tetrahedra  $\Sigma_{\text{ortho}}^3$ .

The pruning stage is parallelized per tetrahedron, triangle, and edge as required. So, computation of  $p_{\text{ortho}}^\sigma$  is done in parallel. The grid data structure is again useful in checking for potential violators of the AC2 condition. Only the atoms belonging to the grid cell corresponding to  $p_{\text{ortho}}^\sigma$  or the those in neighbouring grid cells can violate the AC2 condition.

## 4.6 Handling large data sizes

Typical protein structures consist of up to 100,000 atoms. Our implementation can handle datasets of this size easily for reasonable values of  $\alpha$ . However, the size of datasets is ever increasing. Protein complexes that are available nowadays may consist of millions of atoms, necessitating smart management of GPU memory while handling such data sets.

We propose two strategies and implement one of them. The first strategy is to partition the grid by constructing an octree data structure and choosing an appropriate level in the octree to create partitions. Each partition together with its border cells can be processed independently of other partitions. So, we can copy one partition and its border to the GPU memory, compute its alpha complex, and copy the results back from GPU to CPU memory. After all the partitions are processed, the list of simplices can be concatenated followed by duplicate removal to generate the final alpha complex.

The second strategy is to partition the sorted list of atoms into *chunks* of equal sizes and to process each chunk independently. Here, we assume that the complete list of atoms together with the grid data structure fits in the GPU memory. This is a reasonable assumption considering that datasets containing several million atoms can easily fit on modern GPUs, which typically have at least 2GB video memory. Also, the main difficulty in handling large protein structures is managing the large lists of simplices generated within the intermediate steps of the algorithm, when compared to handling the input list of atoms or the output list of simplices. We compute the alpha complex by executing the algorithm in multiple passes. Each pass computes the alpha complex for a single chunk and copies it back to the CPU memory. We have implemented this second strategy and can handle data sizes of up to 16 million atoms on a GPU with 2GB of memory. Results are reported in the next section.

## 5 Experimental results

We now present results of computational experiments, which demonstrate that the parallel algorithm is fast in practice and significantly better than the state-of-the-art. We also performed runtime profiling to better understand the bottlenecks and effect of the parameter  $\alpha$  on the runtime. We present results for  $\alpha$  in the range 0.0 to 2.0. This range is important for structural analysis of biomolecules as it corresponds to *solvent accessible surface* of the biomolecule for typical solvent molecules like water (van der Waals radius = 1.4Å). The value  $\alpha = 0$  corresponds to the *van der Waals* surface of the biomolecule. All experiments, unless stated otherwise, were performed on a Linux system with an nVidia GTX 660 Ti graphics card running CUDA 8.0 and a 2.0GHz Intel Xeon octa core processor with 16 GB of main memory. The default number of threads per block was set at 512 for all the CUDA kernels.

Mach and Koehl describe two techniques for computing alpha complex of biomolecules called *AlphaVol* and *UnionBall* in their paper [27]. Both approaches construct the weighted Delaunay triangulation of input atoms first followed by a filtering step to obtain the alpha complex. *UnionBall* is the state-of-the-art technique for alpha complex computation for biomolecules on multi-core CPU. It uses heuristics and optimizations specific to biomolecular data to improve upon *AlphaVol*. For biomolecules containing 5 million atoms, *AlphaVol* takes approximately 8600 seconds for computing the alpha complex, while *UnionBall* takes approximately 150 seconds. Our method computes the alpha complex in less than 3 seconds for similar sized data, see Table 1.

### 5.1 Comparison with *gReg3D*

We are not aware of any available software that can compute the alpha complex directly without first constructing the complete Delaunay triangulation. In order to compare the performance, we chose the state-of-the-art parallel algorithm for computing the weighted Delaunay triangulation in 3D, *gReg3D* [5]. The CUDA implementation of *gReg3D* is available in the public domain. Table 1 compares the running times of our proposed algorithm with that of *gReg3D* for twelve different biomolecules at  $\alpha = 0$  and  $\alpha = 1$ . As evident from the table, we consistently observe significant speedup over *gReg3D*. The observed speedup is as high as a factor of 22 for the biomolecule 1X9P at  $\alpha = 0$ , one of the largest molecules in our dataset. Clearly, the speedup goes down for  $\alpha = 1$  when compared to  $\alpha = 0$  because of the increased number of simplices in the output alpha complex. We also report the number of simplices in the alpha complex compared to the total number of simplices in the Delaunay triangulation under the column ‘%Simplex’. This makes it clear why the speedup decreases as  $\alpha$  is increased from 0 to 1. For example, for the protein 1AON, the fraction of alpha complex simplices increases from 15.9% to 30% as  $\alpha$  is increased from 0 to 1. Correspondingly, the speedup decreases from a factor of 13.5 to 5.5. We repeated the experiment on a MS Windows system with an nVidia GTX 980 Ti card running CUDA 8.0 and observed similar speedups. However, the individual runtimes both for our algorithm and for *gReg3D* were higher on the GTX 980 Ti.

The starred entries in Table 1 are results for execution using the data partitioning approach. This is necessitated because these four large molecules generate large intermediate simplex lists that can not fit into the GPU memory if all the atoms in the molecule are processed at once. We observe that *gReg3D* is able to successfully compute the Delaunay complex for only one out of these four large molecules and runs out of GPU memory for the remaining three molecules.

We discuss the run time profiling, the effect of the value of  $\alpha$  and numerical issues in

■ **Table 1** Runtime comparison of the proposed algorithm with *gReg3D* on an nVidia GTX 660 Ti graphics card. Timings are reported in milliseconds. %Simplex refers to the size of the alpha complex as a percentage of the size of the weighted Delaunay triangulation. The last column shows the speedup in runtime of our algorithm over *gReg3D*. ‘\*’ indicates the data was partitioned and processed in chunks. ‘-’ indicates that the code could not execute due to insufficient memory.

$\alpha$	PDB id	#Atoms	$K_\alpha$		<i>gReg3D</i>		%Simplex	Speed up
			#Simplices	Time(ms)	#Simplices	Time(ms)		
0.0	1GRM	260	932	13	6295	117	14.8	<b>9.0</b>
	1U71	1505	5696	13	40878	115	13.9	<b>11.1</b>
	3N0H	1509	5739	14	41244	137	13.9	<b>10.0</b>
	4HHB	4384	38796	29	150141	193	25.8	<b>6.6</b>
	2J1N	8142	29642	18	227719	229	13.0	<b>12.7</b>
	1K4C	16068	62851	27	446383	347	14.1	<b>12.9</b>
	2OAU	16647	123175	56	466586	344	26.4	<b>6.2</b>
	1AON	58674	262244	65	1650841	879	15.9	<b>13.5</b>
	1X9P*	217920	924086	113	6142811	2555	15.0	<b>22.6</b>
	1IHM*	677040	2713083	277	-	-	-	-
	4CWU*	5905140	23450403	2709	-	-	-	-
3IYN*	5975700	24188892	2874	-	-	-	-	
1.0	1GRM	260	1598	15	6295	117	25.4	<b>7.9</b>
	1U71	1505	10828	17	40878	115	26.5	<b>8.5</b>
	3N0H	1509	10965	30	41244	137	26.6	<b>4.6</b>
	4HHB	4384	65987	86	150141	193	44.0	<b>2.2</b>
	2J1N	8142	58205	30	227719	229	25.6	<b>7.6</b>
	1K4C	16068	118467	52	446383	347	26.5	<b>6.7</b>
	2OAU	16647	199101	159	466586	344	42.7	<b>2.2</b>
	1AON	58674	495683	160	1650841	879	30.0	<b>5.5</b>
	1X9P*	217920	1653778	196	6142811	2555	26.9	<b>13.0</b>
	1IHM*	677040	5058507	605	-	-	-	-
	4CWU*	5905140	44411353	5118	-	-	-	-
3IYN*	5975700	45790463	5501	-	-	-	-	

more detail in the full version of the paper [29].

## 6 Conclusions

We proposed a novel parallel algorithm to compute the alpha complex for biomolecular data that does not require prior computation of the complete Delaunay triangulation. The useful characterization of simplices that belong to the alpha complex may be of independent interest. The algorithm was implemented using CUDA, which exploits the characteristics of the atom distribution in biomolecules to achieve speedups of up to a factor of 22 compared to the state-of-the-art parallel algorithm for computing the weighted Delaunay triangulation, and up to a factor of 50 speedup over the state-of-the-art implementation that is optimized for biomolecules. In future work, we plan to further improve the runtime efficiency of the parallel implementation and to resolve the numerical issues using real arithmetic.

Applications of alpha complex outside the domain of biomolecular analysis often require the complete filtration of Delaunay complex. The algorithm as presented here is not best suited for such cases. However, the algorithm may be modified to utilize a previously

computed alpha complex to efficiently compute the alpha complex for higher values of  $\alpha$ . We plan to investigate this extension in future work.

---

## References

- 1 Franz Aurenhammer, Rolf Klein, and Der-Tsai Lee. *Voronoi Diagrams and Delaunay Triangulations*. World Scientific, 2013. doi:10.1142/8685.
- 2 Ulrich Bauer and Herbert Edelsbrunner. The Morse theory of Čech and Delaunay filtrations. In Siu-Wing Cheng and Olivier Devillers, editors, *30th Annual Symposium on Computational Geometry, SOCG'14, Kyoto, Japan, June 08 - 11, 2014*, page 484. ACM, 2014. doi:10.1145/2582112.2582167.
- 3 A Bondi. van der Waals volumes and radii. *The Journal of Physical Chemistry*, 68(3):441–451, 1964.
- 4 Adrian Bowyer. Computing Dirichlet tessellations. *Comput. J.*, 24(2):162–166, 1981. doi:10.1093/comjnl/24.2.162.
- 5 Thanh-Tung Cao, Ashwin Nanjappa, Mingcen Gao, and Tiow Seng Tan. A GPU accelerated algorithm for 3d Delaunay triangulation. In John Keyser and Pedro V. Sander, editors, *Symposium on Interactive 3D Graphics and Games, I3D '14, San Francisco, CA, USA - March 14-16, 2014*, pages 47–54. ACM, 2014. doi:10.1145/2556700.2556710.
- 6 Paolo Cignoni, Claudio Montani, and Roberto Scopigno. DeWall: A fast divide and conquer Delaunay triangulation algorithm in  $E^d$ . *Comput. Aided Des.*, 30(5):333–341, 1998. doi:10.1016/S0010-4485(97)00082-1.
- 7 Nvidia Corporation. CUDA Zone. <https://developer.nvidia.com/cuda-zone>, 2020. [Online; accessed 17-March-2020].
- 8 Nvidia Corporation. Thrust. <https://developer.nvidia.com/thrust>, 2020. [Online; accessed 17-March-2020].
- 9 Tran Kai Frank Da, Sébastien Loriot, and Mariette Yvinec. 3D alpha shapes. In *CGAL User and Reference Manual*. CGAL Editorial Board, 4.11 edition, 2017. URL: <http://doc.cgal.org/4.11/Manual/packages.html#PkgAlphaShapes3Summary>.
- 10 Joe Dundas, Zheng Ouyang, Jeffery Tseng, T. Andrew Binkowski, Yaron Turpaz, and Jie Liang. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research*, 34(Web-Server-Issue):116–118, 2006. doi:10.1093/nar/gkl282.
- 11 H. Edelsbrunner. *Weighted alpha shapes*. University of Illinois at Urbana-Champaign, Department of Computer Science, 1992.
- 12 H. Edelsbrunner and P. Koehl. The geometry of biomolecular solvation. In *Discrete and Computational Geometry*, pages 243–275. MSRI Publications, 2005.
- 13 Herbert Edelsbrunner. *Geometry and Topology for Mesh Generation*, volume 7 of *Cambridge monographs on applied and computational mathematics*. Cambridge University Press, 2001.
- 14 Herbert Edelsbrunner. Alpha shapes – a survey. In R. van de Weygaert, G. Vegter, J. Ritzerveld, and V. Icke, editors, *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*, 2011.
- 15 Herbert Edelsbrunner and John Harer. *Computational Topology - an Introduction*. American Mathematical Society, 2010. URL: <http://www.ams.org/bookstore-getitem/item=MBK-69>.
- 16 Herbert Edelsbrunner, David G. Kirkpatrick, and Raimund Seidel. On the shape of a set of points in the plane. *IEEE Trans. Inf. Theory*, 29(4):551–558, 1983. doi:10.1109/TIT.1983.1056714.
- 17 Herbert Edelsbrunner and Ernst P. Mücke. Simulation of simplicity: a technique to cope with degenerate cases in geometric algorithms. *ACM Trans. Graph.*, 9(1):66–104, 1990. doi:10.1145/77635.77639.
- 18 Herbert Edelsbrunner and Ernst P. Mücke. Three-dimensional alpha shapes. *ACM Trans. Graph.*, 13(1):43–72, 1994. doi:10.1145/174462.156635.

- 19 Herbert Edelsbrunner and Raimund Seidel. Voronoi diagrams and arrangements. *Discret. Comput. Geom.*, 1:25–44, 1986. doi:10.1007/BF02187681.
- 20 Herbert Edelsbrunner and Nimish R. Shah. Incremental topological flipping works for regular triangulations. *Algorithmica*, 15(3):223–241, 1996. doi:10.1007/BF01975867.
- 21 Leonidas J. Guibas, Donald E. Knuth, and Micha Sharir. Randomized incremental construction of delaunay and voronoi diagrams. *Algorithmica*, 7(4):381–413, 1992. doi:10.1007/BF01758770.
- 22 Dan Halperin and Mark H. Overmars. Spheres, molecules, and hidden surface removal. *Comput. Geom.*, 11(2):83–102, 1998. doi:10.1016/S0925-7721(98)00023-6.
- 23 Michael Krone, Barbora Kozlíková, Norbert Lindow, Marc Baaden, Daniel Baum, Július Parulek, Hans-Christian Hege, and Ivan Viola. Visual analysis of biomolecular cavities: State of the art. *Comput. Graph. Forum*, 35(3):527–551, 2016. doi:10.1111/cgf.12928.
- 24 J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, and S. Subramaniam. Analytical shape computation of macromolecules: I. molecular area and volume through alpha shape. *Proteins Structure Function and Genetics*, 33(1):1–17, 1998.
- 25 J. Liang, H. Edelsbrunner, P. Fu, P.V. Sudhakar, and S. Subramaniam. Analytical shape computation of macromolecules: II. inaccessible cavities in proteins. *Proteins Structure Function and Genetics*, 33(1):18–29, 1998.
- 26 J. Liang, H. Edelsbrunner, and C. Woodward. Anatomy of protein pockets and cavities. *Protein Science*, 7(9):1884–1897, 1998.
- 27 Paul Mach and Patrice Koehl. Geometric measures of large biomolecules: Surface, volume, and pockets. *Journal of Computational Chemistry*, 32(14):3023–3038, 2011. doi:10.1002/jcc.21884.
- 28 Talha Bin Masood and Vijay Natarajan. An integrated geometric and topological approach to connecting cavities in biomolecules. In Chuck Hansen, Ivan Viola, and Xiaoru Yuan, editors, *2016 IEEE Pacific Visualization Symposium, PacificVis 2016, Taipei, Taiwan, April 19-22, 2016*, pages 104–111. IEEE Computer Society, 2016. doi:10.1109/PACIFICVIS.2016.7465257.
- 29 Talha Bin Masood, Tathagata Ray, and Vijay Natarajan. Parallel computation of alpha complex for biomolecules. *CoRR*, abs/1908.05944, 2019. URL: <http://arxiv.org/abs/1908.05944>, arXiv:1908.05944.
- 30 Talha Bin Masood, Sankaran Sandhya, Nagasuma R. Chandra, and Vijay Natarajan. CHEXVIS: a tool for molecular channel extraction and visualization. *BMC Bioinform.*, 16:119:1–119:19, 2015. doi:10.1186/s12859-015-0545-9.
- 31 Ashwin Nanjappa. *Delaunay triangulation in  $\mathbb{R}^3$  on the GPU*. PhD thesis, National University of Singapore, 2012.
- 32 Donald R. Sheehy. An output-sensitive algorithm for computing weighted  $\alpha$ -complexes. In *Proceedings of the 27th Canadian Conference on Computational Geometry, CCCG 2015, Kingston, Ontario, Canada, August 10-12, 2015*. Queen’s University, Ontario, Canada, 2015. URL: <http://research.cs.queensu.ca/cccg2015/CCCG15-papers/42.pdf>.
- 33 Raghavendra Sridharamurthy, Talha Bin Masood, Harish Doraiswamy, Siddharth Patel, Raghavan Varadarajan, and Vijay Natarajan. Extraction of robust voids and pockets in proteins. In Lars Linsen, Bernd Hamann, and Hans-Christian Hege, editors, *Visualization in Medicine and Life Sciences III, Mathematics and Visualization*, pages 329–349. Springer, 2016. doi:10.1007/978-3-319-24523-2\_15.
- 34 David F Watson. Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes. *The Computer Journal*, 24(2):167–172, 1981.