Geometry & Topology

Volume 28 (2024)

The signature and cusp geometry of hyperbolic knots

Alex Davies

András Juhász

Marc Lackenby

Nenad Tomašev

msp

msp

Geometry & Topology **28**:5 (2024) 2313–2343
DOI: 10.2140/gt.2024.28.2313
Published: 24 August 2024

The signature and cusp geometry of hyperbolic knots

ALEX DAVIES
ANDRÁS JUHÁSZ
MARC LACKENBY
NENAD TOMAŠEV

We introduce a new real-valued invariant, called the natural slope of a hyperbolic knot in the 3–sphere, which is defined in terms of its cusp geometry. We show that twice the knot signature and the natural slope differ by at most a constant times the hyperbolic volume divided by the cube of the injectivity radius. This inequality was discovered using machine learning to detect relationships between various knot invariants. It has applications to Dehn surgery and to 4–ball genus. We also show a refined version of the inequality, where the upper bound is a linear function of the volume, and the slope is corrected by terms corresponding to short geodesics that link the knot an odd number of times.

57K10, 57K31, 57K32, 68T07

1 Introduction

In low-dimensional topology, there are two very different types of invariant: those derived from hyperbolic structures on 3–manifolds, and those with connections to 4–dimensional manifolds. Of the latter type, one of the most fundamental invariants is the signature of a knot. Our main goal in this paper is to establish a new and unexpected connection between these two fields. We will show that the cusp geometry of a hyperbolic knot in the 3–sphere encodes information about the signature of the knot.

One of the most important geometric features of a hyperbolic knot K is its maximal cusp. The boundary of this cusp is a Euclidean torus that forms the boundary of a regular neighbourhood of K. This torus is isometric to $\mathbb{C}/\Lambda$ for a lattice Λ in $\mathbb{C}$. The meridian and longitude of the knot give generators μ and λ for Λ. The parallelogram in $\mathbb{C}$ spanned by 0, μ, λ and $\mu+\lambda$ forms a fundamental domain for the action of Λ on $\mathbb{C}$. We introduce a new geometric quantity, called the *natural slope*, that measures how far this parallelogram is from being right-angled. It can be defined by the formula

$$\operatorname{slope}(K) = \operatorname{Re}\Big(\frac{\lambda}{\mu}\Big).$$

Alternatively, natural slope can be defined as follows. Pick a geodesic on the torus $\mathbb{C}/\Lambda$ that represents a meridian. Choose any point on such a geodesic and send off a geodesic orthogonally from this point. It runs along the knot and eventually it comes back to the initial meridian; see Figure 1. In doing so, it has

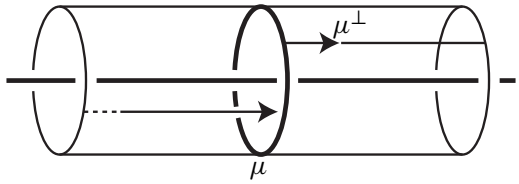

Figure 1: A geodesic running in the direction $\mu^\perp$ that is perpendicular to the meridian μ. By the time it returns to the meridian, it has travelled one longitude minus some multiple s of the meridian. This real number s is the natural slope of K.

gone along a longitude minus some number s of meridians. This number s is not necessarily an integer because the geodesic may return to a different point along the meridian from where it started. This real number s is the natural slope of K.

Quantities with a resemblance to the natural slope have been defined by [Benard et al. 2021; Degtyarev et al. 2022]. However, these other quantities do not seem to be directly related to natural slope, and none of these previous articles seems to provide a connection between hyperbolic geometry and signature.

Experimentally, starting from the plot in Figure 2, we have observed that the natural slope of K is very highly correlated with $2\sigma(K)$, where $\sigma(K)$ is the signature. See Figure 3 for plots of signature versus slope for knots up to 16 crossings in the Regina census [Burton et al. 1999–2021] and for random knots generated by SnapPy [Culler et al. 2021] having 10–80 crossings in their SnapPy-simplified forms. Our goal in this paper is to prove that such a surprising connection holds and to explore its consequences. Our first main result, which we prove in Section 4, establishes that slope(K) is approximately equal to $2\sigma(K)$, but with an additive error that can be bounded by geometric quantities.

Theorem 1.1 *There exists a constant c_1 such that, for any hyperbolic knot K,*

$$|2\sigma(K) - \text{slope}(K)| \leq c_1 \,\text{vol}(K)\,\text{inj}(K)^{-3}.$$

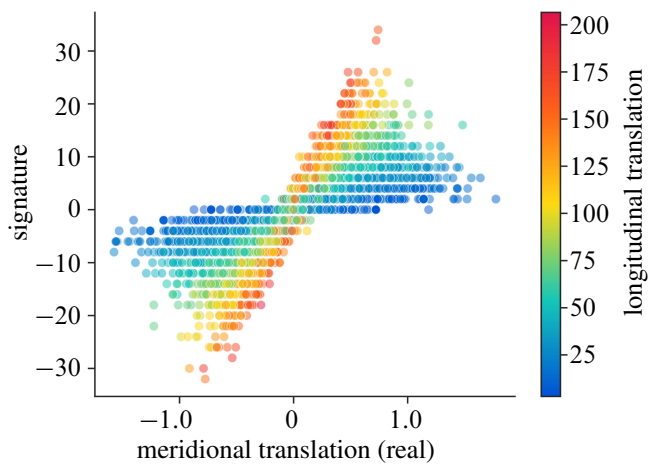

Figure 2: A plot of signature versus the real part of the meridional translation, Re(μ), coloured by longitudinal translation, for a dataset of knots randomly generated by SnapPy.

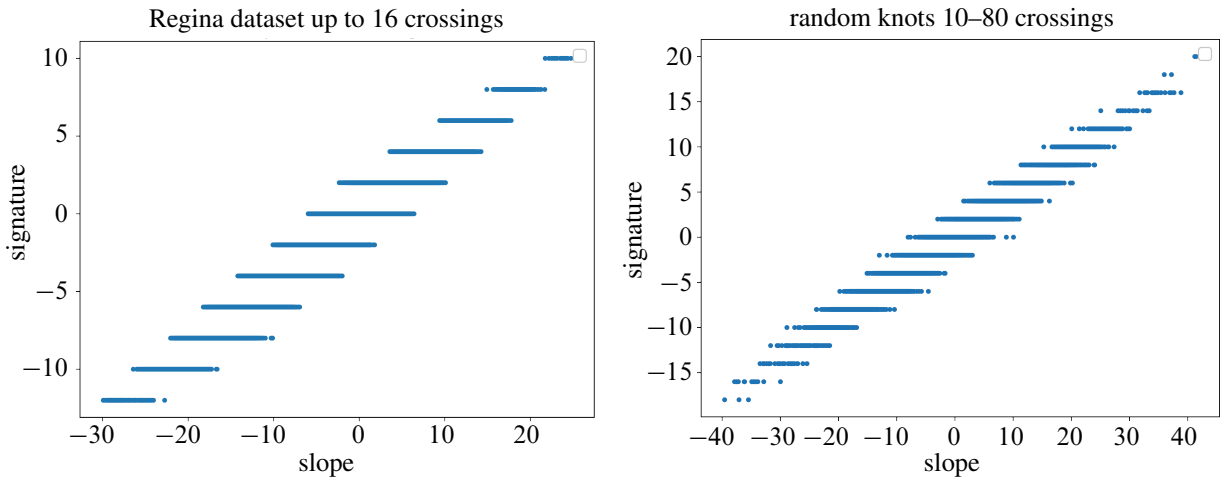

Figure 3: A plot of signature versus slope for knots up to 16 crossings in the Regina census (left) and for a dataset of knots randomly generated by SnapPy having 10–80 crossings in their SnapPy-simplified form (right).

Here $\mathrm{vol}(K)$ is the hyperbolic volume of the complement of K. Also, $\mathrm{inj}(K)$ is the *injectivity radius* of $S^3 \setminus K$, which we define to be

$$\mathrm{inj}(K) = \inf\{\mathrm{inj}_x(S^3 \setminus K) : x \in (S^3 \setminus K) \setminus N\},$$

where N is a maximal cusp and $\mathrm{inj}_x(S^3 \setminus K)$ denotes the injectivity radius of a point x in $S^3 \setminus K$. Note that, although $\mathrm{inj}(K)^{-3}$ appears in the inequality in Theorem 1.1, in practice $\mathrm{inj}(K)$ tends not to be particularly small. (See Figure 12, for example.) Experimental evidence, which we provide in Section 7, suggests that c_1 should be quite small: perhaps $c_1 = 0.3$ suffices. This is based on the largest value 0.234 of $|2\sigma(K) - \mathrm{slope}(K)|\,\mathrm{inj}(K)^3/\mathrm{vol}(K)$ that we managed to obtain by studying a class of knots that are closures of certain braids.

One might wonder whether there is a constant c_2 such that

$$|2\sigma(K) - \mathrm{slope}(K)| \le c_2\,\mathrm{vol}(K)$$

for every hyperbolic knot K. However, we show in Corollary 5.1 that there cannot exist such a constant. We achieve this by exhibiting a sequence of examples that are obtained by twisting three strands of a hyperbolic knot. Nevertheless, we can estimate $\sigma(K)$ in terms of geometric quantities, with an error that is at most a linear function of $\mathrm{vol}(K)$. The main term in this estimate is $\frac{1}{2}\,\mathrm{slope}(K)$, but there are also correction terms that are defined using the complex length of short geodesics. From the complex lengths, the following parameters are computed:

Definition 1.2 Let γ be a geodesic in a hyperbolic 3–manifold with complex length $\mathrm{cl}(\gamma)$. Here $\mathrm{cl}(\gamma)$ is chosen so that $\mathrm{Im}(\mathrm{cl}(\gamma)) \in (-\pi, \pi]$. The *twisting parameter* $\mathrm{tw}(\gamma) = (\mathrm{tw}_p(\gamma), \mathrm{tw}_q(\gamma))$ is the pair (p, q) of coprime integers satisfying the following:

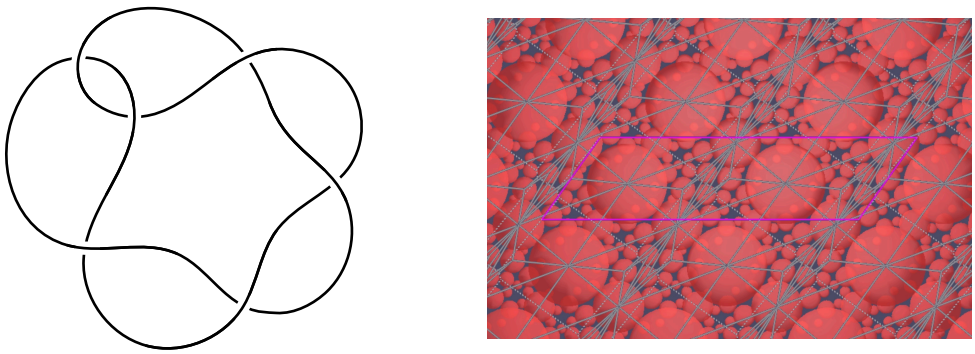

Figure 4: The stevedore knot 6_1 (left), which is a slice knot, and its cusp torus (right), as provided by SnapPy [Culler et al. 2021]. The longitude is 3.9279 and the meridian is $0.7237 + 1.0160i$. Its natural slope is 1.8267 and its signature is 0.

(1) p is even and q is odd and nonnegative.

(2) Subject to this condition, the quantity $|\mathrm{cl}(\gamma)p + 2\pi i q|$ is minimised.

(3) If there are several values of (p, q) for which this quantity is minimised, then choose the pair that is minimal with respect to lexicographical ordering.

Consider a hyperbolic knot K in S^3. For any $\varepsilon \in \mathbb{R}_+$ less than the Margulis constant ε_3, let $\mathrm{OddGeo}\big(\frac{1}{2}\varepsilon\big)$ denote the set of geodesics with length less than $\frac{1}{2}\varepsilon$ and having odd linking number with K. For $p, q \in \mathbb{Z}_+$, the signature correction term $\kappa(p, q)$ is given by Definition 4.2 and satisfies

$$\kappa(p, q) = -\sigma(T(p, q)) - \tfrac{1}{2}pq,$$

where $T(p, q)$ is the (p, q)–torus knot. Then we have the following refinement of Theorem 1.1, which we prove in Section 6, that does not depend on the injectivity radius:

Theorem 1.3 *Let ε_3 be the Margulis constant and let $\varepsilon \in (0, \varepsilon_3)$. Then there is a constant c_4 (depending on ε) such that, for any hyperbolic knot K, the quantities $\sigma(K)$ and*

$$\tfrac{1}{2}\,\mathrm{slope}(K) - \sum_{\gamma \in \mathrm{OddGeo}(\varepsilon/2)} \kappa(\mathrm{tw}_p(\gamma), \mathrm{tw}_q(\gamma))$$

differ by at most $c_4\,\mathrm{vol}(K)$.

Figures 4 and 5 illustrate the relationship between signature and slope in Theorem 1.1 for the knots 6_1 and $12a52$, respectively.

Theorem 1.1 has applications in low-dimensional topology. On the one hand, the signature of K controls the cusp shape, which in turn has consequences for the possible exceptional surgeries on K. On the other hand, the cusp shape controls the signature, which has consequence for the 4–ball genus of K. We now provide these applications.

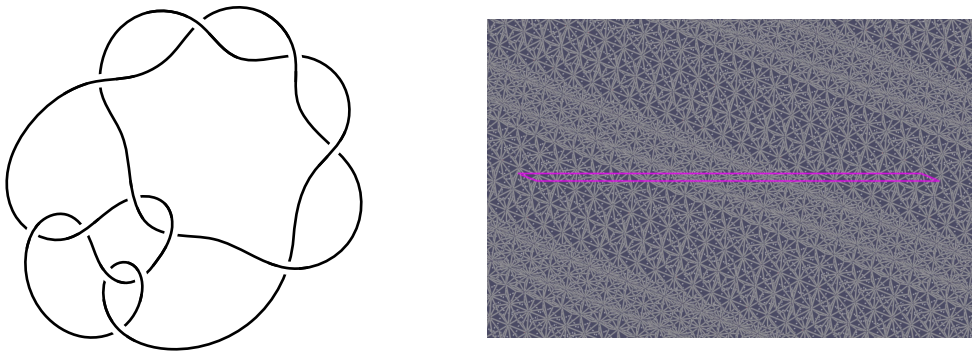

Figure 5: The knot 12a52 (left) and its cusp torus (right). The longitude is 27.7228 and the meridian is $-1.2838 + 0.5145i$. Its natural slope is -18.6064 and its signature is -8. Note how far the parallelogram is from being right-angled; this is the defining feature of having very positive or very negative slope.

1.1 An application to Dehn surgery

Cusp geometry is well known to control the exceptional surgeries on a knot K. Recall that a slope s on $\partial N(K)$ is said to be *exceptional* if the manifold $K(s)$ obtained by Dehn filling along s does not admit a hyperbolic structure.

The *length* of a slope $s = q/p \in \mathbb{Q}$, denoted by $\ell(s)$, is defined to be the length of any geodesic representative of $s = p\lambda + q\mu$ in the boundary of the maximal cusp. A theorem of Agol [2000] and Lackenby [2000] states that, if $\ell(s) > 6$, then s is not exceptional.

We relate slope length to natural slope, using the following simple geometric lemma, which we will prove in Section 2:

Lemma 1.4 *If K is a hyperbolic knot, then the length of the slope q/p satisfies*

$$\ell\Big(\frac{q}{p}\Big) \geq |p \operatorname{slope}(K) + q|.$$

Hence, if q/p is exceptional, then

$$\frac{q}{p} \in \Big[-\operatorname{slope}(K) - \frac{6}{p}, -\operatorname{slope}(K) + \frac{6}{p}\Big].$$

Given that $\operatorname{slope}(K)$ and $2\sigma(K)$ are highly correlated, one would expect that any exceptional slope q/p should lie within a short interval around $-2\sigma(K)$. It is also known that $|p| \leq 8$, by a theorem of Lackenby and Meyerhoff [2013]. Hence, we obtain a bounded set of slopes that contains all the exceptional ones, and that is defined in terms of the signature.

An interesting case is the $(-2, 3, 7)$–pretzel knot $12n242$. This has signature -8 and slope approximately -18.215. It has seven exceptional slopes: 16, 17, 18, $\frac{37}{2}$, 19 and 20. Observe that these slopes are

concentrated in a short interval $[16, 20]$ that contains both $-\text{slope}(K)$ and $-2\sigma(K)$. This close correlation between the exceptional slopes and $-2\sigma(K)$ seems to be a phenomenon that had not previously been observed. Specifically, we have the following consequence of our main theorem:

Corollary 1.5 *If K is a hyperbolic knot and q/p is a slope satisfying*

$$\left|\frac{q}{p} + 2\sigma(K)\right| > \frac{6}{|p|} + c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3} \quad \text{or} \quad |p| > 8,$$

then the manifold $K(q/p)$ obtained by q/p Dehn surgery along K is hyperbolic.

Theorem 1.3 gives a similar bound on slopes resulting in hyperbolic surgeries that does not involve $\operatorname{inj}(K)$.

1.2 An application to 4–ball genus

One of the most important 4–dimensional quantities associated to a knot K is its *4–ball genus* $g_4(K)$. This is defined to be the minimal possible genus of a smoothly embedded compact orientable surface in the 4–ball B^4 with boundary K. One can also define the *topological 4–ball genus* $g_4^{\text{top}}(K)$ by considering locally flat topologically embedded compact orientable surfaces with boundary K. The inequality $g_4(K) \geq g_4^{\text{top}}(K)$ is immediate.

The following result provides a lower bound on $g_4^{\text{top}}(K)$ in terms of purely hyperbolic data. This follows immediately from our main theorem together with the well-known inequality $g_4^{\text{top}}(K) \geq \frac{1}{2}|\sigma(K)|$.

Corollary 1.6 *The topological 4–ball genus $g_4^{\text{top}}(K)$ of a hyperbolic knot K satisfies*

$$g_4^{\text{top}}(K) \geq \tfrac{1}{4}|\text{slope}(K)| - \tfrac{1}{4}c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3}.$$

This corollary seems to be the first time that information about the 4–ball genus has been obtained in terms of hyperbolic geometry. Again, Theorem 1.3 gives a similar lower bound on $g_4^{\text{top}}(K)$ that does not involve $\operatorname{inj}(K)$.

1.3 Spanning surfaces

Theorem 1.1 is proved using a new construction of spanning surfaces with a specified slope. It is of independent interest.

Theorem 1.7 *There is a constant c_3 such that every hyperbolic knot K in S^3 has an unoriented spanning surface F satisfying*

$$|\chi(F)| \leq c_3 \operatorname{vol}(K) \operatorname{inj}(K)^{-3}.$$

Moreover, the boundary slope of this surface is $n/1$, where n is an even integer that is closest to $\text{slope}(K)$.

We prove this in Section 3. The *crosscap number* of a knot K is the minimum of $b_1(F)$ for F an unoriented spanning surface of K. When K is hyperbolic, the above theorem gives an upper bound on a version of the crosscap number where ∂F has slope $n/1$.

Theorem 1.1 is proved by combining this result with a theorem of Gordon and Litherland [1978], which asserts that one can compute the signature of a knot K using any spanning surface F for K; see Theorem 4.1.

Note that slope also gives a lower bound on the Seifert genus:

$$\frac{1}{4\pi}|\text{slope}(K)| + \tfrac{1}{2} \leq g(K);$$

see Proposition 2.5.

1.4 Highly twisted knots

In Section 5, we show the following result for highly twisted knots:

Theorem 1.8 *Let K be a knot in the 3–sphere and let $C_1, \dots, C_n$ be a collection of disjoint simple closed curves in the complement of K that bound disjoint discs. Suppose that $S^3 \setminus (K \cup C_1 \cup \cdots \cup C_n)$ is hyperbolic. Let $K(q_1, \dots, q_n)$ be the knot obtained from K by adding q_i full twists to the strings going through C_i for each $i \in \{1, \dots, n\}$. Let ℓ_i be the linking number between C_i and K, when they are both given some orientation. Suppose that $\ell_1, \dots, \ell_m$ are even and $\ell_{m+1}, \dots, \ell_n$ are odd. Then there is a constant k, depending on K and $C_1, \dots, C_n$, such that, provided each $|q_i|$ is sufficiently large,*

$$\left|\text{slope}(K(q_1, \dots, q_n)) + \sum_{i=1}^{n} \ell_i^2 q_i\right| \leq k, \quad \left|\sigma(K(q_1, \dots, q_n)) + \left(\frac{1}{2}\sum_{i=1}^{m} \ell_i^2 q_i + \frac{1}{2}\sum_{i=m+1}^{n} (\ell_i^2 - 1) q_i\right)\right| \leq k.$$

The slight difference between the behaviour of $\sigma(K(q_1, \dots, q_n))$ and that of $\frac{1}{2}\,\text{slope}(K(q_1, \dots, q_n))$ as the q_i tend to infinity enables us to construct families of knots that show the injectivity radius cannot be dropped from Theorem 1.1.

1.5 Methodology

One of the novel aspects of this work was the use of machine learning. We embarked with the aim of discovering new relationships between various 3–dimensional invariants. By using machine learning, we observed an unexpected nonlinear relationship between $\sigma(K)$ and $\text{Re}(\mu)$, the real part of the meridional translation μ. This led us to define the natural slope, which we observed to have a strong linear correlation with $\sigma(K)$. Theorems 1.1 and 1.3 are the results of our attempts to prove this correlation.

2 Hyperbolic knots and natural slope

A knot K is hyperbolic if its complement $S^3 \setminus K$ admits a complete finite-volume hyperbolic metric. By the Mostow rigidity theorem [1968], the hyperbolic structure is unique up to isometry; hence, every geometric invariant of the hyperbolic structure on $S^3 \setminus K$ is a topological invariant of the knot. For example, the volume $\text{vol}(K) := \text{vol}(S^3 \setminus K)$ and the injectivity radius $\text{inj}(K)$ defined in the introduction are such invariants.

For a pair of coprime integers (p,q), the *torus knot* $T(p,q)$ is one that can be drawn on the surface of the standard torus in the 3–sphere and winds p times in the longitude direction and q times along the meridian. Given a knot K in S^3 and a knot K' in the solid torus $S^1\times D^2$, one can form the satellite of K with pattern K' by mapping the solid torus in a neighbourhood of K, and considering the image of K'. By the work of Thurston [Morgan 1984], a knot is hyperbolic if and only if it is not a torus knot or a satellite knot. In particular, every hyperbolic knot is prime, ie not the connected sum of two nontrivial knots. In other words, one can build all knots from hyperbolic knots and torus knots using satellite operations.

Definition 2.1 For any hyperbolic knot K, the end of $S^3\setminus K$ has a neighbourhood called a *cusp*. The boundary ∂N of a maximal cusp neighbourhood $N\subset S^3\setminus K$ is a Euclidean torus. Identify ∂N with $\mathbb{C}/\Lambda$, where $\mathbb{C}$ is the complex plane and Λ is a lattice in $\mathbb{C}$. We arrange this identification so that the longitude lifts to a straight line in $\mathbb{C}$ starting at 0 and ending at some $\lambda\in\mathbb{R}_{>0}$. This is the knot's *longitudinal translation*. Given this normalisation, the meridian lifts to a straight line starting at 0 and ending at some complex number μ with $\mathrm{Im}(\mu)>0$. This is the *meridional translation* of K.

We remark that the real part of meridional translation $\mathrm{Re}(\mu)$ in the KnotInfo data set [Livingston and Moore 2021] for knots with at most 12 crossings is listed without signs. However, SnapPy [Culler et al. 2021] does compute the sign for hyperbolic knots.

Note that $|\mu|\leq 6$, where $|\mu|$ denotes the length of the meridian. Indeed, by work of Agol [2000] and Lackenby [2000], Dehn filling along a slope longer than 6 gives a hyperbolic 3–manifold, while Dehn filling along the meridian is S^3, which is not hyperbolic. Furthermore, any curve on the cusp torus ∂N has length at least 1. In particular, $|\mu|\geq 1$.

If S is an essential surface with connected boundary in a hyperbolic 3–manifold, then $\ell(\partial S)\leq -2\pi\chi(S)$; see Cooper and Lackenby [1998, Theorem 5.1] or Hass, Rubinstein and Wang [Hass et al. 1999, (6)]. If S is a Seifert surface for a knot K, then $\chi(S)=1-2g(S)$. Hence, if K is hyperbolic, then

$$|\lambda|\leq 4\pi g(K)-2\pi, \tag{2.2}$$

where $g(K)$ is the Seifert genus of K.

For the maximal cusp neighbourhood N, we have

$$\mathrm{vol}(\partial N)=2\,\mathrm{vol}(N)\leq 2\,\mathrm{vol}(K),$$

and $\mathrm{vol}(\partial N)\leq|\lambda||\mu|$. On the other hand, by a result of Lackenby and Purcell [2016], there is a constant C such that, for K alternating,

$$C\,\mathrm{vol}(K)\leq\mathrm{vol}(\partial N).$$

Based on experimental data, one might ask if this also holds for random knots.

Definition 2.3 The *natural slope* $\mathrm{slope}(K)$ of a hyperbolic knot K is defined as follows. Let $\mu^\perp$ be a unit vector at the origin of $\mathbb{C}$ orthogonal to μ. Then some multiple of $\mu^\perp$ is equal to $\lambda-s\mu$ for some $s\in\mathbb{R}$. Then $\mathrm{slope}(K):=s$.

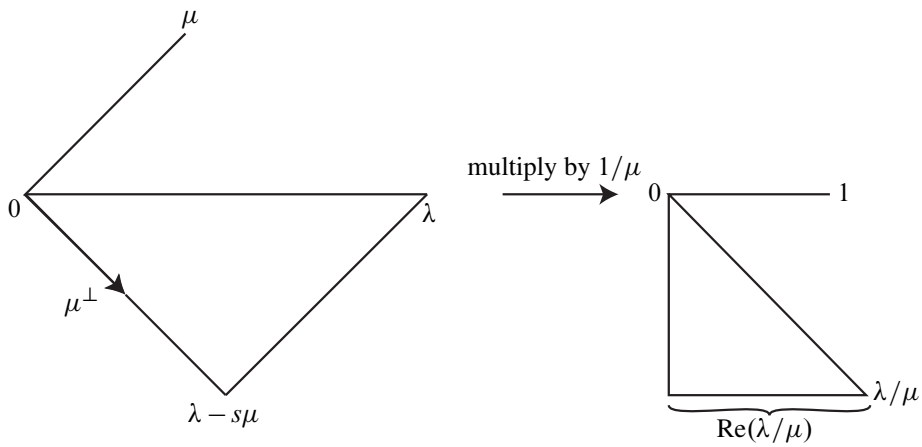

Figure 6: The calculation of natural slope.

Lemma 2.4
$$\operatorname{slope}(K) = \operatorname{Re}\Big(\frac{\lambda}{\mu}\Big) = \frac{\lambda \operatorname{Re}(\mu)}{|\mu|^2}.$$

Proof Figure 6 shows a lift of the cusp torus to the complex plane $\mathbb{C}$. The point $\lambda - s\mu$ is shown (which is a multiple of $\mu^\perp$). If we apply the transformation to $\mathbb{C}$ that is multiplication by $1/\mu$, then $\mu^\perp$ becomes purely imaginary. So $\lambda/\mu - s$ is purely imaginary. Hence, $s = \operatorname{Re}(\lambda/\mu)$. This is also equal to $\lambda \operatorname{Re}(\mu)/|\mu|^2$. □

We are now ready to prove Lemma 1.4 from the introduction:

Proof of Lemma 1.4 We have $\ell(q/p) = |p\lambda + q\mu|$. Since $\lambda \in \mathbb{R}$,
$$\ell\Big(\frac{q}{p}\Big)^2 = p^2\lambda^2 + 2pq\lambda \operatorname{Re}(\mu) + q^2|\mu|^2.$$
On the other hand, by Lemma 2.4, we have $\operatorname{slope}(K) = \lambda \operatorname{Re}(\mu)/|\mu|^2$. Hence,
$$|p \operatorname{slope}(K) + q|^2 = p^2\lambda^2 \frac{\operatorname{Re}(\mu)^2}{|\mu|^4} + 2pq\lambda \frac{\operatorname{Re}(\mu)}{|\mu|^2} + q^2 \le \ell\Big(\frac{q}{p}\Big)^2$$
since $|\mu| \ge 1$. □

Slope gives a lower bound on the Seifert genus:

Proposition 2.5 *If K is a hyperbolic knot in S^3, then*
$$\frac{1}{4\pi}|\operatorname{slope}(K)| + \tfrac{1}{2} \le g(K).$$

Proof By (2.2), we have $|\lambda| \le 4\pi g(K) - 2\pi$. Furthermore, $|\mu| \ge 1$. Together with Lemma 2.4, we obtain that
$$|\operatorname{slope}(K)| = |\lambda| \frac{|\operatorname{Re}(\mu)|}{|\mu|^2} \le \frac{|\lambda|}{|\mu|} \le 4\pi g(K) - 2\pi,$$
and the result follows. □

3 Proof of Theorem 1.7

The key to proving Theorem 1.7 is the construction of a nice triangulation of a hyperbolic knot complement:

Proposition 3.1 *There is a constant c_1 such that, for every hyperbolic knot K in S^3 with embedded cusp neighbourhood N, there is a triangulation $\mathcal{T}$ of $M := S^3 \setminus (K \cup \mathrm{int}(N))$ with the following properties:*

(1) *The number t of tetrahedra of $\mathcal{T}$ is at most $c_1 \,\mathrm{vol}(K)\,\mathrm{inj}(K)^{-3}$.*

(2) *If n is a closest even integer to $\mathrm{slope}(K)$, then $\nu := \lambda - n\mu$ (see Definition 2.3) is a normal curve in ∂M that intersects each edge at most once.*

Proof We remark that the validity of the conclusion in the proposition does not depend on the choice of embedded cusp neighbourhood N. We will pick N as follows. Let $N_{\max}$ be the maximal cusp neighbourhood. Retract this to form the embedded cusp neighbourhood N, so each point of ∂N has distance 0.5 from $\partial N_{\max}$. Note that the Euclidean metric on ∂N is obtained from that of $\partial N_{\max}$ by scaling by the factor $e^{-0.5} = 1/\sqrt{e}$.

Let $\varepsilon := \frac{1}{2}\,\mathrm{inj}(K)$. We use a variation of Jørgensen's and Thurston's method [Thurston 1979, Section 5.11] to build the triangulation $\mathcal{T}$. (See also [Breslin 2009; Kobayashi and Rieck 2011].)

We pick a maximal collection of points in ∂M that are all at least $\frac{1}{8}\varepsilon$ from each other. We will extend this to a collection of points P in M without adding any new points in ∂M. Our aim is to ensure that the Voronoi diagram for P in M restricts to the Voronoi diagram for $P \cap \partial M$ in ∂M, where the latter is given its Euclidean metric. Recall that the Voronoi diagram [1908a; 1908b] corresponding to P is a cell structure of M where the interior of every 3–cell consists of the set of points in M that are closer to a specific point of P than any other point of P. Similarly, the Voronoi diagram for $P \cap \partial M$ is a cell structure of M where the interior of every 2–cell consists of the set of points in ∂M that are closer (in the Euclidean metric) to a specific point of $P \cap \partial M$ than any other point of $P \cap \partial M$.

The Voronoi diagram for M can be constructed as follows. The universal cover $\mathbb{H}^3 \to S^3 \setminus K$ restricts to the universal cover $\widetilde{M} \to M$. This set $\widetilde{M}$ is obtained from $\mathbb{H}^3$ by removing the interior of the inverse image of N. We may arrange that one component of this inverse image is a horoball $N_\infty = \{(x, y, z) : z \geq k\}$ in the upper half-space model for $\mathbb{H}^3$ for some $k > 0$. Let $\widetilde{P}$ denote the inverse image of P in $\widetilde{M}$. Each cell of the Voronoi diagram for M is the image of a cell for the Voronoi diagram for $\widetilde{P}$ in $\widetilde{M}$. Each 2–cell that does not lie in $\partial\widetilde{M}$ is equidistant from two points of $\widetilde{P}$. Hence, it is totally geodesic. Our aim is to ensure that each such 2–cell that intersects the horosphere ∂N_∞ is equidistant between two points of $\widetilde{P} \cap \partial N_\infty$. This will imply that the 2–cell intersects ∂N_∞ in a Euclidean geodesic arc. The union of these arcs forms the 1–skeleton of the Voronoi diagram for $\widetilde{P} \cap \partial N_\infty$ in ∂N_∞. Thus, we can deduce that the Voronoi diagram for P in M restricts to the Voronoi diagram for $P \cap \partial M$ in ∂M.

We now describe how the set P is chosen. We have already picked a maximal collection of points in ∂M that are all at least $\frac{1}{8}\varepsilon$ from each other. This set will be $P \cap \partial M$. We then add points to this set that lie in

the interior of M, but subject to the condition that each of these points in the interior of M has distance at least $\frac{1}{4}\varepsilon$ from the other points in the set. We stop when it is no longer possible to add any further points with this property. Let P be the resulting set of points.

By our choice of P, each point in ∂M has distance less than $\frac{1}{8}\varepsilon$ from some point of $P \cap \partial M$. It also has distance at least $\frac{1}{8}\varepsilon$ from each point of $P \cap \operatorname{int}(M)$. Thus, for each point of ∂M, each of its closest points in P also lies in ∂M.

Now consider a 2–cell of the Voronoi diagram for $\widetilde{M}$ that intersects ∂N_∞ but does not lie in ∂N_∞. This is equidistant between two points p_1 and p_2 of $\widetilde{P}$. The intersection between this 2–cell and ∂N_∞ is an arc. Let x be any point in the interior of this arc. Then x is equidistant between p_1 and p_2, and these are the closest two points of $\widetilde{P}$ to x. As argued above, any point of $\widetilde{P}$ that is closest to x must lie in $\partial\widetilde{M}$. We will show that, in fact, p_1 and p_2 lie in ∂N_∞. Suppose not. Then one of these points lies in $\partial\widetilde{M} \setminus \partial N_\infty$. The shortest arc from x to $\partial\widetilde{M} \setminus \partial N_\infty$ must run through the inverse image of $\partial N_{\max}$. One component of this inverse image is a horosphere about the point at infinity, with distance 0.5 from ∂N_∞. Hence, the length of this arc is at least 0.5. On the other hand, each point in $\partial\widetilde{M}$ has distance less than $\frac{1}{8}\varepsilon$ from some point of $P \cap \partial\widetilde{M}$. We will show below that $\frac{1}{8}\varepsilon < 0.12 < 0.5$, and hence this is a contradiction.

Thus, we have indeed guaranteed that the restriction to ∂M of the Voronoi diagram for P in M is the Voronoi diagram for $P \cap \partial M$ in ∂M, as claimed. We now subdivide each 2–cell of the Voronoi diagram for M into triangles without introducing any new vertices, and subdivide each 3–cell into tetrahedra by coning off from the point of P lying in it, obtaining the triangulation $\mathcal{T}$ of M. Since the restriction of the Voronoi diagram to ∂M agrees with that arising from its Euclidean metric, this implies that each triangle of $\mathcal{T}$ in ∂M is straight.

Since the open balls of radius $\frac{1}{16}\varepsilon$ about the points of P are pairwise disjoint,

$$|P| \operatorname{vol}\big(B\big(\tfrac{1}{16}\varepsilon\big)\big) \leq \operatorname{vol}(S^3 \setminus K),$$

where $B\big(\frac{1}{16}\varepsilon\big)$ is a ball in $\mathbb{H}^3$ of radius $\frac{1}{16}\varepsilon$.

We claim that the number t_p of tetrahedra of $\mathcal{T}$ incident to a point $p \in P$ is at most a universal constant k. Indeed, when p lies in the interior of M, t_p is exactly the number of triangles in the boundary 2–sphere S of the 3–cell of the Voronoi diagram containing p. When p lies in the boundary of M, t_p is the number of triangles in this sphere that are not incident to p. When a vertex of one of these triangles lies in the interior of M, it is equidistant from at least four points of P, one of which is p. When a vertex of the triangles lies on the boundary of M, it is equidistant from at least three points of P, one of which is p. So a vertex in S is specified by choosing two or three other points of P, each of which is at most $\frac{1}{2}\varepsilon$ from p. The ball $B\big(p, \frac{1}{2}\varepsilon\big)$ is embedded in $S^3 \setminus K$, since $\frac{1}{2}\varepsilon = \frac{1}{4}\operatorname{inj}(K)$, and hence lifts to a ball B in $\mathbb{H}^3$. The balls of radius $\frac{1}{16}\varepsilon$ about the inverse image of P in B are disjoint, and lie within $B\big(\frac{9}{16}\varepsilon\big)$. So the number

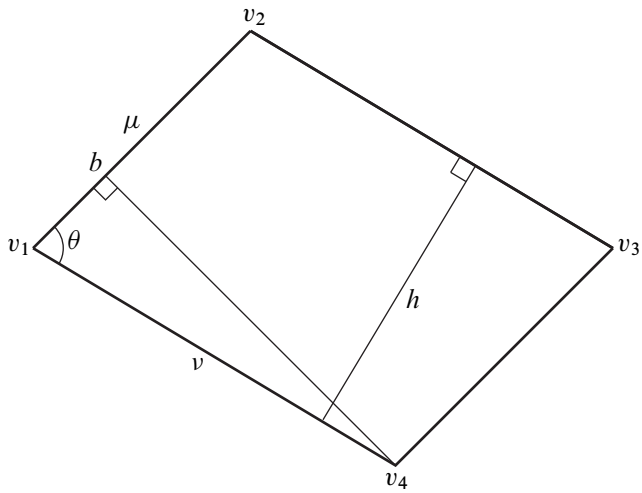

Figure 7: A fundamental domain D in $\partial N_{\max}$ with sides μ and ν.

of points of P at most $\frac{1}{2}\varepsilon$ from p is bounded above by

$$k_0 := \left\lfloor \frac{\operatorname{vol}\big(B\big(\frac{9}{16}\varepsilon\big)\big)}{\operatorname{vol}\big(B\big(\frac{1}{16}\varepsilon\big)\big)} \right\rfloor.$$

It follows that $t_p \leq k := \binom{k_0}{3}$. Then the total number of tetrahedra

$$t \leq k|P| \leq \frac{k \operatorname{vol}(K)}{\operatorname{vol}\big(B\big(\frac{1}{16}\varepsilon\big)\big)} \leq c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3}$$

for a universal constant c_1.

We may pick the Euclidean geodesic representative for the slope ν so that it misses the vertices of $\mathcal{T}$. Hence, ν is a normal curve, because it is a Euclidean geodesic and each triangle of $\mathcal{T}$ in ∂M is straight. We now show ν does not intersect any triangle in ∂M more than once. Let D be a fundamental domain in $\partial N_{\max}$ with sides μ and ν. (See Figure 7.) We will show that the perpendicular distance h between the sides of D that are parallel to ν is at least 0.55. Hence, the perpendicular distance between sides of the corresponding fundamental domain in ∂N_∞ is at least $0.55/\sqrt{e} > 0.33$. On the other hand, we will show that the length of each edge of $\mathcal{T}$ in ∂M is at most 0.23. This will imply that, in the triangulation of ∂M, no triangle can run in D between these opposite sides, and hence that $\mathcal{T}$ satisfies property (2). This will complete the proof.

According to a theorem of Cao and Meyerhoff [2001], the area A of the boundary of the maximal cusp is at least 3.35. Let θ be the angle of two of the four corners of D satisfying $0 < \theta \leq \frac{1}{2}\pi$. Say that this angle is at the vertex v_1 of D, and label the remaining vertices v_2, v_3 and v_4 so that the line joining v_1 to v_2 has slope μ.

Let b be the perpendicular projection of v_4 onto the line joining v_1 and v_2. We claim that b lies between v_1 and v_2, or possibly equals one of these vertices. Place v_1 at the origin in the complex plane. Then

$v_2 = \pm\mu$ and $v_4 = \lambda - n\mu$. Now, by the definition of $s = \text{slope}(K)$, the perpendicular projection of $\lambda - s\mu$ onto the line through v_1 and v_2 is v_1. Hence, the perpendicular projection b of $\lambda - n\mu$ onto this line has distance $|n-s||\mu|$ from v_1. But n is a closest even integer to s, and so $|n-s| \leq 1$. Therefore, b lies between v_1 and v_2, or is equal to one of these points, as claimed.

Hence,

$$\tan\theta \geq \frac{A}{|\mu|^2}$$

and so

$$\sec^2\theta = 1 + \tan^2\theta \geq \frac{A^2 + |\mu|^4}{|\mu|^4}.$$

Therefore,

$$\sin^2\theta = 1 - \cos^2\theta \geq 1 - \frac{|\mu|^4}{A^2 + |\mu|^4} = \frac{A^2}{A^2 + |\mu|^4}.$$

So the distance h satisfies

$$h = |\mu|\sin\theta \geq \frac{|\mu|A}{\sqrt{A^2 + |\mu|^4}}.$$

The square of the reciprocal of this expression is

$$\frac{A^2 + |\mu|^4}{|\mu|^2 A^2} = \frac{1}{|\mu|^2} + \frac{|\mu|^2}{A^2}.$$

It is easy to check that this is a convex function of $|\mu|$ and hence its maximal value over the interval $1 \leq |\mu| \leq 6$ occurs when $|\mu| = 1$ or 6. It also is maximised by taking A as small as possible; in other words, $A = 3.35$. We deduce that h is at least

$$\frac{6 \cdot 3.35}{\sqrt{3.35^2 + 36^2}} \geq 0.55.$$

Hence, the perpendicular distance between sides of the corresponding fundamental domain in ∂N_∞ is at least $0.55/\sqrt{e} > 0.33$.

We now compare this to the maximal length of an edge of $\mathcal{T}$ in ∂M. Each triangle of $\mathcal{T}$ in ∂M lies within a disc centred at a point of $P \cap \partial M$ with radius at most $\frac{1}{8}\varepsilon$. Hence, each triangle has side length at most $\frac{1}{4}\varepsilon = \frac{1}{8}\,\text{inj}(K)$. Now the length L of the shortest slope s on $\partial N_{\max}$ is at most $|\mu| \leq 6$. This gives an upper bound on $\text{inj}(K)$, as follows. By applying an isometry to hyperbolic space, we may arrange that a component of the inverse image of $N_{\max}$ in upper half-space is $\{(x, y, z) : z \geq 1\}$. We may also arrange that a covering transformation corresponding to s is $(x, y, z) \mapsto (x + L, y, z)$. It therefore sends $(0, 0, 1)$ to $(L, 0, 1)$. The hyperbolic distance between these points is at most

$$2\ln\left(\tfrac{1}{2}(6 + \sqrt{40})\right) \leq 3.64.$$

Hence, $\text{inj}(K)$ is at most 1.82 and $\frac{1}{4}\varepsilon$ is at most 0.23. □

Proof of Theorem 1.7 Let the triangulation $\mathcal{T}$ and the curve ν be as in Propositions 3.1. Since $\nu = \lambda - n\mu$ for n even, $[\nu] = [\lambda] \in H_1(\partial M; \mathbb{Z}_2)$, so ν bounds an unoriented surface S in M. If we make S transverse to the 1–skeleton of $\mathcal{T}$, it defines a simplicial 1–cocycle $c \in C^1(M; \mathbb{Z}_2)$ via $c(e) = |S \cap e| \bmod 2$ for each edge e of $\mathcal{T}$. If we connect the midpoints of the edges e of T such that $c(e) = 1$, we obtain a surface F that intersects each tetrahedron T of the triangulation $\mathcal{T}$ in at most one triangle or square. In particular, F is a normal surface. Furthermore, $\partial F = \nu$ as ν is a normal curve that intersects each triangle in ∂M at most once. Discard any closed components of F.

Let t be the number of tetrahedra of $\mathcal{T}$. Furthermore, write v, e and f for the number of vertices, edges and faces of F, respectively. By the above, $f \leq t$. Then $\chi(F) = v - e + f$ and, since F is not a disk, $|\chi(F)| = e - f - v$. Since every face of F is a triangle or a quadrilateral,

$$e \leq \tfrac{1}{2}(4f + e_\partial) \leq t + f + \tfrac{1}{2}e_\partial,$$

where e_∂ is the number of edges of F in ∂M. As $v \geq e_\partial$, we obtain that

$$|\chi(F)| \leq t \leq c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3},$$

where the second inequality is property (1) of $\mathcal{T}$ in Proposition 3.1. □

4 The knot signature

Another fundamental knot invariant is the signature $\sigma(K)$. Given a Seifert surface S for K, ie a compact, oriented and connected surface with boundary K, one can define the Seifert form

$$Q_S \colon H_1(S) \times H_1(S) \to \mathbb{Z}$$

as follows: Given $a, b \in H_1(S)$, we write b^+ for the positive push-off of b into $S^3 \setminus S$. Then $Q_S(a, b) = \operatorname{lk}(a, b^+)$. If V is a matrix of Q_S, then $\sigma(K)$ is the signature of $V + V^T$. The signature is a 4–dimensional invariant, in the sense that it gives a lower bound on the topological 4–ball genus $g_4^{\text{top}}(K)$, which is the minimal genus of a compact, oriented, locally flat, connected surface bounded by K in the 4–ball B^4.

One can also compute the signature of a knot from unoriented surfaces using the work of Gordon and Litherland [1978]. Let F be an unoriented surface bounding a knot K in S^3. Let $\{b_1, \dots, b_n\}$ be a basis of $H_1(F)$, and let b_i' be the double push-off of b_i into $S^3 \setminus F$. Then the *Goeritz matrix* G_F is an $n \times n$ symmetric matrix with $(i, j)^{\text{th}}$ entry $\operatorname{lk}(b_i, b_j')$ for $i, j \in \{1, \dots, n\}$. Furthermore, the *normal Euler number* $e(F)$ of F is defined to be $-\operatorname{lk}(K, K')$, where K' is the framing of K given by F. Gordon and Litherland proved the following:

Theorem 4.1 *Let F be an unoriented surface bounding the knot K in S^3. Then*

$$\sigma(K) = \sigma(G_F) + \tfrac{1}{2}e(F),$$

where $\sigma(G_F)$ is the signature of the Goeritz matrix.

We are now ready to show how Theorem 1.1 follows from Theorem 1.7.

Proof of Theorem 1.1 Let F be the surface provided by Theorem 1.7, with boundary slope $\nu = \lambda - n\mu$, where n is a closest even integer to slope(K). Let G_F be the Goeritz matrix of F. Since

$$|\chi(F)| \leq c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3},$$

we deduce that

$$b_1(F) \leq c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3} + 1,$$

and so $|\sigma(G_F)| \leq c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3} + 1$. Therefore,

$$\begin{aligned} |2\sigma(K) - \operatorname{slope}(K)| &\leq |2\sigma(K) - n| + 1 = |2\sigma(K) + \operatorname{lk}(K, \nu)| + 1 = |2\sigma(G_F)| + 1 \\ &\leq 2c_1 \operatorname{vol}(K) \operatorname{inj}(K)^{-3} + 3 \leq c_2 \operatorname{vol}(K) \operatorname{inj}(K)^{-3} \end{aligned}$$

for the absolute constant

$$c_2 := 2c_1 + \frac{3 \cdot 1.82^3}{2.0298} < 2c_1 + 8.92.$$

Indeed, for any hyperbolic knot K, we have $\operatorname{inj}(K) \leq 1.82$, as shown in the proof of Proposition 3.1, and $\operatorname{vol}(K) > 2.0298$, with the figure eight knot having the smallest volume, by [Cao and Meyerhoff 2001]. □

In the following definition, we introduce the signature correction $\kappa(p, q)$ for integers p and q, which is related to the signature of the (p, q)–torus knot. The correction terms in Theorem 1.3 are defined in terms of $\kappa(p, q)$.

Definition 4.2 For any pair of positive integers (p, q), we define the *signature correction* $\kappa(p, q)$ recursively as follows:

(1) If $p > 2q$ and q is odd, then $\kappa(p, q) = \kappa(p - 2q, q) - 1$.

(2) If $p > 2q$ and q is even, then $\kappa(p, q) = \kappa(p - 2q, q)$.

(3) If $p = 2q$, then $\kappa(p, q) = -1$.

(4) If $q \leq p < 2q$ and q is odd, then $\kappa(p, q) = -\kappa(q, 2q - p) - 1$.

(5) If $q \leq p < 2q$ and q is even, then $\kappa(p, q) = -\kappa(q, 2q - p) - 2$.

(6) If $p < q$, then $\kappa(p, q) = \kappa(q, p)$.

We extend κ to nonzero integers p and q by defining $\kappa(-p, q) = \kappa(p, -q) = -\kappa(p, q)$. When one of p or q is zero, $\kappa(p, q) = 0$.

It is reasonably clear that this gives a well-defined value of $\kappa(p, q)$. This is because it defines $\kappa(p, q)$ uniquely when $p = q$, and, when $p \neq q$, it defines $\kappa(p, q)$ in terms of some $\kappa(p', q')$ where either $q' < q$, or $q' = q$ and $p' < p$. However, the rationale for the definition comes from the following fact, due to Gordon, Litherland and Murasugi [Gordon et al. 1981]:

Theorem 4.3 *The signature of the (p,q)–torus link $T(p,q)$ satisfies*

$$\sigma(T(p,q)) = -\tfrac{1}{2}pq - \kappa(p,q).$$

The signature correction $\kappa(p,q)$ arises naturally as the signature of the Goeritz form of a surface bounding the (p,q)–torus knot, as follows:

Lemma 4.4 *Let V be the standard solid torus in S^3, and let $T(p,q)$ be the curve on ∂V that is the (p,q)–torus knot, where p is even and q is odd. Thus, p is the winding number of $T(p,q)$ in V. Then there is a compact unoriented surface F in V with boundary $T(p,q)$, and $\sigma(G_F) = -\kappa(p,q)$ for any such F.*

Proof Since p is even, $T(p,q)$ is trivial in $H_1(V;\mathbb{Z}_2)$. It therefore bounds an unoriented surface F in V. Applying Gordon and Litherland's signature formula (Theorem 4.1) to F, we deduce that

$$\sigma(T(p,q)) = \sigma(G_F) + \tfrac{1}{2}e(F).$$

The push-off K' of ∂F into F has linking number pq with ∂F. To see this, observe that K' is homologous in V to p times a core curve γ' of V. Similarly, ∂F is homologous in the solid torus $\mathrm{cl}(S^3 \setminus V)$ to q times its core curve γ, which is a meridian of γ'. Thus,

$$\mathrm{lk}(\partial F, K') = pq\,\mathrm{lk}(\gamma,\gamma') = pq;$$

hence, $e(F) = -pq$. So

$$\sigma(G_F) = \sigma(T(p,q)) + \tfrac{1}{2}pq = -\kappa(p,q),$$

where the final equality is Theorem 4.3. □

Lemma 4.5 *Let A be a nonsingular square matrix with real entries. Let A_+ be a nonsingular matrix obtained from A by adding a final row and final column. Then $\sigma(A_+)$ is either $\sigma(A)-1$ or $\sigma(A)+1$.*

Proof Let $\lambda_1 \le \cdots \le \lambda_n$ be the eigenvalues of A, and let $\lambda_1^+ \le \cdots \le \lambda_{n+1}^+$ be the eigenvalues of A_+. Cauchy's interlacing theorem states that

$$\lambda_1^+ \le \lambda_1 \le \lambda_2^+ \le \cdots \le \lambda_n \le \lambda_{n+1}^+.$$

Hence, the number of negative eigenvalues of A_+ is at least the number of negative eigenvalues of A and, similarly, the number of positive eigenvalues of A_+ is at least the number of positive eigenvalues of A. □

Lemma 4.6 *Let V be a solid torus embedded in S^3. Pick a slope λ on ∂V that has winding number 1 in V. Let K be the knot on ∂V that has slope $p\lambda + q\mu$, where μ is the meridian of V, and where p is even and q is odd. Then K bounds a compact unoriented surface F in V with the property that the Goeritz form G_F satisfies $|\sigma(G_F) + \kappa(p,q)| \le 2$.*

Proof Because p is even, K bounds a compact surface F in V. We may pick a basis $e_1,\dots,e_n$ for $H_1(F)$ so that $e_1,\dots,e_{n-1}$ have zero winding number around V. Let V' be an embedding of V in S^3 such that K is sent to $T(p,q)$. Let F' be the image of F.

We claim that the Goeritz forms G_F and $G_{F'}$ agree on the first $n-1$ rows and columns. To prove this, we view V' as the regular neighbourhood of a standard unknot embedded in the horizontal plane. Then, up to isotopy, V can be obtained from V' by applying Reidemeister moves and crossing changes to this unknot. None of these moves affects the first $n-1$ rows and columns of the Goeritz form, for the following reason. Any given entry of the Goeritz form is $\operatorname{lk}(b_i, b_j')$ for a suitable curve b_i in the surface and b_j' the double push-off of another curve in the surface. When the entry of the Goeritz form lies in the first $n-1$ rows and columns, these curves b_i and b_j' have zero winding number around the solid torus. Hence, geometrically, b_i winds an equal number of times around the solid torus in opposite directions, as does b_j'. So, when we perform a Reidemeister move or a crossing change to the solid torus and we compare the resulting projections of $b_i \cup b_j'$ to the horizontal plane, the sum of the signs of the crossings between b_i and b_j' remains unchanged. This sum is $2\operatorname{lk}(b_i, b_j')$. This proves the claim.

Hence, by Lemma 4.5, we have $|\sigma(G_F) - \sigma(G_{F'})| \le 2$. But $\sigma(G_{F'}) = -\kappa(p,q)$ by Lemma 4.4. □

5 Highly twisted knots

The following is Theorem 1.8 from the introduction:

Theorem 1.8 *Let K be a knot in the 3–sphere and let $C_1, \dots, C_n$ be a collection of disjoint simple closed curves in the complement of K that bound disjoint discs. Suppose that $S^3 \setminus (K \cup C_1 \cup \cdots \cup C_n)$ is hyperbolic. Let $K(q_1, \dots, q_n)$ be the knot obtained from K by adding q_i full twists to the strings going through C_i for each $i \in \{1, \dots, n\}$. Let ℓ_i be the linking number between C_i and K, when they are both given some orientation. Suppose that $\ell_1, \dots, \ell_m$ are even and $\ell_{m+1}, \dots, \ell_n$ are odd. Then there is a constant k, depending on K and $C_1, \dots, C_n$, such that, provided each $|q_i|$ is sufficiently large,*

$$\left|\operatorname{slope}(K(q_1,\dots,q_n)) + \sum_{i=1}^{n} \ell_i^2 q_i\right| \le k, \quad \left|\sigma(K(q_1,\dots,q_n)) + \left(\frac{1}{2}\sum_{i=1}^{m} \ell_i^2 q_i + \frac{1}{2}\sum_{i=m+1}^{n} (\ell_i^2 - 1) q_i\right)\right| \le k.$$

One can use this to show that the factor $\operatorname{inj}(K)^{-3}$ cannot simply be dropped from Theorem 1.1 (see Conjecture 7.4 for what we expect for random knots):

Corollary 5.1 *There does **not** exist a constant c_2 such that*

$$|2\sigma(K) - \operatorname{slope}(K)| \le c_2 \operatorname{vol}(K)$$

for every hyperbolic knot K.

Proof Pick $n = 1$ and $\ell_1 = 3$. Then $\operatorname{slope}(K(q_1)) \sim -9q_1$, whereas $2\sigma(K(q_1)) \sim -8q_1$. On the other hand, $\operatorname{vol}(K(q_1))$ is bounded. □

Proof of Theorem 1.8 The knot $K(q_1,\dots,q_n)$ is obtained by performing $-1/q_i$ surgery on C_i for each $i \in \{1,\dots,n\}$. Let L denote the link $K \cup C_1 \cup \cdots \cup C_n$. By Thurston's hyperbolic Dehn surgery theorem, as all the $|q_i|$ tend to infinity, the hyperbolic structures on $S^3 \setminus K(q_1,\dots,q_n)$ tend in the geometric topology to the hyperbolic structure on $S^3 \setminus L$. In fact, more is true. Fix a horoball neighbourhood N of the cusps of $S^3 \setminus L$ that is small enough that the cusp torus T surrounding K lies in the complement of N. Then, if all the $|q_i|$ are sufficiently large, the inclusion $(S^3 \setminus L) \setminus N \to S^3 \setminus K(q_1,\dots,q_m)$ is a bi-Lipschitz homeomorphism onto its image, with bi-Lipschitz constants that tend to 1 as all the $|q_i|$ tend to infinity. (See [Benedetti and Petronio 1992], for instance.)

Let $\lambda(K)$ be the longitude and $\mu(K)$ the meridian of K. These form a basis of the lattice $\Lambda(K)$, where the cusp torus of K in $S^3 \setminus L$ is $\mathbb{C}/\Lambda(K)$. Let γ be the image of $\lambda(K)$ and μ the image of $\mu(K)$ on the cusp torus $\mathbb{C}/\Lambda(K(q_1,\dots,q_n))$ of $K(q_1,\dots,q_n)$. The curves γ and μ form a basis of the lattice $\Lambda(K(q_1,\dots,q_n))$. So we may assume that γ and μ are approximately constant complex numbers when the $|q_i|$ are large. However, we have *not* normalised the lattice so that γ is real. We know that there is some $N \in \mathbb{R}_+$ such that

$$N\mu^\perp = \gamma - s'\mu$$

for some $s' \in \mathbb{R}$. Here N, $\mu^\perp$, γ, s' and μ all depend on $q_1,\dots,q_n$. But N and s' tend to fixed real numbers as the $|q_i|$ go to infinity.

The key observation is that γ is *not* necessarily the longitude λ for $K(q_1,\dots,q_n)$. In fact, the linking number between γ and $K(q_1,\dots,q_n)$ is $\sum_i \ell_i^2 q_i$; see Figure 8. For suppose that the disc bounded by C_i

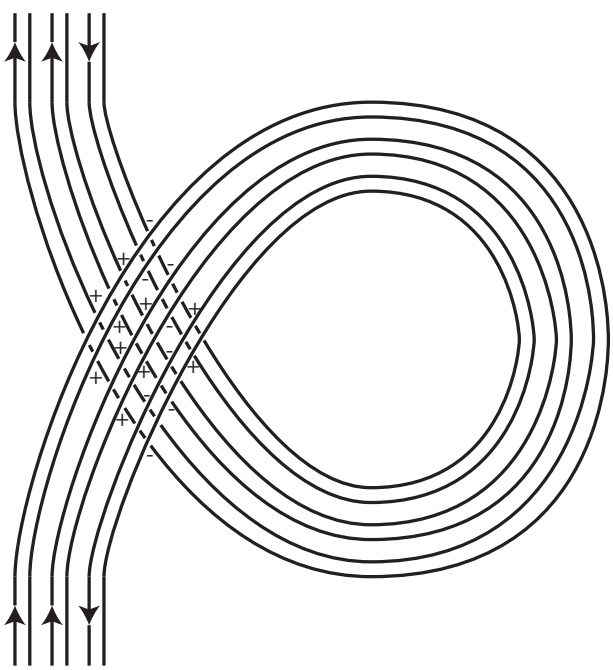

Figure 8: Each full twist about C_i changes the linking number between γ and $K(q_1,\dots,q_n)$ by ℓ_i^2.

intersects K in p_- points of negative sign and p_+ points of positive sign. So $\ell_i = p_+ - p_-$. Then, when we perform a full twist about C_i, we introduce $2(p_+ + p_-)^2$ new crossings between γ and $K(q_1, \dots, q_n)$. Of these, $2(p_+^2 + p_-^2)$ have positive sign and $4p_+p_-$ have negative sign. So the linking number between γ and $K(q_1, \dots, q_n)$ changes by

$$p_+^2 + p_-^2 - 2p_+p_- = \ell_i^2.$$

It follows that

$$\gamma = \lambda + \Big(\sum_{i=1}^{n} \ell_i^2 q_i\Big)\mu,$$

and hence

$$N\mu^\perp = \lambda - \Big(s' - \sum_{i=1}^{n} \ell_i q_i^2\Big)\mu.$$

We conclude that $\text{slope}(K(q_1, \dots, q_n)) = s' - \sum_{i=1}^n \ell_i q_i^2$. On the other hand, there is a constant k such that $|s'| \le k$ if $|q_1|, \dots, |q_n|$ are sufficiently large, which implies the first inequality of the theorem.

Recall that $\ell_{m+1}, \dots, \ell_n$ are odd. Suppose that $q_{m+1}, \dots, q_r$ are even and that $q_{r+1}, \dots, q_n$ are odd. Let $\mu_{m+1}, \dots, \mu_r$ be meridians for $C_{m+1}, \dots, C_r$, respectively. Let F be a spanning surface for

$$K \cup \mu_{m+1} \cup \cdots \cup \mu_r \cup C_{r+1} \cup \cdots \cup C_n.$$

Since this link has even linking number with each component of $C_1 \cup \cdots \cup C_r$, we may choose this spanning surface to be disjoint from these components. We can view this surface as properly embedded in the exterior of $K \cup C_1 \cup \cdots \cup C_n$. It is disjoint from $\partial N(C_1) \cup \cdots \cup \partial N(C_m)$. We have $F \cap \partial N(C_i) = \mu_i$ for $i \in \{m+1, \dots, r\}$. For $i \in \{r+1, \dots, n\}$, the curve $F \cap \partial N(C_i)$ has slope equal to a longitude plus an odd number of meridians. By choosing the surface appropriately, we can ensure that this odd number is 1.

Now perform surgery along $C_1, \dots, C_n$. The surface becomes a surface in the exterior of the new link. On $\partial N(C_i)$ for $i \in \{m+1, \dots, r\}$, it now has slope equal to a meridian plus q_i longitudes. On $\partial N(C_i)$ for $i \in \{r+1, \dots, n\}$, it is a meridian plus $q_i + 1$ longitudes. Since we are assuming that $|q_i|$ is sufficiently large, we can suppose that $q_i \neq 0, -1$ and hence that this slope is not meridional. Within each solid torus $N(C_{m+1}), \dots, N(C_n)$, we can now insert a surface, as shown in Figure 9. Let F' denote the resulting spanning surface of $K(q_1, \dots, q_n)$.

Also shown in Figure 9 is a collection of generators for $H_1(F' \cap N(C_i))$ for $i \ge m+1$. Note that $H_1(F' \cap N(C_i))$ for $i \ge m+1$ form direct summands of $H_1(F')$. So we can extend this set of generators to a basis of $H_1(F')$, by adding further elements of $H_1(F)$. The associated Goeritz form G_F is diagonal when restricted to the rows and columns corresponding to $H_1(F' \cap N(C_1 \cup \cdots \cup C_n))$. Each C_i gives rise to $\frac{1}{2}|q_i|$ diagonal entries when $m+1 \le i \le r$ and $\frac{1}{2}|q_i + 1|$ entries when $r+1 \le i \le n$. These entries are $+1$ when q_i is positive and -1 when q_i is negative. Hence, the signature of this matrix differs from

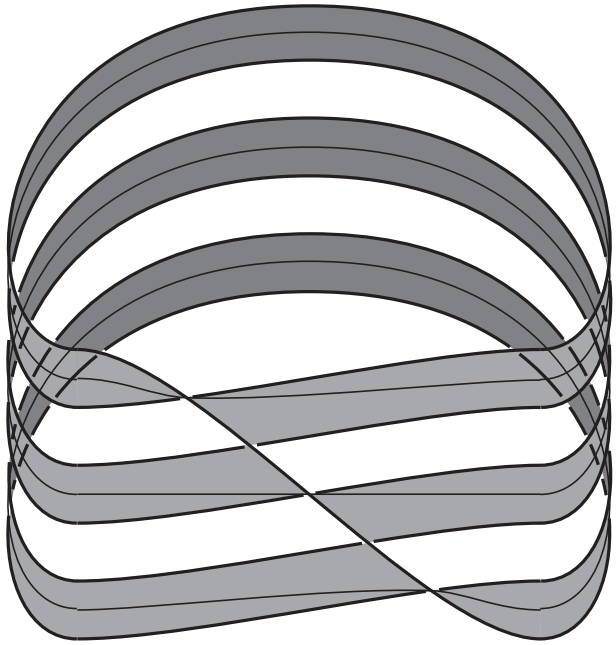

Figure 9: The part of the spanning surface in $N(C_i)$ for $i \geq m+1$. Here $q_i = 5$ or 6.

$\sum_{i=m+1}^{n} \frac{1}{2} q_i$ by at most $\frac{1}{2}(n-r)$. So, applying Lemma 4.5,

$$\left| \sigma(G_{F'}) - \sum_{i=m+1}^{n} \tfrac{1}{2} q_i \right|$$

is bounded.

Theorem 4.1, due to Gordon and Litherland, states that

$$\sigma(K(q_1, \dots, q_n)) = \sigma(G_{F'}) + \tfrac{1}{2} e(F').$$

Here

$$e(F') = -\mathrm{lk}(K(q_1, \dots, q_n), \partial F') = -\mathrm{lk}(K, \partial F') - \sum_{i=1}^{n} \ell_i^2 q_i.$$

The second inequality of the theorem follows immediately. □

6 Proof of Theorem 1.3

In this section, we prove Theorem 1.3 from the introduction:

Theorem 1.3 *Let ε_3 be the Margulis constant and let $\varepsilon \in (0, \varepsilon_3)$. Then there is a constant c_4 (depending on ε) such that, for any hyperbolic knot K, the quantities $\sigma(K)$ and*

$$\tfrac{1}{2} \operatorname{slope}(K) - \sum_{\gamma \in \mathrm{OddGeo}(\varepsilon/2)} \kappa(\mathrm{tw}_p(\gamma), \mathrm{tw}_q(\gamma))$$

differ by at most $c_4 \operatorname{vol}(K)$.

Note that, if we set $\varepsilon = \frac{1}{2}\varepsilon_3$, then c_4 becomes a universal constant. However, given the present uncertainty about the precise value of ε_3, we do not specify ε definitively.

Definition 6.1 Let γ be an embedded closed geodesic in the hyperbolic 3–manifold M, and let $N(\gamma)$ be a regular neighbourhood of γ consisting of points at most a certain distance r from γ. Let $\widetilde{\gamma}$ be a component of the inverse image of γ in $\mathbb{H}^3$, which we can take to be $\{(0,0,z) : z > 0\}$ in the upper half-space model. Let $N(\widetilde{\gamma})$ be the component of the inverse image of $N(\gamma)$ containing $\widetilde{\gamma}$. We let λ be the slope on $\partial N(\gamma)$ that has winding number one around $N(\gamma)$ and that lifts to a path in $N(\widetilde{\gamma})$ starting on the half-plane $\{(x,y,z) : y = 0,\ x \geq 0\}$ and with interior that is disjoint from the half-plane $\{(x,y,z) : y = 0,\ x \leq 0\}$. In the event that this path ends precisely on the half-plane $\{(x,y,z) : y = 0,\ x \leq 0\}$, λ is chosen so that it avoids $\{(x,y,z) : y \leq 0,\ x = 0\}$. Then λ is called the *canonical longitude* of γ. Note that it does not necessarily have zero linking number with γ.

There is the following alternative interpretation of the canonical longitude in terms of the complex length of γ. We give $T = \partial N(\gamma)$ its inherited Riemannian metric. This is homogeneous, since any two points of T differ by an isometry of T. The metric on T therefore has constant curvature, which must be zero by the Gauss–Bonnet theorem. It is therefore Euclidean. We can represent it as the quotient of the Euclidean plane $\mathbb{E}^2$ by a lattice $\mathcal{L}$. Each slope on T corresponds to a lattice point. We can assume that the lattice point corresponding to the meridian is a purely imaginary number μ. As the circumference of a radius r circle in the hyperbolic plane is $2\pi\sinh(r)$, we have

$$\mu = 2\pi\sinh(r)i,$$

where r is the radius of the tube around γ. Let ν be a geodesic in T that is perpendicular to a meridian and that starts and ends on the meridian (but not necessarily at the same point). Then

$$\ell(\nu) = \cosh(r)\operatorname{Re}(\operatorname{cl}(\gamma)),$$

where $\operatorname{cl}(\gamma)$ is the complex length of the geodesic γ and $\operatorname{Re}(\operatorname{cl}(\gamma)) = \ell(\gamma)$; see [Futer et al. 2019, equation (2.2)]. Then the canonical longitude of T is

$$\lambda = \cosh(r)\operatorname{Re}(\operatorname{cl}(\gamma)) + \sinh(r)\operatorname{Im}(\operatorname{cl}(\gamma))i.$$

The significance of the twisting parameter arises from the following lemma:

Lemma 6.2 *Let M be a hyperbolic 3–manifold and $\varepsilon \in (0, \varepsilon_3)$. Let γ be a geodesic in M with $\ell(\gamma) < \frac{1}{2}\varepsilon$. Let T be the toral boundary component of $M_{(0,3\varepsilon/4]}$ that encloses γ, let $\mu \subset T$ be a meridian of γ, and let λ be the canonical longitude. If $(p,q) = (\operatorname{tw}_p(\gamma), \operatorname{tw}_q(\gamma))$, then*

$$\ell(p\lambda + q\mu) \leq c_5 \operatorname{Area}(T)$$

for some constant c_5 depending only on ε.

Proof By the Margulis lemma, the component V of $M_{(0,3\varepsilon/4]}$ containing T is a solid torus, with γ as a core curve. We claim that the tube radius r of V satisfies $r > \frac{1}{8}\varepsilon$. Indeed, note that γ has length $\ell(\gamma) < \frac{1}{2}\varepsilon$, whereas, at each point $y \in T$, the open ball $B\big(y, \frac{3}{8}\varepsilon\big)$ is embedded. If $r < \frac{3}{8}\varepsilon$ and $x \in \gamma$ satisfies $d(x, y) = r$, then $B\big(x, \frac{3}{8}\varepsilon - r\big) \subset B\big(y, \frac{3}{8}\varepsilon\big)$ is an embedded ball about x. So

$$\tfrac{3}{8}\varepsilon - r < \tfrac{1}{2}\ell(\gamma) < \tfrac{1}{4}\varepsilon,$$

and hence $r > \frac{1}{8}\varepsilon$, as claimed.

Suppose that γ_0 is a shortest geodesic on T, and let $L := \ell(\gamma_0)$. We claim that

$$L \in [k_0, k_0']$$

for constants $k_0, k_0' \in \mathbb{R}_+$ depending only on ε. Since $T \subset \partial M_{(0,3\varepsilon/4]}$, every point $p \in T$ has two lifts to $\mathbb{H}^3$ that are exactly $\frac{3}{4}\varepsilon$ apart, and no two lifts of p are less than $\frac{3}{4}\varepsilon$ apart. The meridian of T has length

$$\ell(\mu) = 2\pi \sinh(r) > 2\pi \sinh\big(\tfrac{1}{8}\varepsilon\big).$$

If s is a slope different from the meridian, then $[s] = m[\gamma] \in \pi_1(M)$ for $m \neq 0$. As $[\gamma]$ has infinite order in $\pi_1(M)$, the lift $\tilde{s}$ of s to $\mathbb{H}^3$ satisfies $\tilde{s}(0) \neq \tilde{s}(1)$. Then

$$\ell(s) \geq d_{\mathbb{H}^3}(\tilde{s}(0), \tilde{s}(1)) \geq \tfrac{3}{4}\varepsilon,$$

so we can set $k_0 := \min\big(\frac{3}{4}\varepsilon, 2\pi \sinh\big(\frac{1}{8}\varepsilon\big)\big)$.

We now give an upper bound on L. Let s be a slope on T whose lift $\tilde{s}$ to $\mathbb{H}^3$ satisfies $d_{\mathbb{H}^3}(\tilde{s}(0), \tilde{s}(1)) = \frac{3}{4}\varepsilon$. This is again possible since $T \subset \partial M_{(0,3\varepsilon/4]}$. If $r \leq 2\varepsilon$, then $L \leq |\mu| \leq 2\pi \sinh(2\varepsilon)$. Now suppose that $r > 2\varepsilon$. Let $N(\gamma) \subset V$ be a regular neighbourhood of γ of radius $r - \varepsilon$. Let $\widetilde{\beta}$ be a geodesic in $\mathbb{H}^3$ connecting $\tilde{s}(0)$ and $\tilde{s}(1)$, and let β be its projection to M. Then β is a geodesic homotopic to s of length $\frac{3}{4}\varepsilon$, which hence lies in $V \setminus N(\gamma)$. The nearest point projection $\varphi\colon V \setminus N(\gamma) \to T$ satisfies $\ell(\varphi(\beta)) \leq l_0 \ell(\beta) = l_0\big(\frac{3}{4}\varepsilon\big)$ for a constant l_0 depending only on ε. Hence,

$$L \leq k_0' := \max\big(2\pi \sinh(2\varepsilon), \tfrac{3}{4} l_0 \varepsilon\big),$$

as claimed.

A consequence of $L \geq k_0$ is that $\mathrm{Area}(T) \geq a_0$ for a constant a_0 depending only on ε. Indeed, a disc D of radius $\frac{1}{2}L$ on T about an arbitrary point of T is embedded, so

$$\mathrm{Area}(T) \geq \mathrm{Area}(D) = \big(\tfrac{1}{2}L\big)^2 \pi \geq \big(\tfrac{1}{2}k_0\big)^2 \pi =: a_0.$$

We claim the length of the shortest curve in any nontrivial class in $H_1(T; \mathbb{Z}_2)$ is at most $k_1\,\mathrm{Area}(T)$ for a constant k_1 depending on ε. Indeed, let $\gamma_0^{\perp}\colon I \to T$ be a geodesic arc starting and ending on the shortest geodesic γ_0 and orthogonal to it. Then $\ell(\gamma_0^{\perp}) = \mathrm{Area}(T)/L$. The points $\gamma_0^{\perp}(0)$ and $\gamma_0^{\perp}(1)$ divide γ_0 into two arcs, one of which has length at most $\frac{1}{2}L$. Let γ_1 be a geodesic representative of the closed curve that runs along $\gamma_0^{\perp}$ and then along the shorter of the two arcs in γ_0. We obtain that

$$\ell(\gamma_1) \leq \tfrac{1}{2}L + \frac{\mathrm{Area}(T)}{L}.$$

The curves γ_0 and γ_1 give a basis for $H_1(T;\mathbb{Z}_2)$. Hence, the shortest representative of every nontrivial class in $H_1(T;\mathbb{Z}_2)$ is at most $L+\left(\frac{1}{2}L+\operatorname{Area}(T)/L\right)$. As $L\in[k_0,k_0']$ and $\operatorname{Area}(T)\geq a_0$, we have

$$L+\left(\tfrac{1}{2}L+\frac{\operatorname{Area}(T)}{L}\right)\leq k_1\operatorname{Area}(T)$$

for $k_1:=\frac{3}{2}k_0'/a_0+1/k_0$. Indeed,

$$\tfrac{3}{2}L\leq\tfrac{3}{2}k_0'=\left(k_1-\frac{1}{k_0}\right)a_0\leq\left(k_1-\frac{1}{L}\right)\operatorname{Area}(T).$$

So there is some slope (a,b) on T with a even and b odd such that

$$\ell(a\lambda+b\mu)\leq k_1\operatorname{Area}(T)$$

for some constant k_1 depending on ε.

Let T' be the torus obtained from T by scaling by $\tanh(r)$ in the ν direction. As $r>\frac{1}{8}\varepsilon$, we have $\tanh(r)\in\left(\tanh\left(\frac{1}{8}\varepsilon\right),1\right)$. Since $\tanh(r)<1$, the shortest slope (p,q) on T' with p even and q odd has length at most $k_1\operatorname{Area}(T)$. The lattice that specifies T' is generated by

$$\lambda':=\tanh(r)\cosh(r)\operatorname{Re}(\operatorname{cl}(\gamma))+\sinh(r)\operatorname{Im}(\operatorname{cl}(\gamma))i=\sinh(r)\operatorname{cl}(\gamma)$$

and $\mu=2\pi\sinh(r)i$. So

$$\ell(p\lambda'+q\mu)=|\operatorname{cl}(\gamma)p+2\pi iq||\sinh(r)|.$$

Hence, by Definition 1.2, the slope $p\lambda'+q\mu$ on T' is the shortest among slopes for which p is even and q is odd. Therefore, its length on T' is at most $k_1\operatorname{Area}(T)$. So

$$\ell(p\lambda+q\mu)\leq\frac{k_1}{\tanh(r)}\operatorname{Area}(T)<\frac{k_1}{\tanh\left(\frac{1}{8}\varepsilon\right)}\operatorname{Area}(T).$$

So we can set $c_5:=k_1/\tanh\left(\frac{1}{8}\varepsilon\right)$, which concludes the proof of the lemma. □

Proof of Theorem 1.3 We claim that we can build a triangulation $\mathcal{T}$ of $M_{[3\varepsilon/4,\infty)}$ with the following properties:

(1) The number of tetrahedra of $\mathcal{T}$ is at most $c\operatorname{vol}(K)$, where c depends on ε.

(2) If n is a closest even integer to $\operatorname{slope}(K)$, then some Euclidean geodesic with slope $\lambda-n\mu$ on $\partial N(K)$ is a normal curve in $\partial M_{[3\varepsilon/4,\infty)}$ that intersects each edge of $\mathcal{T}$ at most once.

(3) On the component T of $\partial M_{[3\varepsilon/4,\infty)}$ corresponding to $\partial N(K)$, the edges of $\mathcal{T}$ are Euclidean geodesics with length at most $\frac{1}{15}\varepsilon$.

We follow the construction in the proof of Proposition 3.1, but with different constants. We pick a maximal collection of points in $\partial M_{[3\varepsilon/4,\infty)}$ that are at least $\frac{1}{30}\varepsilon$ apart. We then add points to this collection in the interior of $M_{[3\varepsilon/4,\infty)}$ that have distance at least $\frac{1}{15}\varepsilon$ from each other and from the earlier points. We stop when it is not possible to add any further points, and denote the resulting collection by P. We then form the associated Voronoi diagram, subdivide the 2–cells of this cell structure into triangles without adding

any new vertices, and then triangulate each 3–cell by coning from the relevant point of P. Let $\mathcal{T}$ be the resulting triangulation of $M_{[3\varepsilon/4,\infty)}$.

Exactly the same argument as in the proof of Proposition 3.1 gives that the number of tetrahedra of $\mathcal{T}$ is at most $c\,\mathrm{vol}(K)$, where c depends on ε. The length of each edge in $\partial M_{[3\varepsilon/4,\infty)}$ is now at most $\frac{1}{15}\varepsilon$, because we took points that were at least $\frac{1}{30}\varepsilon$ apart, rather than at least $\frac{1}{8}\varepsilon$ apart. Thus, all that needs to be proved are that the edges of $\mathcal{T}$ in T are Euclidean geodesics and that there is a Euclidean geodesic with slope $\lambda - n\mu$ on $\partial N(K)$ which is a normal curve in $\partial M_{[3\varepsilon/4,\infty)}$ that intersects each edge of $\mathcal{T}$ at most once.

We start by showing that the edges of $\mathcal{T}$ in T are Euclidean geodesics. Following the proof of Proposition 3.1, we need to show that, for each point x on T, its closest points in P all lie in T and have distance at most $\frac{1}{30}\varepsilon$ from x. We also need to show that the shortest geodesic joining x to any of these points remains within the cusp. The first of these statements holds by our choice of P.

Note that T lies within $M_{(0,\varepsilon]}$. By definition of the Margulis constant, $M_{(0,\varepsilon]}$ consists of a cusp and some regular neighbourhoods of geodesics with length at most ε. The Euclidean metrics on T and the cusp component of $\partial M_{(0,\varepsilon]}$ differ by a Euclidean scale factor of $\frac{4}{3}$, and hence are hyperbolic distance $\ln\left(\frac{4}{3}\right) > 0.287$ from each other. On the other hand, the 3–dimensional Margulis constant satisfies $\varepsilon_3 < 0.775$. (See the discussion in [Futer et al. 2022, Section 1.1].) Hence, $\frac{1}{30}\varepsilon < \ln\left(\frac{4}{3}\right)$. We deduce that, for each point x in T, any shortest geodesic to a closest point in P must lie in the cusp. This implies that the restriction to T of the Voronoi diagram for P in M is equal to the Voronoi diagram for $P \cap T$ in T with its Euclidean metric. In particular, the edges of $\mathcal{T}$ in T are Euclidean geodesics, as claimed in (3).

Let $N_{\max}$ be a maximal cusp neighbourhood around K. Then $N_{\max}$ contains T. This torus T is a scaled copy of $\partial N_{\max}$. It is scaled so that, for each point on T, two lifts of this point in $\mathbb{H}^3$ are exactly $\frac{3}{4}\varepsilon$ apart and no two lifts of this point are any closer than this. Say that d is the hyperbolic distance between T and $\partial N_{\max}$. Then the scale factor taking $\partial N_{\max}$ to T is e^{-d}. Now the meridian slope on $\partial N_{\max}$ has length at most 6. Hence, the meridian slope on T has length at most $6e^{-d}$. So any point on T has two lifts to $\mathbb{H}^3$ that are less than $6e^{-d}$ apart, and therefore $\frac{3}{4}\varepsilon \leq 6e^{-d}$. As in the proof of Proposition 3.1, let h be the length in $\partial N_{\max}$ of a Euclidean geodesic that starts and ends on a geodesic with slope $\lambda - n\mu$ and that is orthogonal to this geodesic. It was shown there that $h \geq 0.55$. Hence, the length of the corresponding geodesic on T is at least $0.55e^{-d} \geq \frac{0.55}{6}\left(\frac{3}{4}\varepsilon\right)$. On the other hand, the length of each edge of $\mathcal{T}$ on T is at most $\frac{1}{15}\varepsilon$, and $\frac{1}{15}\varepsilon < \frac{0.55}{6}\left(\frac{3}{4}\varepsilon\right)$. Hence, each such edge can intersect any geodesic with slope $\lambda - n\mu$ at most once. This establishes the claimed properties of $\mathcal{T}$.

Let $T_1, \dots, T_m$ be the components of $\partial M_{[3\varepsilon/4,\infty)}$, where T_i encircles a geodesic $\gamma_i \in \mathrm{OddGeo}\left(\frac{1}{2}\varepsilon\right)$. Let $\mathrm{tw}(\gamma_i) = p\lambda_i + q\mu_i$, where λ_i is the canonical longitude on T_i and μ_i is the meridian, and let C_i be a curve on T_i with this slope. Then

$$\ell(C_i) \leq c_5\,\mathrm{Area}(T_i)$$

by Lemma 6.2. Let

$$C := \bigcup_{i=1}^{m} C_i.$$

Realise each C_i as a Euclidean geodesic in T_i missing the vertices of T_i, and hence as a normal curve in T_i. Since $\ell(C_i) \leq c_5 \operatorname{Area}(T_i)$ and by property (3) of the triangulation $\mathcal{T}$, the normal representative of C_i intersects each edge of $\mathcal{T}$ at most $c_5' \operatorname{Area}(T_i)$ times for a constant c_5' depending only on ε.

We claim that there is a connected normal curve C_i' in T_i for $i \in \{1, \dots, m\}$ with the following properties:

(1) C_i' and C_i are equal in $H_1(T_i; \mathbb{Z}_2)$.

(2) C_i' intersects each edge of $\mathcal{T}$ at most once.

This is constructed as follows. For each edge of $\mathcal{T}$ that intersects C_i an odd number of times, replace this intersection by a single point of intersection. These will be the points of intersection between C_i' and the 1–skeleton of $\mathcal{T}$. Since $|C_i \cap \partial t|$ is even for each triangle t of $\mathcal{T}$, we have $|C_i' \cap \partial t| \in \{0, 2\}$. If $|C_i' \cap \partial t| = 2$, join the two points of $C_i' \cap \partial t$ by a normal arc of C_i'. The result is a collection of simple closed curves in T_i that are mod 2 homologous to C_i. If any of these curves are inessential in T_i, remove them. The resulting curves are essential in T_i. Since they are nontrivial in mod 2 homology, they consist of an odd number of parallel copies of a curve. If this odd number is greater than one, remove all but one of these curves. The result is C_i', and we write

$$C' := \bigcup_{i=1}^{m} C_i'.$$

Let C'' be the union of C' and a normal curve C_K of slope $(1, -n)$ on $\partial N(K)$, where n is a closest even integer to slope(K). We claim that C'' bounds an unoriented surface in $M_{[3\varepsilon/4, \infty)}$. As n is even, there is a compact surface properly embedded in the exterior of K with boundary slope $(1, -n)$. It intersects each geodesic with length at most $\frac{1}{2}\varepsilon$ in a collection of meridians. For a geodesic with odd linking number with K, the number of these meridians is odd. For the others, it is even. As C_i' is homologous to the meridian of T_i over $\mathbb{Z}_2$, we may modify the surface so that its boundary is precisely C''. This proves the claim.

As C'' intersects each edge of $\mathcal{T}$ at most once, we can find a surface F'' in $M_{[\varepsilon/2, \infty)}$ that it bounds such that

$$-\chi(F'') \leq c_6 \operatorname{vol}(K)$$

for some constant c_6, just like in the proof of Theorem 1.7. Now C_i and C_i' are equal in $H_1(T_i; \mathbb{Z}_2)$. Hence, we may insert a compact connected surface F_i into a regular neighbourhood $N(T_i)$ of T_i with $\partial F_i = C_i \cup C_i'$. Since C_i and C_i' intersect each edge of $\mathcal{T}$ at most $c_5' \operatorname{Area}(T_i)$ times, this surface may be chosen so that

$$-\chi(F_i) \leq c_5'' \operatorname{Area}(T_i)$$

for a constant c_5'' depending only on ε. Hence, the surface

$$F := F'' \cup \bigcup_{i=1}^{m} F_i \subset M_{[3\varepsilon/4,\infty)}$$

satisfies $\partial F = C_K \cup C$, and

$$-\chi(F) \leq c_6 \operatorname{vol}(K) + \sum_{i=1}^{m} c_5'' \operatorname{Area}(T_i) \leq c_7 \operatorname{vol}(K) \tag{6.3}$$

for a constant c_7 that depends only on ε. Here the last inequality follows from the observation that $\operatorname{Area}(T_i) \leq c_8 \operatorname{vol}(N(T_i))$ for some constant c_8, where

$$N(T_i) := \{x \in V_i : d(x, T_i) \leq \tfrac{1}{2} r_i\},$$

and V_i is the solid toral component of $M_{(0,3\varepsilon/4]}$ of tube radius r_i that encloses the geodesic $\gamma_i \in \operatorname{OddGeo}(\frac{1}{2}\varepsilon)$.

In each V_i, we construct the surface provided by Lemma 4.6 with boundary $C_i = C \cap V_i$. We attach these surfaces to F to form a surface F_+. We now specify a basis for $H_1(F_+)$. We start by picking a basis for $H_1(V_1 \cap F_+)$. We arrange that all but one of these basis elements have zero winding number around V_1. We then continue to V_2, and so on. We then extend this to a basis for $H_1(F_+)$ by adding some oriented curves in F. We order this basis as follows into $n+1$ blocks. In the first block, we place all the basis elements of $H_1(V_1 \cap F_+)$ that have zero winding number around V_1. In the second block, we do the same for V_2, and so on. In the final block, we place all the remaining basis elements. We saw in the proof of Lemma 6.2 that there is a constant a_0 depending only on ε such that $\operatorname{Area}(T_i) \geq a_0$. As $\sum_{i=1}^{m} \operatorname{Area}(T_i) \leq c_8 \operatorname{vol}(K)$, we have

$$\left|\operatorname{OddGeo}\left(\tfrac{1}{2}\varepsilon\right)\right| \leq \frac{c_8 \operatorname{vol}(K)}{a_0}.$$

This, together with (6.3), implies that the number of elements in this final block is bounded above by a linear function of $\operatorname{vol}(K)$.

Let G be the submatrix of the Goeritz form G_{F_+} consisting of the first n blocks. By Lemma 4.5, $\sigma(G)$ and $\sigma(G_F)$ differ by at most the number of elements in the final block. Note that G is block diagonal. For the block corresponding to V_i, the signature differs from $\sigma(G_{F_+ \cap V_i})$ by at most one by Lemma 4.5. On the other hand,

$$|\sigma(G_{F_+ \cap V_i}) + \kappa(\operatorname{tw}_p(\gamma_i), \operatorname{tw}_q(\gamma_i))| \leq 2$$

by Lemma 4.6. Hence,

$$\left|\sigma(G_{F_+}) + \sum_{\gamma \in \operatorname{OddGeo}(\varepsilon/2)} \kappa(\operatorname{tw}_p(\gamma_i), \operatorname{tw}_q(\gamma_i))\right| \leq c_9 \operatorname{vol}(K)$$

for some constant c_9. By Gordon and Litherland's theorem (Theorem 4.1),

$$\sigma(K) = \sigma(G_{F_+}) + \tfrac{1}{2} e(F_+) = \sigma(G_{F_+}) + \tfrac{1}{2} n.$$

The result follows as n is a closest even integer to $\operatorname{slope}(K)$. □

7 Experimental data and some conjectures about random knots

We set out to find links between hyperbolic and 4–dimensional knot invariants. Initial scatter plots compared some 4–dimensional invariants (the signature and Heegaard Floer invariants τ, ν and ε), the crossing number, and several hyperbolic invariants (volume, meridional and longitudinal translations, and the Chern–Simons invariant). As σ is strongly correlated to τ, ν and ε, we decided to only focus on σ, which is more classical and easier to compute.

The strongest and most surprising correlation was between the signature and the real part of the meridional translation; see Figure 2. There were some more predictable relationships among the hyperbolic invariants.

Figure 10 shows the distribution of

$$c_1(K) := \frac{|2\sigma(K) - \operatorname{slope}(K)|\operatorname{inj}(K)^3}{\operatorname{vol}(K)},$$

which indicates that the constant c_1 appearing in Theorem 1.1 is typically quite small. The largest value of this quantity we managed to obtain is less than 0.234, and we conjecture it is always at most 0.3. The left of Figure 11 shows the maximum and the right the mean of $c_1(K)$ by crossing number for the Regina census of knots of at most 16 crossings. See Figure 12 for a scatter plot of injectivity radius versus volume for random hyperbolic knots of 10–80 crossings. This suggests that the injectivity radius is typically not too small as the volume increases.

We will say that a property P holds asymptotically almost surely, or a.a.s. in short, if the probability that P holds for knots of n crossings tends to 1 as $n \to \infty$.

It is known that there is a constant A such that $\operatorname{vol}(K) \leq Ac(K)$, where $c(K)$ is the crossing number of K. From scatter plots, one might conjecture that there is a constant a such that $ac(K) \leq \operatorname{vol}(K)$ a.a.s. Such

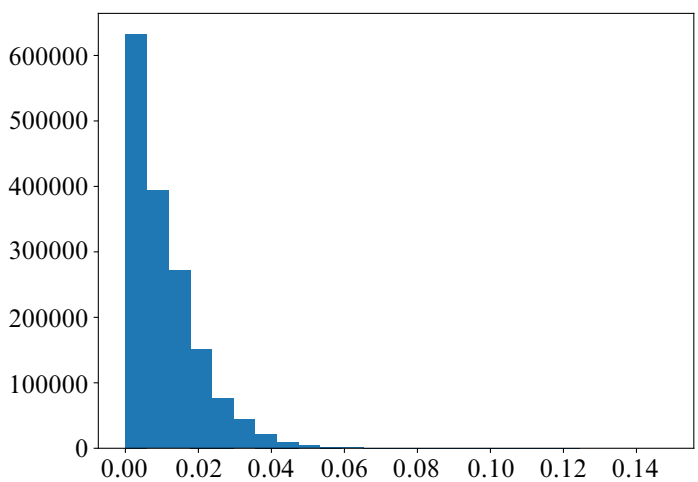

Figure 10: The distribution of $c_1(K) := |2\sigma(K) - \operatorname{slope}(K)|\operatorname{inj}(K)^3/\operatorname{vol}(K)$ for knots up to 16 crossings in the Regina census.

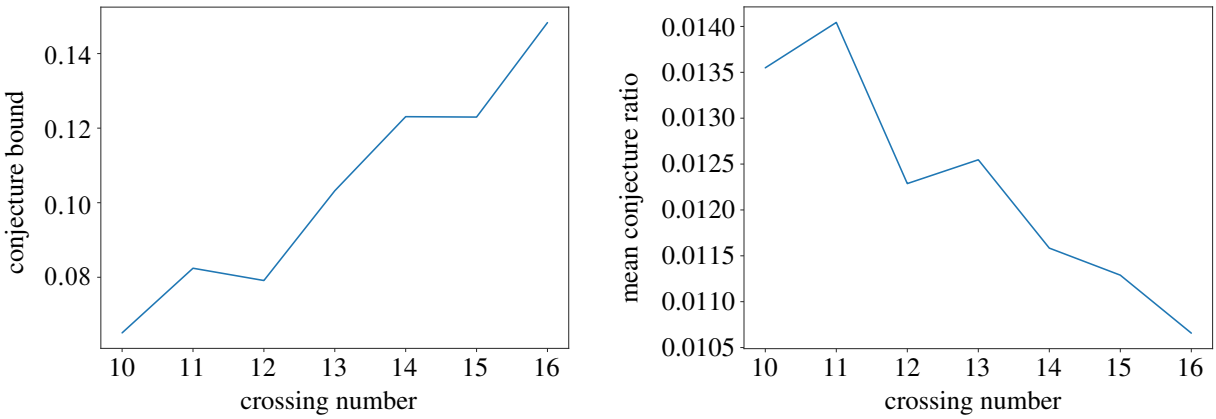

Figure 11: The maximum (left) and the mean (right) of $c_1(K)$ as functions of the crossing number for knots up to 16 crossings in the Regina census.

an inequality cannot hold for all hyperbolic knots K. For example, consider twist knots. More generally, the highly twisted knots considered in Section 5 have bounded volume but unbounded crossing number.

We now consider the behaviour of the signature $\sigma(K)$ for random knots K. By Theorem 4.1, $\sigma(K)$ can be computed from the black surface of a checkerboard colouring of a diagram of K. Hence, it is the signature of a $c(K) \times c(K)$ matrix. If the signs of the eigenvalues of this matrix were independently distributed, then the expected value of $|\sigma(K)|$ would be $C\sqrt{c(K)}$ for some constant C. From computational evidence, it appears the constant is about 2. Based on this heuristic, we introduce the following definition:

Definition 7.1 The *normalised signature* of a hyperbolic knot K is

$$\widehat{\sigma}(K) := \frac{\sigma(K)}{\sqrt{\mathrm{vol}(K)}}.$$

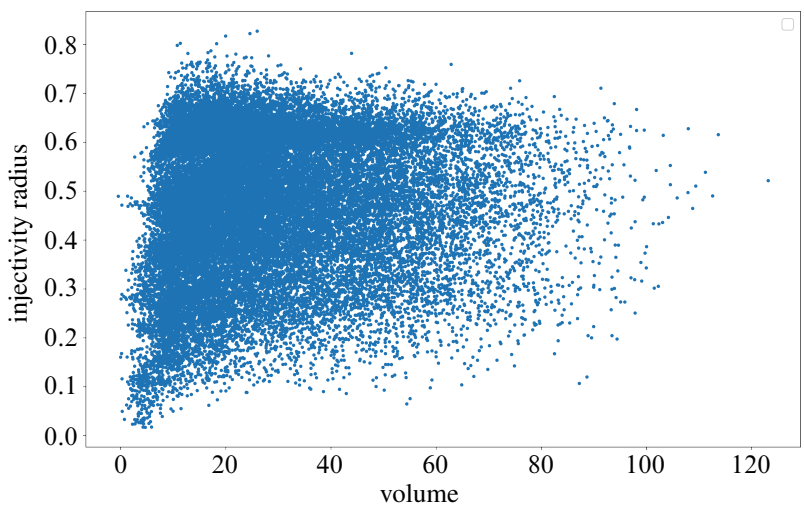

Figure 12: A scatter plot of injectivity radius versus volume for random knots of 10–80 crossings.

We use the volume instead of the crossing number as it is easier to compute using SnapPy and is more regular.

Based on Figure 2, we initially conjectured that, for any hyperbolic knot K in S^3 with $|\widehat{\sigma}(K)| > 1$, the signature $\sigma(K)$ and $\mathrm{Re}(\mu(K))$ have the same sign. However, this turns out not to be true.

Corollary 7.2 *There exists a hyperbolic knot K with $|\widehat{\sigma}(K)| > 1$, but with $\sigma(K)$ and $\mathrm{Re}(\mu(K))$ having opposite signs.*

Proof We start with a hyperbolic link $K \cup C_1 \cup C_2$ in S^3, where C_1 and C_2 bound disjoint embedded discs, and where $\ell_1 = \mathrm{lk}(K, C_1) = 2$ and $\ell_2 = \mathrm{lk}(K, C_2) = 3$. We then build the highly twisted knots $K(q_1, q_2)$ as in Theorem 1.8. Set $q_1 = 17q$ and $q_2 = -8q$, where q is a large positive integer. Then

$$\mathrm{slope}(K(q_1, q_2)) \sim -4 \cdot 17q + 9 \cdot 8q = 4q, \quad \text{whereas} \quad \sigma(K(q_1, q_2)) \sim -2 \cdot 17q + 4 \cdot 8q = -2q.$$

Hence, for q sufficiently large, $\sigma(K(q_1, q_2))$ and $\mathrm{slope}(K(q_1, q_2))$ have opposite signs, and hence $\sigma(K(q_1, q_2))$ and $\mathrm{Re}\big(\mu(K(q_1, q_2))\big)$ also have opposite signs by Lemma 2.4. Note that $\widehat{\sigma}(K(q_1, q_2)) > 1$ if q is sufficiently large, because $|\sigma(K(q_1, q_2))|$ tends to infinity whereas $\mathrm{vol}(K(q_1, q_2))$ is bounded. □

However, we do conjecture the following:

Conjecture 7.3 *If K is a hyperbolic knot in S^3 with $|\widehat{\sigma}(K)| > 1$, then $\sigma(K)$ and $\mathrm{Re}(\mu(K))$ have the same sign asymptotically almost surely.*

We also state the following conjecture, which proposes a more precise relationship between slope and signature:

Conjecture 7.4 *There are constants b and c such that, for any hyperbolic knot K in S^3, we have*

$$|2\sigma(K) - \mathrm{slope}(K)| \le b\sqrt{\mathrm{vol}(K)} + c \tag{7.5}$$

asymptotically almost surely.

By Corollary 5.1, this does not hold for all knots either. In fact, there are families of hyperbolic knots for which $|2\sigma(K) - \mathrm{slope}(K)|$ is not bounded by a linear function of the volume.

The proof of Theorem 1.1 provides some heuristic for Conjecture 7.4. Indeed, if we assume that the signs of the eigenvalues of the Goeritz matrix G_F are independent, then the signature on average is of order $\sqrt{c(K)}$. This justifies the factor $\sqrt{\mathrm{vol}(K)}$ in the upper bound.

If $b < 2$ (and the data supports this; see Figure 13), then Conjecture 7.4 implies Conjecture 7.3 for knots K with sufficiently large volume a.a.s. This is because (7.5) is equivalent to the inequality

$$\left|2\widehat{\sigma}(K) - \frac{\mathrm{slope}(K)}{\sqrt{\mathrm{vol}(K)}}\right| \le b + \frac{c}{\sqrt{\mathrm{vol}(K)}}.$$

If $b < 2$, then $b + c/\sqrt{\mathrm{vol}(K)} < 2$ for all knots with sufficiently large volume. So, if $|\widehat{\sigma}(K)| > 1$, then $\widehat{\sigma}(K)$ and $\mathrm{slope}(K)$ have the same sign.

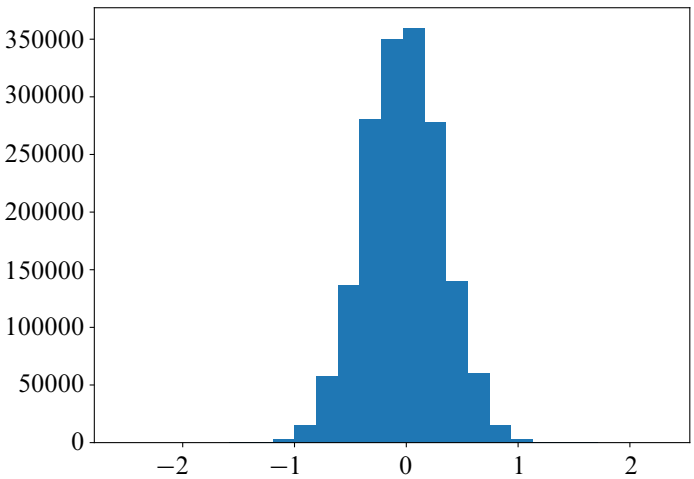

Figure 13: The distribution of the normalised residual $(2\sigma(K) - \text{slope}(K))/\sqrt{\text{vol}(K)}$ for knots up to 16 crossings in the Regina census.

References

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

[Benard et al. 2021] **L Benard**, **V Florens**, **A Rodau**, *A slope invariant and the A–polynomial of knots*, preprint (2021) arXiv 2103.14151

[Benedetti and Petronio 1992] **R Benedetti**, **C Petronio**, *Lectures on hyperbolic geometry*, Springer (1992) MR Zbl

[Breslin 2009] **W Breslin**, *Thick triangulations of hyperbolic n–manifolds*, Pacific J. Math. 241 (2009) 215–225 MR Zbl

[Burton et al. 1999–2021] **B A Burton**, **R Budney**, **W Pettersson**, et al., *Regina: software for low-dimensional topology* (1999–2021) Available at `https://regina-normal.github.io/`

[Cao and Meyerhoff 2001] **C Cao**, **G R Meyerhoff**, *The orientable cusped hyperbolic 3–manifolds of minimum volume*, Invent. Math. 146 (2001) 451–478 MR Zbl

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

[Culler et al. 2021] **M Culler**, **N M Dunfield**, **M Goerner**, **J R Weeks**, *SnapPy, a computer program for studying the geometry and topology of 3–manifolds* (2021) Available at `http://snappy.computop.org`

[Degtyarev et al. 2022] **A Degtyarev**, **V Florens**, **A G Lecuona**, *Slopes and signatures of links*, Fund. Math. 258 (2022) 65–114 MR Zbl

[Futer et al. 2019] **D Futer**, **J S Purcell**, **S Schleimer**, *Effective distance between nested Margulis tubes*, Trans. Amer. Math. Soc. 372 (2019) 4211–4237 MR Zbl

[Futer et al. 2022] **D Futer**, **J S Purcell**, **S Schleimer**, *Effective bilipschitz bounds on drilling and filling*, Geom. Topol. 26 (2022) 1077–1188 MR Zbl

[Gordon and Litherland 1978] **C M Gordon**, **R A Litherland**, *On the signature of a link*, Invent. Math. 47 (1978) 53–69 MR Zbl

[Gordon et al. 1981] **C M Gordon**, **R A Litherland**, **K Murasugi**, *Signatures of covering links*, Canad. J. Math. 33 (1981) 381–394 MR Zbl

[Hass et al. 1999] **J Hass**, **J H Rubinstein**, **S Wang**, *Boundary slopes of immersed surfaces in 3–manifolds*, J. Differential Geom. 52 (1999) 303–325 MR Zbl

[Kobayashi and Rieck 2011] **T Kobayashi**, **Y Rieck**, *A linear bound on the tetrahedral number of manifolds of bounded volume (after Jørgensen and Thurston)*, from "Topology and geometry in dimension three" (W Li, L Bartolini, J Johnson, F Luo, R Myers, J H Rubinstein, editors), Contemp. Math. 560, Amer. Math. Soc., Providence, RI (2011) 27–42 MR Zbl

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

[Lackenby and Meyerhoff 2013] **M Lackenby**, **R Meyerhoff**, *The maximal number of exceptional Dehn surgeries*, Invent. Math. 191 (2013) 341–382 MR Zbl

[Lackenby and Purcell 2016] **M Lackenby**, **J S Purcell**, *Cusp volumes of alternating knots*, Geom. Topol. 20 (2016) 2053–2078 MR Zbl

[Livingston and Moore 2021] **C Livingston**, **A H Moore**, *KnotInfo*: *table of knot invariants*, electronic reference (2021) Available at `https://knotinfo.math.indiana.edu/`

[Morgan 1984] **J W Morgan**, *On Thurston's uniformization theorem for three-dimensional manifolds*, from "The Smith conjecture" (J W Morgan, H Bass, editors), Pure Appl. Math. 112, Academic, Orlando, FL (1984) 37–125 MR Zbl

[Mostow 1968] **G D Mostow**, *Quasi-conformal mappings in n–space and the rigidity of hyperbolic space forms*, Inst. Hautes Études Sci. Publ. Math. 34 (1968) 53–104 MR Zbl

[Thurston 1979] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at `https://url.msp.org/gt3m`

[Voronoi 1908a] **G Voronoi**, *Nouvelles applications des paramètres continus à la théorie des formes quadratiques, deuxième mémoire: Recherches sur les parallélloèdres primitifs*, J. Reine Angew. Math. 134 (1908) 198–287 MR Zbl

[Voronoi 1908b] **G Voronoi**, *Nouvelles applications des paramètres continus à la théorie des formes quadratiques, premier mémoire: sur quelques propriétés des formes quadratiques positives parfaites*, J. Reine Angew. Math. 133 (1908) 97–102 MR Zbl

DeepMind
London, United Kingdom

Mathematical Institute, University of Oxford
Oxford, United Kingdom

Mathematical Institute, University of Oxford
Oxford, United Kingdom

DeepMind
London, United Kingdom

`adavies@google.com`, `juhasza@maths.ox.ac.uk`, `lackenby@maths.ox.ac.uk`, `nenadt@deepmind.com`

Proposed: David Gabai
Seconded: Cameron Gordon, Stavros Garoufalidis

Received: 20 April 2022
Revised: 26 August 2022

GEOMETRY & TOPOLOGY

msp.org/gt

See inside back cover or msp.org/gt for submission instructions.

The subscription price for 2024 is US $805/year for the electronic version, and $1135/year (+$70, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP. Geometry & Topology is indexed by Mathematical Reviews, Zentralblatt MATH, Current Mathematical Publications and the Science Citation Index.

Geometry & Topology (ISSN 1465-3060 printed, 1364-0380 electronic) is published 9 times per year and continuously online, by Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840. Periodical rate postage paid at Oakland, CA 94615-9651, and additional mailing offices. POSTMASTER: send address changes to Mathematical Sciences Publishers, c/o Department of Mathematics, University of California, 798 Evans Hall #3840, Berkeley, CA 94720-3840.

GT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

GEOMETRY & TOPOLOGY

Volume 28 Issue 5 (pages 1995–2482) 2024