

# Computational topology and normal surfaces: Theoretical and experimental complexity bounds\*

Benjamin A. Burton<sup>†</sup>      João Paixão<sup>‡</sup>      Jonathan Spreer<sup>§</sup>

Author's self-archived version

Available from <http://www.maths.uq.edu.au/~bab/papers/>

## Abstract

In three-dimensional computational topology, the theory of normal surfaces is a tool of great theoretical and practical significance. Although this theory typically leads to exponential time algorithms, very little is known about how these algorithms perform in “typical” scenarios, or how far the best known theoretical bounds are from the real worst-case scenarios. Here we study the combinatorial and algebraic complexity of normal surfaces from both the theoretical and experimental viewpoints. Theoretically, we obtain new exponential lower bounds on the worst-case complexities in a variety of settings that are important for practical computation. Experimentally, we study the worst-case and average-case complexities over a comprehensive body of roughly three billion input triangulations. Many of our lower bounds are the first known exponential lower bounds in these settings, and experimental evidence suggests that many of our theoretical lower bounds on worst-case growth rates may indeed be asymptotically tight.

## 1 Introduction

In three-dimensional computational topology, many important problems are solved by exponential-time algorithms: key examples include Haken’s algorithm for recognising the unknot [13], or breaking down a triangulated 3-manifold into its prime decomposition [17, 18]. This is in contrast to two dimensions in which many problems are solved in polynomial time, or higher dimensions in which important topological problems can

become undecidable [10, 23].

A common feature of many such three-dimensional algorithms, and the source of both their solvability and their exponential running times, is their use of *normal surfaces*. In essence, normal surfaces are embedded 2-dimensional surfaces that intersect the surrounding 3-dimensional triangulation in a simple fashion. Most importantly, they describe topological features using combinatorial data, and are thereby well-suited for algorithmic enumeration and analysis.

Amongst the most important normal surfaces are the *vertex normal surfaces*. These correspond to the vertices of a high-dimensional polytope (called the *projective solution space*), and together they generate the space of all possible normal surfaces within the input triangulation. Many topological algorithms begin by enumerating all vertex normal surfaces in the input triangulation, and for many problems (such as unknot recognition and prime decomposition) this enumeration is in fact the main bottleneck for the entire algorithm.

One remarkable feature of many algorithms in three-dimensional computational topology is that, although they have extremely large theoretical worst-case complexity bounds, they appear to be *much* easier to solve in practice than these bounds suggest. For example:

- In 1980, Thurston asked if the Weber-Seifert dodecahedral space is Haken (the precise meaning of this is not important here) [1]. This long-standing question became a symbolic benchmark for computational topology, and was only resolved by computer proof after 30 years [9]. At the heart of the proof was an enumeration of all vertex normal surfaces in an  $n = 23$ -tetrahedron triangulation: despite a prohibitive  $O(16^n \times \text{poly}(n))$ -time enumeration algorithm (the best available at the time) and a best known bound of  $O(3.303^n)$  vertex normal surfaces [5], the enumeration ran in just  $5\frac{1}{2}$  hours with only 1751 vertex normal surfaces in total.

---

\*The first author is grateful to the Australian Research Council for their support under the Discovery Projects funding scheme (projects DP1094516 and DP110101104). This work was done while the second author was visiting The University of Queensland and he is thankful for the funding by FAPERJ for this visit. Computational resources used in this work were provided by the Queensland Cyber Infrastructure Foundation.

<sup>†</sup>School of Mathematics and Physics, The University of Queensland, Brisbane, Australia. [bab@maths.uq.edu.au](mailto:bab@maths.uq.edu.au)

<sup>‡</sup>Department of Mathematics, Pontifícia Universidade Católica, Rio de Janeiro, Brazil, [jpaixao@mat.puc-rio.br](mailto:jpaixao@mat.puc-rio.br)

<sup>§</sup>School of Mathematics and Physics, The University of Queensland, Brisbane, Australia. [j.spreer@uq.edu.au](mailto:j.spreer@uq.edu.au)

- The problem of unknot recognition is of particular interest. Modern derivatives of Haken’s original algorithm [18] have an exponential time complexity [15], but there is a growing discussion as to whether a faster algorithm might exist [11, 14]. Certainly unknot recognition lies in **NP** [15], and also **co-NP** if we assume the generalised Riemann hypothesis [20]; moreover, recent algorithmic developments based on linear programming now exhibit an experimental polynomial-time behaviour [6]. Deciding whether unknot recognition has a worst-case polynomial-time solution is now a major open problem in computational topology.

This severe gap between theory and practice is still poorly understood. There appear to be two causes: (i) the best theoretical complexity bounds are far from tight; (ii) “pathological” inputs that exhibit high-complexity behaviour are rare, with “typical” inputs often far easier to work with.

Proving such claims mathematically remains extremely elusive. Obtaining tight complexity bounds requires a deep interaction between topology, normal surfaces and polytope theory, and it is difficult to avoid making very loose estimates in at least one of these areas. Understanding “typical” behaviour (such as average- or generic-case complexity) is hampered by our very limited understanding of random 3-manifold triangulations: even the simple task of generating a random 3-manifold triangulation with  $n$  tetrahedra has no known sub-exponential-time solution [26]. In this setting, experimental work plays a crucial role in understanding the realistic performance of algorithms, as well as the innate difficulty of the problems that they aim to solve.

In this paper we focus our attention on the problem of enumerating all vertex normal surfaces within a given  $n$ -tetrahedron input triangulation: as mentioned earlier, this is a central component—and often the main bottleneck—of many algorithms in computational 3-manifold topology. Enumeration algorithms are still evolving [6, 8], and they are often hand-tailored to a particular topological problem of interest. For this reason we do not focus on the complexity of any specific algorithm, but instead we study two aspects of normal surface theory that affect and constrain all of these algorithms:

- *Combinatorial complexity:* We study the *total number* of vertex normal surfaces within the input triangulation  $\mathcal{T}$ , which we denote by  $\sigma(\mathcal{T})$ . This is our main quantity of interest. It yields an immediate *lower bound* for the time complexity of any enumeration algorithm, since it determines the out-

put size.<sup>1</sup> Moreover,  $\sigma(\mathcal{T})$  also factors into *upper bounds*, since modern enumeration algorithms are designed to run faster in situations where  $\sigma(\mathcal{T})$  is small [8].<sup>2</sup>

- *Algebraic complexity:* As detailed in Section 2, each normal surface is described by a non-negative integer vector in  $\mathbb{R}^{7n}$  (or in some settings,  $\mathbb{R}^{3n}$ ). We investigate the *maximum coordinate* of this vector of any vertex normal surface within the input triangulation  $\mathcal{T}$ , which we denote by  $\kappa(\mathcal{T})$ . This quantity is important for the *implementation* of enumeration algorithms, since it affects whether we can work with fast native machine integer types or whether we must fall back to significantly more expensive arbitrary-precision integer arithmetic [8]. Moreover,  $\kappa(\mathcal{T})$  features in algorithms that extend or even avoid the enumeration problem:

- Some algorithms, such as recognising small Seifert fibred spaces [27], require the complete enumeration of not just vertex normal surfaces but a much larger “lattice” of normal surfaces whose size is a function of  $\kappa(\mathcal{T})$ .
- Some algorithms, such as determining the crosscap number of a knot [7], avoid vertex enumeration entirely by solving an integer program instead; here the bounds on  $\kappa(\mathcal{T})$  feature as coefficients in the integer program, and directly affect whether the program can be solved using off-the-shelf integer programming software.

In summary, by focusing our attention on the quantities  $\sigma(\mathcal{T})$  and  $\kappa(\mathcal{T})$ , we learn not only about the behaviour of current enumeration algorithms, but also about the intrinsic limits and behaviour of the problem that they seek to solve.

We approach these combinatorial and algebraic quantities  $\sigma(\mathcal{T})$  and  $\kappa(\mathcal{T})$  through both theory and experiment. Theoretically, we construct infinite “pathological” families of triangulations in Section 3 that establish *exponential lower bounds* on the worst-case scenario for both  $\sigma(\mathcal{T})$  and  $\kappa(\mathcal{T})$ . Experimentally, we examine both the *worst case* and *average case* behaviour of these quantities in Section 4, using a comprehensive census of billions of input triangulations.

Such results are highly important for practitioners in three-dimensional computational topology, particularly given the exponential nature of many key algo-

<sup>1</sup>Specifically, since each vertex normal surface can be described in  $O(n)$  space [15], the output size is  $O(\sigma(\mathcal{T}) \times n)$ .

<sup>2</sup>The tree traversal enumeration algorithm (the current state of the art) has running time  $O(4^n \sigma(\mathcal{T}) \times \text{poly}(n))$ .

rithms. Despite this, just one preliminary study of this type appears in the literature [3]. This scarcity of results has two causes:

- (i) *The lack of large, comprehensive censuses of both “typical” and “atypical” triangulations.*

There are many censuses of 3-manifold triangulations in the literature, but these typically focus on well-structured triangulations with special properties (such as minimal triangulations, or irreducible manifolds). Such triangulations are often easy to work with [17], and offer little insight into an algorithm’s worst-case (or even average-case) behaviour.

It is only recently that large, comprehensive bodies of census data have been developed to study *all* triangulations of a given input size [3, 4]. By using such censuses for our experimental data, we ensure that we identify pathological cases, and also gain a clear understanding of how common or rare they are.

- (ii) *The intense computation required to study normal surfaces with such large bodies of data.*

Normal surface enumeration algorithms have enjoyed significant advances in recent years, and modern algorithms now run many orders of magnitude faster than their earlier counterparts [6, 8]. The experimental work in this paper required several years of combined CPU time, and without recent algorithmic advances [2, 4, 6] this work would not have been possible.

The preliminary study in [3] examines only the combinatorial complexity  $\sigma(\mathcal{T})$ , and works with a data set of roughly 150 million triangulations of closed 3-manifolds. The study in this paper is significantly richer, both in scope and detail:

- We examine the algebraic complexity  $\kappa(\mathcal{T})$  in addition to the combinatorial complexity  $\sigma(\mathcal{T})$ ;
- We work with a comprehensive data set of almost *three billion* triangulations, spanning both *closed* manifolds (which are important for algorithms such as prime decomposition) and *bounded* manifolds (which are important for knot algorithms);
- We also examine these quantities in “optimised” settings that arise in practical computation—in particular, *one vertex triangulations* (a common optimisation used in many topological algorithms), and the restricted problem of enumerating only *vertex normal discs* (which is important for unknot

recognition, or testing surfaces for incompressibility).

Our pathological families yield the first known explicit exponential lower bounds on worst-case complexity for the computationally important settings of bounded triangulations, closed 1-vertex triangulations, and normal discs. In many settings our pathological families match the experimental worst-case growth rates precisely, and we conjecture that the resulting exponential bounds are in fact exact.

Of related note is a result of Hass et al. [16], who establish an exponential lower bound on the worst-case complexity of a triangulated disc spanned by the unknot in  $\mathbb{R}^3$  (this has particular relevance for the complexity of unknot recognition). Their result operates under stricter geometric constraints, and it is not yet known how it translates to the more flexible setting of normal surfaces.

We emphasise again that our experimental data sets use exhaustive censuses of *all* possible input triangulations below a given size. This paper introduces the first such census of bounded 3-manifold triangulations in the literature, totalling over 20 billion triangulations of size  $n \leq 9$ .

We use exhaustive censuses because there is no known efficient algorithm for randomly sampling large triangulations [12, 26], and although there are other methods for generating random 3-manifolds [12, 22], nothing is known about the bias of the resulting sample of triangulations. As a result, although our *census* is large, the triangulations it contains are all relatively small. Nevertheless, there are strong reasons to believe that our experimental results are indicative of behaviour for larger inputs; we discuss this further in Section 5.

## 2 Preliminaries

By a *triangulation*  $\mathcal{T}$ , we mean a collection of  $n$  abstract tetrahedra  $\Delta_i = i(0123)$ ,  $1 \leq i \leq n$ , some or all of whose faces are affinely identified or “glued together” in pairs; here (0123) refers to the four vertices of tetrahedron  $\Delta_i$ . As a consequence of these face gluings, many tetrahedron edges may become identified together; we refer to the result as a single *edge of the triangulation*, and likewise with vertices. The gluings must be in a way that no edge is identified with itself in reverse as a result. Moreover, each tetrahedron face must be identified with one and only one partner (we call these *internal faces*), or with nothing at all (we call these *boundary faces*). The set of boundary faces is called the *boundary of the triangulation* and denoted by  $\partial\mathcal{T}$ . If  $\partial\mathcal{T} = \emptyset$  then  $\mathcal{T}$  is called a *closed triangulation*, otherwise it is said to be *bounded*. Not all triangulations (closed



Figure 1: Normal triangles and quadrilaterals within a tetrahedron.

or bounded) represent 3-manifolds. However, unless otherwise specified, this is true for all triangulations presented in this paper.

Throughout this article, the gluings of the triangles are given by a bijection of their vertices  $i(abc) \mapsto j(def)$  where the symbol  $i(abc)$  denotes triangle  $(abc)$  from tetrahedron  $i$  and the order of the vertices determines the gluing. A triangulation as defined above is sometimes referred to as *generalised triangulation*; these are more general and flexible than simplicial complexes. An important case is a *1-vertex triangulation*, in which all tetrahedron vertices become identified together.

The *face pairing graph* of a triangulation  $\mathcal{T}$  is the multigraph whose nodes represent tetrahedra, and whose arcs represent pairs of tetrahedron faces that are glued together. A face pairing graph may contain loops (if two faces of the same tetrahedron are glued together), and/or multiple edges (if two tetrahedra are joined together along more than one face). See Figure 2 for examples.

A *properly embedded surface* in  $\mathcal{T}$  is a surface  $\mathbf{s} \subseteq \mathcal{T}$  with no self-intersections, and whose boundary lies entirely within  $\partial\mathcal{T}$ . A *normal surface* in  $\mathcal{T}$  is a properly embedded surface that meets each tetrahedron  $\Delta$  of  $\mathcal{T}$  in a disjoint collection of triangles and quadrilaterals, each running between distinct edges of  $\Delta$ , as illustrated in Figure 1. There are four *triangle types* and three *quadrilateral types* in  $\Delta$  according to which edges they meet. Within each tetrahedron there may be several triangles or quadrilaterals of any given type; collectively these are referred to as *normal pieces*. The intersection of a normal piece of a tetrahedron with one of its faces is called *normal arc*; each face has three *arc types* according to which two edges of the face an arc meets.

Counting the number of pieces of each type for a normal surface  $\mathbf{s}$  gives rise to a 7-tuple per tetrahedron of  $\mathcal{T}$  and hence a  $7n$ -tuple of non-negative integers describing  $\mathbf{s}$  as a point in  $\mathbb{R}_{\geq 0}^{7n}$ , called its *normal coordinates*. Such a point must satisfy a set of linear homogeneous *matching equations* (one for each arc type of each internal face). These equations are necessary but not sufficient: the normal coordinates must also satisfy a set of combinatorial constraints called the *quadrilateral constraints*, which we discuss further in the journal version of this paper.

The solution set to the matching equations in  $\mathbb{R}_{\geq 0}^{7n}$

is a polyhedral cone (the cross-section polytope of this cone is also known as the *projective solution space*). A *vertex normal surface* is one whose normal coordinates lie on an extremal ray of this polyhedral cone and, in addition, its normal coordinates are minimal for all integer points on this ray. Thus, there are only finitely many such vertex normal surfaces; every normal surface can then be expressed as a positive rational linear combination of these surfaces just like every point in a polyhedral cone is a positive rational linear combination of points on its extremal rays. This is why, when enumerating normal surfaces in a triangulation, we typically just consider the finite set of vertex normal surfaces.

### 3 Theoretical lower bounds

Here we establish lower bounds for the worst-case values of  $\sigma(\mathcal{T})$  and  $\kappa(\mathcal{T})$ , i.e., the maximum possible  $\sigma(\mathcal{T})$  or  $\kappa(\mathcal{T})$  for an  $n$ -tetrahedron triangulation  $\mathcal{T}$ . Recall that  $\sigma(\mathcal{T})$  measures the combinatorial complexity (the number of vertex normal surfaces within  $\mathcal{T}$ ), and that  $\kappa(\mathcal{T})$  measures the algebraic complexity (the maximum coordinate of any vertex normal surface of  $\mathcal{T}$ ).

Few such results are known: there are no explicit lower bounds on the worst-case  $\kappa(\mathcal{T})$  in the literature, and the only explicit lower bound on the worst-case  $\sigma(\mathcal{T})$  is given by a family of closed triangulations with  $\sigma(\mathcal{T}) \in \Theta(17^{n/4}) \simeq \Theta(2.03^n)$  [3]. In this section, we give new exponential lower bounds for  $\sigma(\mathcal{T})$  and  $\kappa(\mathcal{T})$  in a variety of settings that hold particular relevance for key algorithms in computational geometry and topology. We sketch the main constructions and results here; see the journal version for detailed proofs.

#### 3.1 Closed triangulations with many normal surfaces

Important 3-manifold algorithms such as prime decomposition often begin by converting the input triangulation to a 1-vertex triangulation, whereupon the subsequent processing becomes significantly easier [17]. The  $\Theta(2.03^n)$  family of [3] is not of this type (each triangulation has  $n+1$  vertices), which raises the question of how such bounds behave in a 1-vertex setting:

**THEOREM 3.1.** *There is a family  $\mathcal{A}_n$ ,  $n \geq 1$ , of closed 1-vertex triangulations with  $n$  tetrahedra and  $\sigma(\mathcal{A}_n) = 2^n$  vertex normal surfaces.*

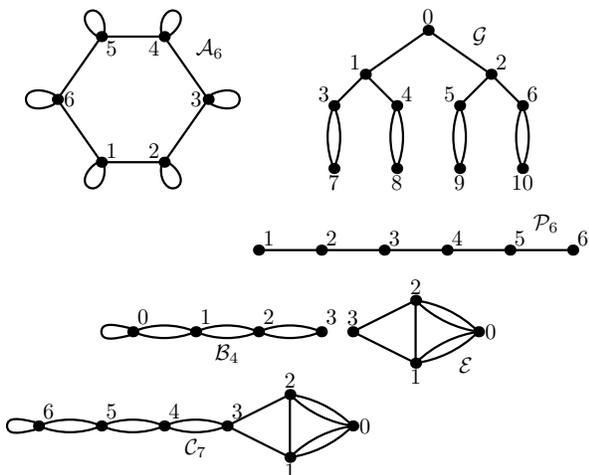


Figure 2: Face pairing graphs of the binomial family  $\mathcal{A}_6$ , triangulation  $\mathcal{G}$ , the path family  $\mathcal{P}_6$ , the bounded family  $\mathcal{B}_4$ , triangulation  $\mathcal{E}$ , and the closed family  $\mathcal{C}_7$ .

We call these triangulations *binomial triangulations*, because more precisely they have  $\binom{n}{k}$  vertex normal surfaces of genus  $k$  for each  $k = 0, \dots, n$  (whereby  $\sigma(\mathcal{A}_n) = \sum \binom{n}{k} = 2^n$ ). We construct each  $\mathcal{A}_n$  from  $n$  tetrahedra  $\Delta_1, \dots, \Delta_n$  in the following manner.

We begin by folding together two faces of the same tetrahedron  $\Delta_i = i(0123)$  by the gluing  $i(012) \mapsto i(013)$  for  $i = 1, \dots, n$ . Then we identify tetrahedra  $\Delta_i$  and  $\Delta_{i+1}$  by  $i(123) \mapsto (i+1)(230)$  (where  $\Delta_{n+1} = \Delta_1$ ). These gluings identify all the vertices to a single vertex, therefore  $\mathcal{A}_n$  is a 1-vertex triangulation. See Figure 2 for a picture of the face pairing graph of  $\mathcal{A}_n$  for  $n = 6$ . It can be shown that each  $\mathcal{A}_n$  is a closed 1-vertex triangulation of the 3-sphere (see the journal version of this article for the details).

To see why  $\sigma(\mathcal{A}_n) = 2^n$ , we observe that in each tetrahedron there are two normal “subsurfaces” which are compatible with any other normal surface of the triangulation. One of these adds genus to the overall surface and the other does not. We show that the vertex normal surfaces are precisely combinations of these subsurfaces, whereby the binomial coefficients and  $2^n$  growth rate easily follow.

REMARK. Experimentation suggests that this family  $\mathcal{A}_n$  might in fact yield a tight upper bound for closed 1-vertex triangulations; see Section 4 for details.

**3.2 Bounded triangulations with many normal surfaces** The number of vertex normal surfaces in a bounded triangulation has a direct impact on algorithms such as unknot recognition and incompressibility testing [18]. Here we give the first explicit exponential

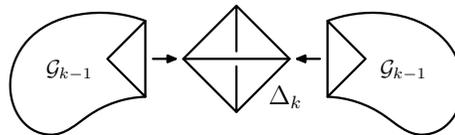


Figure 3: Attaching two copies of  $\mathcal{G}_{k-1}$  to the tetrahedron  $\Delta$ .

lower bound on the worst-case growth rate of this quantity. The proof is based on a general construction principle (Lemma 3.1) which, for an arbitrary bounded triangulation  $\mathcal{G}_0$  satisfying certain weak constraints, uses a recursive squaring argument to obtain a family of triangulations  $\{\mathcal{G}_k\}$  with  $\Omega(\beta^n)$  vertex normal surfaces, where the exponential base  $\beta$  is derived from  $\mathcal{G}_0$ . By choosing a suitable starting triangulation  $\mathcal{G}_0$ , we obtain the explicit base  $\beta \simeq 2.3715$  (Corollary 3.1).

For both unknot recognition and incompressibility testing, we can improve the underlying algorithms by only considering vertex normal *discs* (vertex normal surfaces that are topologically trivial). In Theorem 3.2 we show that this restricted quantity is also worst-case exponential: we build a family of triangulations with  $\Theta(2^n)$  vertex normal discs.

LEMMA 3.1. *Suppose  $\mathcal{G}_0$  is a bounded triangulation with  $n_0$  tetrahedra,  $f_0$  is a boundary face of  $\mathcal{G}_0$  such that not all vertices of  $f_0$  are identified in  $\mathcal{G}_0$ , and  $c_0$  is one of the three normal arc types on  $f_0$ . If there are  $\alpha_0$  vertex normal surfaces in  $\mathcal{G}_0$  that meet  $f_0$  in at least one arc of type  $c_0$  but in no other normal arc types, then  $\mathcal{G}_0$  can be extended to a family of triangulations  $\{\mathcal{G}_k\}$  in which the number of vertex normal surfaces grows at a rate of  $\Omega(\beta^n)$ , where  $\beta = \alpha_0^{1/(n_0+1)}$ .*

In the proof (for a detailed proof see the journal version), we recursively construct  $\mathcal{G}_k$  by joining two copies of  $\mathcal{G}_{k-1}$  to an additional tetrahedron  $\Delta_k$  along their faces  $f_{k-1}$  (see Figure 3). Suppose there are  $\alpha_{k-1}$  vertex normal surfaces in  $\mathcal{G}_{k-1}$  that meet  $f_{k-1}$  in only arcs of type  $c_{k-1}$ . For each pair of such surfaces in the two copies of  $\mathcal{G}_{k-1}$ , we can combine these surfaces in a way that extends through  $\Delta_k$  to meet one of the free faces  $f_k$  of  $\Delta_k$  in just one chosen normal arc type  $c_k$ , and this extension yields a vertex normal surface of  $\mathcal{G}_k$ . There are  $\alpha_{k-1}^2$  such pairings, and therefore  $\alpha_k \geq \alpha_{k-1}^2$  such vertex normal surfaces of  $\mathcal{G}_k$ . This recurrence yields the final growth rate of  $\Omega(\beta^n)$  where  $\beta = \alpha_0^{1/(n_0+1)}$ . We need the assumption that not all vertices of the boundary face  $f_0$  are identified in  $\mathcal{G}_0$  to show that each triangulation  $\mathcal{G}_k$  represents a bounded 3-manifold.

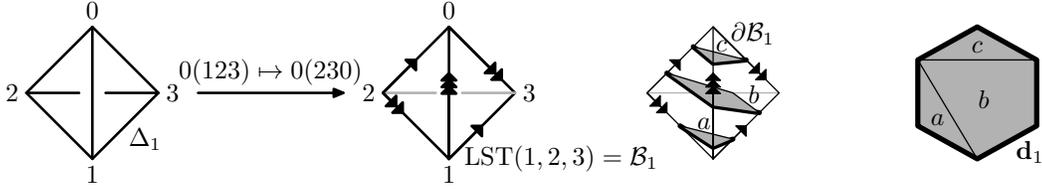


Figure 4: Left: construction of the 1-tetrahedron triangulation of  $\text{LST}(1, 2, 3)$ . Right: the meridian disc  $\mathbf{d}_1$  which is a 6-gon with 1 normal quad, 2 normal triangles and a maximum of 3 edge intersections.

**COROLLARY 3.1.** *There is a triangulation  $\mathcal{G} = \mathcal{G}_0$  that is the starting point of a family of bounded triangulations  $\{\mathcal{G}_k\}$ ,  $k \geq 0$ , with  $n_k = 12 \cdot 2^k - 1$  tetrahedra and  $\sigma(\mathcal{G}_k) \geq 2.3715^{n_k}$ .*

We prove this using Lemma 3.1 by choosing a starting triangulation  $\mathcal{G} = \mathcal{G}_0$  with  $n_0 = 11$  tetrahedra and a choice of face  $f_0$  and arc type  $c_0$  with  $\alpha_0 = 31\,643$  corresponding vertex normal surfaces. This yields a growth rate of  $\Omega(\beta^n)$  with  $\beta = 31\,643^{1/12} \simeq 2.3715$ . The face pairing graph of  $\mathcal{G}$  is shown in Figure 2; for a detailed construction see the journal version.

For our final result, we construct the *path triangulation*  $\mathcal{P}_n$  from  $n$  tetrahedra  $\Delta_i = i(0123)$ ,  $i = 1, \dots, n$ , by joining tetrahedra  $\Delta_i$  and  $\Delta_{i+1}$  by the map  $i(012) \mapsto (i+1)(013)$ . It can be shown that each  $\mathcal{P}_n$  is a bounded triangulation whose underlying 3-manifold is the 3-ball (for details see the journal version).

**THEOREM 3.2.** *For each  $n \geq 1$ ,  $\mathcal{P}_n$  has  $2^{n+1} + \frac{(n+1)(n+2)}{2} \in \Theta(2^n)$  vertex normal discs.*

We prove this by obtaining explicit recurrences for the number of vertex normal surfaces with different normal arcs based on the matching equations. We have essentially two choices for each normal arc, and either a triangle or quadrilateral can be added to the previous triangulation giving the  $\Theta(2^n)$  growth rate. Again, see the journal version for details.

**3.3 Lower bounds for the size of normal coordinates** Here we give exponential lower bounds on the worst-case algebraic complexity  $\kappa(\mathcal{T})$ . Our bounds follow a Fibonacci growth rate of  $\Omega([(1 + \sqrt{5})/2]^n) \simeq \Omega(1.618)^n$ . Understanding the worst-case  $\kappa(\mathcal{T})$  is important for improving the time and space complexity of normal surface enumeration algorithms due to a better handling of the integer arithmetic involved (see Section 1).

To obtain such lower bounds, we first construct a family of bounded triangulations, each containing a vertex normal disc with coordinates growing exponentially

in the number of tetrahedra. We then close the bounded family using a constant number of additional tetrahedra so that the vertex normal surface with exponential coordinates is preserved. In this way, we are able to construct two families of triangulations, bounded and closed, with Fibonacci type growth rates for  $\kappa(\mathcal{T})$ .

The key objects of the construction are so-called *layered solid tori* [17]. These are parameterised triangulations of the solid torus: the layered solid torus denoted  $\text{LST}(a, b, a + b)$  has as its boundary a triangulation of the torus with exactly three boundary edges, such that the *meridian disc* (the unique disc of the solid torus meeting the boundary in a non-contractible closed curve) intersects the boundary edges in  $a$ ,  $b$  and  $a + b$  points.

Layered solid tori are very common tools when constructing triangulations of a given type of 3-manifold (see [19, 21, 25] for more about constructing 3-manifolds). The most prominent example of a layered solid torus, the one tetrahedron triangulation of  $\text{LST}(1, 2, 3)$ , is shown in Figure 4.

**THEOREM 3.3.** *There is a family  $\mathcal{B}_n$  of bounded 1-vertex triangulations with  $n$  tetrahedra, where  $\mathcal{B}_n$  contains a vertex normal disc  $\mathbf{d}_n$  with maximum coordinate  $F_{n+1}$ , where  $F_k$  denotes the  $k$ -th Fibonacci number.*

The family  $\mathcal{B}_n$  consists of layered solid tori of type  $\text{LST}(F_{n+1}, F_{n+2}, F_{n+3})$  and each vertex normal surface  $\mathbf{d}_n$  is the corresponding meridian disc. See the journal version for details of the proof.

The key idea for the construction of a closed family  $\mathcal{C}_n$  of  $n$ -tetrahedron triangulations containing a vertex normal surface with exponentially growing coordinates is to find a small  $m$ -tetrahedron triangulation  $\mathcal{E}$  with the same boundary as the family of layered solid tori  $\mathcal{B}_n$  acting like a type of plug. By this we mean that  $\mathcal{E}$  contains a normal surface intersecting  $\partial\mathcal{E}$  in the same way as  $\mathbf{d}_n$  intersects  $\partial\mathcal{B}_n$  for all  $n \geq 1$ . This gives rise to a vertex normal surface in  $\mathcal{C}_n$  with maximum coordinate greater than or equal to  $F_{n+1}$ . Since  $m$  is constant, this gives the same asymptotic lower bound on  $\kappa(\mathcal{T})$  for closed triangulations as for bounded triangulations.

Input size $n$	Closed	Closed and 1-vertex	Bounded
1	4	3	3
2	17	12	17
3	81	63	156
4	577	433	2 308
5	5 184	3 961	45 046
6	57 753	43 584	995 920
7	722 765	538 409	25 225 447
8	9 787 509	7 148 483	695 134 018
9	139 103 032	99 450 500	19 933 661 871
10	2 046 869 999	1 430 396 979	
Total	$n \leq 10$ : 2 196 546 921	$n \leq 10$ : 1 537 582 427	$n \leq 8$ : 721 402 915 $n \leq 9$ : 20 655 064 786

Table 1: The number of 3-manifold triangulations in the census

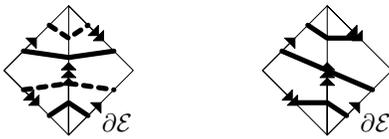


Figure 5: Intersection of  $\mathbf{s}$  (left) and  $\mathbf{t}$  (right) with  $\partial\mathcal{E}$

REMARK. [The triangulation  $\mathcal{E}$ ] There is a 4-tetrahedron bounded triangulation  $\mathcal{E}$  with  $\partial\mathcal{E} = \partial\mathcal{B}_m$  for all  $m$  which contains two vertex normal surfaces  $\mathbf{s}$  and  $\mathbf{t}$  that can be combined into a normal surface intersecting  $\partial\mathcal{E}$  in the same pattern as  $\mathbf{d}_m$ . The face pairing graph of  $\mathcal{E}$  is shown in Figure 2, and the intersection of  $\mathbf{s}$  and  $\mathbf{t}$  with its boundary  $\partial\mathcal{E}$  is shown in Figure 5. A detailed description of  $\mathcal{E}$  can be found in the journal version.

THEOREM 3.4. *There is a family  $\mathcal{C}_n$  of closed 1-vertex triangulations with  $n$  tetrahedra,  $n \geq 5$ , each containing a vertex normal surface with maximum coordinate at least  $F_{n-3}$  if  $n \equiv 2 \pmod{3}$  or at least  $2F_{n-3}$  otherwise.*

As outlined above, we construct  $\mathcal{C}_n$  by gluing  $\mathcal{B}_{n-4}$  and  $\mathcal{E}$  along their boundary tori. If  $n \equiv 2 \pmod{3}$ , the meridian disc  $\mathbf{d}_{n-4}$  glued to a combination of  $\mathbf{s}$  and  $\mathbf{t}$  yields a vertex normal projective plane or, if  $n \equiv 0, 1 \pmod{3}$ , twice  $\mathbf{d}_{n-4}$  with a combination of  $\mathbf{s}$  and  $\mathbf{t}$  yields a vertex normal sphere; the maximum coordinates are then as stated. See the journal version for details.

We note that the vertex normal sphere from above (in the case  $n \equiv 0, 1 \pmod{3}$ ) is the only non-vertex linking normal sphere in  $\mathcal{C}_n$ . Detecting these normal surface types is one of the key tasks in important 3-manifold problems such as prime decomposition. Hence, the family  $\mathcal{C}_n$  is an example for a case where, in order to prove the existence of such a normal sphere, dealing

with exponentially large normal coordinates cannot be avoided. This is a hint towards the conjecture that these problems are intrinsically hard to solve using normal surface enumeration methods.

#### 4 Experimental behaviour

We turn now to an experimental study of the combinatorial and algebraic complexities of vertex normal surfaces. Our experimental data consists of *all closed 3-manifold triangulations* with  $n \leq 10$  tetrahedra, and *all bounded 3-manifold triangulations* with  $n \leq 8$  tetrahedra (each appearing precisely once up to relabelling). As shown in Table 1, this yields almost 3 billion triangulations in total ( $2\,196\,546\,921 + 721\,402\,915 = 2\,917\,949\,836$ ). We also extract the  $\sim 1.5$  billion *1-vertex* triangulations from the closed census for additional study.

Generating exhaustive censuses of all possible inputs requires sophisticated algorithms and significant computational resources. The  $n \leq 10$ -tetrahedron census of all closed triangulations first appeared in [4] (which also describes some of the underlying algorithms). The  $n \leq 8$ -tetrahedron census of all bounded triangulations is new to this paper. Moreover, we have constructed this bounded census for  $n = 9$ , with over 20 billion triangulations; however, we only use  $n \leq 8$  for our experiments because the subsequent analysis of normal surfaces for  $n = 9$  remains out of our computational reach (for  $n = 8$  this analysis already consumed years of CPU time).

Table 2 summarises our experimental results, and gives worst-case and average-case measurements (labelled *Max* and *Avg* respectively) for the quantities  $\sigma(\mathcal{T})$  and  $\kappa(\mathcal{T})$  in our various settings. Each measurement is taken over the relevant census of triangulations from Table 1. The *Closed* and *Bounded* columns refer to all closed or bounded triangulations respectively;

Input size $n$	Combinatorial complexity $\sigma(\mathcal{T})$								Algebraic complexity $\kappa(\mathcal{T})$			
	Closed		Closed 1-vertex		Bounded		Bounded, discs		Closed		Bounded	
	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Avg	Max	Avg
1	3	2.0	2	1.7	7	5.0	7	4.0	1	1.0	1	1.0
2	7	3.9	4	3.3	14	8.2	14	5.2	2	1.2	2	1.3
3	11	5.5	8	4.9	35	14.0	27	7.0	3	1.5	3	1.7
4	18	8.8	16	7.8	85	31.3	69	11.6	4	1.8	6	2.5
5	36	13.3	32	12.0	236	69.5	176	20.2	7	2.1	12	3.4
6	70	20.8	64	18.6	688	152.6	440	34.8	10	2.3	20	4.3
7	144	32.2	128	28.8	1943	376.6	1109	61.7	16	2.6	36	5.8
8	291	50.2	256	44.7	5725	947.4	2768	112.4	26	2.9	65	7.5
9	584	78.5	512	69.4					42	3.2		
10	1175	123.2	1024	108.2					68	3.6		
Growth	$2.03^n$	$1.56^n$	$2^n$	$1.56^n$	$2.73^n$	$2.23^n$	$2.45^n$	$1.69^n$	$1.62^n$	$1.14^n$	$1.81^n$	$1.34^n$

Table 2: Experimental worst-case and average-case results

	Combinatorial complexity $\sigma(\mathcal{T})$			Algebraic complexity $\kappa(\mathcal{T})$		
	Lower bound	Experimental growth	Upper bound	Lower bound	Experimental growth	Upper bound
Closed	$\Omega(2.03^n)$	$\leftarrow$	$O(14.556^n)$	$\Omega(1.62^n)$	$\leftarrow$	$O(3.17^n)$
Closed 1-vertex	$\Omega(2^n)$	$\leftarrow$	$O(4.852^n)$	$\Omega(1.62^n)$	$\leftarrow$	$O(3.17^n)$
Bounded	$\Omega(2.37^n)$	$\simeq 2.73^n$	$O(64^n)$	$\Omega(1.62^n)$	$\simeq 1.82^n$	$O(31.63^n)$
Bounded, discs only	$\Omega(2^n)$	$\simeq 2.45^n$	$O(64^n)$	$\Omega(1.62^n)$	$\leftarrow$	$O(31.63^n)$

Table 3: Summary of worst-case theoretical and experimental results

in the *Closed 1-vertex* column we restrict our attention to 1-vertex triangulations of closed manifolds, and in the *Bounded, discs* column we only count vertex normal discs (not all vertex normal surfaces). We have also measured the algebraic complexity  $\kappa(\mathcal{T})$  in the 1-vertex and discs-only settings, but we omit the details due to space constraints; see the journal version of this paper for the details.

The final row of Table 2 gives a “best estimate” of the exponential growth rate of each quantity with respect to  $n$  (we just list the base of the exponential, ignoring any coefficients or polynomial factors). Most growth rates are estimated by linear regression<sup>3</sup>, though for cases where the worst cases matches a known family of triangulations (see below) we give the corresponding known rate.

We can make some broad observations from Table 2:

- The average-case scenarios grow at a significantly slower rate than the worst-case scenarios, sometimes astonishingly so. This is consistent with past observations in which “typical” triangulations exhibit significantly smaller complexity properties than expected (see Section 1).

- For closed manifolds, 1-vertex triangulations only give a very slight improvement: the worst case drops from the  $\Theta(17^{n/4}) \simeq \Theta(2.03^n)$  family described in [3] down to the  $\Theta(2^n)$  family of Theorem 3.1. The closeness of these results is surprising, since the theoretical bounds on  $\sigma(\mathcal{T})$  for 1-vertex triangulations are much smaller than the general bounds (see below), and algorithms for working with them are often much simpler [17].

- Bounded triangulations exhibit higher complexity properties than their closed counterparts. For the combinatorial complexity  $\sigma(\mathcal{T})$  this discrepancy is very pronounced—even the average case for bounded triangulations is well above the worst case for closed triangulations. This is again consistent with past experiences in working with normal surface algorithms [9]. Restricting our attention to normal discs (e.g., for unknot recognition) does alleviate this problem somewhat.

Table 3 compares the experimental behaviour of the worst-case  $\sigma(\mathcal{T})$  against its best known theoretical lower and upper bounds (here “lower bounds” refers to families of pathological triangulations with the highest known growth rate of  $\sigma(\mathcal{T})$ , such as those constructed in Section 3).

Regarding  $\sigma(\mathcal{T})$ : the lower bound of  $\Theta(17^{n/4}) \simeq \Theta(2.03^n)$  is known from [3], and the remaining three

<sup>3</sup>Specifically, we take a weighted linear regression of  $\log \sigma$  or  $\log \kappa$  as a function of  $n$ . The weights are taken to be  $1, \dots, n$ , in order to limit the influence of anomalous small cases.

lower bounds are new to this paper (Theorem 3.1, Corollary 3.1 and Theorem 3.2). The first two upper bounds are taken from [5]; the final two  $O(64^n)$  bounds are well known but do not appear in the literature (for a proof we refer to the journal version). Regarding  $\kappa(\mathcal{T})$ : all four lower bounds of  $\Theta([(1 + \sqrt{5})/2]^n) \simeq \Theta(1.62^n)$  are new to this paper (Theorems 3.3 and 3.4), and all four upper bounds are taken from [8].

Here we see that the experimental growth rates are much closer to the lower bounds than the upper bounds; in particular, an arrow ( $\leftarrow$ ) indicates that the experimental worst-case growth rate is *identical* to the best lower bound (up to a constant factor). This invites the following conjectures:

**CONJECTURE 4.1.** *For closed triangulations, the maximum number of vertex normal surfaces  $\sigma(\mathcal{T})$  for any given  $n$  grows at an asymptotic rate of  $\Theta(17^{n/4}) \simeq \Theta(2.03^n)$ , and for closed 1-vertex triangulations this reduces to  $\Theta(2^n)$ .*

*For closed triangulations as well as closed 1-vertex triangulations, the maximum coordinate of any vertex normal surface  $\kappa(\mathcal{T})$  for any given  $n$  grows at an asymptotic rate of  $\Theta([(1 + \sqrt{5})/2]^n) \simeq \Theta(1.62^n)$ . For bounded triangulations, the maximum coordinate of any vertex normal disc for any given  $n$  likewise grows at an asymptotic rate of  $\Theta([(1 + \sqrt{5})/2]^n) \simeq \Theta(1.62^n)$ .*

For closed 1-vertex triangulations, our experimental data gives an even stronger result:

**THEOREM 4.1.** *For closed 1-vertex triangulations, the maximum number of vertex normal surfaces  $\sigma(\mathcal{T})$  for any given  $n \leq 10$  is precisely  $2^n$ , and is attained by the binomial triangulations  $\mathcal{A}_n$  as described in Section 3.1.*

**CONJECTURE 4.2.** *Theorem 4.1 is true for all positive integers  $n$ .*

## 5 Discussion

As noted in the introduction, although we consider close to 3 billion distinct triangulations, they are all relatively small with less than or equal to 10 tetrahedra. Despite this, there are reasons to believe that our experimental results might generalise. Because we allow more flexible triangulations (not just simplicial complexes), this census contains a rich diversity of 3-manifolds (including 5114 distinct closed  $\mathbb{P}^2$ -irreducible manifolds [4, 24]). Moreover, several of the patterns that we see in Table 2 (in particular Theorem 4.1) are established early on and are remarkably consistent.

As an exception, in the case of arbitrarily bounded manifolds, a greater number of tetrahedra seems to allow a better choice for the starting triangulation  $\mathcal{G}_0$

in Lemma 3.1 in order to obtain a higher exponential base.

As seen in Table 3, there is still a long way to go before lower bounds and upper bounds on worst-case complexities converge. Not only does this paper produce the first explicit lower bounds in several computationally important settings, but it also gives strong experimental evidence that these lower bounds are close to (or even exactly) tight. This suggests that it is now the upper bounds that require significant improvement, inviting new directions of research with a rich interplay between topology, polytopes and complexity theory.

## References

- [1] Joan S. Birman, *Problem list: Nonsufficiently large 3-manifolds*, Notices Amer. Math. Soc. **27** (1980), no. 4, 349.
- [2] Benjamin A. Burton, *Converting between quadrilateral and standard solution sets in normal surface theory*, Algebr. Geom. Topol. **9** (2009), no. 4, 2121–2174.
- [3] ———, *The complexity of the normal surface solution space*, SCG '10: Proceedings of the Twenty-Sixth Annual Symposium on Computational Geometry, ACM, 2010, pp. 201–209.
- [4] ———, *Detecting genus in vertex links for the fast enumeration of 3-manifold triangulations*, ISSAC 2011: Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation, ACM, 2011, pp. 59–66.
- [5] ———, *Maximal admissible faces and asymptotic bounds for the normal surface solution space*, J. Combin. Theory Ser. A **118** (2011), no. 4, 1410–1435.
- [6] ———, *Computational topology with Regina: Algorithms, heuristics and implementations*, To appear in Geometry & Topology Down Under, Amer. Math. Soc., arXiv:1208.2504, 2012.
- [7] Benjamin A. Burton and Melih Ozlen, *Computing the crosscap number of a knot using integer programming and normal surfaces*, ACM Trans. Math. Software **39** (2012), no. 1, 4:1–4:18.
- [8] ———, *A tree traversal algorithm for decision problems in knot theory and 3-manifold topology*, Algorithmica **65** (2013), no. 4, 772–801.
- [9] Benjamin A. Burton, J. Hyam Rubinstein, and Stephan Tillmann, *The Weber-Seifert dodecahedral space is non-Haken*, Trans. Amer. Math. Soc. **364** (2012), no. 2, 911–932.
- [10] Maurice Chiodo, *Finding non-trivial elements and splittings in groups*, J. Algebra **331** (2011), 271–284.
- [11] Nathan M. Dunfield and Anil N. Hirani, *The least spanning area of a knot and the optimal bounding chain problem*, SCG '11: Proceedings of the Twenty-Seventh Annual Symposium on Computational Geometry, ACM, 2011, pp. 135–144.

- [12] Nathan M. Dunfield and William P. Thurston, *Finite covers of random 3-manifolds*, *Invent. Math.* **166** (2006), no. 3, 457–521.
- [13] Wolfgang Haken, *Theorie der Normalflächen*, *Acta Math.* **105** (1961), 245–375.
- [14] Joel Hass, *New results on the complexity of recognizing the 3-sphere*, *Oberwolfach Rep.* **9** (2012), no. 2, 1425–1426.
- [15] Joel Hass, Jeffrey C. Lagarias, and Nicholas Pippenger, *The computational complexity of knot and link problems*, *J. Assoc. Comput. Mach.* **46** (1999), no. 2, 185–211.
- [16] Joel Hass, Jack Snoeyink, and William P. Thurston, *The size of spanning disks for polygonal curves*, *Discrete Comput. Geom.* **29** (2003), no. 1, 1–17.
- [17] William Jaco and J. Hyam Rubinstein, *0-efficient triangulations of 3-manifolds*, *J. Differential Geom.* **65** (2003), no. 1, 61–168.
- [18] William Jaco and Jeffrey L. Tollefson, *Algorithms for the complete decomposition of a closed 3-manifold*, *Illinois J. Math.* **39** (1995), no. 3, 358–406.
- [19] Robion Kirby, *A calculus for framed links in  $S^3$* , *Invent. Math.* **45** (1978), no. 1, 35–56.
- [20] Greg Kuperberg, *Knottedness is in NP, modulo GRH*, Preprint, [arXiv:1112.0845](https://arxiv.org/abs/1112.0845), November 2011.
- [21] W. B. Raymond Lickorish, *An introduction to knot theory*, *Graduate Texts in Mathematics*, no. 175, Springer, New York, 1997.
- [22] Joseph Maher, *Random Heegaard splittings*, *J. Topol.* **3** (2010), no. 4, 997–1025.
- [23] A. A. Markov, *Insolubility of the problem of homeomorphy*, *Proc. Internat. Congress Math. 1958*, Cambridge Univ. Press, New York, 1960, pp. 300–306.
- [24] Sergei Matveev, *Algorithmic topology and classification of 3-manifolds*, *Algorithms and Computation in Mathematics*, no. 9, Springer, Berlin, 2003.
- [25] Peter Orlik, *Seifert manifolds*, *Lecture Notes in Mathematics*, no. 291, Springer-Verlag, Berlin, 1972.
- [26] Vic Reiner and John Sullivan, *Triangulations: Minutes of the open problem session*, *Oberwolfach Rep.* **9** (2012), no. 2, 1478–1481.
- [27] J. Hyam Rubinstein, *An algorithm to recognise small Seifert fiber spaces*, *Turkish J. Math.* **28** (2004), no. 1, 75–87.