THE COMPUTATIONAL COMPLEXITY OF DETERMINING KNOT GENUS IN A FIXED 3-MANIFOLD

MARC LACKENBY, MEHDI YAZDI

Abstract. We show that the problem of determining the genus of a knot in a fixed compact, orientable three-dimensional manifold lies in **NP**. This answers a question asked by Agol, Hass, and Thurston in 2002. Previously, this was known for rational homology three-spheres, by the work of the first author.

1. Introduction

Let M be a compact 3-manifold, and let K be a knot inside M. Since the work of Dehn in 1910 [9], deciding whether K can be unknotted has been a major question in low-dimensional topology. Dehn formulated the word and the isomorphism problems for groups in an attempt to solve this question (The isomorphism problem was stated by Tietze [32] in 1908 as well.) This in turn led to Novikov's discovery of the undecidability of the word problem for finitely presented groups [24] and the undecidability of the isomorphism problem for finitely presented groups by Adian [1] and Rabin [25]. Haken was the first person to prove that the unknot recognition problem is decidable using the theory of normal surfaces, introduced previously by Kneser [16].

Seifert defined the *genus* of a knot in the 3-sphere [29]. Consider all connected, compact, embedded, orientable surfaces in M whose boundary coincides with K, and let the genus, $g(K)$, be the minimum genus of the surfaces in this family. If there is no such surface, then we define $g(K) = \infty$ in this case. An easy observation is that $g(K) < \infty$ if and only if K represents the trivial element in the first homology group $H_1(M; \mathbb{Z})$. Furthermore, $g(K) = 0$ if and only if K is the unknot.

Thus, one of the most basic decision problems in low-dimensional topology is as follows: given a knot K in a 3-manifold M and a non-negative integer g, is the genus of K equal to g? The manifold M is provided via a triangulation in which K is a specified subcomplex. We term this problem 3-manifold knot genus. There are good reasons for believing this to be hard. Agol, Hass and Thurston [2] considered the related problem Upper bound on 3-manifold knot genus, which asks whether $g(k) \leq g$, and they proved that this problem is **NP**-complete. A consequence is that if 3-manifold knot genus were to be in **NP**, then **NP** = **co-NP**, contradicting a basic conjecture in complexity theory. (See Theorem 1.4 of [18] for this deduction).

It is natural to ask whether the difficulty of 3-manifold knot genus is consequence of the fact that K and M can both vary. What if we fix the manifold M, and only allow K to vary? In [2], Agol, Hass and Thurston asked about the computational complexity of this problem. The specific case where M is the 3-sphere was addressed by the first author. He showed [18] that, in this restricted setting, deciding whether a knot has genus g *is* in **NP**. More generally, if we are given a triangulation of a rational homology 3-sphere M, a knot K as a subcomplex and an integer g, then the question 'does $g(K)$ equal g?' lies in **NP**.

Let $N(K)$ be a tubular neighbourhood of K with interior $N^\circ(K)$. The reason why knots in rational homology 3-spheres seem to be so much more tractable than in general 3-manifolds is that, in this situation, there can be only one possible homology class in $H_2(M - N^\circ(K), \partial N(K))$, up to sign, for a compact oriented spanning surface. This suggests that knots in more complicated 3-manifolds M might be difficult to analyse, since as soon as $b_1(M) \geq 1$, there may be infinitely many possibilities for the homology class of a spanning surface. However, the main result of this paper is that, provided we consider knots in a fixed 3-manifold M, then the problem of determining their genus lies in **NP**.

In order to state this result more precisely, we need to explain how the knots K in M are presented. Any closed orientable 3-manifold is obtained by integral surgery on a framed link L in the 3-sphere [20, 36]. When M is closed, we fix such a surgery description of M, by fixing a diagram D for L where the framing of L is diagrammatic framing and this specifies the surgery slopes. We specify knots K in M by giving a diagram for $K \cup L$ that contains D as a sub-diagram. The *total number of crossings* of K is defined as the number of crossings in this diagram between K and itself and between K and L.

Problem: Determining knot genus in the fixed closed 3-manifold M.
Input: A diagram of $K \cup L$ that contains D as a subdiagram, and an integer $g \geq 0$ in binary.
Question: Is the genus of K equal to g?

The size of the input is equal to sum of the number of digits of g in binary and the total number of crossings of K. Strictly speaking, there are infinitely many decision problems here, one for each 3-manifold M and surgery diagram D.

Theorem 1.1. *Let M be a closed, orientable 3-manifold given by integral surgery on a framed link in the 3-sphere. The problem* Determining knot genus in the fixed closed 3-manifold *M lies in **NP**.*

This can be generalised to compact orientable 3-manifolds with non-empty boundary, as follows. Any compact orientable 3-manifold M can be specified by means of the disjoint union of a graph Γ and a framed link L in the 3-sphere. The manifold M is obtained from S^3 by removing an open regular neighbourhood of Γ and performing surgery along L. We fix a diagram D for $\Gamma \cup L$, where again the surgery slopes on L agree with the diagrammatic framing. We can then specify a knot K in M by giving a diagram for $K \cup \Gamma \cup L$ that contains D as a sub-diagram. Again, the *total crossing number* of K is the number of crossings in this diagram between K and itself and between K and $\Gamma \cup L$. We say that Determining knot genus in a fixed 3-manifold M is the decision problem asking whether the genus of K is equal to a given non-negative integer.

Theorem 1.2. *Let M be a compact, orientable 3-manifold given as above. The problem* Determining knot genus in the fixed 3-manifold *M lies in **NP**.*

1.1. **Ingredients of the proof.** (1) One of the key technical tools in the paper is the use of non-standard measures of complexity for various objects. We introduce the relevant terminology now.

For an integer n, let $C_{\text{nat}}(n) = |n|$ and let $C_{\text{dig}}(n)$ be the number of digits of n when expressed in binary. In the case of negative n, we view the minus sign at the front as

an extra digit. For a list of integers $(n_1, \cdots, n_k)$, let $C_{\text{nat}}(n_1, \cdots, n_k)$ be $\sum_i C_{\text{nat}}(n_i)$. Similarly, let $C_{\text{dig}}(n_1, \cdots, n_k)$ be $\sum_i C_{\text{dig}}(n_i)$. For a matrix A with integer entries A_{ij}, let $C_{\text{nat}}(A)$ be $\sum_{ij} C_{\text{nat}}(A_{ij})$ and let $C_{\text{dig}}(A)$ be $\sum_{ij} C_{\text{dig}}(A_{ij})$. For a rational number p/q, with p and q in their lowest terms, let $C_{\text{nat}}(p/q) = C_{\text{nat}}(p) + C_{\text{nat}}(q)$ and let $C_{\text{dig}}(p/q) = C_{\text{dig}}(p) + C_{\text{dig}}(q)$.

The C_{dig} notions of size are the most natural ones and the ones that are most widely used in complexity theory, because they reflect the actual amount of memory required to store the number, list or matrix. However, we will also find the C_{nat} versions useful.

(2) One of the main ingredients is the following result, proved by the first author in [18].

Theorem 1.3 (Lackenby)**.** Thurston norm of a homology class *is in* ***NP****.*

The decision problem Thurston norm of a homology class takes as its input a triangulation $\mathcal{T}$ for a compact orientable 3-manifold M, a simplicial 1-cocycle c and an integer g, and it asks whether the Thurston norm of the dual of c is equal to g. (For the definition of the Thurston norm, see Section 2.2.) The measure of complexity of $\mathcal{T}$ is its number of tetrahedra, denoted $|\mathcal{T}|$. The measure of complexity of c is $C_{\text{dig}}(c)$, where we view c as a list of integers, by evaluating it against all the edges of $\mathcal{T}$ (when they are oriented in some way). The measure of complexity of g is $C_{\text{dig}}(g)$.

(3) Thus, one can efficiently certify the Thurston norm of the dual of a *single* cohomology class. However, in principle, a minimal genus Seifert surface for the knot K could be represented by one of infinitely many classes. To examine all possible classes simultaneously, one needs a good picture of the Thurston norm ball.

Theorem 1.4. *Fix an integer $B \geq 0$. The problem* Thurston norm ball for $b_1 \leq B$ *lies in* ***FNP****, where b_1 denotes the first Betti number.*

Recall that **FNP** is the generalisation of **NP** from decision problems (where a yes/no answer is required) to function problems (where more complicated outputs might be required). A formal definition is given in Section 2.1.

The problem Thurston norm ball for $b_1 \leq B$ is as follows. It takes, as its input, a triangulation $\mathcal{T}$ of a compact orientable 3-manifold X with $b_1(X) \leq B$, and a list of simplicial integral cocycles $\phi_1, \cdots, \phi_b$ that form a basis for $H^1(X; \mathbb{R})$. The output is all the information that one needs to compute the Thurston norm ball:

(1) a collection of integral cocycles that forms a basis for the subspace W of $H^1(X; \mathbb{R})$ with Thurston norm zero;
(2) a list $V \subset H^1(X; \mathbb{Q})$ of points that project to the vertices of the unit ball of $H^1(X; \mathbb{R})/W$, together with a list of subsets of these vertices that form faces. These are given as rational linear combinations of $\phi_1, \cdots, \phi_b$.

At first sight, this theorem seems to lead easily to the proof of Theorem 1.1. However, its power is blunted by the non-standard notion of complexity that it uses. This is defined to be $|\mathcal{T}| + \sum_i C_{\text{nat}}(\phi_i)$. Thus, it only works well when $C_{\text{nat}}(\phi_i)$ is 'small' for each i. That such a collection of simplicial cocycles exists in our setting is a consequence of the following surprising result.

(4) Constructing an efficient basis for the second homology of a knot complement, for a fixed ambient manifold.

Theorem 1.5. *Let M be a compact orientable 3-manifold given by removing an open regular neighbourhood of a graph Γ in S^3 and performing integral surgery on a framed link L in the complement of Γ. Let D be a fixed diagram for $\Gamma \cup L$ where the surgery slopes on L coincide with the diagrammatic framing. Let K be a homologically trivial knot in M, given by a diagram of $K \cup \Gamma \cup L$ that contains D as a sub-diagram. Let c be the total crossing number of K. Set $X = M - N^\circ(K)$ as the exterior of K in M. There is an algorithm that builds a triangulation of X with $O(c)$ tetrahedra, together with simplicial 1-cocycles $\phi_1, \cdots, \phi_b$ that form an integral basis for $H^1(X;\mathbb{Z})$ with $\sum_i C_{\text{nat}}(\phi_i)$ at most $O(c^2)$. The algorithm runs in time polynomial in c. All the above implicit constants depend on the manifold M and not the knot K.*

(5) Controlling the number of faces and vertices of the Thurston norm ball polyhedron, in the presence of an efficient basis for the second homology.

A crucial step in the proof of Theorem 1.4 is to bound the number of vertices and faces of the Thurston norm ball of the manifold X. The following result gives this, assuming that we have a good bound on the Thurston norm of a collection of surfaces that form a basis for $H_2(X, \partial X; \mathbb{R})$.

Theorem 1.6. *Let X be a compact orientable 3-manifold, and let m be a natural number. Assume that there exist properly immersed oriented surfaces $S_1, \cdots, S_b$ in X such that their homology classes form a basis for $H_2(X, \partial X; \mathbb{R})$, and for each $1 \leq i \leq b$ we have $|\chi_-(S_i)| \leq m$. Denote by W the subspace of $H^1(X;\mathbb{R})$ with trivial Thurston norm. The number of facets of the unit ball for the induced Thurston norm on $H^1(X;\mathbb{R})/W$ is at most $(2m+1)^b$, where $b = b_1(X)$ is the first Betti number of X. Hence, the number of vertices is at most $(2m+1)^{b^2}$ and the number of faces is at most $b(2m+1)^{b^2}$.*

For the definition of $\chi_-(S)$, see Section 2.2. The proof of Theorem 1.6 uses the fact, due to Thurston [31], that the vertices of the dual unit ball of the Thurston norm are integral. See Theorem 4.1 for a result that gives an upper bound on the number of these integral points.

(6) Constructing a basis for the subspace of the second homology with trivial Thurston norm.

Theorem 1.7. *Let $\mathcal{T}$ be a triangulation of a compact orientable irreducible 3-manifold X. If X has any compressible boundary components, suppose that these are tori. Then there is a collection $w_1, \cdots, w_r$ of integral cocycles that forms a basis for the subspace W of $H^1(X;\mathbb{R})$ consisting of classes with Thurston norm zero with $\sum_i C_{\text{dig}}(w_i)$ at most $O(|\mathcal{T}|^2)$.*

This is proved by showing that $W \cap H^1(X;\mathbb{Z})$ is spanned by fundamental normal surfaces, which is a consequence of work of Tollefson and Wang [33].

(7) In Theorem 1.7, it is assumed that X is irreducible and that every component of ∂X is toroidal or incompressible. In Section 8, we explain how we may ensure this. We cut along a maximal collection of compression discs and essential normal spheres to decompose X into pieces, and we construct a new simplicial basis for the cohomology of the pieces. We also use the following result from [18].

Theorem 1.8 (Lackenby)**.** *The following decision problem lies in **NP**. The input is a triangulation of a compact orientable 3-manifold M with (possibly empty) toroidal boundary and $b_1(M) > 0$, and the problem asks whether M is irreducible.*

Corollary 1.9. *The following decision problem lies in* ***NP****. The input is a triangulation of a compact orientable 3-manifold M with $b_1(M) > 0$, and the problem asks whether M is irreducible and has incompressible boundary.*

This is an immediate consequence of Theorem 1.8. This is because a compact orientable 3-manifold M is irreducible with incompressible boundary if and only if its double DM is irreducible. This follows from the equivariant sphere theorem [22]. Moreover, $b_1(DM) > 0$ if and only if $b_1(M) > 0$.

1.2. **Varying M and K.** As mentioned above, it seems very unlikely that Theorems 1.1 and 1.2 remain true if M and K are allowed to vary, because of the following result of Agol, Hass and Thurston [2].

Theorem 1.10. *The following problem is* ***NP****-complete. The input is a triangulation of a closed orientable 3-manifold M, a knot K in its 1-skeleton and an integer g, and the problem asks whether the genus of K is at most g.*

However, what if we allow M to vary but fix $b_1(M)$ in advance? It is unclear to the authors whether the problem of determining knot genus in such manifolds M is likely to lie in **NP**.

We believe that in this more general setting, Theorem 1.5 does not hold. Certainly, the proof of Theorem 1.5 required M to be fixed. This bound on $\sum_i C_{\text{nat}}(\phi_i)$ was used to bound $\chi_-(S_i)$, where S_i is a representative surface for the Poincaré dual of c_i. In the absense of such a bound, it is not clear that one can find a good upper bound on the number of faces and vertices of the Thurston norm ball for $H^1(X;\mathbb{R})$. In particular, it is an interesting question whether there is a sequence of 3-manifolds X with bounded first Betti number and triangulations $\mathcal{T}$, where the number of vertices of the Thurston norm ball of X grows faster than any polynomial function of $|\mathcal{T}|$.

2. Preliminaries

Notation 2.1. *For a metric space X and $A \subset X$, denote the metric completion of $X - A$ with the induced path metric by $X \setminus\setminus A$.*

For a subset A of a topological space Y, the interior of A is shown by A°.

The first Betti number of a manifold M is indicated by $b_1(M)$.

2.1. **Complexity Theory.** The material in this section is borrowed from [3, 26], and we refer the reader to them for a more thorough discussion.

Let $\{0,1\}^*$ be the set of all finite strings in the alphabet $\{0,1\}$. A *problem* P is defined as a function from $\{0,1\}^*$ to $\{0,1\}^*$. Here the domain is identified with the *inputs* or *instances*, and the range is identified with the *solutions*. A *decision problem* is a problem whose range can be taken to be $\{0,1\} \subset \{0,1\}^*$. Intuitively a decision problem is a problem with yes or no answer.

A (deterministic) *Turing machine* is a basic computational device that can be used as a model of computation. We refer the reader to Page 12 of [3] for a precise definition. By an *algorithm* for the problem P, we mean a Turing machine M that given any instance I of the problem on its tape, computes and halts exactly with the solution $P(I)$. We say M runs in time $T : \mathbb{N} \longrightarrow \mathbb{N}$, if for any instance I of binary length $|I|$, if we start the Turing machine M with I on its tape, the machine halts after at most $T(|I|)$ steps.

The *complexity class* **P** consists of all decision problems P for which there exists a Turing machine M and positive constants c, d such that M answers the problem in time cn^d.

The complexity class **NP** consists of decision problems such that their yes solutions can be efficiently *verified*. By this we mean that there is a Turing machine that can verify the yes solutions in polynomial time. This is possibly a larger complexity class than the class **P**, which was described as the set of decision problems that can be efficiently *solved*. In other words, $\mathbf{P} \subseteq \mathbf{NP}$. The precise definition is as follows. By a *language*, we mean a subset of $\{0,1\}^*$. In our context, we have a decision problem $P\colon \{0,1\}^* \longrightarrow \{0,1\}$ and we take L as the set of instances whose solutions are equal to 1 (yes answer).

Definition 2.2. A language $L \subset \{0,1\}^*$ is in **NP**, if there exists a polynomial $p\colon \mathbb{N} \longrightarrow \mathbb{N}$ and a Turing machine M that runs in polynomial time (called the *verifier* or *witness* for L) such that for every instance $x \in \{0,1\}^*$

$$x \in L \iff \exists u \in \{0,1\}^{p(|x|)} \text{ such that } M(x,u) = 1.$$

If $x \in L$ and $u \in \{0,1\}^{p(|x|)}$ and $M(x,u) = 1$, we call u a *certificate* for x.

A decision problem is called **NP**-*hard* if it is at least as hard as any other problem in **NP**. More specifically, every problem in **NP** is *Karp-reducible* to any **NP**-hard problem. (See Page 42 of [3] for a definition of Karp-reducibility.) In particular, if any **NP**-hard problem is solvable in polynomial time, then $\mathbf{P} = \mathbf{NP}$.

Now instead of restricting our attention to decision problems, we consider the computational complexity of more general problems. Recall that a *problem* P is just a function $P\colon \{0,1\}^* \to \{0,1\}^*$. We say that P is in **FNP** if there is a deterministic polynomial time verifier that, given an arbitrary input pair (x,y) where $x, y \in \{0,1\}^*$, determines whether $P(x) = y$.

2.2. **Thurston norm.** Let M be any compact orientable 3-manifold. Thurston [31] defined a semi-norm on the second homology group $H_2(M, \partial M; \mathbb{R})$. This norm generalises the notion of knot genus, and for any homology class measures the minimum 'complexity' between all properly embedded orientable surfaces representing that homology class. More precisely, for any properly embedded connected orientable surface S, define

$$\chi_-(S) := \max\{0, -\chi(S)\}.$$

If S has multiple components, define $\chi_-(S)$ as the sum of the corresponding values for the components of S. Now for any integral homology class $a \in H_2(M, \partial M; \mathbb{R})$ define the *Thurston norm* of a, $x(a)$, as

$$x(a) = \min\{\chi_-(S) \mid [S] = a, \quad S \text{ is compact, oriented and properly embedded}\}.$$

This defines the norm for integral homology classes. One can extend this linearly to rational homology classes, and then extend it continuously to all real homology classes.

Consider the special case that K is a knot of genus g in S^3, and $M := S^3 - N^\circ(K)$, where $N(K)$ is a tubular neighbourhood of K. The second homology group $H_2(M, \partial M; \mathbb{R})$ is isomorphic to $\mathbb{R}$ and the Thurston norm of a generator for the integral lattice

$$H_2(M, \partial M; \mathbb{Z}) \subset H_2(M, \partial M; \mathbb{R})$$

is equal to $2g - 1$ if $g \geq 1$, and 0 otherwise.

In general this might be a semi-norm as opposed to a norm, since one might be able to represent some non-trivial homology classes by a collection of spheres, discs, tori or annuli. However, if W denotes the subspace of $H_2(M, \partial M; \mathbb{R})$ with trivial Thurston norm, then one gets an induced norm on the quotient vector space $H_2(M, \partial M; \mathbb{R})/W$.

Thurston proved that the unit ball of this norm is a convex polyhedron. Given any norm on a vector space V, there is a corresponding dual norm on the dual vector space, that is the space of functionals on V. In our case, the dual space to $H_2(M, \partial M; \mathbb{R})$ is $H^2(M, \partial M; \mathbb{R})$. Thurston showed that the unit ball for the corresponding dual norm x^* is a convex polyhedron with *integral* vertices. For a thorough exposition of Thurston norm and examples see [31, 6].

Finally, it is possible to define a norm x_s using singular surfaces and allowing real coefficients. Thus, if $S_1, \cdots, S_k$ are oriented singular surfaces in a 3-manifold M, and if $S = \sum a_i S_i$ is a real linear combination, representing a homology class a, we may define

$$\chi_-(S) = \sum |a_i| \chi_-(S_i).$$

The singular norm x_s is defined as

$$x_s(a) = \inf\{\chi_-(S) \mid [S] = a\}.$$

Gabai [10] proved the equivalence of the two norms x and x_s, previously conjectured by Thurston [31].

Theorem 2.3 (Gabai)**.** *Let M be a compact oriented 3-manifold. Then on $H_2(M)$ or $H_2(M, \partial M)$, $x_s = x$ where x_s denotes the norm on homology based on singular surfaces.*

2.3. **Bareiss algorithm for solving linear equations.** Gaussian elimination is a useful method for solving a system of linear equations with integral coefficients, computing determinants and calculating their echelon form. The algorithm uses $O(n^3)$ arithmetic operations, where n is the maximum of the number of variables and the number of equations. One caveat is that the intermediate values for the entries during the process can get large. An algorithm due to Bareiss resolves this issue. If the maximum number of bits for entries of the input is L, then the running time of the algorithm is at most a polynomial function of $n + L$. Moreover, no intermediate value (including the final answer) needs more than $O(n \log(n) + nL)$ bits [4].

2.4. **Mixed integer programming.** This refers to the following decision problem. Let $n \geq 0$ and $m > 0$ be integers, and let k be a positive integer satisfying $k \geq n$. Let A be an $m \times k$ matrix with integer coefficients, and let $b \in \mathbb{Z}^m$. Then the problem asks whether there is an $x = (x_1, \ldots, x_k)^T \in \mathbb{R}^k$ such that

$$Ax \leq b$$

$$x_i \in \mathbb{Z} \text{ for all } i \text{ satisfying } 1 \leq i \leq n.$$

The size of the input is given by $k + m + C_{\text{dig}}(A) + C_{\text{dig}}(b)$. Lenstra [19] provided an algorithm to solve this problem that runs in polynomial time for any fixed value of n.

It is also shown in [19], using estimates of von zur Gathen and Sieveking [34], that if the above instance of Mixed Integer Programming does have a positive solution x, then it has one for which $C_{\text{dig}}(x)$ is bounded above by a polynomial function of the size of the input.

Figure 1 shows an example of Mixed Integer Programming where

(1) the shaded region is the *feasible region* namely $Ax \leq b$ where $x \in \mathbb{R}^2$;
(2) the dots indicate integral points inside the feasible region.

In this example, there is at least one integral point inside the feasible region and hence the answer is yes.

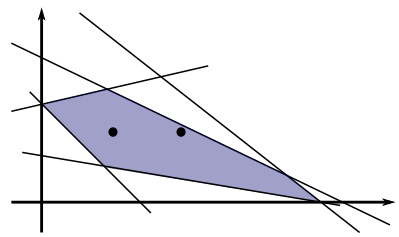

FIGURE 1. Integer Linear Programming

2.5. **Polyhedra and their duals.** Our exposition is from [5] and we refer the reader to that for more details and proofs. A set of points $\{y_0, y_1, \cdots, y_m\} \subset \mathbb{R}^d$ is *affinely independent* if the vectors $y_1 - y_0, \cdots, y_m - y_0$ are linearly independent. A *polytope* P is the convex hull of a non-empty finite set $\{x_1, \cdots, x_n\}$ in $\mathbb{R}^d$. We say P is k-dimensional if some $(k+1)$-subfamily of $\{x_1, \cdots, x_n\}$ is affinely independent, and k is maximal with respect to this property. A convex subset F of P is called a *face* of P if for any two distinct points $y, z \in P$ such that $]y, z[\, \cap F$ is non-empty, we have $[y, z] \subset F$. Here $]y, z[$ and $[y, z]$ denote the open and closed segments connecting y and z respectively. A face F is *proper* if $F \neq \emptyset, P$. A point $x \in P$ is a *vertex* if $\{x\}$ is a face. A *facet* F of P is a face of P with $\dim(F) = \dim(P) - 1$. Every face F of P is itself a polytope, and coincides with the convex hull of the set of vertices of P that lie in F. Every proper face F of P is the intersection of facets of P containing F (see Theorem 10.4 of [5]).

The intersection of any family of faces of P is again a face. For any family $\mathcal{A}$ of faces of P, there is a largest face contained in all members of $\mathcal{A}$ denoted by $\inf \mathcal{A}$, and there is a smallest face that contains all members of $\mathcal{A}$ denoted by $\sup \mathcal{A}$. Denote the set of faces of P by $\mathcal{F}(P)$ and let $\subset$ denote inclusion. Therefore, the partially ordered set $(\mathcal{F}(P), \subset)$ is a *complete lattice* with the lattice operations $\inf \mathcal{A}$ and $\sup \mathcal{A}$. The pair $(\mathcal{F}(P), \subset)$ is called the *face-lattice* of P.

Let P be a d-dimensional polytope in $\mathbb{R}^d$ containing the origin. Define the *dual* of P as

$$P^* := \{y \in \mathbb{R}^d \mid \sup_{x \in P} \langle x, y \rangle \leq 1\}.$$

For any face F of P, define the dual face $F^\triangle$ as

$$F^\triangle := \{y \in P^* \mid \sup_{x \in F} \langle x, y \rangle = 1\}.$$

We have $(P^*)^* = P$, and $(F^\triangle)^\triangle = F$. There is a one-to-one correspondence between faces F of P and faces $F^\triangle$ of P^*, and

$$\dim(F) + \dim(F^\triangle) = d - 1.$$

Moreover, the mapping $F \mapsto F^\triangle$ defines an *anti-isomorphism* of face-lattices (see Corollary 6.8 of [5])

$$(\mathcal{F}(\mathcal{P}), \subset) \to (\mathcal{F}(\mathcal{P}^*), \subset).$$

A subset Q of $\mathbb{R}^d$ is called a *polyhedral set* if Q is the intersection of a finite number of closed half-spaces or $Q = \mathbb{R}^d$. Polytopes are precisely the non-empty bounded polyhedral sets.

2.6. **Pseudo-manifolds, orientability and degree of mappings.** At some point in this article, we need to talk about the degree of a mapping between two topological spaces that a-priori are not manifolds. They are similar to manifolds, but with particular types of singularities. The following discussion is from [28]. 'A *closed pseudo-manifold* is defined as follows:

PM1) It is a *pure*, finite n-dimensional simplicial complex ($n \geq 1$); by pure we mean that each k-simplex is a face of at least one n-simplex (*purity condition*).

PM2) Each $(n-1)$-simplex is a face of exactly two n-simplices (*non-branching condition*).

PM3) Every two n-simplexes can be connected by means of a series of alternating n- and $(n-1)$-simplexes, each of which is incident with its successor (*connectivity condition*).

A closed pseudo-manifold is said to be *orientable* if each of its n-simplices can be oriented coherently, that is, oriented so that opposite orientations are induced in each $(n-1)$-simplex by the two adjoining n-simplices.

A closed n-chain on an orientable and coherently oriented closed pseudo-manifold is completely determined whenever one knows how often a single, arbitrarily chosen, oriented n-simplex appears in the chain. This is so because each of the n-simplices adjoining this simplex must appear equally often, from (PM3). One can reach each n-simplex by moving successively through adjoining simplices; hence all n-simplices must appear equally often. Consequently the n-th homology group is the free cyclic group. In other words, the n-th Betti number is equal to 1. A basis for this group is one of the two chains which arise by virtue of the coherent orientation of the pseudo-manifold.'

A choice for one of these chains is an *orientation* of the pseudo-manifold and its homology class is then called the *fundamental class*.

Let D and R be oriented pseudo-manifolds of dimension n, and $f\colon D \longrightarrow R$ be a continuous map. The n-th homology groups of D and R are both isomorphic to a free cyclic group. Denote by $[D]$ and $[R]$ the fundamental homology classes of D and R respectively. Then there is an integral number d such that $f_*([D])$ is homologous to $d[R]$. This number d is defined as the *degree* of f. Similar to the case of manifolds, one can compute the degree by counting signed preimages of a generic point, where the sign depends on whether the map is locally orientation-preserving or not.

2.7. **Normal surfaces.** The theory of normal surfaces was introduced by Kneser in [16] where he proved a prime decomposition theorem for compact 3-manifolds, and was extended by Haken in his work on algorithmic recognition of the unknot [11]. Let $\mathcal{T}$ be a triangulation of a compact 3-manifold M. A surface S properly embedded in M is said to be *normal* if it intersects each tetrahedron in a collection of disjoint triangles and squares, as shown in Figure 2.

In each tetrahedron, there are 4 types of triangle and 3 types of square. Thus, in total, there are $7t$ types of triangles and squares in $\mathcal{T}$, where t is the number of tetrahedra in $\mathcal{T}$. A normal surface S determines a list of $7t$ non-negative integers, which count the number of triangles and squares of each type in S. This list is called the *vector* for S and is denoted by (S).

The normal surface S is said to be *fundamental* if (S) cannot be written as $(S_1)+(S_2)$ for non-empty properly embedded normal surfaces S_1 and S_2. It is said to be a *vertex surface* if no non-zero multiple of (S) can be written as $(S_1) + (S_2)$ for non-empty

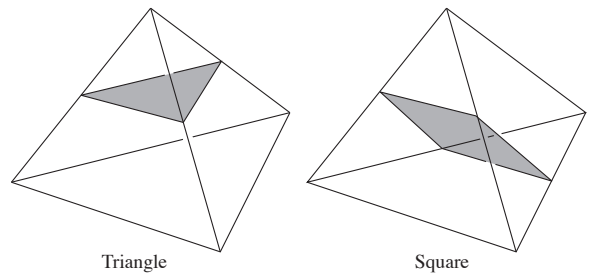

FIGURE 2. A triangle and a sqaure

properly embedded normal surfaces S_1 and S_2. This has an alternative interpretation in terms of the normal solution space, as follows.

The *normal solution space* $\mathcal{N}(\mathcal{T})$ is a subset of $\mathbb{R}^{7t}$. The co-ordinates of $\mathbb{R}^{7t}$ correspond to the $7t$ types of triangles and squares in $\mathcal{T}$. The subset $\mathcal{N}(\mathcal{T})$ consists of those points in $\mathbb{R}^{7t}$ where every co-ordinate is non-negative and that satisfy the normal *matching equations* and *compatibility conditions*. There is one matching equation for each type of normal arc in each face of $\mathcal{T}$ not lying in ∂M. The equation asserts that in each of the two tetrahedra adjacent to that face, the total number of triangles and squares that intersect the given face in the given arc type are equal. The compatibility conditions assert that for different types of square within a tetrahedron, at least one of the corresponding co-ordinates is zero. For any properly embedded normal surface S, its vector (S) lies in $\mathcal{N}(\mathcal{T})$. Indeed, the set of points in $\mathcal{N}(\mathcal{T})$ that are a vector of a properly embedded normal surface is precisely $\mathcal{N}(\mathcal{T}) \cap \mathbb{Z}^{7t}$.

The *projective solution space* $\mathcal{P}(\mathcal{T})$ is the intersection

$$\mathcal{N}(\mathcal{T}) \cap \{(x_1, \cdots, x_{7t}) : x_1 + \cdots + x_{7t} = 1\}.$$

It is shown in [21] that $\mathcal{P}(\mathcal{T})$ is a union of convex polyhedra. A normal surface S is *carried* by a face of $\mathcal{P}(\mathcal{T})$ if its vector (S) lies on a ray through the origin of $\mathbb{R}^{7t}$ that goes through that face. When a normal surface S is carried by a face C, and $(S) = (S_1) + (S_2)$ for normal surfaces S_1 and S_2, then S_1 and S_2 are also carried by C. The reason for this is that C is the intersection between $\mathcal{P}(\mathcal{T})$ and some hyperplanes of the form $\{x_i = 0\}$. Since $(S) = (S_1) + (S_2)$, then (S_1) and (S_2) also lie in these hyperplanes and hence also are carried by C.

A normal surface S is a vertex surface exactly when some non-zero multiple of (S) is a vertex of one of the polyhedra of $\mathcal{P}(\mathcal{T})$. Using this observation, it was shown by Hass and Lagarias (Lemma 3.2 in [12]) that each co-ordinate of the vector of a vertex normal surface is at most 2^{7t-1}. Hence, the number of points of intersection between a vertex normal surface and the 1-skeleton of $\mathcal{T}$ is at most $28t2^{7t-1}$. They also showed that each co-ordinate of a fundamental normal surface in $\mathcal{T}$ has modulus at most $t2^{7t+2}$.

A common measure of complexity for a normal surface S is its *weight* $w(S)$ which is its number of intersections with the 1-skeleton of $\mathcal{T}$.

3. MAIN THEOREM

In this section, we give a proof of the main theorem, assuming various ingredients that will be proved in later sections.

Theorem 1.2. *Let M be a compact, orientable 3-manifold given by a fixed diagram D for $\Gamma \cup L$, where M is obtained from S^3 by removing an open regular neighbourhood of the graph Γ and performing surgery on the framed link L. The problem* Determining knot genus in the fixed manifold *M lies in **NP**.*

Proof. We are given a diagram for $K \cup \Gamma \cup L$, which contains D as a sub-diagram, where K is our given knot. Let c be the total crossing number of this diagram of K. Recall that this is the number of crossings between K and itself, and between K and $\Gamma \cup L$. Set $X := M - N^\circ(K)$ to be the exterior of a tubular neighbourhood of K in M.

Step 1: By Theorem 1.5, we can construct a triangulation $\mathcal{T}$ of $X = M - N^\circ(K)$ and simplicial 1-cocycles $\phi_1, \cdots, \phi_b$ such that

(1) the number of tetrahedra $|\mathcal{T}|$ of $\mathcal{T}$ is at most a linear function of c;
(2) the cocycles $\phi_1, \cdots, \phi_b$ form an integral basis for $H^1(X;\mathbb{Z})$;
(3) the complexity $\sum_i C_{\text{nat}}(\phi_i)$ is at most a polynomial function of c.

Moreover, the construction of $\mathcal{T}$ and $\phi_1, \cdots, \phi_b$ can be done in polynomial time in c.

Step 2: We check whether K is homologically trivial in M, as otherwise there is no Seifert surface for K and the genus of K is ∞. We do this by attaching a solid torus to $\mathcal{T}$ to form a triangulation of M in which K is simplicial and then determining whether K is the boundary of a simplicial 2-chain. This can be done in time that is polynomial in c, using the Bareiss algorithm for solving linear equations.

If K is homologically trivial, it has a *longitude*, which is defined as the boundary of any Seifert surface S for K in X. The longitude is unique up to sign, for the following reason. If ℓ' is any other longitude, the intersection number $[\ell'].[\ell]$ on ∂X equals the intersection number $[\ell'].[S]$ in X, but this is zero because ℓ' is homologically trivial in X.

Again using the Bareiss algorithm, the longitude ℓ on $\partial N(K)$ can be determined in time that is polynomial in c and, when it is represented as a simplicial 1-cycle, its complexity $C_{dig}(\ell)$ is at most a polynomial function of c.

Step 3: Note the first Betti number of X is bounded above by the constant $B = b_1(M) + 1$, since we are drilling a knot from M. Therefore, we can use Theorem 1.4 to compute the unit ball of the Thurston norm on $H^1(X;\mathbb{R})$, using a non-deterministic Turing machine. Here $H^1(X;\mathbb{R})$ has been identified with $H_2(X, \partial X;\mathbb{R})$ using Poincaré duality.

Note that the size of the input, that is the sum of the number of tetrahedra and $\sum_i C_{\text{nat}}(\phi_i)$, is at most a polynomial function of c. Therefore, this can be done in time that is at most polynomially large in c. Hence, we can construct the following:

(1) A basis $\{w_1, \cdots, w_r\}$ for the subspace W of $H^1(X;\mathbb{R})$ with trivial Thurston norm. Each w_i is an integral cocycle and is written as a linear combination of the given cocycles $\{\phi_1, \cdots, \phi_b\}$. Denote by p the projection map from $H^1(X;\mathbb{R})$ to $H^1(X;\mathbb{R})/W$.
(2) A set of points $V \subset H^1(X;\mathbb{Q})$ such that $p(V)$ is the set of vertices of the unit ball of $H^1(X;\mathbb{R})/W$, together with a list $\mathcal{F}$ of subsets of V. Each element of V is written in terms of the basis $\{\phi_1, \cdots, \phi_b\}$. We think of $\mathcal{F}$ as the list of faces of the unit ball for $H^1(X;\mathbb{R})/W$, in the sense that for $F \in \mathcal{F}$ the set

$$\{p(v) \mid v \in F\},$$

forms the set of vertices of some face of the unit ball for $H^1(X;\mathbb{R})/W$. Moreover, this covers all the faces as we go over all the elements F of $\mathcal{F}$.

Because the problem Thurston norm ball for $b_1 \leq B$ lies in **FNP**, the number of digits of the output is at most a polynomial function of the complexity of the input. Hence, $\sum_i C_{\text{dig}}(w_i)$, $\sum_{v\in V} C_{\text{dig}}(v)$ and $|\mathcal{F}|$ are all bounded above by polynomial functions of c.

Step 4: There is an identification between $H^1(X;\mathbb{Z})$ and $H_2(X,\partial X;\mathbb{Z})$ using Poincaré duality, and there is a boundary map

$$H_2(X,\partial X;\mathbb{Z}) \longrightarrow H_1(\partial X;\mathbb{Z}).$$

This in turn induces a boundary map

$$\partial\colon H^1(X;\mathbb{Z}) \longrightarrow H_1(\partial X;\mathbb{Z}).$$

For each facet $F = \{u_1, \cdots, u_s\}$ of the unit ball of the Thurston norm on $H^1(X;\mathbb{R})/W$, denote by $\text{Cone}(F)$ the cone over the face F:

$$\text{Cone}(F) := \{r_1u_1 + \cdots + r_su_s \mid r_1, \cdots, r_s \in \mathbb{R}_{\geq 0}\}.$$

For each facet F, consider the following minimum:

$$m_F := \min\,\{x(h) \;\mid\; h \in (W + \text{Cone}(\mathcal{F})) \cap H^1(X;\mathbb{Z}) \text{ and } \partial(h) = \pm[\ell]\}.$$

This integer m_F is given to us non-deterministically. We will show below that the number of digits of m_F is bounded above by a polynomial function of c. We verify that m_F is indeed the above minimum, in time that is at most polynomially large in c, by the following argument.

Let h be any element of $W + \text{Cone}(F)$. We may write

$$h = \beta_1 w_1 + \cdots + \beta_r w_r + \alpha_1 u_1 + \cdots + \alpha_s u_s, \tag{1}$$

for some $\beta_1, \cdots, \beta_r \in \mathbb{R}$ and $\alpha_1, \cdots, \alpha_s \in \mathbb{R}_{\geq 0}$. We also require that h lies in $H^1(X;\mathbb{Z})$ which is the condition

$$h = \gamma_1\phi_1 + \cdots + \gamma_b\phi_b, \tag{2}$$

for some $\gamma_1, \cdots, \gamma_b \in \mathbb{Z}$. The condition $\partial(h) = \pm[\ell]$ translates into

$$\beta_1\partial(w_1) + \cdots + \beta_r\partial(w_r) + \alpha_1\partial(u_1) + \cdots + \alpha_s\partial(u_s) = \pm[\ell]. \tag{3}$$

Since each of the vertices $\{u_1, \cdots, u_s\}$ has Thurston norm equal to one and they all lie on the same face, we have

$$x(h) = \alpha_1 + \cdots + \alpha_s.$$

Therefore, we are looking to find the minimum m_F of the linear functional $\alpha_1 + \cdots + \alpha_s$ under the conditions

i) $\alpha_i \in \mathbb{R}_{\geq 0}$, $\beta_i \in \mathbb{R}$, $\gamma_i \in \mathbb{Z}$;

ii) (1) = (2) (by which we mean putting the right hand sides of the equations equal) and (3).

Since K is homologically trivial, there is a solution to this system of constraints. Hence, as explained in Section 2.4, there is a solution where $\sum_i C_{\text{dig}}(\alpha_i)$ is bounded above by a polynomial function of the number of real variables, the number of equations and the number of bits encoding the coefficients of the linear constraints. Hence, m_F is bounded above by a polynomial function of c.

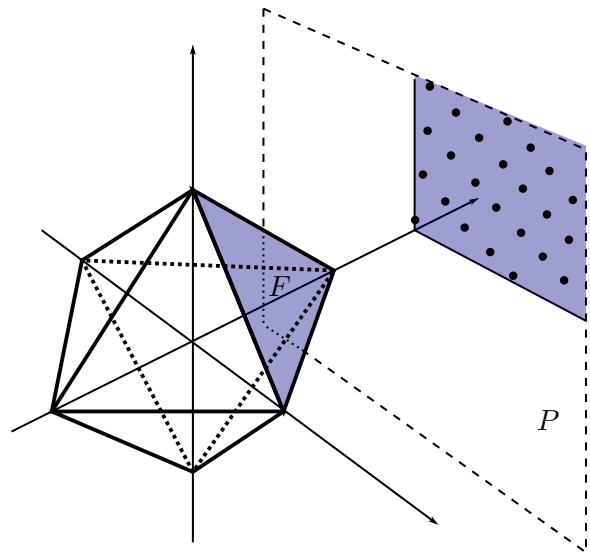

FIGURE 3. Finding the minimum genus between Seifert surfaces coming from a single face of the Thurston norm ball

Our certificate includes the value of m_F. Thus, we must verify that the condition

$$x(h) = m_F$$

is satisfied for some h, whereas the condition

$$x(h) \leq m_F - 1$$

has no solution. (Note that the Thurston norm takes only integer values on elements of $H^1(X;\mathbb{Z})$ and so we do not need to consider the possibility that $x(h)$ might lie strictly between $m_F - 1$ and m_F.) These are instances of Mixed Integer Programming. Since the number b of integer variables is fixed, Lenstra's algorithm provides a solution in polynomial time, as a function of the number of real variables, the number of equations and the number of bits encoding the coefficients of the linear constraints. By hypothesis, these are at most polynomially large in c. See Figure 3 for an example, with the following properties.

(1) The octahedron is the unit ball of the Thurston norm ($W = \{0\}$).
(2) The affine plane P is the location of points with boundary equal to $[\ell]$. In this example, P is disjoint from the unit ball.
(3) The shaded region on P is the intersection of P with the cone over the shaded face F of the unit ball. Here the shaded face F is a triangle, and its projection to P is a degenerate (non-compact) triangle since one edge of F happened to be parallel to P.
(4) The dots on P indicate the integral points on P.
(5) The face F determines the equation for the linear functional $x(h)$. The Mixed Integer Programming problem asks whether there is an integral point h on P that satisfies $x(h) \leq m_F$ (respectively $m_F - 1$) and the constraints i) and ii) above.

Step 5: The minimum of m_F over all facets F of the unit ball of the Thurston norm for $H^1(X;\mathbb{R})/W$ is equal to the Thurston complexity of K. Moreover, the genus can

be easily read from the Thurston complexity. Therefore, we can check if this minimum is equal to the given integer g or not. On the other hand, the number of facets is at most polynomially large in c by the combination of Theorems 1.5 and 1.6. Hence, the algorithm runs in time that is at most polynomially large in c.

This finishes the non-deterministic algorithm for finding the genus of a knot in a fixed 3-manifold. □

4. The number of faces and vertices of the Thurston norm ball

In this section, we prove Theorem 1.6, which provides an upper bound on the number of faces and vertices of the Thurston norm ball. The key to this is the following result, which controls the number of integral points in the dual norm ball.

Theorem 4.1. *Let X be a compact orientable 3-manifold, and let m be a natural number. Assume that there exist properly immersed surfaces $S_1, \cdots, S_b$ in X such that their homology classes form a basis for $H_2(X, \partial X; \mathbb{R})$, and for each $1 \leq i \leq b$ we have $|\chi_-(S_i)| \leq m$. Define the set $\mathcal{A}$ as the set of integral points inside $H^2(X, \partial X; \mathbb{Z}) \otimes \mathbb{Q}$ whose dual norm is at most one. The size of $\mathcal{A}$ is at most $(2m+1)^b$, where $b = b_1(X)$ is the first Betti number of X.*

Proof. Let $\langle \cdot, \cdot \rangle$ be the pairing between cohomology and homology. Define dual elements $e^1, \cdots, e^b \in H^2(X, \partial X; \mathbb{Z}) \otimes \mathbb{Q}$ as

$$\langle e^i, [S_j] \rangle = \delta_{ij},$$

where $1 \leq i, j \leq b$, and δ_{ij} is the Kronecker function. Every integral point in $u \in H^2(X, \partial X; \mathbb{Z}) \otimes \mathbb{Q}$ can be written as

$$u = \alpha_1 e^1 + \cdots + \alpha_b e^b,$$

where α_i are integers. This is because u being integral means that its evaluation against each element of $H_2(X, \partial X; \mathbb{Z})$ is an integer. In particular, $\alpha_i = \langle u, [S_i] \rangle$ is an integer. Assume that the dual norm of u is at most one. By definition of the dual norm, for each $1 \leq i \leq b$ we have:

$$|\langle u, [S_i] \rangle| \leq x([S_i]) = x_s([S_i]),$$

where $x([S_i])$ and $x_s([S_i])$ are the Thurston norm and the singular Thurston norm of $[S_i]$, and the last equality is by Theorem 2.3. Since $|\chi_-(S_i)| \leq m$, we have

$$x_s([S_i]) \leq m.$$

Combining the two inequalities implies that

$$|\alpha_i| = |\langle u, [S_i] \rangle| \leq x([S_i]) = x_s([S_i]) \leq m.$$

Since $-m \leq \alpha_i \leq m$ is an integer, there are at most $2m+1$ possibilities for each coordinate of the tuple $(\alpha_1, \cdots, \alpha_b)$. Therefore the number of possibilities for u is at most $(2m+1)^b$. □

Theorem 1.6. *Let X be a compact orientable 3-manifold, and let m be a natural number. Assume that there exist properly immersed oriented surfaces $S_1, \cdots, S_b$ in X such that their homology classes form a basis for $H_2(X, \partial X; \mathbb{R})$, and for each $1 \leq i \leq b$ we have $|\chi_-(S_i)| \leq m$. Denote by W the subspace of $H^1(X; \mathbb{R})$ with trivial Thurston norm. The number of facets of the unit ball for the induced Thurston norm on $H^1(X; \mathbb{R})/W$ is at most $(2m+1)^b$, where $b = b_1(X)$ is the first Betti number of X.*

Hence, the number of vertices is at most $(2m+1)^{b^2}$ *and the number of faces is at most* $b(2m+1)^{b^2}$.

Proof. Note we have identified $H_2(X, \partial X; \mathbb{Z})$ with $H^1(X; \mathbb{Z})$ using Poincaré duality. Facets of the unit ball for $H^1(X; \mathbb{R})/W$ correspond to the vertices of the dual ball. As the vertices of the dual ball are integral and have dual norm equal to one, the number of them is at most $(2m+1)^b$ by Theorem 4.1. This proves the first part of the theorem.

Let d be the dimension of $H^1(X; \mathbb{R})/W$; hence $d \leq b$. Every m-dimensional face of the unit ball is the intersection of $(d-m)$ facets. Hence, the number of m-dimensional faces is at most

$$\binom{(2m+1)^b}{d-m}.$$

As a result, the total number of faces is at most

$$\binom{(2m+1)^b}{1} + \binom{(2m+1)^b}{2} + \cdots + \binom{(2m+1)^b}{b},$$

which is bounded above by $b(2m+1)^{b^2}$. In particular, the number of vertices is at most $(2m+1)^{b^2}$. □

5. A basis for the homology of a knot complement with small Thurston complexity

Let Γ be a graph embedded in S^3 and let L be a framed link in the complement of Γ. Let M be the compact orientable 3-manifold given by removing an open regular neighbourhood of Γ and performing surgery along L. We are considering a knot K in M given by a diagram for $K \cup \Gamma \cup L$. In this section, we show how to compute bases for $H^1(M)$ and $H^1(M - N^\circ(K))$. From these, we will be able to construct a basis for $H_2(M - N^\circ(K), \partial M \cup \partial N(K))$ with relatively small Thurston complexity.

Our first step is to construct bases for $H^1(S^3 - N^\circ(\Gamma \cup L))$ and $H^1(S^3 - N^\circ(K \cup \Gamma \cup L))$ using the following lemma.

Lemma 5.1. *Let G be a graph in S^3, possibly with multiple edges between vertices and edge loops. Then $H^1(S^3 - N^\circ(G))$ has the following basis. Pick a maximal forest F in G. For each edge $e \in G - F$, orient it in some way and let L_e be the knot that starts at the initial vertex of e, runs along e and then back to the start of e through an embedded path in F. Let ψ_e be the homomorphism $\pi_1(S^3 - N^\circ(G)) \to \mathbb{Z}$ that sends a loop in $S^3 - N^\circ(G)$ to its linking number with L_e. Then $\{\psi_e : e$ is an edge of $G - F\}$ forms an integral basis for $H^1(S^3 - N^\circ(G))$.*

Proof. Note first that ψ_e really is a homomorphism $\pi_1(S^3 - N^\circ(G)) \to \mathbb{Z}$ since any homotopically trivial loop is sent to zero. These homomorphisms form linearly independent elements of $H^1(S^3 - N^\circ(G))$ because ψ_e evaluates to 1 on the meridian of e, but evaluates to 0 on the meridian of any other edge of $G - F$. By Alexander duality, $b_1(S^3 - N^\circ(G)) = b_1(G)$, which is equal to the number of edges in $G - F$. So, $\{\psi_e : e$ is an edge of $G - F\}$ forms a rational basis for $H^1(S^3 - N^\circ(G); \mathbb{Q})$. In fact, it forms an integral basis for $H^1(S^3 - N^\circ(G))$, for the following reason. Any element $\psi \in H^1(S^3 - N^\circ(G))$ is a linear combination $\sum_e \lambda_e \psi_e$, where each $\lambda_e \in \mathbb{Q}$. Since ψ is integral, its evaluation on the meridian of an edge e of $G - F$ is integral. But this number is λ_e. □

We now consider how to compute $H^1(M)$. We obtain M from $S^3 - N^\circ(\Gamma \cup L)$ by attaching solid tori, each of which can be viewed as a 2-cell and a 3-cell. So, $H^1(M)$ can be viewed as the subgroup of $H^1(S^3 - N^\circ(\Gamma \cup L))$ consisting of those classes that evaluate to zero on each of the surgery slopes of the framed link L. This subgroup can be expressed in terms of the *generalised linking matrix* of $\Gamma \cup L$.

Recall that the *linking matrix* for the oriented framed link $L = L_1 \cup \cdots \cup L_{|L|}$ is defined as the $|L| \times |L|$ symmetric matrix whose (i, j) entry is equal to $\ell k(L_i, L_j)$ when $i \neq j$, and is equal to the framing of L_i when $i = j$. Here $\ell k(L_i, L_j)$ is the linking number of L_i and L_j, where L_i and L_j are considered as disjoint knots in S^3. More generally, we define the *generalised linking matrix* A of $\Gamma \cup L$ to have rows given by the components of L and columns given by the components of L and also by the edges of $\Gamma - F$, where F is a maximal forest in Γ. For a component L_i of L and an edge e of $\Gamma - F$, the corresponding entry of A is $\ell k(L_i, L_e)$, where L_e is the knot defined as in Lemma 5.1. Similarly, for each component L_i of L and each component L_j of L with $i \neq j$, the corresponding entry of A is $\ell k(L_i, L_j)$. Finally, when $L_i = L_j$, the corresponding entry of A is the framing of L_i.

Let k be the number of columns of the generalised linking matrix. Thus, k is the sum of the number of components of L and the number of edges of $\Gamma - F$. In other words, $k = b_1(\Gamma \cup L)$. To make the notation more uniform, we identify the edges in $\Gamma - F$ by numbers $|L| + 1 \leq j \leq k$ and refer to L_e for $e \in \Gamma - F$ by L_j.

Lemma 5.2. *The cohomology group $H^1(M)$ is isomorphic to the subgroup of $H^1(S^3 - N^\circ(\Gamma \cup L)) = \mathbb{Z}^k$ given by the kernel of the generalised linking matrix.*

Proof. We have already identified $H^1(S^3 - N^\circ(\Gamma \cup L))$ by specifying an integral basis for it in Lemma 5.1. The subgroup $H^1(M)$ consists of those classes that evaluate to zero on each of the surgery slopes of the framed link L. When we write the surgery slope of L_i as a linear combination of the basis elements, the coefficients are precisely the entries of the ith row of the generalised linking matrix. □

We now wish to compute $H^1(M - N^\circ(K))$. This can be viewed as containing $H^1(M)$ as a subgroup, using the long exact sequence of the pair $(M, M - N^\circ(K))$:

$$0 = H^1(M, M - N^\circ(K)) \to H^1(M) \to H^1(M - N^\circ(K)) \to H^2(M, M - N^\circ(K)) \to H^2(M).$$

Now, using excision and Poincaré duality,

$$H^2(M, M - N^\circ(K)) \cong H^2(N(K), \partial N(K)) \cong H_1(N(K))$$

and $H^2(M) \cong H_1(M, \partial M)$. So, under our assumption that K is homologically trivial, the map $H^2(M, M - N^\circ(K)) \to H^2(M)$ is the trivial map. Therefore, $H^1(M - N^\circ(K))$ can be viewed as adding on a $\mathbb{Z}$ summand to $H^1(M)$. This summand is isomorphic to $H^2(M, M - N^\circ(K))$.

We now wish to construct an integral basis for $H^1(M - N^\circ(K))$. Such a basis can be found by starting with an integral basis for $H^1(M)$ and taking its image in $H^1(M - N^\circ(K))$, and then adding one more element. This extra element must map to a generator for $H^2(M, M - N^\circ(K))$. We will build explicit cocycles representing this basis.

Theorem 1.5. *Let M be a compact orientable 3-manifold given by removing an open regular neighbourhood of a graph Γ in S^3 and performing integral surgery on a framed link L in the complement of Γ. Let D be a fixed diagram for $\Gamma \cup L$ where the surgery slopes*

on L coincide with the diagrammatic framing. Let K be a homologically trivial knot in M, given by a diagram of $K \cup \Gamma \cup L$ that contains D as a sub-diagram. Let c be the total crossing number of K. Set $X = M - N^\circ(K)$ as the exterior of K in M. There is an algorithm that builds a triangulation of X with $O(c)$ tetrahedra, together with simplicial 1-cocycles $\phi_1, \cdots, \phi_b$ that form an integral basis for $H^1(X;\mathbb{Z})$ with $\sum_i C_{\text{nat}}(\phi_i)$ at most $O(c^2)$. The algorithm runs in time polynomial in c. All the above implicit constants depend on the manifold M and not the knot K.

Proof. **Step 1:** Building a triangulation of $S^3 - N^\circ(K \cup \Gamma \cup L)$.

We view S^3 as the union of $\mathbb{R}^3$ and a point at infinity. We will arrange for $K \cup \Gamma \cup L$ to sit inside $\mathbb{R}^3$ as specified by the diagram. Thus, the vertical projection map $\mathbb{R}^3 \to \mathbb{R}^2$ onto the first two co-ordinates will project $K \cup \Gamma \cup L$ onto the underlying planar graph specified by the diagram. Our triangulation will have the following properties:

(1) The number of tetrahedra is bounded above by a linear function of c.
(2) Each edge of the triangulation is straight in $\mathbb{R}^3$.
(3) The meridian of each component of $K \cup L$ and of each edge of Γ is simplicial, as is the surgery slope of each component of L.

There are many possible ways to build this triangulation. We will follow the recipe given by Coward and the first author in Section 4 of [7]. The triangulation provided by Theorem 4.3 of [7] has all the required properties when $\Gamma = \emptyset$. We only need to generalise to the situation where $\Gamma \neq \emptyset$ and show that the triangulation can be constructed algorithmically in polynomial time, as a function of c. We briefly review the steps in Section 4 of [7].

Step 1 is to embed the underlying planar graph G of the diagram into $\mathbb{R}^2$ as a union of straight arcs, as follows. We first modify Γ by expanding each vertex of Γ with valence more than 3 into a tree, so that each of the new vertices has valence exactly 3. This does not change the exterior of $K \cup \Gamma \cup L$. Let $\overline{G}$ be the graph obtained from G by collapsing parallel edges to a single edge and removing edge loops. Fáry's theorem says that $\overline{G}$ has an embedding in $\mathbb{R}^2$ where each edge is straight [13]; this was proved independently by Wagner [35] and Stein [30] as well. Such an embedding can be found in polynomial time using, for example, the algorithms of [8] or [27]. We place this embedded graph into the interior of a square Q. Now reinstate the parallel edges of G with 2 straight arcs each and the edge loops of G with 3 straight arcs. Step 2 is to replace each edge of G by 4 parallel edges, replace each 2-valent vertex of G by 4 vertices joined by 3 edges and replace each 4-valent vertex of G by 16 vertices arranged in a grid. Furthermore, each 3-valent vertex of G is replaced by a triangle. This is triangulated by placing a vertex in its interior and coning from that point. The result is a graph G_+ where each edge is still straight (see Figure 4) .

In Step 3, we triangulate the complementary regions of $G_+ \cup \partial Q$ by adding new edges. Let E be its 1-skeleton. In Step 4, we insert 4 copies of Q into the cube $Q \times I$, one being the top face, one the bottom face, and two parallel copies between them. Insert $E \times I$ into the cube $Q \times I$. This divides the cube into convex balls. We triangulate each face of each ball by adding a vertex to its interior and then coning off, and we then triangulate each ball by adding a vertex to its interior and then coning. We now modify this triangulation as follows. Near each crossing, there are 27 little cubes. We remove these and insert a new triangulation of the cube. A portion of this is shown in Figure 5. It is clear that if these are inserted in the correct way, we obtain a regular

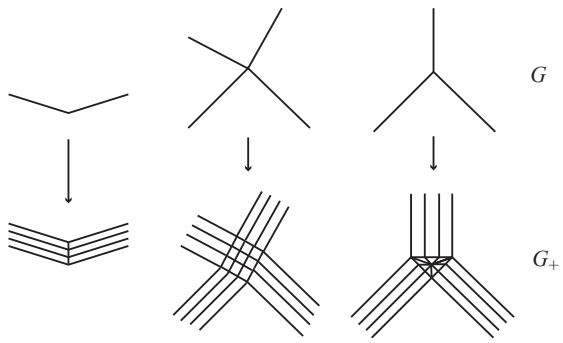

FIGURE 4. Forming G_+ from G

neighbourhood of $K \cup \Gamma \cup L$ as a simplicial subset of this triangulation. Removing the interior of this gives the required triangulation of the exterior of $K \cup \Gamma \cup L$. It is clearly constructible in polynomial time and has the required properties.

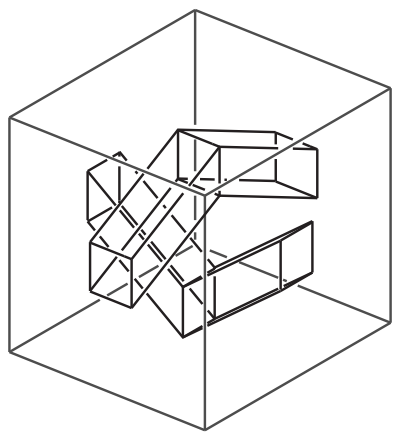

FIGURE 5. The triangulation near each crossing

Step 2: Building the triangulation of $M - N^\circ(K)$.

The above triangulation of $S^3 - N^\circ(K \cup \Gamma \cup L)$ extends to a triangulation of $M - N^\circ(K)$ in an obvious way. We need to attach a solid torus to each component of $\partial N(L)$. We do this by attaching a triangulated meridian disc along the surgery slope. We then attach on a 3-ball, which is triangulated as a cone on its boundary. This process is again completed in polynomial time, and the number of tetrahedra remains bounded above by a linear function of c.

As described above, we can form a basis for $H^1(M - N^\circ(K))$ by

(1) picking a basis for $H^1(M)$ and taking its image under the homomorhpism $H^1(M) \to H^1(M - N^\circ(K))$ induced by inclusion;
(2) adding one extra element that maps to a generator for $H^2(M, M - N^\circ(K))$.

Step 3: Defining $b_1(M)$ simplicial 1-cocycles on $S^3 - N^\circ(K \cup \Gamma \cup L)$.

We have already identified elements of $H^1(M)$ with integral solutions to the equation $A\beta = 0$, where A is the generalised linking matrix for $\Gamma \cup L$. Therefore, consider an integral solution $\beta = (\beta_1, \cdots, \beta_k)^T$ to the equation $A\beta = 0$. The corresponding cocycle $\sum_{i=1}^k \beta_i \psi_i$ is a 1-cocycle on $H^1(S^3 - N^\circ(\Gamma \cup L))$ that evaluates to zero on each surgery slope of the framed link. We can restrict this to a cocycle on $S^3 - N^\circ(K \cup \Gamma \cup L)$, which represents an element of $H^1(M - N^\circ(K))$.

More specifically, define the 1-cocycle c_β on $S^3 - N^\circ(K \cup \Gamma \cup L)$ as follows. Let T be a maximal tree in the 1-skeleton of the triangulation. For every edge $e \in T$ define $\langle c_\beta, e\rangle = 0$. For any oriented edge $e \notin T$, construct a loop ℓ_e that starts at the initial vertex of e, runs along e and then back to the start of e through an embedded path in T. Since we are assigning 0 to every edge contained in T, it should be clear that the numbers assigned to e and ℓ_e are the same. Define

$$\langle c_\beta, e\rangle := \sum_{i=1}^{k} \beta_i \, \ell k(\ell_e, L_i),$$

where β_i are integers. It is clear that this forms a 1-cocycle since each term $\ell k(\ell_e, L_i)$ is a 1-cocycle.

Step 4: Extending the simplicial 1-cocycles c_β to the triangulation of $M - N^\circ(K)$.

The triangulation of $M - N^\circ(K)$ is obtained by gluing triangulated solid tori to the triangulation of $S^3 - N^\circ(K \cup \Gamma \cup L)$, such that the restrictions of both triangulations to their common boundary, $\partial N(L)$, agree with each other. The manifold X is obtained by Dehn filling along L_i for $1 \le i \le |L|$. We can extend the cocycles over the attached solid tori since we started with β satisfying $A\beta = 0$. This can be achieved with linear size control over the values of newly added edges. It is also easy to see that $\langle c_\beta, m_K\rangle = 0$, where m_K is the meridian of K.

Step 5: Constructing the extra cocycle.

We construct an extra 1-cocycle on $M - N^\circ(K)$ that will form a generator for the summand of $H^1(M - N^\circ(K))$ corresponding to $H^2(M, M - N^\circ(K))$. This extra element, together with the cocycles that formed from a basis for $H^1(M)$, will provide the required basis for $H^1(M - N^\circ(K))$.

Denote by $\kappa = (\kappa_1, \cdots, \kappa_k)^T$ with $\kappa_i := \ell k(K, L_i)$ the vector encoding the linking numbers of K with L_i. We claim that the condition on K being homologically trivial in M, is equivalent to the linear equation $A\beta = -\kappa$ having an integral solution. The homology group $H_1(S^3 - N^\circ(\Gamma \cup L); \mathbb{Z})$ is freely generated by the meridians $\mu_1, \cdots, \mu_k$ encircling $L_1, \cdots, L_k$. For $1 \le i \le |L|$, denote by λ_i the longitude of L_i that has zero linking number with L_i. Then $H_1(M; \mathbb{Z})$ is obtained by adding the relations $a_{ii}\, \mu_i + \lambda_i = 0$, one for each component L_i of L. The latter relation is equivalent to

$$a_{ii}\, \mu_i + \sum_{j \neq i} \ell k(L_i, L_j) \mu_j = \sum_j a_{ij}\, \mu_j = 0.$$

In other words if we set $\mu = (\mu_1, \cdots, \mu_k)$ then the relations are obtained by putting the entries of the vector $A\mu$ equal to 0. Therefore, K being trivial in $H_1(M;\mathbb{Z})$ is exactly the condition that $A\beta = -\kappa$ has an integral solution for β.

Let θ be any integral solution to the linear equation $A\theta = -\kappa$. Define the 1-cocycle c_θ similar to c_β but with slight modification to make the evaluation on the meridian of K non-zero. More precisely

$$\langle c_\theta, e\rangle := \ell k(\ell_e, K) + \sum_{i=1}^{k} \theta_i \, \ell k(\ell_e, L_i).$$

The evaluation of c_θ on each surgery curve is zero, and the evaluation on the meridian of K is equal to 1. It therefore is sent to the generator of $H^2(M, M - N^\circ(K))$, under the map $H^1(M - N^\circ(K)) \to H^2(M, M - N^\circ(K))$.

Step 6: Analysing the computational cost of the algorithm.

The number of edges of the triangulation of $S^3 - N^\circ(K \cup \Gamma \cup L)$ is $O(c)$. A spanning tree T and the loops ℓ_e for $e \notin T$ can be found in polynomial time in the number of edges. The numbers $\ell k(\ell_e, L_i)$ can be computed as follows. We can construct the diagram $L \cup \Gamma \cup \ell_e$, since ℓ_e is a union of edges of the triangulation. Each edge of the triangulation is straight, and so when it is projected to the plane of the diagram, the image of ℓ_e is a union of straight arcs. We compute the linking number $\ell k(\ell_e, L_i)$ using the usual signed count over the crossings of ℓ_e with L_i. Each of the linking numbers is at most linear in the number of crossings c. This is because the triangulation of $S^3 - N^\circ(K \cup \Gamma \cup L)$ has $O(c)$ edges and each edge can contribute at most a consant number of crossings. Moreover the coordinates θ_i are at most linear in c and can be computed in polynomial time, as A is a fixed matrix. Therefore the constructed 1-cocycles on $S^3 - N^\circ(K \cup \Gamma \cup L)$ have C_{nat} at most $O(c^2)$ and can be computed in polynomial time. Extending these 1-cocycles over the attached triangulated solid tori can be done in polynomial time. Moreover, the extension keeps the total number of tetrahedra linear in c, and the total C_{nat} at most quadratic in c. □

6. Surfaces with trivial Thurston norm

Recall from Section 2.7 the definition of a fundamental normal surface. In this section, we will prove the following result.

Theorem 6.1. *Let $\mathcal{T}$ be a triangulation of a compact orientable irreducible 3-manifold X. If X has any compressible boundary components, suppose that these are tori. The subspace of $H_2(X, \partial X; \mathbb{Z})$ with trivial Thurston norm is spanned by a collection of fundamental normal tori, annuli and discs.*

As an immediate consequence, we obtain the following.

Theorem 1.7. *Let $\mathcal{T}$ be a triangulation of a compact orientable irreducible 3-manifold X. If X has any compressible boundary components, suppose that these are tori. Then there is a collection $w_1, \cdots, w_r$ of integral cocycles that forms a basis for the subspace W of $H^1(X;\mathbb{R})$ consisting of classes with Thurston norm zero with $\sum_i C_{\text{dig}}(w_i)$ at most $O(|\mathcal{T}|^2)$.*

Proof. By Theorem 6.1, there is a collection of fundamental normal surfaces that forms a generating set for $W \cap H^1(X;\mathbb{Z})$. Some subset of this collection therefore forms a

basis for W. Any fundamental surface S intersects each edge of $\mathcal{T}$ at most $t2^{7t+2}$ times, where $t = |\mathcal{T}|$, the number of tetrahedra of $\mathcal{T}$. Hence, when S is oriented, the cocycle w dual to S has evaluation at most $t2^{7t+2}$ on each edge. So $C_{\text{dig}}(w)$ is at most $O(t^2)$. □

We will prove Theorem 6.1 using results of Tollefson and Wang [33]. In that paper, X was required to be irreducible and its compressible boundary components were required to be tori. It is for this reason that these are also hypotheses of Theorem 6.1.

Definition 6.2. Let X be a compact orientable 3-manifold with a triangulation $\mathcal{T}$. A compact oriented normal surface F properly embedded in X is *lw-taut* if

(1) its homology class $[F]$ is non-trivial in $H_2(X, \partial X)$;
(2) it is χ_- minimising;
(3) there is no union of components of F that is homologically trivial;

and furthermore it has smallest weight among all normal surfaces in its homology class satisfying the above conditions.

Recall from Section 2.7 the definition of the projective solution space $\mathcal{P}(\mathcal{T})$. Recall also the notion of a normal surface being carried by a face of $\mathcal{P}(\mathcal{T})$.

Definition 6.3. A face of $\mathcal{P}(\mathcal{T})$ is *lw-taut* if every surface carried by the face is lw-taut.

The following result was proved by Tollefson and Wang (Theorem 3.3 and Corollary 3.4 in [33]).

Theorem 6.4. *Let $\mathcal{T}$ be a triangulation of a compact orientable irreducible 3-manifold X. If X has any compressible boundary components, suppose that these are tori. Let F be an lw-taut surface and let C be the minimal face of $\mathcal{P}(\mathcal{T})$ that carries F. Then C is lw-taut. Furthermore, there are unique orientations assigned to the surfaces carried by C such that if G and H are carried by C, then the normal sum $G + H$ satisfies $[G + H] = [G] + [H] \in H_2(X, \partial X)$ and $x([G + H]) = x([G]) + x([H])$.*

Proof of Theorem 6.1. Let T consist of those fundamental annuli, tori and discs that lie in some lw-taut face. Consider any element of $H_2(X, \partial X; \mathbb{Z})$ with trivial Thurston norm. This is represented by an lw-taut surface F. Let C be the minimal face of $\mathcal{P}(\mathcal{T})$ that carries F. By Theorem 6.4, C is an lw-taut face. Now, F is a normal sum of fundamental surfaces $G_1, \cdots, G_n$ that are also carried by C. By Theorem 6.4, they are all oriented surfaces. Since they are lw-taut and X is irreducible, no G_i is a sphere. By Theorem 6.4, $[F] = [G_1] + \cdots + [G_n]$ in $H_2(X, \partial X)$ and $0 = x([F]) = x([G_1]) + \cdots + x([G_n])$. Since Thurston norm is always non-negative, this implies that $x([G_i]) = 0$ for each i. As the G_i are oriented and lw-taut, they are discs, annuli and tori, and hence they are elements of T. □

7. Computational complexity of Thurston norm ball

In this section, we analyse the decision problem Thurston norm ball for $b_1 \leq B$ that was mentioned in the Introduction. We now define it precisely. The input is a triangulation $\mathcal{T}$ for a compact orientable 3-manifold X with first Betti number $b_1(X) \leq B$, and a collection of integral simplicial 1-cocycles $\{\phi_1, \cdots, \phi_b\}$ that forms a basis for $H^1(X; \mathbb{R})$. The problem asks to compute the unit ball for the Thurston semi-norm. Here we have identified $H^1(X; \mathbb{R})$ with $H_2(X, \partial X; \mathbb{R})$ using Poincaré duality. The output consists of the following two sets of data:

1) A collection of integral cocycles that forms a basis for the subspace W of $H^1(X;\mathbb{R})$ with Thurston norm zero. These are written as rational linear combinations of the given cocycles $\{\phi_1, \cdots, \phi_b\}$. Denote by p the projection map from $H^1(X;\mathbb{R})$ to $H^1(X;\mathbb{R})/W$.

2) A finite set of points $V \subset H^1(X;\mathbb{Q})$ such that $p(V)$ is the set of vertices of the unit ball of $H^1(X;\mathbb{R})/W$, together with a list $\mathcal{F}$ of subsets of V. The set $\mathcal{F}$ is the list of faces of the unit ball for $H^1(X;\mathbb{R})/W$. In other words, for $F \in \mathcal{F}$ the set

$$\{p(v) \ \mid \ v \in F\},$$

forms the set of vertices of some face of the unit ball for $H^1(X;\mathbb{R})/W$. Moreover, this covers all the faces as we go over all the elements of $\mathcal{F}$. Thus, the unit ball of $H^1(X;\mathbb{R})$ is the inverse image of the unit ball of $H^1(X;\mathbb{R})/W$ under the projection map p.

The complexity of the input is defined to be $|\mathcal{T}| + \sum_i C_{\text{nat}}(\phi_i)$. Recall that $|\mathcal{T}|$ is the number of tetrahedra of $\mathcal{T}$. As discussed in the Introduction, the fact that the complexity of ϕ_i is measured using C_{nat} rather than C_{dig} is definitely not standard. In order to simplify the notation a little, we let Φ be the matrix with columns $\phi_1, \cdots, \phi_b$. More specifically, it has b columns and has a row for each oriented edge of $\mathcal{T}$, and its (i,j) entry is the evaluation of ϕ_j on the ith edge. So $C_{\text{nat}}(\Phi) = \sum_i C_{\text{nat}}(\phi_i)$.

Theorem 1.4. *Fix an integer $B \geq 0$. The problem* Thurston norm ball for $b_1 \leq B$ *lies in* ***FNP****, where b_1 denotes the first Betti number.*

We will prove this over the next two sections. In this section, we will consider the following restricted version of the problem.

In Thurton norm ball for irreducible boundary-irreducible 3-manifolds with $b_1 \leq B$, we consider compact, orientable, irreducible, boundary-irreducible 3-manifolds X. We allow X to be disconnected. Thus, the input is a triangulation $\mathcal{T}$ for X with first Betti number $b_1(X) \leq B$, and a collection of simplicial integral 1-cocycles $\{\phi_1, \cdots, \phi_b\}$ that forms a basis for $H^1(X;\mathbb{R})$. The output is the data in (1) and (2) above.

Theorem 7.1. Thurton norm ball for irreducible boundary-irreducible 3-manifolds with $b_1 \leq B$ *is in* ***FNP****.*

Proof. Let $d = \dim(H^1(X;\mathbb{R})/W)$, and denote by $B_{\bar{x}}$ the unit ball of the induced Thurston norm $\overline{x}$ on $H^1(X;\mathbb{R})/W$:

$$B_{\bar{x}} = \{v \in H^1(X;\mathbb{R})/W \ \mid \ \overline{x}(v) \leq 1\}.$$

Then $B_{\bar{x}}$ is a convex polyhedron. The boundary, $\partial B_{\bar{x}}$, inherits a facial structure from $B_{\bar{x}}$, where the faces of $\partial B_{\bar{x}}$ correspond to faces of $B_{\bar{x}}$ except for the face $B_{\bar{x}}$ itself. In particular, top-dimensional faces of $\partial B_{\bar{x}}$ correspond to facets of $B_{\bar{x}}$, and from now on a top-dimensional face refers to a top-dimensional face of $\partial B_{\bar{x}}$. The plan of the proof is as follows:

(1) A basis for the subspace W consisting of classes with Thurston norm zero is given to us non-deterministically.
(2) The list of vertices V and faces $\mathcal{F}$ is given to us non-deterministically.
(3) We verify that for each face $F \in \mathcal{F}$, the vertices of F actually lie on the same face.

(4) Let P be the space obtained by patching together geometric realisations of given top-dimensional faces of $\partial B_{\bar{x}}$ along their common boundaries. We have the maps

$$P \xrightarrow{i} \partial B_{\bar{x}} \xrightarrow{\pi} S^{d-1},$$

where i is the inclusion (well-defined by (3)) and π is the radial projection onto the $(d-1)$-dimensional sphere S^{d-1}. We verify that the composition $\pi \circ i$ is a homeomorphism.

(5) We verify that the list of faces of $\partial B_{\bar{x}}$ is complete.

Step 1: A basis for W

By Theorem 1.7, there is a collection $w_1, \cdots, w_r$ of integral cocycles that forms a basis for the subspace W of $H^1(X;\mathbb{R})$ consisting of classes with Thurston norm zero and that satisfies $\sum_i C_{\text{dig}}(w_i) \leq O(|\mathcal{T}|^2)$. We assume that these simplicial cocycles are given to us non-deterministically. We can certify that the elements $w_1, \cdots, w_r$ have Thurston norm zero, using Theorem 1.3.

We express each w_i as a linear combination of the given cocycles $\phi_1, \cdots, \phi_b$, as follows. There is a coboundary map $\partial^* \colon C^0(\mathcal{T}) \to C^1(\mathcal{T})$ from 0-cochains to 1-cochains. There is a natural basis $x_1, \cdots, x_m$ for $C^0(\mathcal{T})$ where x_i is the 0-cocycle that evaluates to 1 on the ith vertex of $\mathcal{T}$ and evaluates to zero on the other vertices. We wish to solve

$$\alpha_1\phi_1 + \cdots + \alpha_b\phi_b + \beta_1\partial^*(x_1) + \cdots + \beta_k\partial^*(x_m) = w_i.$$

Using the Bareiss algorithm, this can be done in polynomial time as a function of $C_{\text{dig}}(\Phi)$ and $|\mathcal{T}|$. The resulting coefficients $\alpha_1, \cdots, \alpha_b$ have $C_{\text{dig}}(\alpha_i)$ at most a polynomial function of $C_{\text{dig}}(\Phi)$ and $|\mathcal{T}|$. We can also verify whether the cocycles $w_1, \cdots, w_r$ are linearly independent in $H^1(X;\mathbb{R})$.

In the remaining steps, we will certify that the induced Thurston semi-norm on $H^1(X;\mathbb{R})/W$ is indeed a norm, hence the basis elements actually generate W.

Step 2A: Bounding the number of faces and vertices of the Thurston unit ball

We are given the simplicial integral cocycles $\{\phi_1, \cdots, \phi_b\}$. From these, we can construct properly embedded oriented surfaces $S_1, \cdots, S_b$ that are Poincaré dual to $\phi_1, \cdots, \phi_b$, and whose total complexity, $\sum_i \chi_-(S_i)$, is at most $O(C_{\text{nat}}(\Phi))$. To see this geometrically, fix a 1-cocycle ϕ_i and consider an arbitrary simplicial triangle Δ in the triangulation. Assume that the numbers that ϕ_i associates to the edges of Δ are $a, b, c \geq 0$ such that $a = b + c$. We can draw $a = b + c$ normal arcs in Δ that intersect the edges of Δ in respectively a, b and c points. Given any tetrahedron, we can look at the drawn normal curves on its boundary triangles and place normal disks (triangles or squares) inside the tetrahedron with the given boundary curves. Construct an embedded surface S_i by putting together the normal disks together glued along the common boundaries of the tetrahedra. The constructed surface is Poincaré dual to the starting 1-cocycle, and $\chi_-(S_i)$ is at most a linear multiple of $C_{\text{nat}}(\phi_i)$.

By Theorem 1.6, the total number of faces and vertices of the Thurston unit ball for $H^1(X;\mathbb{R})/W$ are at most polynomial functions of $C_{\text{nat}}(\Phi)$. Note the degrees of these polynomials are bounded above by B^2, which is a fixed constant by our assumption.

Step 2B: Bounding the number of bits encoding the coefficients of the vertices of the Thurston unit ball

Lemma 7.2. *There is a set of points $V \subset H^1(X;\mathbb{Q})$ such that*

$$\{p(v) \mid v \in V\},$$

is the set of vertices of the unit ball for the Thurston norm on $H^1(X,\mathbb{R})/W$ with the following properties:

(1) *$|V|$ is at most a polynomial function of $C_{\mathrm{nat}}(\Phi)$;*
(2) *each element of V is $\gamma_1\phi_1 + \cdots + \gamma_b\phi_b$, for rational numbers $\gamma_1, \cdots, \gamma_b$ such that $\sum_i C_{\mathrm{dig}}(\gamma_i)$ is at most a polynomial function of $\log(C_{\mathrm{nat}}(\Phi)))$.*

Proof. Define $\mathcal{A}$ as the set of integral points in $H^2(X, \partial X; \mathbb{Z}) \otimes \mathbb{Q}$ with dual norm at most one. By the previous step, we can construct surfaces $S_1, \cdots, S_b$ Poincaré dual to $\phi_1, \cdots, \phi_b$ whose total complexity, $\sum \chi_-(S_i)$, is at most $O(C_{\mathrm{nat}}(\Phi))$. By Theorem 4.1, the size of $\mathcal{A}$ is at most a polynomial function of $O(C_{\mathrm{nat}}(\Phi))$. Let $v \in H^1(X;\mathbb{Q})$ be such that $p(v)$ is a vertex of the unit ball for $H^1(X;\mathbb{R})/W$. Then there are points $a_1, \cdots, a_r \in \mathcal{A}$ such that the set of points $z \in H^1(X;\mathbb{R})$ satisfying the equations

$$\langle a_1, PD(z)\rangle = 1,$$

$$\vdots$$

$$\langle a_r, PD(z)\rangle = 1,$$

coincides with the affine space $v + W$. Here $PD(z)$ is the Poincaré dual to z, and $a_1, \cdots, a_r$ can be chosen to be the set of vertices spanning the face of the dual unit ball that is dual to the vertex $p(v)$. Moreover, since $z \in H^1(X;\mathbb{R})$ lies inside a b-dimensional space, at most b of the above equations can be linearly independent; hence we may assume that $r \leq b$ by choosing a suitable subset of $\{a_1, \cdots, a_r\}$. Recall that the dual basis $\{e^1, \cdots, e^b\}$ for $H^2(X, \partial X; \mathbb{Z}) \otimes \mathbb{Q}$ is defined as

$$\langle e^i, [S_j]\rangle = \delta_{ij},$$

where δ_{ij} is the Kronecker function. From the proof of Theorem 4.1 we know that, if we write a_i in the basis $\{e^1, \cdots, e^b\}$ then the coefficients are integral, and their absolute values are bounded above by $O(C_{\mathrm{nat}}(\Phi))$. Hence for each $1 \leq i \leq r$ we can write

$$a_i = \sum_j \eta^i_j \, e^j,$$

with $|\eta^i_j| \leq O(C_{\mathrm{nat}}(\Phi))$. Since $\{\phi_1, \cdots, \phi_b\}$ is a basis for $H^1(X;\mathbb{R})$ we can write

$$z = \gamma_1\phi_1 + \cdots + \gamma_b\phi_b,$$

for real numbers γ_j. Now for $1 \leq i \leq r$ we have

$$1 = \langle a_i, PD(z)\rangle = \langle \sum_j \eta^i_j \, e^j, \sum_s \gamma_s [S_s]\rangle = \eta^i_1\gamma_1 + \cdots + \eta^i_b\gamma_b.$$

This gives a set of r linear equations for $\gamma_1, \cdots, \gamma_b$. The number of variables and the number of equations are bounded above by the constant B, and the total absolute values of the coefficients η^i_j is bounded above by a polynomial function of $C_{\mathrm{nat}}(\Phi)$. Therefore, there exists a rational solution, $z = \gamma_1\phi_1 + \cdots + \gamma_b\phi_b$, where the total number of bits for $(\gamma_1, \cdots, \gamma_b)$ is at most a polynomial function of $\log(C_{\mathrm{nat}}(\Phi))$, for example by the Bareiss algorithm. $\square$

Step 2C: The list of vertices, V, and faces $\mathcal{F}$
By Lemma 7.2, there is a set of points $V \subset H^1(X;\mathbb{Q})$ such that

$$\{p(v) \mid v \in V\},$$

is the set of vertices of the unit ball for the Thurston norm on $H^1(X,\mathbb{R})/W$, $|V|$ is at most a polynomial function of $C_{\text{nat}}(\Phi)$, and the total number of bits for writing each element of V in terms of the basis $\{\phi_1, \cdots, \phi_b\}$ is at most a polynomial function of $\log(C_{\text{nat}}(\Phi))$. Likewise, the number of faces of the unit ball for $H^1(X;\mathbb{R})/W$, that is $|\mathcal{F}|$, is bounded above by a polynomial function of $C_{\text{nat}}(\Phi)$. The sets V and $\mathcal{F}$ are part of the certificate, and are given to us non-deterministically. We use Theorem 1.3 to certify that each element of V has Thurston norm one.

Step 3: Certifying that the vertices of each face in $\mathcal{F}$ actually lie on the same face
Recall that the number of elements of $\mathcal{F}$ is at most polynomially large in $C_{\text{nat}}(\Phi)$. For any element $F \in \mathcal{F}$ with $F = \{u_1, \cdots, u_s\}$, we check that

$$x(K_1 u_1 + \cdots + K_s u_s) = K_1\, x(u_1) + \cdots + K_s\, x(u_s) = K_1 + \cdots + K_s,$$

for some positive integral choices of $K_1, \cdots, K_s$. Here x represents the Thurston norm on $H^1(X;\mathbb{R})$. It is clear that once proven, this implies that $\{u_1, \cdots, u_s\}$ lie on the same face.

We would like to choose each K_i such that the 1-cocycle $K_i u_i$ is integral. First we need to check that there is a choice of K_i that is not too large. To see this, we can write u_i in the integral basis $\phi_1, \cdots, \phi_b$ as

$$u_i = \alpha_1^i\, \phi_1 + \cdots + \alpha_b^i\, \phi_b,$$

and take K_i to be the product of denominators of α_j^i for $1 \leq j \leq b$. From Step 2 we know that the total number of digits of K_i is bounded above by a polynomial function of $C_{\text{nat}}(\Phi)$. The numbers K_i are part of the certificate, and are given to us non-deterministically.

Therefore, we can use Theorem 1.3 to certify that the Thurston norm of

$$K_1\, u_1 + \cdots + K_s\, u_s,$$

is $K_1 + \cdots + K_s$. This finishes the certification for each face F. Since the total number of faces is bounded above by a polynomial function of $C_{\text{nat}}(\Phi)$, we are done.

Step 4A: Decomposing each top-dimensional face in $\mathcal{F}$ into simplices.
The dimension of a face F is the maximum number m such that F has $m+1$ affinely independent vertices. Hence, we can compute the dimension of each face from the list of its vertices, in time that is bounded above by a polynomial function of $C_{\text{nat}}(\Phi)$. The boundary of the polytope can be subdivided to a triangulation without adding any new vertices. This can easily be proved by induction on the dimension of the faces. In particular, each top-dimensional face in $\mathcal{F}$ can be subdivided in this way so that along incident faces, their triangulations agree. Such a subdivision will be provided to us non-deterministically. Thus, for each top-dimensional face F in $\mathcal{F}$, we are provided with a collection of subsets of F, each consisting of d vertices, where d is the dimension of $H^1(X;\mathbb{R})/W$. The number of such subsets is at most $|F|^d$, which is at most $|V|^b$. Let Σ denote the collection of all these subsets, as we run over all top-dimensional faces.

Then the number of elements of Σ is at most a polynomial function of $C_{\text{nat}}(\Phi)$.

Step 4B: Certifying that the composition $\pi \circ i$ is injective.
It is enough to verify that any pair of simplices $\sigma_1, \sigma_2 \in \Sigma$ have disjoint interiors. Here and afterwards, we slightly abuse the notation by denoting the geometric realisation of σ_i by σ_i again. The condition $\sigma_1^\circ \cap \sigma_2^\circ$ is equivalent to $\text{Cone}(\sigma_1^\circ) \cap \text{Cone}(\sigma_2^\circ) = \emptyset$, since the restriction of the Thurston norm to both σ_1 and σ_2 is equal to 1. We now show how to verify the condition $\text{Cone}(\sigma_1^\circ) \cap \text{Cone}(\sigma_2^\circ) = \emptyset$ using Linear Programming. Assume that $\{u_1, \cdots, u_d\}$ forms the list of vertices of σ_1, and $\{y_1, \cdots, y_d\}$ forms the list of vertices of σ_2. We would like to check that

$$\alpha_1 u_1 + \cdots + \alpha_d u_d = \beta_1 y_1 + \cdots + \beta_d y_d \tag{4}$$

has no solution for $\alpha_i > 0$ and $\beta_j > 0$. However, Equation (4) has a solution with $\alpha_i > 0$ and $\beta_j > 0$ if and only if it has a solution with $\alpha_i \geq 1$ and $\beta_j \geq 1$, essentially by scaling. This is an instance of Linear Programming. Since no variables are required to be integers, it can be solved in polynomial time as a function of $C_{\text{nat}}(\Phi)$.

Step 4C: Certifying that the composition $\pi \circ i$ is a homeomorphism.
The certificate provides the vertices and faces of the boundary of the unit norm ball, and it provides a collection Σ of $(d-1)$-dimensional simplices. We have checked that the interiors of these simplices are disjoint. But we need to check that their union is the entire boundary of the unit norm ball. We check that P is a closed, oriented, pseudo-manifold as follows. For this purpose, we need to check the conditions of *purity*, *non-branching*, *connectivity* and *orientability*.

Purity condition: For each element σ of Σ, we check that $p(\sigma)$ actually forms the vertices of a $(d-1)$-dimensional simplex. To do this, we verify that its vertices form a linearly independent set in $H^1(X;\mathbb{R})/W$. This can be done in polynomial time in $|\mathcal{T}|$ and $C_{\text{nat}}(\Phi)$.

Non-branching condition: We check that every $(d-2)$-dimensional simplex appears in exactly two $(d-1)$-dimensional simplices. In other words, for each $(d-2)$-dimensional face of Σ, we check that it lies in exactly one other $(d-1)$-dimensional simplex in Σ. Since $|\Sigma|$ is bounded by a polynomial in $C_{\text{nat}}(\Phi)$, this can be checked in time that is polynomially bounded in $C_{\text{nat}}(\Phi)$.

Connectivity condition: For every pair of simplices σ_1 and σ_2 in Σ, we check that σ_1 and σ_2 can be connected by a path consisting of $(d-2)$- and $(d-1)$-dimensional simplices. We may assume that such a (minimal) path is given to us non-deterministically. Since $|\Sigma|$ is bounded by polynomials in $C_{\text{nat}}(\Phi)$, this can be checked in time that is polynomially bounded in $C_{\text{nat}}(\Phi)$.

Orientability: We specify an orientation of each simplex in Σ by specifying an ordering of its vertices. We check that this orientation is compatible with its orientation from $H^1(X;\mathbb{R})/W$, by checking that the matrix with columns given by the elements of Σ has positive determinant. For every two top-dimensional faces that share a $(d-2)$-dimensional face, we check that they are glued by an orientation-reversing map along their intersection.

We have now established that P is a closed, oriented, pseudo-manifold. We check that the map $\pi \circ i$ is injective and surjective, and hence a homeomorphism. Injectivity was established in Step 4B. To prove the surjectivity, it is enough to show that the degree

of the map ϕ is non-zero. Here we are using the fact that the degree is well-defined between compact, oriented pseudo-manifolds of the same dimension. Moreover, any such map that is not surjective has degree 0, since $S^{d-1} - \{\text{point}\}$ is contractible and the degree is invariant under homotopy. To check that the degree is non-zero in our case, note that the degree can be computed as the signed count of the points in P that map to a generic but fixed element in S^{d-1}. Since all of the signs agree by our construction, this signed count is always non-zero. This finishes Step 4 of the certification.

Step 5: Certifying that the list of faces of $\partial B_{\bar{x}}$ is complete. The maps $\pi \circ i$ and π are both homeomorphisms, hence so is the map $i\colon P \to \partial B_{\bar{x}}$. This implies that the list of top-dimensional faces used to construct the space P is the complete list of top-dimensional faces of $\partial B_{\bar{x}}$, otherwise the inclusion map i would not have been surjective.

For every face $F \in \mathcal{F}$ there are top-dimensional faces $F_1, \cdots, F_r \in F$ with $r \leq d$ such that

$$F = \bigcap_{i=1}^{r} F_i,$$

where we have considered faces as subsets of the vertices V. Moreover, any intersection as above determines a face. Hence, we may go over all subsets of size at most d of the set of top-dimensional faces, and verify that our list of faces is complete. □

Remark 7.3. Note that in defining the complexity of the basis for cohomology, we used C_{nat} rather C_{dig}. Although this was enough for the current application, it would be interesting to know if Thurston norm ball for $b_1 \leq B$ still lies in **NP** if we change the definiton of the complexity of the cohomology basis to C_{dig}.

8. Decomposing a triangulated manifold along spheres and discs

In the previous section, we proved Theorem 7.1. This established the main theorem in the case of irreducible boundary-irreducible 3-manifolds. In this section, we start to tackle the general case, by decomposing our given 3-manifold along spheres and discs.

The following is Theorem 11.4 and Addendum 11.5 of [18]. It provides a method for building a triangulation of the connected summands of a 3-manifold X. The input is a triangulation $\mathcal{T}$ of X, together with normal spheres S that specify the connected sum. The running time of the algorithm is bounded above in terms of the weight $w(S)$. Recall that this is the number of intersection points between S and the 1-skeleton of $\mathcal{T}$.

Theorem 8.1. *There is an algorithm that takes, as its input, the following data:*

1. *a triangulation $\mathcal{T}$ with t tetrahedra of a compact orientable 3-manifold X;*
2. *a vector (S) for a normal surface S in $\mathcal{T}$ that is a union of disjoint spheres;*
3. *a simplicial 1-cocycle c on $\mathcal{T}$.*

The output is a triangulation $\mathcal{T}'$ of $X \backslash\backslash S$ and a simplicial 1-cocycle c' on $\mathcal{T}'$ with the following properties:

1. *the number of tetrahedra in $\mathcal{T}'$ is at most $200t$;*
2. *the classes $[c']$ and $i^*([c])$ in $H^1(X \backslash\backslash S)$ are equal, where $i\colon X \backslash\backslash S \to X$ is the inclusion map;*
3. $C_{\text{nat}}(c') \leq 1200t\, C_{\text{nat}}(c)$.

The algorithm runs in time that is bounded above by a polynomial function of

$$t(\log w(S))(\log(C_{\mathrm{nat}}(c)+1)).$$

We need a small extension of this result.

Theorem 8.2. *Theorem 8.1 remains true if S is a union of disjoint spheres and discs.*

The proof is essentially identical, and we therefore only sketch it.

When S is a normal surface properly embedded in a compact orientable 3-manifold X with a triangulation $\mathcal{T}$, then $X\backslash\backslash S$ inherits a handle structure, as follows. One first dualises $\mathcal{T}$ to form a handle structure $\mathcal{H}$ for X. The normal surface S then determines a surface that is standard in $\mathcal{H}$, which means that it is disjoint from the 3-handles, it intersects each handle in discs and, in the case of a 1-handle or 2-handle, these discs respect the handle's product structure. Then, by cutting along this surface, each i-handle of $\mathcal{H}$ is decomposed into i-handles in the required handle structure. We call this the *induced* handle structure on $X\backslash\backslash S$.

We do not actually construct this handle structure in the proof of Theorem 8.1. The reason is that the number of handles (when S is closed) is at least $w(S)$. So it is not possible to build this handle structure in time that is bounded above by a polynomial function of $\log w(S)$. In the next definition, it is useful to think about $\mathcal{H}'$ as the induced handle structure on $X' = X\backslash\backslash S$ and where S' is the copies of S in $\partial X'$.

Definition 8.3. Let $\mathcal{H}'$ be a handle structure for a compact 3-manifold X'. Let S' be a compact subsurface of $\partial X'$ such that $\partial S'$ is disjoint from the 2-handles and respects the product structure on the 1-handles. A handle H of $\mathcal{H}'$ is a *parallelity handle* for (X', S') if it admits a product structure $D^2 \times I$ such that

(1) $D^2 \times \partial I = H \cap S'$;
(2) each component of intersection between H and another handle of $\mathcal{H}'$ is $\beta \times I$, for an arc β in ∂D^2.

The union of the parallelity handles is the *parallelity bundle*.

We will typically view the product structure $D^2 \times I$ on a parallelity handle as an I-bundle over D^2. It is shown in Lemma 5.3 of [17] that these I-bundle structures patch together to form an I-bundle structure on the parallelity bundle.

Definition 8.4. Let $\mathcal{B}$ be an I-bundle over a compact surface F. Its *horizontal boundary* $\partial_h\mathcal{B}$ is the (∂I)-bundle over F. Its *vertical boundary* $\partial_v\mathcal{B}$ is the I-bundle over ∂F. We say that a subset of $\mathcal{B}$ is *vertical* if it is a union of fibres, and that it is *horizontal* if it is a surface transverse to the fibres.

The main step in the proof of Theorem 11.4 in [18] was an application of the following result (Theorem 9.3 in [18]).

Theorem 8.5. *There is an algorithm that takes, as its input,*

(1) *a triangulation $\mathcal{T}$, with t tetrahedra, for a compact orientable 3-manifold X;*
(2) *a vector (S) for an orientable normal surface S;*

and provides as its output, the following data. If S' is the two copies of S in $\partial(X\backslash\backslash S)$, and $\mathcal{B}$ is the parallelity bundle for the pair $(X\backslash\backslash S, S')$ with its induced handle structure, then the algorithm produces a handle structure for $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ and, for each component B of $\mathcal{B}$, it determines:

(1) *the genus and number of boundary components of its base surface;*

(2) *whether B is a product or twisted I-bundle; and*
(3) *for each component A of $\partial_v B$, the location of A in $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$.*

It runs in time that is bounded by a polynomial in $t\log(w(S))$.

In the above, the meaning of the *location* of A is as follows. The intersection between A and each handle of $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ is a union of fibres in the I-bundle structure on A, and hence is a square. In the case when A lies entirely in $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$, then A is a union of these squares, and in this case, the algorithm provides these squares in the order they appear as one travels around A. However, A need not lie entirely in $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$. This arises in the situation where S has boundary. For example, if D and D' are normally parallel discs of S that are incident to the boundary of X, then the space between them becomes a parallelity handle $D^2\times I$ such that $\partial D^2\times I$ intersects ∂X. Thus, in this situation, A is decomposed into a union of squares, which are the components of intersection between A and the handles of $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ and also components of intersection between A and ∂X. The algorithm provides the squares lying in $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ in the order they appear as one travels around A.

Thus, the triangulation $\mathcal{T}'$ is constructed by decomposing each of the handles of $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ into tetrahedra and by giving a compatible triangulation of $\mathcal{B}$. The number of handles of $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ is bounded above by a linear function of t and each of these handles can intersect its neighbours in a very limited number of possibilities. Thus, it is not hard to triangulate $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$ using at most $100t$ tetrahedra. In addition, we may ensure that the intersection with $\partial_v\mathcal{B}$ is simplicial. The horizontal boundary of each component B of $\mathcal{B}$ is a planar surface, since S is a union of spheres and discs. Thus, the topology of B is determined entirely by the number of boundary components of its base surface and whether it is a twisted I-bundle or a product. It is shown that the total number of boundary components of the base surface of $\mathcal{B}$ is at most $10t$. Hence, it is not hard to construct the triangulation on $\mathcal{B}$ with at most $100t$ tetrahedra.

We now explain briefly how the cocycle c' is constructed. This is explained in Addendum 11.5 in [18].

For each oriented edge e in $\mathcal{T}'$, we need to define $c'(e)$. It is convenient to dualise c to form an oriented surface F properly embedded in X. We may assume that F is transverse to S and that the intersection between F and $\mathcal{B}$ is vertical in $\mathcal{B}$. If e lies in $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$, then we define $c'(e)$ to be the algebraic intersection number between e and $F\backslash\backslash S$. This therefore defines the restriction of c' to $\partial_v\mathcal{B}$. In the proof of Addendum 11.5 in [18], we replace F by any compact oriented surface F' that equals F in $(X\backslash\backslash S)\backslash\backslash\mathcal{B}$, that is vertical in $\mathcal{B}$ and that satisfies $\partial_v\mathcal{B}\cap F'=\partial_v\mathcal{B}\cap F$. It is shown how to do this while maintaining control over the number of intersections with the edges of $\mathcal{T}'$. In particular, the cocycle c' dual to F' satisfies $C_{\mathrm{nat}}(c')\leq 1200t\,C_{\mathrm{nat}}(c)$. Now, F' and F differ by a class that is represented by a vertical surface in $\mathcal{B}$ disjoint from $\partial_v\mathcal{B}$. In our situation, any such surface is dual to the trivial class in $H^1(X\backslash\backslash S)$, since S is a union of spheres and discs. Thus, in fact, $[c']$ and $i^*([c])$ are equal.

This completes the outline of the proof of Theorem 8.2. We will first apply it to essential spheres in X with the following property.

Definition 8.6. A collection of disjoint essential spheres S properly embedded in a 3-manifold X is *complete* if the manifold obtained from $X\backslash\backslash S$ by attaching a 3-ball to each spherical boundary component is irreducible.

The following was proved by King (Lemma 4 in [15]). King's result is stated for closed orientable 3-manifolds, but his argument extends immediately to compact orientable 3-manifolds with boundary. (See also Lemma 2.6 in [23]).

Theorem 8.7. *Let $\mathcal{T}$ be a triangulation of a compact orientable 3-manifold X with t tetrahedra. Then there is a complete collection of disjoint essential normal spheres in $\mathcal{T}$ with weight at most 2^{185t^2}.*

It might be possible to improve this estimate. It was shown by Jaco and Tollefson (Theorem 5.2 in [14]) that, when X is closed, it contains a complete collection of essential spheres, each of which is a vertex normal surface. (See Section 2.7 for the definition of a vertex normal surface.) By Lemma 3.2 in [12] a vertex normal surface has weight at most $28t2^{7t-1}$. However, the generalisation of Jaco and Tollefson's argument to manifolds with non-empty boundary does not seem so straightforward. In any case, Theorem 8.7 is sufficient for our purposes.

Jaco and Tollefson also proved the following result dealing with compression discs for the boundary (Theorem 6.2 in [14]). It refers to a *complete* collection of compressing discs, which means that the manifold obtained by compressing along these discs has incompressible boundary.

Theorem 8.8. *Let $\mathcal{T}$ be a triangulation of a compact orientable irreducible 3-manifold X. Then X has a complete collection of disjoint compressing discs, each of which is a vertex normal surface. Hence, each such disc has weight at most $28t2^{7t-1}$, and their total weight is at most $280t^2 2^{7t-1}$.*

The final estimate is a consequence of the well known result, essentially due to Kneser [16], that in any collection of more than $10t$ disjoint normal surfaces, at least two of the surfaces are parallel.

Proof of Theorem 1.4. We are given a triangulation $\mathcal{T}$ of the compact orientable 3-manifold X and a collection of integral simplicial cocycles $\phi_1, \cdots, \phi_b$ that forms a basis for $H^1(X;\mathbb{R})$. Our goal is to compute the Thurston norm ball. Recall that the required output is:

(1) A collection of elements that are integral linear combinations of $\phi_1, \cdots, \phi_b$. These will form a basis $\mathcal{B}$ for the subspace W of $H^1(X;\mathbb{R})$ with Thurston norm zero.
(2) A collection V of rational linear combinations of $\phi_1, \cdots, \phi_b$ that project to the vertices of the norm ball in $H^1(X;\mathbb{R})/W$.
(3) A collection $\mathcal{F}$ of subsets of V that form the faces.

These will all be part of our certificate. In addition, the following will also form our certificate:

(1) A normal surface S in $\mathcal{T}$, given via its vector (S), that is in fact a complete collection of disjoint essential spheres. It has weight at most 2^{185t^2} where $t = |\mathcal{T}|$.
(2) A triangulation $\mathcal{T}'$ for the manifold X' obtained by cutting along S and then attaching a 3-ball to each spherical boundary component.
(3) A collection of simplicial 1-cocycles $\phi'_1, \cdots, \phi'_b$ that are the images of $\phi_1, \cdots, \phi_b$ in $H^1(X')$ under the map $H^1(X) \to H^1(X \backslash\backslash S) \cong H^1(X')$.
(4) A normal surface D in $\mathcal{T}'$, given via its vector (D), that is in fact a complete collection of disjoint compression discs for $\partial X'$. It has weight at most $280|\mathcal{T}'|^2 2^{7|\mathcal{T}'|-1}$.

(5) A triangulation $\mathcal{T}''$ for $X'' = X'\backslash\backslash D$.
(6) A collection of simplicial 1-cocycles $\phi''_1, \cdots, \phi''_b$ that are the images of $\phi'_1, \cdots, \phi'_b$ in $H^1(X'')$.
(7) A certificate for the decision problem Thurton norm ball for irreducible boundary-irreducible 3-manifolds with $b_1 \leq B$, which provides the data for the Thurston norm ball of $H^1(X'')$. This data is a basis for the subspace W'' of $H^1(X''; \mathbb{R})$ with Thurston norm zero, together with the vertices V'' and faces $\mathcal{F}''$ for the norm ball in $H^1(X''; \mathbb{R})/W''$.

The certificate is verified as follows:

(1) Verification that S is a collection of spheres using the algorithm in [2].
(2) Verification that $\mathcal{T}'$ is a triangulation of X' and that $\phi'_1, \cdots, \phi'_b$ are the images of $\phi_1, \cdots, \phi_b$ in $H^1(X')$, using Theorem 8.1.
(3) Verification that D is a collection of discs using [2].
(4) Verification that $\mathcal{T}''$ is a triangulation of X'' and that $\phi''_1, \cdots, \phi''_b$ are the images of $\phi'_1, \cdots, \phi'_b$ in $H^1(X'')$ using Theorem 8.1.
(5) Verification that each component of X'' either is irreducible and boundary-irreducible or is a rational homology 3-sphere, using Corollary 1.9.
(6) Verification of the certificate for Thurton norm ball for irreducible boundary-irreducible 3-manifolds with $b_1 \leq B$ for the manifold X''. The components of X'' that are (possibly reducible) rational homology 3-spheres play no role here.
(7) Verification that we may write $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$ such that
i) the elements of $\mathcal{B}_1$ form a basis for the kernel of the map

$$H^1(X) \to H^1(X - N^\circ(S)) \cong H^1(X'),$$

where $N(S)$ is a tubular neighbourhood of S;
ii) the elements of $\mathcal{B}_2$ project to a basis for the kernel of the map

$$H^1(X') \to H^1(X'\backslash\backslash D) = H^1(X''),$$

and this projection is one-to-one;
iii) the elements of $\mathcal{B}_3$ project to a basis of W'' and this projection is one-to-one.
(8) Verification that the map $H^1(X) \to H^1(X'')$ sets up a bijection $V \to V''$ and a bijection $\mathcal{F} \to \mathcal{F}''$.

The input to Thurton norm ball for irreducible boundary-irreducible 3-manifolds with $b_1 \leq B$ requires a collection of integral cocycles that forms a basis for $H^1(X''; \mathbb{R})$. Although $\phi''_1, \cdots, \phi''_b$ might not form a basis, they do form a spanning set, and therefore some subset of them (which can easily be found) forms a basis.

The output provides integral cocycles that form a basis for the subspace W'' of norm zero. It also consists of a set of points V'' in $H^1(X''; \mathbb{Q})$ that give the vertices of the norm ball and a collection $\mathcal{F}''$ of subsets of V'' that give the faces. Looking at the long exact sequence of the pair $(X, X - N^\circ(S))$ we have

$$H^1(X, X - N^\circ(S)) \to H^1(X) \to H^1(X - N^\circ(S)) \to H^2(X, X - N^\circ(S)) \to \cdots$$

By excision and the Poincaré duality we have

$$H^1(X, X - N^\circ(S)) \cong H^1(N(S), \partial N(S)) \cong H_2(N(S)) \cong H_2(S).$$

Similarly $H^2(X, X - N^\circ(S)) \cong H_1(S) \cong 0$. Therefore, the above long exact sequence takes the form

$$H_2(S) \to H^1(X) \xrightarrow{p} H^1(X - N^\circ(S)) \to 0.$$

Thus, p is surjective. It is Thurston norm-preserving and its kernel is generated by spheres in S. Similarly, the map $H^1(X') \to H^1(X'')$ is surjective, norm-preserving and its kernel is generated by discs in D. Thus, we let $\mathcal{B}_1$ be a basis for the kernel of p. We let $\mathcal{B}_2$ be a collection of elements that are sent by p to a basis for the kernel of $H^1(X') \to H^1(X'')$. Finally, assume that $\mathcal{B}_3$ is a subset of $H^1(X)$ that projects to a basis for the subspace W'' of $H^1(X''; \mathbb{R})$ with Thurston norm zero. Then $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \mathcal{B}_3$ is a basis for the subspace W of $H^1(X; \mathbb{R})$ with Thurston norm zero. Now, there is an induced isomorphism from $H^1(X; \mathbb{R})/W$ to $H^1(X''; \mathbb{R})/W''$ which is norm-preserving. Thus, we may obtain the points V in $H^1(X; \mathbb{Q})$ by running through each element of V'' in $H^1(X''; \mathbb{Q})$ and picking a point in its inverse image. A set of points in V spans a face if and only if the corresponding points in V'' do. Thus, we obtain the required output for Thurton norm ball for $b_1 \leq B$.

We need to show that the certificate exists and can be verified in polynomial time.

By Theorem 8.7, there is a complete collection of disjoint essential normal spheres, S, in $\mathcal{T}$ with weight at most 2^{185t^2} where $t = |\mathcal{T}|$, the number of tetrahedra in $\mathcal{T}$. The normal coordinates of elements of S are part of the certificate, and are given to us non-deterministically. Now we may decompose the manifold along S and then attach balls to any resulting spherical boundary components. Let X' be the resulting irreducible 3-manifold. Theorem 8.1 guarantees that we may build a triangulation $\mathcal{T}'$ of X' with no more than $O(|\mathcal{T}|)$ tetrahedra, and simplicial 1-cocycles $\phi'_j \in H^1(X'; \mathbb{Z})$ such that the cohomology classes $i^*([\phi_j])$ and $[\phi'_j]$ are equal and $C_{\text{nat}}(\phi'_j)$ is bounded above by a polynomial function of $|\mathcal{T}|$ and $C_{\text{nat}}(\phi_j)$. Moreover, this procedure can be done in time that is a polynomial function of $bt(\log w(S))(\log(C_{\text{nat}}(\phi_j)) + 1))$, which is bounded above by a polynomial function of $|\mathcal{T}|$ and $C_{\text{nat}}(\Phi)$ by our assumption on the weight of S and the complexity of the homology basis.

By Theorem 8.8, there is a complete collection of compression discs for X' that are normal in $\mathcal{T}'$ and with weight at most $280|\mathcal{T}'|^2 2^{7|\mathcal{T}'|-1}$. Applying Theorem 8.2, we may cut along these discs, forming a 3-manifold X'' and obtain a triangulation $\mathcal{T}''$ and cocycles $\phi''_1, \cdots, \phi''_b$. As above, the number of tetrahedra is $O(|\mathcal{T}'|)$ and therefore $O(|\mathcal{T}|)$. The cocycles ϕ''_j have C_{nat} that is bounded above by a polynomial function of $|\mathcal{T}|$ and $C_{\text{nat}}(\Phi)$. The procedure may be completed in polynomial time.

Finally, the certificate for Thurton norm ball for irreducible boundary-irreducible 3-manifolds with $b_1 \leq B$ is verified in polynomial time. □

9. Other representations of the manifold and knot

In the decision problem Determining knot genus in the fixed 3-manifold M, the manifold M is given to us by means of a diagram D for $\Gamma \cup L$, where Γ is a graph in S^3 and L is a framed link, and K is specified by giving a diagram for $K \cup \Gamma \cup L$ that contains D as a subdiagram. This method of representing M and K is a natural one. However, it also played a critical role in the proof of Theorem 1.2, as the construction of an efficient basis for $H_2(M - N^\circ(K), \partial M \cup \partial N(K))$ relied on this presentation of M and K. So it is reasonable to consider other methods for representing M and K, and to ask whether the resulting decision problems still lie in **NP**.

For simplicity, we will focus on closed orientable 3-manifolds M, although much of our discussion does generalise to the case of non-empty boundary.

One way of specifying a closed orientable 3-manifold is by giving a Heegaard splitting for it. Here, we are given a closed orientable surface S, a union α of disjoint simple closed curves $\alpha_1, \cdots, \alpha_g$ in S and another collection β of disjoint simple closed curves $\beta_1, \cdots, \beta_g$ in S, with the property that $S - N^\circ(\alpha)$ and $S - N^\circ(\beta)$ are both planar and connected. We also assume that each component of $S - N^\circ(\alpha \cup \beta)$ is a disc. We suppose that M is obtained by attaching two handlebodies to S so that the curves α bound discs in one handlebody and the curves β bound discs in the other handlebody. We think of this presentation of M as fixed and given to us in some way, for example by specifying a triangulation of S in which the curves α and β are all simplicial.

We now wish to add K to the picture. We do this by specifying a diagram for K in S, in other words an immersed curve with generic double points at which under/over crossing information is specified. We also assume that this immersed curve intersects the α and β curves transversely. We call this a *diagram* for K. This specifies an embedding of K into $S \times [-1, 1]$ and hence into M, once we have agreed on the convention that the handlebody with discs attached to the α curves lies on the $S \times \{-1\}$ side. We say that the *total crossing number* of K is the sum of the number of crossings of K with itself and its number of intersections with the α and β curves. This is our measure of complexity for K.

Note that every knot K in M is specified by such a diagram, as follows. Each handlebody is a regular neighbourhood of a graph. We can isotope K off a small open regular neighbourhood of these two graphs. It then lies in the complement of this open neighbourhood, which is a copy of $S \times [-1, 1]$. The projection $S \times [-1, 1] \to S$ onto the first factor specifies the diagrammatic projection map. After a small isotopy, the image of K has only generic double point singularities, which form the crossings of K with itself.

Thus, we can phrase the following decision problem. We fix a Heegaard diagram for M in a closed orientable surface S, as above.

Problem: Determining knot genus in the fixed closed orientable 3-manifold M via a Heegaard diagram.
Input: A diagram of K in S, as above, and an integer $g \geq 0$ in binary.
Question: Is the genus of K equal to g?

Theorem 9.1. Determining knot genus in the fixed closed orientable 3-manifold M via a Heegaard diagram *lies in* ***NP***.

Remark 9.2. We briefly discuss the above requirement that each component of $S - N^\circ(\alpha \cup \beta)$ is a disc. This almost always happens automatically anyway. Indeed, if some component of $S - N^\circ(\alpha \cup \beta)$ is not a disc, then it contains an essential simple closed curve that bounds a disc in both handlebodies. The Heegaard splitting is then reducible. However, we can always ensure that each component of $S - N^\circ(\alpha \cup \beta)$ is a disc, by performing an isotopy to β. For if $S - N^\circ(\alpha \cup \beta)$ is not a union of discs, then we can pick a properly embedded essential arc in some component joining the β curves to the α curves, and then isotope the relevant β curve along it, to introduce two new intersection points between the α curves and the β curves. We call this a *finger move*. Repeating this process if necessary, we end with the required collection of α and β curves.

The reason for making this requirement is that it avoids the following scenario. Suppose that some component P of $S - N^\circ(\alpha \cup \beta)$ is not a disc. Then we could choose a diagram of some knot K to wind many times around P, plus possibly intersect ∂P. In this way, we would get infinitely many distinct diagrams, all with the same total crossing number. Thus, in this case, the total crossing number would not become a reasonable measure for the complexity of the diagram.

We will prove Theorem 9.1 by reducing Determining knot genus in the fixed closed orientable 3-manifold M via a Heegaard diagram to Determining knot genus in the fixed closed orientable 3-manifold M. In order to this, we need an algorithm to translate a diagram for a knot K in a Heegaard surface to a planar diagram for K lying in the complement of some surgery curves. This is provided by the following result.

Theorem 9.3. *Let S be a closed orientable surface with curves $\alpha = \alpha_1 \cup \cdots \cup \alpha_g$ and $\beta = \beta_1 \cup \cdots \cup \beta_g$ specifying a Heegaard splitting of M. Suppose that $S - N^\circ(\alpha \cup \beta)$ is a union of discs. Then there is a diagram D of a framed link L in S^3 that specifies a surgery description of M and that has the following property. Let K be a knot in M given via a diagram of K in S with total crossing number c. Then there is a diagram of a knot in the complement of L that is isotopic to K, that contains D as a subdiagram and that has total crossing number $O(c^2)$. This may be constructed in polynomial time as a function of c. Here, the implied constant depends only on M and the Heegaard splitting, and not on K.*

We start with the case of the *standard* Heegaard splitting for S^3. This has curves $\alpha_1, \cdots, \alpha_g$ and $\beta_1, \cdots, \beta_g$ satisfying $|\alpha_i \cap \beta_j| = \delta_{ij}$.

Lemma 9.4. *Let S be a closed orientable surface with genus g, equipped with curves that give the standard genus g Heegaard splitting for the 3-sphere. Let K be a knot given by a diagram in S with total crossing number c. Then there is a diagram for K in the plane with crossing number at most c^2. This may be constructed in polynomial time as a function of c. This remains true if K is a link with several components. Furthermore, some of its components may be framed via surface framing in S, in which case we can also require that the resulting planar diagram specifies the same framing on these components.*

Proof. Let c_K be the number of crossings in S between K and itself. Then the total crossing number c of K is

$$c_K + \sum_i |K \cap \alpha_i| + \sum_i |K \cap \beta_i|.$$

We will modify the given diagram of K in S so that it becomes disjoint from the α curves. So consider a curve α_i. We may isotope its intersection points with K so that they all lie in a small neighbourhood of the point $\alpha_i \cap \beta_i$. We may then isotope K across the disc bounded by β_i. This has the effect of removing these points of $\alpha_i \cap K$, but possibly introducing new crossings of K. Near each point of $K \cap \beta_i$, we get new $|\alpha_i \cap K|$ new crossings of K. Thus, after these modifications, the total crossing number of K is

$$c_K + \sum_i |K \cap \alpha_i|(|K \cap \beta_i| - 1)$$

which is clearly at most c^2. We now use this to create a diagram for K in the plane. We compress S along the curves $\alpha_1, \cdots, \alpha_g$. Since the diagram for K is now disjoint from these curves, the result is a diagram for K in the 2-sphere, and hence the plane. □

We now extend this to slightly more general Heegaard splittings for S^3.

Lemma 9.5. *Let S be a closed orientable surface with genus g. Let $\alpha = \alpha_1 \cup \cdots \cup \alpha_g$ be a union of disjoint simple closed curves that cut S to a planar connected surface. Let $\beta = \beta_1 \cup \cdots \cup \beta_g$ be another collection of disjoint simple closed curves with the same property. Suppose that there is an isotopy taking β to curves that, with α, form the standard Heegaard splitting for S^3. Let K be a knot given by a diagram in S with total crossing number c. Then there is a planar diagram for K in S^3 with crossing number at most c^2. This diagram may be constructed in a polynomial time as a function of c. Here, the implied constants depend on the curves α and β but not K. This remains true if K is a link with several components, some of which may be framed.*

Proof. We are assuming that there is an isotopy taking $\beta_1, \cdots, \beta_g$ to curves $\beta'_1, \cdots, \beta'_g$ satisfying $|\alpha_i \cap \beta'_j| = \delta_{ij}$. This isotopy may be performed by performing a sequence of *bigon* moves. Here, one has a disc D in S with the interior of D disjoint from α and β, and with ∂D consisting of a sub-arc of an α curve and a sub-arc of a β curve. The isotopy slides this β arc across D. We shall show how to create a new diagram for K in S when such a move is performed. This will have the property that the total crossing number of the new diagram is at most the total crossing number of the old diagram. Hence, after these moves are performed, we may construct a diagram for K in the plane with crossing number at most c^2, using Lemma 9.4.

Within the disc D, there is a portion of the diagram for K. We will pull this portion of the diagram entirely through α or through β, so that after this, the arcs of K within D run directly from α to β without any crossings. The choice of whether to slide this portion of the diagram through α or β is made so that it does not increase the number of crossings. Thus, if there are c_α crossings between K and α along ∂D, and c_β crossings between K and β along ∂D, then after this operation, the number of crossings between K and ∂D is $2\min\{c_\alpha, c_\beta\}$. Thus, the total crossing number of K has not gone up. After this, we may isotope β across D without changing the number of crossings. □

Lemma 9.6. *Let S be a closed orientable surface with genus g. Let α be disjoint simple closed curves that cut S to a planar connected surface. Let β be another collection of disjoint simple closed curves with the same property. Suppose that each component of $S - N^\circ(\alpha \cup \beta)$ is a disc. Let C be an essential simple closed curve in S. Then there is a constant $\lambda \geq 1$ with the following property. Let K be a link, some components of which may be framed, given by a diagram in S with total crossing number c. Let K' be obtained from K by Dehn twisting about C, and let β' also be obtained from β by Dehn twisting about C. Then the total crossing number of the diagram on S given by $K' \cup C$ with respect to the curves α and β' is at most $\lambda c + \lambda$. Moreover, this diagram may be constructed in polynomial time as a function of c.*

Proof. By assumption, each component of $S - N^\circ(\alpha \cup \beta)$ is a disc. We realise this as a convex Euclidean polygon with straight sides, where each side is parallel to an arc of intersection with α or β. We may assume that C intersects $\alpha \cup \beta$ minimally, and hence that its intersection with this disc consists of straight arcs. We isotope the diagram of K within this disc so that most of it lies very close to one of the edges of the polygon and is distant from C. We also ensure that the remainder of the diagram consists of

straight arcs. Each intersection point between K and C lies in a straight arc of K, and this straight arc has an endpoint on $\alpha \cup \beta$. Thus, there is a constant $\lambda_1 > 0$, depending on α, β and C, such that the number of crossings between K and C is at most $\lambda_1 c$. We now perform the Dehn twist about C, giving the link K' and the curves β'. The intersection points between K' and β' correspond to the intersection points between K and β. The crossings of K' with itself correspond to the crossings of K with itself. Each crossing between K and C gives $|C \cap \alpha|$ extra crossings between K' and α. Thus, the total crossing number of K goes up by a factor of at most $1 + \lambda_1 |C \cap \alpha|$. We also need to consider the crossings involving C. There are at most $\lambda_1 c$ of these with K', and at most a constant number with $\alpha \cup \beta'$. The required bound then follows. □

Remark 9.7. In the above lemma, we made the hypothesis that each component of $S - N^\circ(\alpha \cup \beta)$ is a disc. We would like to ensure that α and β' have the same property, in other words that each component of $S - N^\circ(\alpha \cup \beta')$ is a disc. However, this might not be the case. Near C, there are various components of $S - N^\circ(\alpha \cup \beta)$. The components of $S - N^\circ(\alpha \cup \beta')$ are obtained by cutting along C and then possibly gluing some of these together in a different way. An example is shown in Figure 6, where a component of $S - N^\circ(\alpha \cup \beta')$ is obtained from two components of $S - N^\circ(\alpha \cup \beta \cup C)$ glued together. However, if this process does create some components of $S - N^\circ(\alpha \cup \beta')$ that are not discs, they may be cut into discs using finger moves, as in Remark 9.2. The number of finger moves that are needed is at most $|C \cap (\alpha \cup \beta)|$. This has the effect of increasing the total crossing number of $K' \cup C$ by at most $2|K \cap C|$, which is at most a constant times c.

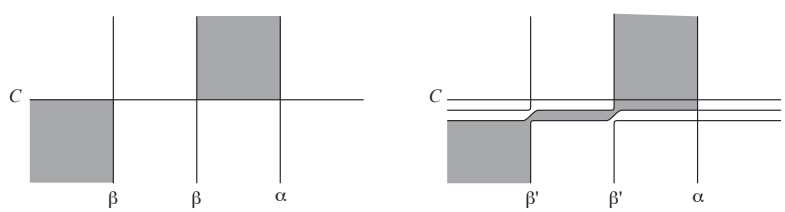

FIGURE 6. Dehn twisting β along C

Proof of Theorem 9.3. We are given a closed orientable surface S with curves α and β specifying a fixed Heegaard splitting of M. We are also given a diagram in S of a knot K with total crossing number c. We will change the diagram and the Heegaard splitting in a sequence of modifications. There is an orientation-preserving homeomorphism of S taking the curves $\beta = \beta_1 \cup \cdots \cup \beta_g$ to curves $\beta'' = \beta''_1 \cup \cdots \cup \beta''_g$ that satisfy $\alpha_i \cap \beta''_j = \delta_{ij}$. This homeomorphism is obtained by a product of Dehn twists about simple closed curves in S, followed by an isotopy. We can apply such a Dehn twist if we also add a surgery curve C that undoes it. Thus, we can replace the knot K and curves β by a knot K' together with the framed surgery curve C, and the curves β' obtained by Dehn twisting along C. By Lemma 9.6, the new knot K' and surgery curve C' have total crossing number bounded above by a constant times c. (The additive constant in the lemma can be subsumed into the multiplicative constant since we can assume that the total crossing number is non-zero.) By Remark 9.7, we can also ensure that each component of $S - N^\circ(\alpha \cup \beta')$ is a disc, at a cost of increasing the total crossing number of $K' \cup C$ by at most a constant factor. Repeating this for each Dehn twist in the sequence, we

end with curves β', a diagram for K and the framed link L specifying the surgery. This has total crossing number that is at most $O(c)$. The curves β' are isotopic to β'', and so by Lemma 9.5, we obtain a planar diagram for $K \cup L$ with total crossing number that is at most $O(c^2)$. □

This completes the proof of Theorem 9.1.

Remark 9.8. There is another possible way of representing M and K using triangulations. We could be simply given a triangulation for M containing K as a subcomplex. We would be told that this was indeed a triangulation of M. But in the absense of an efficient method converting this to a *fixed* triangulation of M, it is hard to see how this could be useful.

References

[1] Sergei Ivanovich Adian. The unsolvability of certain algorithmic problems in the theory of groups. *Tr. Mosk. Mat. Obs.*, 6:231–298, 1957.

[2] Ian Agol, Joel Hass, and William Thurston. The computational complexity of knot genus and spanning area. *Transactions of the American Mathematical Society*, 358(9):3821–3850, 2006.

[3] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[4] Erwin H Bareiss. Sylvester's identity and multistep integer-preserving Gaussian elimination. *Mathematics of computation*, 22(103):565–578, 1968.

[5] Arne Brondsted. *An introduction to convex polytopes*, volume 90. Springer Science & Business Media, 2012.

[6] A Candel and L Conlon. *Foliations II, Graduate Studies in Mathematics*, volume 60. American Mathematical Society, 2003.

[7] Alexander Coward and Marc Lackenby. An upper bound on Reidemeister moves. *American Journal of Mathematics*, 136(4):1023–1066, 2014.

[8] H. de Fraysseix, J. Pach, and R. Pollack. How to draw a planar graph on a grid. *Combinatorica*, 10(1):41–51, 1990.

[9] Max Dehn. Über die topologie des dreidimensionalen raumes. *Mathematische Annalen*, 69(1):137–168, 1910.

[10] David Gabai. Foliations and the topology of 3-manifolds. *Journal of Differential Geometry*, 18(3):445–503, 1983.

[11] Wolfgang Haken. Theorie der normal flachen. *Acta Math.*, 105:245–375, 1961.

[12] Joel Hass and Jeffrey C. Lagarias. The number of Reidemeister moves needed for unknotting. *J. Amer. Math. Soc.*, 14(2):399–428, 2001.

[13] Fáry István. On straight-line representation of planar graphs. *Acta scientiarum mathematicarum*, 11(229-233):2, 1948.

[14] William Jaco and Jeffrey L. Tollefson. Algorithms for the complete decomposition of a closed 3-manifold. *Illinois J. Math.*, 39(3):358–406, 1995.

[15] Simon A. King. The size of triangulations supporting a given link. *Geom. Topol.*, 5:369–398, 2001.

[16] Hellmuth Kneser. Geschlossene flächen in dreidimensionalen mannigfaltigkeiten. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 38:248–259, 1929.

[17] Marc Lackenby. The crossing number of composite knots. *Journal of Topology*, 2(4):747–768, 2009.

[18] Marc Lackenby. The efficient certification of knottedness and Thurston norm. *arXiv preprint arXiv:1604.00290*, 2016.

[19] Hendrik W Lenstra Jr. Integer programming with a fixed number of variables. *Mathematics of operations research*, 8(4):538–548, 1983.

[20] WB Raymond Lickorish. A representation of orientable combinatorial 3-manifolds. *Ann. Math*, 76(2):531–540, 1962.

[21] Sergei Matveev. *Algorithmic topology and classification of 3-manifolds*, volume 9. Springer, 2007.

[22] William Meeks, III, Leon Simon, and Shing Tung Yau. Embedded minimal surfaces, exotic spheres, and manifolds with positive Ricci curvature. *Ann. of Math. (2)*, 116(3):621–659, 1982.

[23] Aleksandar Mijatović. Simplifying triangulations of S^3. *Pacific J. Math.*, 208(2):291–324, 2003.

[24] Petr Sergeevich Novikov. On the algorithmic unsolvability of the word problem in group theory. *Trudy Mat. Inst. Steklov.*, 44:3–143, 1955.
[25] Michael O. Rabin. Recursive unsolvability of group theoretic problem. *Annals of Math.*, 67(1):172–194, 1958.
[26] Elaine Rich. *Automata, computability and complexity: theory and applications*. Pearson Prentice Hall Upper Saddle River, 2008.
[27] W. Schnyder. Embedding planar graphs on the grid. In *Proc. 1st ACM-SIAM Sympos. Discrete Algorithms*, pages 138–148, 1990.
[28] H Seifert and W Threlfall. *A textbook of topology*, volume 89. Academic Press, 1980.
[29] Herbert Seifert. Über das geschlecht von knoten. *Mathematische Annalen*, 110(1):571–592, 1935.
[30] Sherman K Stein. Convex maps. *Proceedings of the American Mathematical Society*, 2(3):464–466, 1951.
[31] William P Thurston. A norm for the homology of 3-manifolds. *Mem. Amer. Math. Soc.*, 339:99–130, 1986.
[32] Heinrich Tietze. Über die topologischen invarianten mehrdimensionaler mannigfaltigkeiten. *Monatshefte für Mathematik und Physik*, 19(1):1–118, 1908.
[33] Jeffrey L. Tollefson and Ningyi Wang. Taut normal surfaces. *Topology*, 35(1):55–75, 1996.
[34] Joachim von zur Gathen and Malte Sieveking. A bound on solutions of linear integer equalities and inequalities. *Proc. Amer. Math. Soc.*, 72(1):155–158, 1978.
[35] Klaus Wagner. Bemerkungen zum vierfarbenproblem. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 46:26–32, 1936.
[36] Andrew H Wallace. Modifications and cobounding manifolds. *Canadian Journal of Mathematics*, 12:503–528, 1960.

Mathematical Institute, University of Oxford,
Woodstock Road, Oxford OX2 6GG, United Kingdom