msp

Geometry & Topology 20 (2016) 2053–2078

Cusp volumes of alternating knots

MARC LACKENBY
JESSICA S PURCELL

We show that the cusp volume of a hyperbolic alternating knot can be bounded above and below in terms of the twist number of an alternating diagram of the knot. This leads to diagrammatic estimates on lengths of slopes, and has some applications to Dehn surgery. Another consequence is that there is a universal lower bound on the cusp density of hyperbolic alternating knots.

57M25, 57M50

1 Introduction

A major goal in knot theory is to relate geometric properties of the knot complement to combinatorial properties of the diagram. In this paper, we relate the diagram of a hyperbolic alternating knot to its cusp area, answering a question asked by Thistlethwaite. He observed using computer software such as SnapPea [29; 11] and other programs, including one developed by Thistlethwaite and Tsvietkova [26], that all alternating knots with small numbers of crossings have cusp area related to their twist number. This led him to ask if all alternating knots have cusp area bounded below by a linear function of their twist number. This question is finally settled in the affirmative in this paper. Specifically, we show the following.

Theorem 1.1 *Let D be a prime, twist reduced alternating diagram for some hyperbolic knot K, and let* $\mathrm{tw}(D)$ *be its twist number. Let C denote the maximal cusp of the hyperbolic 3–manifold $S^3 \setminus K$. Then*

$$A(\mathrm{tw}(D) - 2) \leq \mathrm{Area}(\partial C) < 10\sqrt{3}(\mathrm{tw}(D) - 1),$$

where A is at least 2.278×10^{-19}.

The terms twist reduced, twist number, and maximal cusp are defined below.

Theorem 1.1 has the following almost immediate corollary concerning the length of slopes on a cusp of an alternating knot.

Published: 15 September 2016 DOI: 10.2140/gt.2016.20.2053

Corollary 1.2 *Every slope, except possibly the meridional slope, on the maximal cusp of a hyperbolic alternating knot has length at least* $B(\mathrm{tw}(D)-2)$, *where* B *is at least* 7.593×10^{-20} *and* D *is a prime, twist reduced alternating diagram of the knot.*

This result has several applications. We mention the following, purely topological consequence.

Theorem 1.3 *For any closed* 3*–manifold* M *with sufficiently large Gromov norm, there are at most finitely many prime alternating knots* K *and fractions* p/q *such that* M *is obtained by* p/q *surgery along* K.

This is related to [17, Problem 3.6] in Kirby's list, which is attributed to Clark: is there a closed 3–manifold M that can be obtained by Dehn surgery along infinitely many distinct knots in the 3–sphere? This was answered affirmatively by Osoinach [23], who provided examples of such 3–manifolds M (see also Abe, Jong, Luecke and Osoinach [1]). However, Theorem 1.3 gives a negative answer to the version of Kirby's problem that restricts to prime alternating knots and 3–manifolds with large Gromov norm.

Alternating diagrams

We now recall various definitions concerning alternating diagrams that we will use throughout the paper.

In a diagram of a knot or link, a *twist region* is a (possibly empty) string of bigon regions in the 4–valent diagram graph, such that the knot is alternating within this region, and such that the string is maximal, in the sense that no other bigons of the diagram lie on either side. A single crossing adjacent to no bigons is also a twist region.

The *twist number* $\mathrm{tw}(D)$ of a diagram D is its number of twist regions. See Figure 1.

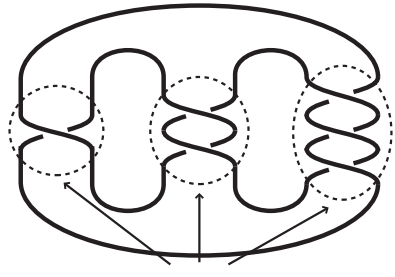

Figure 1: A diagram with twist number 3 (figure adapted from [19])

Two crossings are *twist equivalent* if there is a simple closed curve in the projection plane of the diagram that meets the diagram exactly at the two crossings, and at each of these crossings, runs between opposite regions.

Note that we can apply a sequence of flypes to an alternating diagram to put all twist equivalent crossings together into a single twist region. When this is done, we say the diagram is *twist reduced*. More precisely, we say a diagram is *twist reduced* if for any simple closed curve meeting the diagram at precisely two crossings, and such that, at each of these crossings, it runs between opposite regions, the curve encloses a string of bigons on one side. By the above observation, every alternating diagram has a twist reduced alternating diagram. See Figure 2.

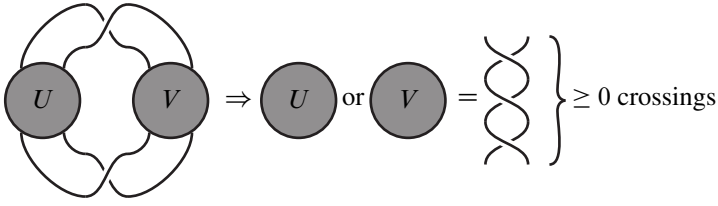

Figure 2: A twist reduced diagram (figure from [19])

The *twist number* $\mathrm{tw}(K)$ of a prime alternating knot K is the number of twist regions in a prime, twist reduced diagram. The twist number can be seen to be a well-defined invariant of an alternating knot in several ways, for example by using the invariance of "characteristic squares" under flyping as in Lackenby [19, Section 4] and then applying the solution to the Tait flying conjecture by Menasco and Thistlethwaite [22], or by relating twist number to the Jones polynomial as in Dasbach and Lin [12].

Cusps

Menasco [21] showed that any nonsplit prime alternating link that is not a $(2,q)$–torus link is hyperbolic, using results of Thurston [28] on the geometrisation of Haken manifolds. One important feature of a hyperbolic knot or link complement is its cusp(s), and several key geometric invariants arise from the study of cusps of hyperbolic 3–manifolds.

Definition 1.4 If M is a finite-volume hyperbolic 3–manifold, then it has ends of the form $T^2 \times [1,\infty)$. Each end can be realised geometrically as the image $p(H)$ of some horoball H in $\mathbb{H}^3$, where $p\colon \mathbb{H}^3 \to M$ is the covering map. This is known as a *cusp*. For each end, there is a 1–parameter family of cusps that are obtained from each other by changing the choice of horoball H, but keeping the same limiting point on

the sphere at infinity. If the cusps are expanded in the hyperbolic manifold until each is tangent to itself or another, the result is called a choice of *maximal cusps*. When the manifold has just one end, for example in the case of a hyperbolic knot complement, then there is a unique maximal cusp. When we speak of *the* cusp volume of a hyperbolic alternating knot, we mean the volume of this maximal cusp. Similarly, the *cusp area* is the Euclidean area of the torus that is the boundary of the maximal cusp. By an exercise in hyperbolic geometry, the cusp area is twice the cusp volume.

Finally, the *cusp density* of M is defined to be the largest possible volume in a choice of maximal cusps divided by the total volume of the manifold.

Cusp density

Our results have interesting consequences for cusp density. By work of Böröczky [8], the cusp density of a hyperbolic knot or link complement is at most $\sqrt{3}/(2v_3)$, where $v_3 \approx 1.01494$ is the volume of an ideal regular tetrahedron. The cusp density of the figure–8 knot realises this bound (see Thurston [27]). It follows from work of Eudave-Muñoz and Luecke [13] that general knots in S^3 may have arbitrarily small cusp density. However, Theorem 1.1 implies that alternating knots have universally bounded cusp density.

Corollary 1.5 *The cusp density of an alternating knot has a universal lower bound.*

This is a consequence of Theorem 1.1, because the cusp volume of $S^3 \setminus K$ is at least $(A/2)(\mathrm{tw}(K)-2)$, whereas it is a theorem of Lackenby, with improvements due to Agol and D Thurston [19], that the hyperbolic volume of $S^3 \setminus K$ is at most $10v_3(\mathrm{tw}(K)-1)$.

Slope length

Definition 1.6 A *slope* on a torus is an isotopy class of essential simple closed curves. If M is a compact 3–manifold with interior admitting a complete, finite volume hyperbolic structure, and C is a choice of maximal cusps, then a slope σ on ∂M has an associated *length*. This is defined to be the length of any Euclidean geodesic representative for σ on ∂C.

Cusp area and slope length are closely related, since high cusp area forces every slope, except possibly one, to be long.

In the case of the meridional slope for a knot in the 3–sphere, its length is well known to be bounded, but nevertheless, several interesting questions about meridional length remain open. It was shown by Thurston [27] that with respect to any maximal cusp,

the length of any slope is at least one. By work of Agol [6] and Lackenby [18], the length of the meridian of a knot in S^3 can be no more than six. For an alternating knot, the meridian has length less than three (see Adams, Colestock, Fowler, Gillam and Katerman [5]). In fact, no alternating knots have been found with meridional length more than two (see Adams [4]).

While the meridian of a knot in S^3 has universally bounded length, the next shortest curve on a cusp does not. Several different results in the past gave evidence that for alternating knots, its length is bounded below by the twist number. For example, Lackenby [18] showed that, in a combinatorial metric on the cusp torus, the second shortest curve had length bounded below by a linear function of the twist number. Unfortunately, the combinatorial metric that Lackenby used cannot be applied to the Euclidean metric on the torus in any known way, and so Lackenby's result does not give geometric information on the cusp. Purcell [24] showed that for knots with a high number of crossings in each twist region, including alternating knots, the second shortest curve had length bounded below by a constant times the twist number. However, this result gave no information about knots with a small number of crossings in any twist region. Futer, Kalfagianni and Purcell [15] showed the cusp area is bounded below by twist number for 2–bridge knots, which are a subset of alternating knots.

Corollary 1.2 gives that the length of every nonmeridional slope for a hyperbolic alternating knot must be long, unless the knot lies in one of the limited families with bounded twist number. Corollary 1.2 is an immediate consequence of Theorem 1.1, as follows. For suppose that σ is some nonmeridional slope on the boundary of a maximal cusp C, and let μ be the meridian. Let $L(\sigma)$ and $L(\mu)$ be their lengths, and denote the modulus of their intersection number by $\Delta(\sigma,\mu)$. Then, an elementary calculation in Euclidean geometry (for example as in the proof of [10, Lemma 2.1]) gives that

$$L(\sigma)L(\mu) \geq \text{Area}(\partial C)\Delta(\sigma,\mu).$$

The length of μ is at most 3, by [5], and so we deduce that

$$L(\sigma) \geq \frac{\text{Area}(\partial C)\Delta(\sigma,\mu)}{L(\mu)} \geq B(\text{tw}(K)-2),$$

where B is greater than 7.593×10^{-20}.

Applications to Dehn surgery

The above results on slope length are significant because slope length has important consequences for Dehn surgery. It is a well-known consequence of the 6–theorem of Agol [6] and Lackenby [18], together with Perelman's proof of the geometrisation conjecture, that if M is a finite-volume hyperbolic 3–manifold and $s_1,\dots,s_k$ are slopes on

distinct components of ∂M, such that, with respect to some horoball neighbourhood C of the cusps, each s_i has length more than 6, then the manifold $M(s_1, \dots, s_k)$ obtained by Dehn filling along these slopes is hyperbolic. It was shown by Futer, Kalfagianni and Purcell [14] that if these slopes all have length at least $\ell_{\min} > 2\pi$ on ∂C, then the hyperbolic volume of $M(s_1, \dots, s_k)$ satisfies

$$\text{Volume}(M(s_1, \dots, s_k)) \geq \left(1 - \left(\frac{2\pi}{\ell_{\min}}\right)^2\right)^{3/2} \text{Volume}(M).$$

Therefore, applying Corollary 1.2, we deduce the second inequality in the following result.

Theorem 1.7 *Let K be a hyperbolic alternating knot. If $\text{tw}(K) \geq 1.361 \times 10^{20}$, then any manifold M obtained by a nontrivial surgery along K is hyperbolic and satisfies*

$$\begin{aligned} \tfrac{1}{2} v_8(\text{tw}(K)/2 - 1) &\leq \tfrac{1}{2} \text{Volume}(S^3 \setminus K) \\ &\leq \text{Volume}(M) \\ &< \text{Volume}(S^3 \setminus K) \leq 10 v_3(\text{tw}(K) - 1), \end{aligned}$$

where $v_8 \approx 3.66386$ is the volume of a regular hyperbolic ideal octahedron, and $v_3 \approx 1.01494$ is the volume of a regular hyperbolic ideal tetrahedron.

Two of the remaining inequalities are upper and lower bounds on the volume of hyperbolic alternating link complements, due to Lackenby [19], with improvements by Agol and D Thurston [19], and Agol, Storm and W Thurston [7]. The third inequality is the fact that hyperbolic volume decreases when Dehn filling is performed (see Thurston [27]).

We can also deduce Theorem 1.3 from Theorem 1.1, together with a theorem of Cooper and Lackenby [10] which asserts that, for any $\epsilon > 0$ and any closed orientable 3–manifold M, there are at most finitely many cusped hyperbolic 3–manifolds X and slopes s on ∂X with length at least $2\pi + \epsilon$, such that M is obtained by Dehn filling X along s.

Let M be a closed 3–manifold with Gromov norm more than 8.561×10^{20}. Suppose that M is obtained by p/q surgery along a prime alternating knot K. Let X be the exterior of K. Then X is hyperbolic, and since Gromov norm does not increase when Dehn filling is performed, the Gromov norm of X is also at least 8.561×10^{20}. Since Gromov norm and hyperbolic volume are proportional, the volume of X is at least $8.561 \times 10^{20} v_3$. Hence, by Lackenby's theorem, with improvements due to

Agol and D Thurston [19], the twist number of K is at least 8.561×10^{19}. So, by Corollary 1.2, the length of the filling slope p/q is at least

$$(7.593 \times 10^{-20}) \times (8.561 \times 10^{19} - 2) > 6.5.$$

Setting $\epsilon = 6.5 - 2\pi$, and applying the theorem of Cooper and Lackenby, we obtain Theorem 1.3.

Crossing arcs

In his experimental analysis of the hyperbolic structures of alternating knot complements, Thistlethwaite observed several geometric features that are related to cusp area, and formulated various conjectures. To explain one of these conjectures, we need the following definitions.

Definition 1.8 Let α be an arc in S^3, such that $\alpha \cap K = \partial\alpha$. Suppose that α is homotopic in $S^3 \setminus K$ to a geodesic γ. A maximal cusp for $S^3 \setminus K$ intersects γ in two half-open intervals which contain the ends of γ, together possibly with some closed intervals. Define the *length* of α to be the length of the arc that results from removing the two half-open intervals from γ.

Definition 1.9 A *crossing arc* is defined to be an embedded arc in S^3 with endpoints on K, which projects to a single point lying at a crossing in the diagram of K.

Thistlethwaite conjectured that crossing arcs in alternating knots have universally bounded length, with length bounded above by $\log 8$. Since two crossing arcs are homotopic in $S^3 \setminus K$ only if they belong to the same twist region (as shown in Lackenby and Purcell [20, Proposition 7.12]), and since short arcs contribute to cusp area, Thistlethwaite's conjecture would imply Theorem 1.1. In fact, our proof of Theorem 1.1 hinges on the fact that many distinct geodesic arcs in $S^3 \setminus K$ have bounded length. However, we cannot show that crossing arcs are short, but expect this to be the case. Therefore, we offer Theorem 1.1 as additional evidence for Thistlethwaite's conjecture.

Organisation

The results in this paper are organised as follows. In Section 2, we prove the main result for alternating knots with a bounded number of crossings per twist region. This case is simpler to explain than the general case, and yet introduces the main tools and techniques we will use in general, particularly Theorem 2.7. We apply this theorem to

the checkerboard surfaces of the diagram. In Section 3, we remind the reader of an essential immersed surface, originally appearing in Lackenby and Purcell [20], which will play the role of the checkerboard surfaces in the general case. We use this to give a proof of Theorem 1.1.

Acknowledgements Purcell is supported in part by NSF grant DMS-1252687, and by Australian Research Council grant DP160103085.

2 Bounded crossings in each twist region

In this section, we prove a version of the main theorem in the case that our alternating knot has a bounded number of crossings per twist region. That is, we show that such knots have cusp volume bounded above and below by linear functions of the twist number.

We work through this case separately because it is somewhat simpler than the most general case, and also because it allows us to develop and apply many of the tools we will use in the general setting.

Checkerboard surfaces

Recall the checkerboard surfaces of an alternating link complement are obtained by colouring the regions of the diagram in a checkerboard manner, in blue and red, say, and then attaching twisted bands at crossings to connect regions of the same colour. Note that the checkerboard surfaces intersect only at crossings, and then in a single arc running from the strand of the link at the top of the crossing to the strand of the link at the bottom. This is a crossing arc, as in Definition 1.9.

The following result is due to Menasco and Thistlethwaite [22].

Lemma 2.1 *The checkerboard surfaces arising from a prime alternating diagram are essential.*

Any essential surface in a hyperbolic 3–manifold can be given a pleating (see Thurston [27], or for a more detailed proof, see Canary, Epstein and Green [9], or Lackenby [18]). We pleat both checkerboard surfaces in $S^3 \setminus K$. This gives an induced hyperbolic metric on the disjoint union of the two checkerboard surfaces.

Lemma 2.2 *Let S denote the disjoint union of the two checkerboard surfaces for the alternating knot K with a prime, twist reduced diagram, which we also denote by K.*

Then the Euler characteristic of S is $\chi(S) = 2 - \mathrm{cr}(K)$, where $\mathrm{cr}(K)$ is the number of crossings of the diagram. Thus under the hyperbolic metric obtained from pleating S in $S^3 \smallsetminus K$, the area of S is $\mathrm{Area}(S) = 2\pi(\mathrm{cr}(K) - 2)$.

This result is not hard, and appears elsewhere (see, for example [5]). We offer a particularly simple proof.

Proof In the disjoint union of red and blue surfaces, each crossing arc appears twice: once in the red, and once in the blue. In $S^3 \smallsetminus K$, replace each of the crossing arcs with two homotopic arcs, each with an endpoint on the overstrand of the crossing, but with separated endpoints on the understrand of the crossing. Attach red and blue regions of the diagram to these arcs. The result is just the plane of projection of the diagram, punctured once by each overcrossing; hence it has Euler characteristic $2 - \mathrm{cr}(K)$. But by construction, this Euler characteristic is exactly the Euler characteristic of the disjoint union of the two checkerboard surfaces. The area formula follows immediately from the Gauss–Bonnet theorem. □

Lemma 2.3 *Let S denote the disjoint union of two essential properly immersed surfaces S_1 and S_2 in $S^3 \smallsetminus K$. Give S a hyperbolic metric, by pleating S and pulling back the path metric on $S^3 \smallsetminus K$. Let H denote the maximal cusp of $S^3 \smallsetminus K$. Suppose that, for $i = 1$ and 2, the boundary of S_i is a single curve on $\partial N(K)$ with integral slope σ_i. Then S has an embedded cusp H_S with cusp area at least $|\sigma_1 - \sigma_2|$, such that H_S maps into H under the map from S to $S^3 \smallsetminus K$. Moreover, if K is neither the figure-eight knot nor the knot 5_2, the area of H_S is at least $\sqrt[4]{2}\,|\sigma_1 - \sigma_2|$.*

Proof For $i = 1, 2$, there exists an embedded cusp H_i for S_i with area at least the length of a geodesic representative γ_i for ∂S_i on ∂H, and which has image in $S^3 \smallsetminus K$ contained in H, for example by a lemma of Futer and Schleimer [16, Lemma 2.5].

Orient γ_1 and γ_2 so that their intersection numbers with the meridian of K are opposite. If we resolve the intersections between γ_1 and γ_2 so that they respect these orientations, the result is a collection of meridians with the same length as $\gamma_1 \cup \gamma_2$. The number of these meridians is $|\sigma_1 - \sigma_2|$. Each has length at least 1, by standard arguments in hyperbolic geometry. So, the length of $\gamma_1 \cup \gamma_2$ is at least $|\sigma_1 - \sigma_2|$, and this forms a lower bound for the area of $H_S = H_1 \cup H_2$.

In fact, Adams [2] proved that the meridian of any hyperbolic knot in the 3–sphere has length at least $\sqrt[4]{2}$, with the exception of the figure-eight and 5_2 knots. Hence, using this bound, the length of $\gamma_1 \cup \gamma_2$ is at least $\sqrt[4]{2}\,|\sigma_1 - \sigma_2|$, and again this forms a lower bound for the area of $H_S = H_1 \cup H_2$. □

Lemma 2.4 *Let S denote the disjoint union of the two checkerboard surfaces for the hyperbolic alternating knot K. Then S has an embedded cusp H_S such that the area of H_S is at least twice the number of crossings in the prime, twist reduced diagram of K, and such that H_S maps into the maximal cusp for $S^3 \setminus K$. That is,*

$$\text{Area}(H_S) \geq 2\,\text{cr}(K).$$

Moreover, if K is neither the figure-eight knot nor the knot 5_2, the area of H_S is at least $2^{5/4}\,\text{cr}(K)$.

Proof Colour the two checkerboard surfaces red and blue, and denote them R and B. Denote their disjoint union by S. Note that R and B are properly embedded in the exterior of K, each with boundary a single curve with integral slope.

We claim that the difference in slopes of ∂R and ∂B is twice the crossing number. This is because R and B intersect exactly in crossing arcs, with boundary slopes intersecting once at an overcrossing, once at an undercrossing. See Figure 3.

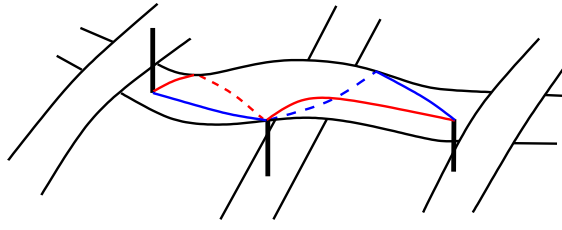

Figure 3: The red and blue surfaces run over two meridians per overcrossing.

Thus Lemma 2.3 gives the required cusp for S, with area at least $2\,\text{cr}(K)$. Moreover, when K is neither the figure-eight knot nor the knot 5_2, the area of the cusp is at least $2^{5/4}\,\text{cr}(K)$. □

The above lemma is simple, but quite striking. It is easy to construct hyperbolic surfaces with large area but with small cusp area, but Lemma 2.4 asserts that this type of geometry never occurs among checkerboard surfaces of alternating knots, when in pleated form.

This lower bound on the cusp area of S is central to our argument. We will use it in the next subsection to establish the existence a large collection of disjoint properly embedded essential arcs in S, each of which has bounded length. By Lemma 2.1, these can be homotoped to geodesics in $S^3 \setminus K$ without increasing their length. We will see in the next subsection that each such arc gives a definite contribution to the cusp volume of $S^3 \setminus K$. However, it may be the case that distinct geodesics in S are

homotoped to the same geodesic in $S^3 \smallsetminus K$. In the remainder of this subsection, we investigate this phenomenon more thoroughly.

Consider a twist region of the diagram. A small regular neighbourhood of this twist region in the diagram is a disc, and a regular neighbourhood of this disc is a 3–ball in S^3. The intersection between this ball and a checkerboard surface is the subsurface that is *associated* with this twist region. Note that if the twist region has c crossings, then one of these subsurfaces has Euler characteristic $2-c$ and the other is a disc. See Figure 4.

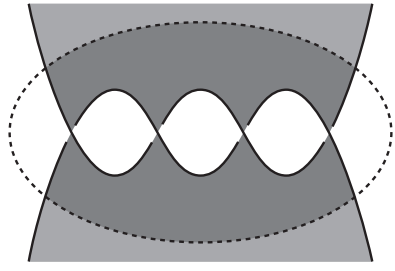

Figure 4: The subsurface of a checkerboard surface associated with a twist region

The following is [20, Proposition 7.12].

Lemma 2.5 *Let S denote the disjoint union of the two checkerboard surfaces of a prime, twist reduced, alternating diagram of a hyperbolic knot K. Suppose a_1 and a_2 are disjoint essential embedded arcs in S that are not homotopic in S, but are homotopic in $S^3 \smallsetminus K$ after including S into $S^3 \smallsetminus K$. Then either a_1 and a_2 are isotopic in S to crossing arcs in the same twist region of the diagram, or they both lie on the same checkerboard surface and both are isotopic in that checkerboard surface to arcs in the same subsurface associated with some twist region.*

The proof of this result does not require all of the machinery of [20], but mostly just the techniques developed in Section 7 of that paper. The proof is rather reminiscent of the arguments in [18] and [19]. Since the arcs a_1 and a_2 are homotopic in $S^3 \smallsetminus K$, there is a map of a rectangle into $S^3 \smallsetminus \operatorname{int}(N(K))$, with the left and right edges mapped to a_1 and a_2, and the top and bottom mapped to $\partial N(K)$. One then considers the inverse image of the checkerboard surfaces in this rectangle, which divides the rectangle into regions. Each region is mapped into the complement of the checkerboard surfaces, and so one can view its boundary as specifying a curve in the alternating diagram. Lemma 2.5 is then proved by analysing these curves and by ruling out various trivial configurations, until the map of the rectangle into $S^3 \smallsetminus \operatorname{int}(N(K))$ is of a very restricted

form. The conclusion of the lemma then follows quickly. For more details, see [20, Section 7].

Corollary 2.6 *Let S denote the disjoint union of the two checkerboard surfaces for a prime, twist reduced, alternating diagram of the hyperbolic knot K. Suppose that each twist region contains at most N crossings. Then, for any collection of disjoint embedded essential nonparallel arcs in S, all of which are homotopic in $S^3 \smallsetminus K$, the number of such arcs is most $3N-1$.*

Proof We may assume that, in this collection, there is more than one arc lying in the same checkerboard surface, B, say. By Lemma 2.5, all the arcs lying in B can be isotoped to lie in the subsurface associated with some twist region. We may arrange that, after this isotopy, the arcs remain disjoint, embedded and nonparallel. The number of disjoint essential arcs, no two of which are parallel, that can lie in such a subsurface is at most $3N-2$. Also, by Lemma 2.5, any arc in the collection in R can be isotoped to lie in the subsurface associated to this twist region. This can support at most one such arc, up to isotopy. □

Short arcs and cusp volume

The following result will be used to show that there are many geodesic arcs of bounded length in a hyperbolic alternating knot complement.

Recall that, in Definition 1.8, we defined the length of an arc in S^3, with both endpoints on K. We present a variation of this definition now. Let F be a finite-volume hyperbolic surface, and let H be an embedded horoball neighbourhood of its cusps. Let α be a bi-infinite geodesic in F, with both its ends in H. Then the intersection between α and H consists of two half-open intervals containing these ends, plus perhaps some closed bounded intervals. Define the *length* of α with respect to H to be the length of the arc obtained by removing the two half-open intervals. Thus, if α also intersects H in some closed bounded intervals, then these do contribute to the length of α.

Theorem 2.7 *Let F be a (possibly disconnected) finite-area hyperbolic surface. Let H be an embedded horoball neighbourhood of the cusps of F. Let $k = \mathrm{Area}(H)/\mathrm{Area}(F)$ and let $d > 0$. Then there is a collection of at least*

$$\frac{(ke^d - 1)\pi}{(e^d - 1)(\sinh(d) + 2\pi)} |\chi(F)|$$

embedded disjoint bi-infinite geodesic arcs, each with both ends in H, and each having length at most $2d$ with respect to H.

Before we give the proof of this theorem, we make a couple of observations about it. On the one hand, the lower bound on the number of geodesic arcs that it provides, with length at most $2d$, can be viewed as somewhat inefficient. This is because, as d tends to infinity, the lower bound decreases to zero, whereas clearly the number of geodesics with length at most $2d$ is an increasing function of d. However, the real utility of the theorem is its universal nature. The maximal number of disjoint embedded geodesic arcs in a cusped hyperbolic surface F is $3|\chi(F)|$. Theorem 2.7 asserts that one may find a collection of such arcs, with cardinality which is a definite fraction of this maximum, and where all the arcs have bounded length. The fraction depends little on the topology and geometry of F: just the cusp density of F and the length of the arcs determine this fraction. It is the general nature of this estimate that is crucial for our proof of Theorem 1.1.

Proof Pick a maximal collection of embedded disjoint bi-infinite geodesic arcs G, each with both ends in H, and each having length at most $2d$ with respect to H. Use the upper half space model of $\mathbb{H}^2$. Let $\widetilde{H}$ be the inverse image of H in $\mathbb{H}^2$. We may arrange that one component H_∞ of $\widetilde{H}$ is $\{y \geq 1\}$. For each component g of G, pick lifts $\widetilde{g}$ and $\widetilde{g}'$ to $\mathbb{H}^2$, each of which starts on H_∞, but which do not differ by a covering transformation that preserves H_∞. These run to other components $H_{\widetilde{g}}$ and $H_{\widetilde{g}'}$ of $\widetilde{H}$. Let $I(\widetilde{g})$ be an open interval on ∂H_∞ centred at the point where $\widetilde{g}$ exits, with length 2π. Define $I(\widetilde{g}')$ similarly. Each of these projects to an interval in ∂H with length also 2π (or possibly to all of ∂H).

For each component g of G, its intersection with H is two half-open intervals, plus possibly some closed intervals (which may be points). For each closed interval, pick a lift of this to $\mathbb{H}^2$ that lies in H_∞. This lift lies in a lift $\widehat{g}$ of g in $\mathbb{H}^2$. Consider the vertical projection of $\widehat{g} \cap H_\infty$ onto ∂H_∞. Its image in ∂H is termed the *vertical projection of the interval onto* ∂H.

We define I to be the union of the following subsets of ∂H:

(1) for each geodesic g in G, the images of $I(\widetilde{g})$ and $I(\widetilde{g}')$ in ∂H;

(2) for each geodesic g in G and for each closed interval of $g \cap H$, the vertical projection of this interval onto ∂H.

Claim 1 *The length of I is at most $|G|(2\sinh(d) + 4\pi)$.*

Clearly, for each geodesic g in G, the intervals $I(\widetilde{g})$ and $I(\widetilde{g}')$ contribute at most 4π to the length of I. So, we consider the projection of the closed intervals in $g \cap H$ onto ∂H. Let $\ell_1, \ldots, \ell_n$ be the lengths of these intervals. Then the projections of

these intervals onto ∂H have lengths $2\sinh(\ell_1/2), \ldots, 2\sinh(\ell_n/2)$. Since sinh is increasing and convex on the positive real line, and $\ell_1 + \cdots + \ell_n \leq 2d$, we deduce that the total length of these projected intervals is at most $2\sinh d$. This proves the claim.

Claim 2 *For each lift $\hat{g}$ of a component of G, if $\hat{g}$ intersects ∂H_∞, then the vertical projection of $\hat{g}$ onto ∂H_∞ lies entirely in the inverse image of I.*

We may assume that $\hat{g}$ has maximal projection onto ∂H_∞, in the sense that this projection is not contained within the projection of some other lift of G. Let x be one of the two points of $\hat{g} \cap \partial H_\infty$. Let $\overline{\alpha}$ be the geodesic ray starting at x running vertically into H_∞. Then $\overline{\alpha} - x$ is disjoint from the inverse image of G, by our assumption about the maximality of the projection of $\hat{g}$. Let $\overline{\beta}$ denote the subset of $\hat{g}$ that is a geodesic ray starting at x and exiting H_∞. Then $\overline{\alpha} \cup \overline{\beta}$ is a piecewise geodesic, and after a small isotopy, it can be made disjoint from the inverse image of G. Hence, it maps to an arc in F that is disjoint from G. This is embedded, because $\overline{\alpha}$ and $\overline{\beta}$ both have images that are embedded in F, and these images have disjoint interiors. Straighten $\overline{\alpha} \cup \overline{\beta}$ to a geodesic γ that hits ∂H_∞ orthogonally. Then the image of γ is also embedded and is isotopic to an arc disjoint from G. Its length outside the cusp is less than $2d$. So, the image of γ lies in G, by the maximality of G. Therefore $I(\gamma)$ maps to I.

Now the triangle with sides $\overline{\alpha}$, $\overline{\beta}$ and γ has area less than π. On the other hand, the area contained in the portion of H_∞ lying over $I(\gamma)$ on the side of γ containing x is π. Hence $I(\gamma)$ contains the interval in ∂H_∞ between $\gamma \cap \partial H_\infty$ and x. Hence, this entire interval lies in the inverse image of I. Since this argument applies to both points of $\hat{g} \cap \partial H_\infty$, the claim is proved.

For the parts of ∂H not lying in I, expand H by vertical distance d, giving a (closed) neighbourhood H' of the cusps. See Figure 5.

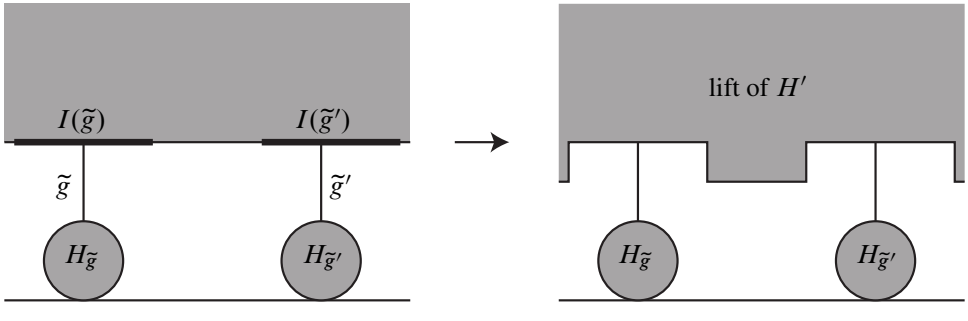

Figure 5: Expanding H to form H'

Claim 3 *H' is disjoint from $G-H$ and is embedded in F.*

Suppose that H' intersects $G-H$ or is not embedded in F. We view H' as obtained from H by slowly expanding it vertically. There is some first moment in time when H' intersects G or when H' becomes nonembedded.

Let us suppose that H' intersects $G-H$ first. Emanating from this point of intersection, there is a geodesic segment which runs vertically in H' to the boundary of the cusp H. Pick a lift α of this geodesic segment in $\mathbb{H}^2$ with one endpoint in ∂H_∞. Let x be the other endpoint. Note by choice of H', the geodesic segment α has length at most d. The starting point x of α lies on a lift $\hat{g}$ of a component of G. This bi-infinite geodesic $\hat{g}$ has ends in horoballs H_1 and H_2 that are not H_∞. The length of $\hat{g}$ between H_1 and H_2 is at most $2d$, and so one of the components β of $\hat{g}-x$ has length at most d outside of H_1, say. Extend α and β to infinite geodesic rays $\overline{\alpha}$ and $\overline{\beta}$ running into H_∞ and H_1, respectively, and meeting at x. Note that $\overline{\beta}$ misses H_∞, because otherwise, by Claim 2, the entire vertical projection of $\hat{g}$ onto ∂H_∞ has image in I, whereas, by assumption, $\alpha \cap \partial H_\infty$ does not have image in I. See Figure 6. Straighten $\overline{\alpha} \cup \overline{\beta}$ to a geodesic γ that hits ∂H_∞ and ∂H_1 orthogonally. Clearly, the image of γ in F has length at most $2d$ with respect to H. Also, the image of γ is embedded in F, because it is a geodesic homotopic to $\overline{\alpha} \cup \overline{\beta}$, which has image embedded in F. So the image of γ lies in G, by the maximality of G. Consider the triangle formed by γ, $\overline{\alpha}$ and $\overline{\beta}$. This has area less than π. Since α avoids $I(\gamma)$, and $\overline{\beta}$ misses H_∞, we deduce that the region lying between $\overline{\alpha}$ and γ in H_∞ has area at least π, which is a contradiction, proving the claim in this case.

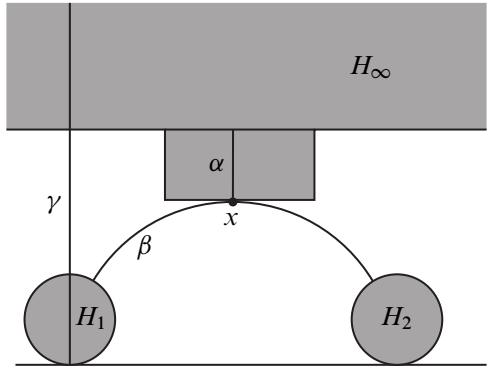

Figure 6: The case where H' intersects $G-H$ first

Suppose now that when H' is formed by expanding H, it becomes nonembedded first. Then at this moment in time, the neighbourhood fails to be embedded at some point.

Emanating from this point, there are two geodesic segments which run to the boundary of the horoball neighbourhood H, and which lift to geodesic segments α_1 and α_2 in $\mathbb{H}^2$ running to the boundary of distinct horoballs in $\mathbb{H}^2$. We may take one of these to be H_∞. The geodesic segments α_1 and α_2 each have length at most d, by choice of H'. So extending to infinite rays $\overline{\alpha}_1$ and $\overline{\alpha}_2$ and concatenating, we obtain a piecewise geodesic that has length at most $2d$ with respect to lifts of H. Its image in F is embedded (because this moment in time was the first time when H' bumped into itself). It also misses G. So, it is isotopic to an embedded geodesic g that has length at most $2d$ with respect to the cusp. Thus, g lies in G by the maximality of G. So at the two points where g intersects ∂H, there are the images of the two intervals $I(\widetilde{g})$ and $I(\widetilde{g}')$. Now, neither of these intervals is expanded when forming H', and so at least one of $\overline{\alpha}_1$ and $\overline{\alpha}_2$ misses both intervals, say $\overline{\alpha}_1$ does. Consider the triangle formed by g, $\overline{\alpha}_1$ and $\overline{\alpha}_2$. This has area less than π. But the region lying between $\overline{\alpha}_1$ and g inside H has area at least π. This contradiction proves the claim.

The claim implies the theorem, because then

$$\begin{aligned}\operatorname{Area}(H') &\geq \operatorname{Area}(H) + (e^d - 1)\operatorname{Length}(\partial H - I)\\ &\geq \operatorname{Area}(H) + (e^d - 1)\operatorname{Length}(\partial H) - (e^d - 1)|G|(2\sinh(d) + 4\pi)\\ &= e^d \operatorname{Area}(H) - (e^d - 1)|G|(2\sinh(d) + 4\pi).\end{aligned}$$

Thus

$$\begin{aligned}\operatorname{Area}(F) &\geq \operatorname{Area}(H')\\ &\geq e^d \operatorname{Area}(H) - (e^d - 1)|G|(2\sinh(d) + 4\pi)\\ &= ke^d \operatorname{Area}(F) - (e^d - 1)|G|(2\sinh(d) + 4\pi).\end{aligned}$$

So,

$$\begin{aligned}|G| &\geq \frac{ke^d - 1}{(e^d - 1)} \frac{\operatorname{Area}(F)}{(2\sinh(d) + 4\pi)}\\ &= \frac{(ke^d - 1)\pi}{(e^d - 1)(\sinh(d) + 2\pi)} |\chi(F)|.\end{aligned}$$

□

Recall from Definition 1.8, the definition of the length of an arc in $S^3 \setminus K$, measured with respect to a maximal cusp.

Lemma 2.8 *Suppose that a one-cusped hyperbolic 3–manifold M contains at least p homotopically distinct essential arcs, each with length at most L measured with respect to the maximal cusp H of M. Then the cusp area* $\operatorname{Area}(\partial H)$ *is at least* $p\sqrt{3}e^{-2L}$.

Proof Consider the universal cover $\mathbb{H}^3$ of M, arranged so that one component of the lift of the maximal cusp H is the horoball $H_\infty = \{z \geq 1\}$. Let T be a fundamental domain for the cusp torus ∂H on ∂H_∞.

Let G be our collection of essential arcs, with $|G| \geq p$. Each arc a in G has geodesic representative in M which lifts to give *two* vertical geodesics meeting T: for oriented a, one such lift runs from infinity to a point on $\mathbb{C}$ below T, and the other lift runs from a point on $\mathbb{C}$ below T to infinity. One end of each arc is at infinity. The other runs into a horoball centred at some other point on $\mathbb{C}$. Since the maximal cusp is embedded, each of these horoballs is disjoint in $\mathbb{H}^3$. Hence each gives a circular "shadow" on ∂H. We estimate the size of the shadows and apply a circle packing argument to obtain our area estimate.

Because the (hyperbolic) length of each arc a is at most L, the horoball at the end of each lift of a in T has (Euclidean) diameter at least e^{-L}. Each such horoball projects onto ∂H_∞, giving a disk of diameter e^{-L} on ∂H. Note that while larger horoballs may partially cover smaller ones, if we shrink all horoballs to have diameter exactly e^{-L}, then the collection of horoballs will remain embedded in $\mathbb{H}^3$, and so the projections of these smaller disks onto ∂H are all disjoint. Hence each arc a in G corresponds to two embedded disks on ∂H of diameter e^{-L}.

It follows that the area of the cusp torus ∂H is at least $2p \cdot \pi(e^{-L}/2)^2 = p\pi e^{-2L}/2$. We can improve this estimate using a disk packing argument (see [8, Theorem 1]): the area of ∂H must be at least $2\sqrt{3}/\pi$ times the sum of the areas of the disks, or

$$\text{Area}(\partial H) \geq \frac{2\sqrt{3}}{\pi} \cdot \frac{p\pi e^{-2L}}{2} = p\sqrt{3}e^{-2L}. \qquad \square$$

Theorem 2.9 *Suppose K is a hyperbolic alternating knot with a prime, twist reduced diagram, which we also call K, with no more than N crossings in each twist region. Let H be a maximal cusp of the hyperbolic manifold $S^3 \smallsetminus K$. Then the cusp area satisfies*

$$\frac{1.844 \times 10^{-4}}{(3N-1)}(\text{tw}(K) - 2) \leq \text{Area}(\partial H) < 10\sqrt{3}(\text{tw}(K) - 1).$$

Proof The upper bound comes from the upper bound on the hyperbolic volume of an alternating knot complement, proved by Lackenby and improved by Agol and D Thurston in [19]. From this result, we know

$$\text{Volume}(S^3 \smallsetminus K) \leq 10v_3(\text{tw}(K) - 1),$$

where $v_3 \approx 1.01494$ is the volume of a regular hyperbolic ideal tetrahedron.

Böröczky [8] proved that the cusp volume of a hyperbolic 3–manifold is at most $\sqrt{3}/(2v_3)$ times the volume of the manifold. Hence we have

$$\text{Volume}(H) < 5\sqrt{3}(\text{tw}(K) - 1).$$

Finally, by basic hyperbolic geometry, we have $\text{Area}(\partial H) = 2\,\text{Volume}(H)$.

For the lower bound, we first note that it trivially holds in the case of the figure-eight knot and the knot 5_2, because these both have $\text{tw}(K) = 2$. So, we can assume that K is neither of these knots. Let S denote the disjoint union of the two checkerboard surfaces. Since S is essential, it may be pleated in $S^3 \smallsetminus K$, and the pleating pulls back to give a hyperbolic metric on S. Let H_S be an embedded cusp for S coming from Lemma 2.4, with $\text{Area}(H_S) \geq 2^{5/4}\,\text{cr}(K)$. By Lemma 2.2, S has area $2\pi(\text{cr}(K)-2))$. Set $k = \text{Area}(H_S)/\text{Area}(S)$. Then, $k > \sqrt[4]{2}/\pi$. Set $d = \log(2/k) < \log(2^{3/4}\pi)$. Then by Theorem 2.7, there are at least

$$\frac{(ke^d - 1)\pi}{(e^d - 1)(\sinh(d) + 2\pi)}|\chi(S)| > (0.083)|\chi(S)|$$

disjoint geodesic arcs in S of length at most $2d$ with respect to H_S. Map these into $S^3 \smallsetminus K$. Because the checkerboard surfaces are essential, each such arc remains essential under the mapping. Moreover, the length of each such arc may only decrease in $S^3 \smallsetminus K$, either by straightening, or by cutting off more length since the image of H_S is contained in H. Hence each has length at most $2d$ with respect to the maximal cusp of $S^3 \smallsetminus K$.

By Corollary 2.6, at most $3N-1$ of these arcs may be homotopic to the same arc. Thus in $S^3 \smallsetminus K$, we have at least $(0.083)|\chi(S)|/(3N-1)$ homotopically distinct essential arcs of length at most $2d$.

Now Lemma 2.8 implies that the cusp area satisfies

$$\begin{aligned}
\text{Area}(\partial H) &\geq \frac{(0.083)|\chi(S)|}{3N-1}\sqrt{3}e^{-4d} \\
&> \frac{(0.083)\sqrt{3}\,|\chi(S)|}{8\pi^4(3N-1)} \\
&\geq \frac{1.844\times 10^{-4}}{(3N-1)}(\text{cr}(K)-2) \\
&\geq \frac{1.844\times 10^{-4}}{(3N-1)}(\text{tw}(K)-2).
\end{aligned}$$

□

3 An essential twisted surface

In the previous section, we showed that the complement of a hyperbolic alternating knot contains essential arcs of bounded length by considering the checkerboard surfaces of the knot. We want the number of such arcs to be bounded below by a linear function of the twist number of the knot. When there is a bounded number of crossings per twist region, then the checkerboard surfaces give us this result. However, in general, if there are a high number of crossings in a twist region, then the arcs with bounded length on the checkerboard surface may all be homotopic within this twist region. We need to find arcs that we can guarantee are not all concentrated within a constant number of twist regions.

In order to do so, we use an essential immersed surface in the knot complement, called a *twisted checkerboard surface*. Two variations of this surface are defined in [20], and proved to be π_1–injective and boundary-π_1–injective. A map $f\colon S \to M$ of a surface S into a compact 3–manifold M with $f(\partial S) \subset \partial M$ is said to be *boundary-π_1–injective* if, for any arc $\alpha\colon I \to S$ with endpoints in ∂S, the existence of a homotopy (rel endpoints) of $f \circ \alpha$ into ∂M implies the existence of a homotopy (rel endpoints) of α into ∂S.

In this section, we review the definitions of these surfaces, and remind the reader of key results in [20]. Using these results, we give a proof of Theorem 1.1.

Twisted surfaces

Recall that K is a hyperbolic alternating knot with a prime, twist reduced diagram, which we also refer to by K. For each twist region of K with at least N crossings (N will be determined explicitly below), augment the diagram, in the sense of Adams [3]. That is, add an unknotted circle to the diagram encircling the two strands of the twist region. This circle is called a *crossing circle*. It bounds a disk, the *crossing disk*, which meets the knot transversely in two points.

The result of adding crossing circles to all twist regions with at least N crossings yields the diagram of a link. We will denote this link and its diagram by L.

Recall that $S^3 \smallsetminus L$ is homeomorphic to $S^3 \smallsetminus L'$, where L' is any link obtained from L by removing any even number of crossings from each twist region encircled by a crossing circle in the diagram of L (eg [25]). We will work with the diagram in which each twist region encircled by a crossing circle has 1 or 2 crossings remaining. We denote this link (and its diagram) by L_2. Finally, for L_2, remove the crossing circles from the diagram. The result is a knot which we denote by K_2. An example is shown in Figure 7.

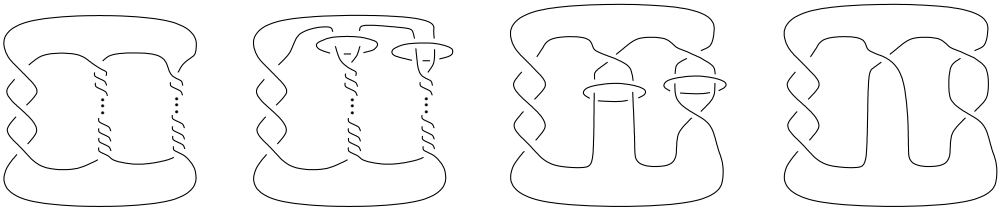

Figure 7: Examples of diagrams for K (far left), L (middle left), L_2 (middle right) and K_2 (far right) (figure from [20])

To form the twisted checkerboard surface, start with the checkerboard surfaces of K_2. These are embedded surfaces in $S^3 \setminus \nu(K_2)$, where $\nu(K_2)$ denotes a small regular neighbourhood of K_2. Colour these surfaces *red* and *blue*.

A small regular neighbourhood of each crossing circle in L_2 will intersect either the red or the blue checkerboard surface in two meridional disks. Thus when we remove these disks from the surfaces, we obtain two surfaces embedded in $S^3 \setminus \nu(L_2)$, which we also colour red and blue, and denote by R_2 and B_2.

Because $S^3 \setminus \nu(L_2)$ is homeomorphic to $S^3 \setminus \nu(L)$, the surfaces R_2 and B_2 are also embedded in $S^3 \setminus \nu(L)$. Notice that under the homeomorphism, a meridional curve on the boundary of a regular neighbourhood of the j^{th} crossing circle will be sent to a curve with slope $\pm 1/n_j$, where $2n_j$ crossings were removed from the corresponding twist region to go from the diagram of L to that of L_2.

Finally, to obtain $S^3 \setminus \nu(K)$ from $S^3 \setminus \nu(L)$, we perform meridional Dehn filling on each crossing circle of L. That is, we reinsert a regular neighbourhood of the crossing circle so that the meridian of the crossing circle bounds a disk in the attached solid torus.

To obtain the twisted checkerboard surfaces, when we reinsert the regular neighbourhood of the crossing circles, we also attach annuli or Möbius bands to the surfaces R_2 and B_2, as follows. Let C_j denote the j^{th} crossing circle in L. Recall that one of R_2 or B_2 meets $\partial\nu(C_j)$ in two curves, corresponding to two meridians of a crossing circle of L_2. Because the boundary slopes are each $\pm 1/n_j$ on $\partial\nu(C_j)$, a meridional disk of $\nu(C_j)$ meets the surface $2n_j$ times. For each meridional disk, attach an interval I to opposite points of intersection on the disk, ensuring that in $S^3 \setminus \nu(K)$, the interval runs through the core of the solid torus $\nu(C_j)$. When we take such an interval for each disk $D^2 \times \{\theta\}$ in $D^2 \times S^1$, the result is that we have attached an I–bundle to the surface R_2 or the surface B_2. When n_j is odd, so that we are attaching opposite boundary components when we attach the interval, the I–bundle is an annulus, connecting the two boundary components of R_2 (or B_2) on $\partial\nu(C_j)$. When n_j is even, the intervals connect one

boundary component on $\partial\nu(C_j)$ to itself, hence we are attaching two Möbius bands to R_2 (or B_2). This is illustrated in Figure 8.

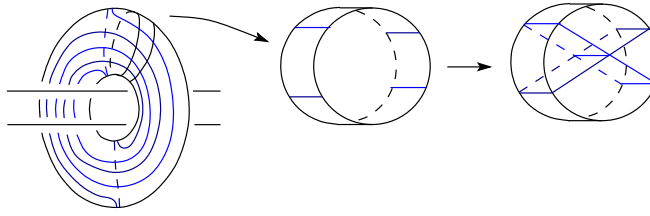

Figure 8: A cross section of the solid torus added to $S^3 \smallsetminus \nu(L)$, and how the surface extends into it (figure from [20])

When we do this process for each crossing circle, we obtain two surfaces, denoted $S_{R,2}$ and $S_{B,2}$, and immersions $\phi_{R,2}\colon S_{R,2} \to S^3 \smallsetminus K$ and $\phi_{B,2}\colon S_{B,2} \to S^3 \smallsetminus K$. Again we colour the surfaces red and blue.

Lemma 3.1 *Let* $\mathrm{tw}_N(K)$ *denote the number of twist regions of* K *with at least* N *crossings, and let* $\mathrm{cr}(\cdot)$ *denote the number of crossings of a diagram. Then the Euler characteristics of the red and blue twisted checkerboard surfaces satisfy*

$$|\chi(S_{R,2}) + \chi(S_{B,2})| = \mathrm{cr}(K_2) + 2\,\mathrm{tw}_N(K) - 2.$$

Proof By Lemma 2.2, the absolute value of the Euler characteristic of the disjoint union of the checkerboard surfaces in K_2 is $\mathrm{cr}(K_2) - 2$. To obtain surfaces R_2 and B_2, remove two disks for each crossing circle, so that

$$|\chi(R_2) + \chi(B_2)| = \mathrm{cr}(K_2) + 2\,\mathrm{tw}_N(K) - 2.$$

Finally, connect pairs of boundary components by annuli or attach pairs of Möbius bands for each crossing circle. This has no effect on Euler characteristic. □

The next theorem is the main result of [20].

Theorem 3.2 *Let* $S_{R,2}$ *and* $S_{B,2}$ *be the surfaces constructed as above, by twisting checkerboard surfaces around crossing circles encircling each twist region with at least* N *crossings. Let* $\phi_{R,2}\colon S_{R,2} \to S^3 \smallsetminus K$ *be the immersion of* $S_{R,2}$ *into* $S^3 \smallsetminus K$, *and similarly for* $S_{B,2}$. *Then* $\phi_{R,2}$ *and* $\phi_{B,2}$ *are* π_1*–injective and boundary-*π_1*–injective, provided* $N \geq 91$.

We also showed the following result, analogous to Lemma 2.5, in [20, Theorem 7.1]. As in Lemma 2.5, this also refers to subsurfaces associated to twist regions, but we

need to clarify what this means in the case of the twisted checkerboard surfaces. Each arises from a twist region of the knot K_2. We choose the associated subsurfaces of the checkerboard surface for K_2 as follows. If the twist region is encircled by a crossing circle of L_2, we choose the subsurface in R_2 or B_2 so that it is punctured twice by this crossing circle. However, the subsurfaces are disjoint from all other crossing circles. Hence, they form subsurfaces of R_2 and B_2. Since there are inclusions $R_2 \subset S_{R,2}$ and $B_2 \subset S_{B,2}$, we obtain subsurfaces of $S_{R,2}$ and $S_{B,2}$ which are the *subsurfaces associated with a twist region of* K_2.

Theorem 3.3 *Suppose $N \geq 121$ and suppose that two distinct essential arcs in the surface $S_{B,2}$ have homotopic images in $S^3 \setminus K$, but are not homotopic in $S_{B,2}$. Then the two arcs are homotopic in $S_{B,2}$ into the same subsurface associated with some twist region of K_2.*

Just as in Corollary 2.6, we obtain the following result.

Corollary 3.4 *Suppose $N \geq 121$. Let $\mathcal{C}$ be a collection of disjoint embedded essential nonparallel arcs in $S_{B,2} \sqcup S_{R,2}$ which are all homotopic in $S^3 \setminus K$. Then the number of arcs in $\mathcal{C}$ is at most $2(3N-2)$.*

Proof By Theorem 3.3, all the arcs of $\mathcal{C}$ that lie in $S_{B,2}$ can be homotoped in $S_{B,2}$ to lie in a subsurface associated with a twist region of K_2. Note that after this homotopy, the arcs remain disjoint, embedded and nonparallel. One way of proving this is to use a hyperbolic structure on $S_{B,2}$, and to realise the arcs that separate the subsurface from the remainder of $S_{B,2}$ as geodesics, and also to realise the arcs in $\mathcal{C}$ as geodesics. Then the geodesics in $\mathcal{C}$ are disjoint, embedded and nonparallel, and lie in the subsurface associated with the twist region of K_2. The number of disjoint embedded essential nonparallel arcs in such a subsurface is at most $3N-2$. Note that when a twist region of K_2 is encircled by a crossing circle, it has at most 2 crossings, and so in this case, the maximal number of disjoint embedded essential nonparallel arcs in a corresponding subsurface is at most 6, which is less than $3N-2$. By switching roles of $S_{B,2}$ and $S_{R,2}$, we obtain the same result for all the arcs of $\mathcal{C}$ that lie in $S_{R,2}$. Hence there are at most $2(3N-2)$ arcs in total. □

Because each twisted checkerboard surface is boundary-π_1–injective in $S^3 \setminus K$, in the hyperbolic structure on $S^3 \setminus K$ it can be given a pleating. (This is proved, for instance, in [18, Lemma 2.2].) The pleating pulls back to give a hyperbolic structure on the surface. The following, analogous to Lemma 2.4, gives information on that hyperbolic structure.

Proposition 3.5 *Let S denote the disjoint union of $S_{B,2}$ and $S_{R,2}$, with immersion $f\colon S \to S^3 \smallsetminus K$. Consider the hyperbolic structure on S given by pulling back a pleating of its image in the hyperbolic structure on $S^3 \smallsetminus K$. Let H be the maximal cusp for $S^3 \smallsetminus K$. Then there is an embedded cusp for S contained in the preimage $f^{-1}(H)$ with area at least $2\,\mathrm{tw}(K)$. Moreover, if K is neither the figure-eight knot nor the knot 5_2, then the area of this cusp is at least $2^{5/4}\,\mathrm{tw}(K)$.*

Proof The boundary of the red surface on $\nu(K)$ is a single curve with integral slope, and similarly for the boundary of the blue surface. The difference in slope $|\partial S_{B,2} - \partial S_{R,2}|$ is twice the crossing number of K_2, because the surfaces R_2 and B_2 intersect exactly in the crossing arcs of K_2, with boundary slopes intersecting once at an overcrossing, and once at an undercrossing, just as in Figure 3. Thus Lemma 2.3 implies there is an embedded cusp for S contained in the preimage $f^{-1}(H)$, with cusp area at least $2\,\mathrm{cr}(K_2)$. Moreover, if K is neither the figure-eight knot nor the knot 5_2, then the area of this cusp is at least $2^{5/4}\,\mathrm{cr}(K_2)$.

Finally, because the diagram of K_2 contains at least one or two crossings per twist region of K, we obtain the required lower bound on cusp area. □

Bounding cusp volume

We are now ready to give the proof of the main theorem.

Theorem 1.1 *Let D be a prime, twist reduced alternating diagram for some hyperbolic knot K, and let $\mathrm{tw}(D)$ be its twist number. Let C denote the maximal cusp of the hyperbolic 3–manifold $S^3 \smallsetminus K$. Then*

$$A(\mathrm{tw}(D) - 2) \leq \mathrm{Area}(\partial C) < 10\sqrt{3}(\mathrm{tw}(D) - 1),$$

where A is at least 2.278×10^{-19}.

Proof The proof of the upper bound is identical to that of Theorem 2.9, so we only need prove the lower bound. Again, this is trivial in the case where K is the figure-eight knot or the knot 5_2, and so we will assume that it is neither of these knots.

Form the surfaces $S_{R,2}$ and $S_{B,2}$ as above, first by removing all but one or two crossings from twist regions of K with at least $N = 121$ crossings to yield the knot K_2, and then twisting the checkerboard surfaces of K_2 and attaching annuli and Möbius bands. Now let S denote the disjoint union of $S_{R,2}$ and $S_{B,2}$, and let $f\colon S \to S^3 \smallsetminus K$ denote the immersion. By Theorem 3.2, S is π_1–injective and boundary-π_1–injective, hence it can be pleated. Give S the hyperbolic metric induced by the pleating.

Let H denote the maximal cusp of $S^3 \setminus K$. By Proposition 3.5 there exists an embedded cusp H_S for S contained in the preimage $f^{-1}(H)$ such that $\text{Area}(H_S) \geq 2^{5/4}\,\text{tw}(K)$. By Lemma 3.1,

$$\text{Area}(S) = 2\pi(\text{cr}(K_2) + 2\,\text{tw}_N(K) - 2) < 240\pi\,\text{tw}(K).$$

Set $k = \text{Area}(H_S)/\text{Area}(S) \geq \sqrt[4]{2}/(120\pi)$, and set $d = \log(2/k)$. By Theorem 2.7, there exist at least

$$\frac{(ke^d - 1)\pi}{(e^d - 1)(\sinh(d) + 2\pi)}|\chi(S)| > \frac{|\chi(S)|}{65143}$$

disjoint geodesic arcs in S that have length at most $2d$ with respect to H_S. Map these arcs into $S^3 \setminus K$. Because S is boundary-π_1–injective, each such arc remains essential under the mapping. Moreover, its length may only decrease in $S^3 \setminus K$, either by straightening the arc in $S^3 \setminus K$, or by cutting off more length on one end, since H_S may be a strict subset of $f^{-1}(H)$. Hence each such arc has length at most $2d$ with respect to the maximal cusp of $S^3 \setminus K$.

By Corollary 3.4, for any subcollection $\mathcal{C}$ of these arcs, all of which are homotopic in $S^3 \setminus K$, the number of arcs in $\mathcal{C}$ is at most $2(3N - 2) = 722$. So, in $S^3 \setminus K$, we will have at least $|\chi(S)|/(722 \times 65143)$ homotopically distinct embedded essential arcs of length at most $2d$.

Now, as in the proof of Theorem 2.9, Lemma 2.8 implies that the cusp area satisfies

$$\begin{aligned}\text{Area}(\partial H) &\geq \frac{\sqrt{3}\,|\chi(S)|}{722 \times 65143} e^{-4d} \\ &\geq \frac{2\sqrt{3}}{(240\pi)^4(722 \times 65143)}(\text{tw}(K) - 2) \\ &\geq 2.278 \times 10^{-19}(\text{tw}(K) - 2).\end{aligned}$$

□

References

[1] **T Abe**, **I D Jong**, **J Luecke**, **J Osoinach**, *Infinitely many knots admitting the same integer surgery and a 4–dimensional extension*, preprint (2014) arXiv

[2] **C C Adams**, *Waist size for hyperbolic 3–manifolds, II*, unpublished manuscript

[3] **C C Adams**, *Augmented alternating link complements are hyperbolic*, from: "Low-dimensional topology and Kleinian groups", (D B A Epstein, editor), London Math. Soc. Lecture Note Ser. 112, Cambridge Univ. Press (1986) 115–130 MR

[4] **C Adams**, *Hyperbolic knots*, from: "Handbook of knot theory", (W Menasco, M Thistlethwaite, editors), Elsevier B. V., Amsterdam (2005) 1–18 MR

[5] **C Adams**, **A Colestock**, **J Fowler**, **W Gillam**, **E Katerman**, *Cusp size bounds from singular surfaces in hyperbolic 3–manifolds*, Trans. Amer. Math. Soc. 358 (2006) 727–741 MR

[6] **I Agol**, *Bounds on exceptional Dehn filling*, Geom. Topol. 4 (2000) 431–449 MR

[7] **I Agol**, **P A Storm**, **W P Thurston**, *Lower bounds on volumes of hyperbolic Haken 3–manifolds*, J. Amer. Math. Soc. 20 (2007) 1053–1077 MR

[8] **K Böröczky**, *Packing of spheres in spaces of constant curvature*, Acta Math. Acad. Sci. Hungar. 32 (1978) 243–261 MR

[9] **R D Canary**, **D B A Epstein**, **P Green**, *Notes on notes of Thurston*, from: "Analytical and geometric aspects of hyperbolic space", (D B A Epstein, editor), London Math. Soc. Lecture Note Ser. 111, Cambridge Univ. Press (1987) 3–92 MR

[10] **D Cooper**, **M Lackenby**, *Dehn surgery and negatively curved 3–manifolds*, J. Differential Geom. 50 (1998) 591–624 MR

[11] **M Culler**, **N M Dunfield**, **J R Weeks**, *SnapPy, a computer program for studying the geometry and topology of 3–manifolds* Available at `http://snappy.computop.org`

[12] **O T Dasbach**, **X-S Lin**, *A volumish theorem for the Jones polynomial of alternating knots*, Pacific J. Math. 231 (2007) 279–291 MR

[13] **M Eudave-Muñoz**, **J Luecke**, *Knots with bounded cusp volume yet large tunnel number*, J. Knot Theory Ramifications 8 (1999) 437–446 MR

[14] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Dehn filling, volume, and the Jones polynomial*, J. Differential Geom. 78 (2008) 429–464 MR

[15] **D Futer**, **E Kalfagianni**, **J S Purcell**, *Cusp areas of Farey manifolds and applications to knot theory*, Int. Math. Res. Not. 2010 (2010) 4434–4497 MR

[16] **D Futer**, **S Schleimer**, *Cusp geometry of fibered 3–manifolds*, Amer. J. Math. 136 (2014) 309–356 MR

[17] **R Kirby**, **editor**, *Problems in low-dimensional topology*, from: "Geometric topology", (W H Kazez, editor), AMS/IP Stud. Adv. Math. 2, Amer. Math. Soc., Providence, RI (1997) 35–473 MR

[18] **M Lackenby**, *Word hyperbolic Dehn surgery*, Invent. Math. 140 (2000) 243–282 MR

[19] **M Lackenby**, *The volume of hyperbolic alternating link complements*, Proc. London Math. Soc. 88 (2004) 204–224 MR With an appendix by Ian Agol and Dylan Thurston

[20] **M Lackenby**, **J S Purcell**, *Essential twisted surfaces in alternating link complements*, preprint (2014) arXiv

[21] **W Menasco**, *Closed incompressible surfaces in alternating knot and link complements*, Topology 23 (1984) 37–44 MR

[22] **W Menasco**, **M Thistlethwaite**, *The classification of alternating links*, Ann. of Math. 138 (1993) 113–171 MR

[23] **J K Osoinach, Jr**, *Manifolds obtained by surgery on an infinite number of knots in S^3*, Topology 45 (2006) 725–733 MR

[24] **J S Purcell**, *Cusp shapes under cone deformation*, J. Differential Geom. 80 (2008) 453–500 MR

[25] **D Rolfsen**, *Knots and links*, corrected reprint of 1st edition, Mathematics Lecture Series 7, Publish or Perish, Houston, TX (1990) MR

[26] **M Thistlethwaite**, **A Tsvietkova**, *An alternative approach to hyperbolic structures on link complements*, Algebr. Geom. Topol. 14 (2014) 1307–1337 MR

[27] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at http://msri.org/publications/books/gt3m

[28] **W P Thurston**, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. 6 (1982) 357–381 MR

[29] **J R Weeks**, *SnapPea* Available at http://www.geometrygames.org/SnapPea/

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

School of Mathematical Sciences, Monash University
9 Rainforest Walk, Victoria 3800, Australia

lackenby@maths.ox.ac.uk, jpurcell@mathematics.byu.edu

Proposed: Ian Agol
Seconded: Martin Bridson, Walter Neumann

Received: 26 November 2014
Revised: 19 August 2015