Illuminating new and known relations between knot invariants

IOP Publishing Mach. Learn.: Sci. Technol. 5 (2024) 045061 https://doi.org/10.1088/2632-2153/ad95d9

MACHINE LEARNING
Science and Technology

PAPER

Illuminating new and known relations between knot invariants

Jessica Craven[1], **Mark Hughes**[2,3,*], **Vishnu Jejjala**[4] **and Arjun Kar**[5]

[1] Division of Physics, Mathematics, and Astronomy (PMA), California Institute of Technology, Pasadena, CA 91125, United States of America
[2] Department of Mathematics, Brigham Young University, 275 TMCB, Provo, UT 84602, United States of America
[3] Max Planck Institute for Mathematics, Vivatsgasse 7, 53111 Bonn, Germany
[4] Mandelstam Institute for Theoretical Physics, School of Physics, NITheCS, and CoE-MaSS, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg WITS 2050, South Africa
[5] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada
[*] Author to whom any correspondence should be addressed.

E-mail: hughes@mathematics.byu.edu, jcraven@caltech.edu, v.jejjala@wits.ac.za and arjunkar@phas.ubc.ca

Keywords: knots, knot invariants, machine learning, neural networks

OPEN ACCESS

RECEIVED
24 May 2024

REVISED
12 October 2024

ACCEPTED FOR PUBLICATION
21 November 2024

PUBLISHED
10 December 2024

Abstract

We automate the process of machine learning correlations between knot invariants. For nearly 200 000 distinct sets of input knot invariants together with an output invariant, we attempt to learn the output invariant by training a neural network on the input invariants. Correlation between invariants is measured by the accuracy of the neural network prediction, and bipartite or tripartite correlations are sequentially filtered from the input invariant sets so that experiments with larger input sets are checking for true multipartite correlation. We rediscover several known relationships between polynomial, homological, and hyperbolic knot invariants, while also finding novel correlations which are not explained by known results in knot theory. These unexplained correlations strengthen previous observations concerning links between Khovanov and knot Floer homology. Our results also point to a new connection between quantum algebraic and hyperbolic invariants, similar to the generalized volume conjecture.

1. Introduction

A knot K is the image of an embedding $S^1 \hookrightarrow S^3$. The same knot can, however, be drawn in $\mathbb{R}^2$ with over/undercrossing information in numerous ways. These different knot diagrams are related to each other by enacting sequences of Reidemeister moves. In general, it can be difficult to determine whether two knot diagrams represent the same knot because the required sequence of local moves may be complicated. The ambiguity in representation is partially resolved by appealing to topological invariants, which are certain numbers, polynomials, or other algebraic structures that may be computed using a knot diagram but that are independent of the particular diagram chosen for a specific knot. If two diagrams have different topological invariants, they represent different knots. The converse is not necessarily true.

The most common invariants have various mathematical and physical origins. For example, the Jones polynomial is an algebraic invariant associated to the Hecke algebra of the braid group [1]. It is equivalently described by skein relations and the Kauffman bracket [2] or, as Witten showed, in quantum field theory as the unknot normalized vacuum expectation value of the Wilson loop operator in $SU(2)$ Chern–Simons gauge theory [3]. Evaluating the trace of the Wilson loop in representations of $SU(2)$ other than the fundamental one, we obtain colored Jones polynomials. Promoting $SU(2)$ to $SU(N)$ generalizes the Jones polynomial to the HOMFLY-PT polynomial [3–5]. There are other polynomial invariants as well [6, 7].

Polynomial invariants in knot theory are typically Laurent polynomials with integer coefficients. This fact about the Jones polynomial is explained by associating the coefficients to the dimensions of certain bigraded homology groups, thus realizing the Jones polynomial as the graded Euler characteristic of this homology theory. The bigrading in this Khovanov homology yields a new polynomial knot invariant, the Khovanov polynomial, a two variable polynomial whose powers are the homological and quantum gradings [8, 9]. Just as Khovanov homology categorifies the Jones polynomial, knot Floer homology categorifies the Alexander

polynomial [10]. These homology theories lead to a new class of integer-valued topological invariants such as the Rasmussen s-invariant [11], the Ozsváth–Szabó τ-invariant [10], and Hom's ε-invariant [12][6].

Other invariants are associated only to hyperbolic knots, namely those knots K for which the complement of K in S^3 admits a unique complete constant curvature hyperbolic metric[7]. Examples of such hyperbolic invariants include the volume of the complement and the lengths of certain longitude and meridian cycles in this metric. In addition, there are various three-dimensional numerical invariants such as the bridge index and Turaev genus and, since every knot bounds orientable embedded surfaces, invariants that rely on these surfaces such as the determinant, Arf invariant, and the four-dimensional smooth slice genus.

Just as the Jones polynomial has a meaning in three-dimensional Chern–Simons theory [3], the Khovanov polynomial is also described in the language of quantum field theory. The homological and quantum gradings correspond to fermion and instanton numbers F and I that weight classical supersymmetric solutions to a four-dimensional super-Yang–Mills theory [22]. Similarly, from a five-dimensional super-Yang–Mills gauge theory perspective, the bigradings in the Khovanov polynomial arise in a Q-cohomology in the Hilbert space and are associated to a $U(1) \times U(1)$ symmetry generated by F and I [23]. In six-dimensional $(0, 2)$-theory, one can characterize the Khovanov polynomial in terms of the cohomology of a supercharge $\mathcal{Q}$ on a certain Cauchy slice with a surface operator whose topology is set by the knot [23]. The physics interpretation of various knot invariants are natural in diverse dimensions, and correlations between the knot invariants point to relationships between different quantum field theories.

Given this vast zoo of knot invariants, we would like to formulate general theories that explicate the relationships between the various quantities. The generalized volume conjecture, which relates evaluations of the n-colored Jones polynomial at roots of unity $J_n(e^{2\pi i/n}; K)$ with the hyperbolic volume $\mathrm{Vol}(K)$ and the Chern–Simons invariant $\mathrm{CS}(K)$ [24–27] provides an example of a non-trivial relation and states that

$$\lim_{n\to\infty} \frac{2\pi \log J_n\left(e^{2\pi i/n}; K\right)}{n} = \mathrm{Vol}\,(K) + 2\pi^2 i\,\mathrm{CS}\,(K)\ , \tag{1.1}$$

where n labels the irreducible representation of $SU(2)$ with dimension n.

As an initial step in the endeavor, machine learning supplies a practical tool for identifying how knot invariants are correlated. For example, neural networks can predict quasipositivity, the s-invariant, and the τ-invariant from braid words together with other input invariants [28]. Inspired by the volume conjecture and numerical investigations by Dunfield [29] and Khovanov [30], work in [31] showed that the volume of the knot complement of hyperbolic knots can be machine learned from the Jones polynomial with better than 97% accuracy[8]. This success is explained physically through the analytic continuation of Chern–Simons theory [33]. Indeed, based on an evaluation of the Jones polynomial at $t = e^{3\pi i/4}$, there is a simple formula that approximates the volume to a similar accuracy as the trained neural network [32]. This constitutes a 'reverse engineering' of the neural network behavior in [31], a task which is known to be difficult in general. The phase at which the Jones polynomial is evaluated is determined using layer-wise relevance propagation [34] to ascertain the input feature most important to the neural network's performance. Likewise, the polynomial invariants also predict the s-invariant and the slice genus to 98% accuracy [35].

The unknot decision problem [36] and the slice ribbon conjecture [37] have also recently been attacked using machine learning. The use of machine learning as a tool to understand the structure of knots more generally has gained popularity [38–45]. Perhaps the most impressive result to date in this arena is a theorem that relates the signature σ, slope, volume, and injectivity radius of hyperbolic knots [46, 47]:

Theorem 1. *There exists a real constant c such that*

$$|2\sigma\,(K) - \mathrm{slope}\,(K)\,| \ \leqslant\ c\,\frac{\mathrm{Vol}\,(K)}{\mathrm{inj}\,(K)^3}\ . \tag{1.2}$$

The four topological invariants appearing in theorem 1 were determined from a saliency map that assigned an attribution score to each of a dozen inputs using gradient methods. Just as the four color theorem [48] demonstrated the utility of computers for proving theorems, (1.2) establishes the utility of machine learning for deducing exact results in mathematics.

In the spirit of these earlier efforts, in this work we systematize the study of how well various numerical knot invariants can be predicted either from a polynomial invariant or from up to three other numerical

[6] The HOMFLY-PT polynomial is categorified by a triply graded homology theory [13, 14]. New numerical invariants are likewise associated to this Khovanov–Rozansky homology [15–20].

[7] All but 32 of the 1701 936 knots up to 16 crossings are hyperbolic [21].

[8] The are approximately 840 000 unique Jones polynomials for knots up to 16 crossings. When knots with different volumes have the same Jones polynomial, the volumes differ on average by about 3%, so the neural network's performance is nearly optimal on this dataset [32].

invariants. We discover surprising correlations among knot invariants; targeting experiments to search for bipartite and multipartite correlations reveals that the most interesting of our results are actually relations between individual invariants, similar in spirit to the volume conjecture (1.1). This is rather surprising, as it means that multipartite correlations among knot invariants are either uncommon or subtle compared to relations between single invariants, and may require other techniques to detect. Although we primarily used feedforward neural networks in our experiments, more sophisticated architectures were explored but did not produced improved prediction accuracies (see section 2.2 for more details). The results obtained involve many of the well-known invariants we described earlier, including the Jones polynomial, Floer homological invariants τ and ε, and a hyperbolic longitude invariant. These correlations point toward the existence of both known and unknown structures relating algebraic and geometric knot invariants.

Two sections follow. In section 2, we describe our automated exploration of bipartite, tripartite, and more general multipartite correlations within a knot database, including a number of interesting case studies. In section 3, we discuss the potential mathematical and physical implications of these case studies. Two appendices contain information on running experiments and viewing results (appendix A) and short definitions of selected database invariants (appendix B).

2. Experiments

2.1. Data

All the experiments were run using the KnotInfo database [49], which we expanded to a larger dataset when possible. A full list of the 53 invariants which were used in the experiments is included in appendix B. Neural networks were trained to predict given invariants from others, thereby identifying interesting relationships between invariants in the database. We considered experiments with between one and three inputs to the neural network. This resulted in over 700 000 possible experiments. However, not all invariant values are calculated for all knots in the KnotInfo database. A neural network was only trained if, for a specific combination of invariants, there were more than 1000 knots to work with. The final set of results includes data from around 199 330 experiments. In some case studies where additional data was available, we supplemented the KnotInfo dataset with all 313 199 hyperbolic knots up to 15 crossings. In order to determine how our experimental results depend on the crossing number of the knots, we also used SnapPy [50] to generate a dataset of 2866 187 distinct knots with crossing numbers between 15 and 40 along with a collection of computable invariant values.

In order to be used in the experiments, some data had to be processed. All polynomials were simply represented as flattened versions of the vectors appearing in the KnotInfo database. Any invariants represented by Yes/No or Y/N were converted to ones and zeros, respectively. For regression tasks, any entries where the output invariant is zero were removed[9]. For classification tasks, the output invariants were rescaled so that the lowest number in the output data set is zero.

2.2. Machine learning architecture

For each initial experiment, the neural networks were trained on 80% of the data. The ReLu (Rectified Linear Unit) activation function $r(x) \equiv \max(0, x)$ was used on the hidden layers, and the networks were trained using the Adam optimizer. The output layer was either a softmax $\vec{m}(\vec{v})$ with components $m(\vec{v})_k \equiv e^{v_k}/(\sum_j e^{v_j})$ or a ReLu activation function, depending on whether the task was classification or regression[10]. Similarly, the loss function was either sparse categorical cross-entropy[11] or mean squared error, depending on the task. Each network was trained for 100 'epochs,' each of which involved a single pass through the entire training set. The validation loss was monitored during training to prevent over-fitting.

During the initial experimentation the neural networks all had three hidden layers, each consisting of one hundred nodes[12]. These networks were trained on all the candidate combinations of input and target invariants (for a total of 199 330 experiments), and experiments where the output invariants were successfully predicted were flagged for further study. Additional feedforward neural networks of varying depths and widths were trained on these input/output pairs over the largest SnapPy random dataset to

[9] These knots were removed to avoid dividing by zero in the relative mean squared error percentage calculation. Only a handful of knots are removed in these cases, so this removal makes little difference to the experiments.

[10] All regression tasks we performed were predicting non-negative quantities, so the ReLu activation could be applied.

[11] The categorical cross-entropy is generally applied after a softmax layer, in which case it is equal to $-\log v_k - \sum_{i \neq k} \log(1 - v_i)$ where v_i are the softmax outputs and k represents the true label of the input. The sparse variant simply allows for the true label dataset to be loaded as integers rather than vectors with a one in the jth component representing the jth class.

[12] For more detailed discussions of architecture, see [31, 35].

Table 1. Descriptions of the neural network architectures used during the secondary round of experimentation.

Network	Depth (# of hidden layers)	Width (neurons/hidden layer)	Training epochs
Smallest	3	100	100
Narrow	9	100	300
Shallow	3	500	300
Large	6	200	300
Largest	6	500	300

determine how the choice of neural network architecture affected the accuracies of the resulting predictions. These neural networks are described in table 1.

Other neural network architectures were also considered for use in these experiments. Transformer neural networks [51] have achieved great success in working with sequential data, including large language models and machine translation. Although we implemented limited transformer experiments as part of this work, they did not result in improved prediction accuracy over the standard feedforward neural networks. This is likely due to the fact that our goal involved predicting numerical invariants, and thus the extra overhead of a transformer network was not needed in this setting to achieve accurate predictions.

Another class of neural network architectures that can be considered in this context are Kolmogorov–Arnold networks (KANs), introduced recently in [52]. In similar (albeit limited) invariant prediction tasks, we also found that KANs did not produce improved prediction results, though more extensive hyperparameter tuning may yield better results. As a result of the increased interpretability of these trained networks they may warrant a more extensive investigation beyond the scope of this paper.

2.3. Case studies

In this section, we discuss a handful of interesting results. In the cases where the input was a polynomial invariant, in addition to training networks on the full polynomial we also performed a search to find good evaluations of the polynomial. This was done by evaluating the polynomial at points on the square in $\mathbb{C}$ with corners $\pm 1 \pm i$ and training neural networks on each of the evaluations. We found that a single evaluation of a polynomial invariant is often sufficient to predict the desired output with high accuracy[13]. In other cases, we found single integer invariants that predicted real-valued quantities, hinting at a very strong underlying correlation. The accuracies of the neural network predictions in each of these case studies are summarized in table 2.

We note that in some cases (for example, predicting the longitude from τ) the precise architecture used had little impact on the accuracy of the predictions. In some cases the smaller networks performed better, for example when predicting ε from the Jones polynomial. In instances where the larger networks achieved higher accuracies the increase in performance was often modest. Taken together, these results may suggest that smaller network architectures provide a good starting point when performing invariant prediction tasks.

2.3.1. Ozsváth–Szabó $\tau \longrightarrow \varepsilon$

In an experiment using the expanded SnapPy random dataset, neural networks were able to predict ε from τ with essentially perfect accuracy (over 99%). This is an expected result, as the sign of τ is equal to ε for the majority of these knots. We include this case as a simple example of how the experiments rediscover known results.

2.3.2. Jones polynomial $\longrightarrow \varepsilon$

With the exception of a relation between the Jones polynomial and Rasmussen s-invariant (which was also uncovered by machine learning techniques [35]), the polynomial invariants are not known to relate directly to homological integer invariants like ε and τ. In this and the next case study, we find such correlations and discuss their implications further in section 3.

[13] For regression tasks, the quantity we refer to as accuracy is obtained from the mean relative error:

$$\text{accuracy} = 1 - \frac{1}{N}\sum \left| \frac{\text{predicted value} - \text{actual value}}{\text{actual value}} \right|,$$

where N is the size of the dataset. For classification tasks, it is instead defined simply as (number of correct classifications)/(total classifications). We modeled discrete invariant prediction as a classification task and continuous invariant prediction as a regression task.

Table 2. Summary of specific case study accuracy results, using various neural network architectures. The accuracy listed in each task is obtained as the average accuracy over five training runs. The highest average accuracy obtained in each task is underlined.

Output	Input	Smallest	Shallow	Narrow	Large	Largest
Epsilon	Tau	0.9997	0.9997	0.9997	0.9997	0.9997
Epsilon	Jones Polynomial	0.9056	0.9009	0.6950	0.8152	0.6810
Epsilon	Jones at $-0.6+0.1i$	0.7337	0.7678	0.3879	0.7134	0.3889
Epsilon	Real Jones at $-0.6+0.1i$	0.6983	0.7538	0.7857	0.7334	0.7166
Tau	Jones Polynomial	0.7757	0.7705	0.7808	0.7846	0.3002
Tau	Jones at $-0.7+0.1i$	0.5418	0.5265	0.4533	0.5495	0.5507
Longitude	Tau	0.8473	0.8478	0.8433	0.8469	0.8468
Longitude	Jones Polynomial	0.8962	0.9021	0.9061	0.9043	0.9014
Longitude	Jones at $-1-0.2i$	0.7957	0.7963	0.7960	0.7961	0.7958
Longitude	Alexander Polynomial	0.8936	0.8941	0.7780	0.8991	0.7864
Longitude	Determinant	0.8271	0.8267	0.8270	0.8264	0.8032
Longitude	Seifert Genus	0.8361	0.8306	0.8010	0.8214	0.8376

The Jones polynomial predicted ε to 96.22% accuracy in the smaller dataset, while the accuracy from predicting the modal value of ε gives 46.69% accuracy[14]. In the larger dataset of knots up to 15 crossings, the Jones polynomial predicted ε to 94.51% accuracy. Using the largest random SnapPy dataset the Jones polynomial predicted ε with an accuracy of 90.56%. All experiments that follow in this case study were performed using the KnotInfo dataset.

We are also interested in whether evaluations of the Jones polynomial can be used to predict ε.[15] To test this, we chose five random complex points on which to train the neural network. We then used layerwise relevance propagation (LRP) [34] to identify relevant inputs. LRP is a technique used to assign a relevance score to each input feature in a neural network, in order to explain how the neural network makes it predictions. It achieves this by modeling relevance as a sort of graph-theoretic flow which propagates backwards through a network, starting at the output layer, and redistributing the flow (relevance scores) into the previous layers while enforcing flow conservation so that the total relevance in each layer is always equal to the original value at the output source[16].

In [32], LRP identified $t = e^{3\pi i/4}$ as a phase at which the evaluation of the Jones polynomial predicts the hyperbolic volume to 97% accuracy. To pinpoint this phase, the Jones polynomial was evaluated at various roots of unity, in part referencing the volume conjecture and also because of the success of the neural network when given information about the coefficients only and not the degrees. Here, where there is no *a priori* reason to suspect that phases are important, we consider evaluations at random complex numbers close to the origin in the upper half plane.

The results of the LRP experiment are shown in figure 1. Since one evaluation is consistently more relevant than others, we trained the neural network on this single evaluation of the Jones polynomial. Training on the Jones polynomial at $t = -0.98 + 0.88i$, the neural network achieved an accuracy of 82.01% at predicting ε over five runs.

The LRP experiments indicate that a single evaluation of the Jones polynomial may be sufficient to predict ε. To find the optimal evaluation, we trained the neural network on single Jones polynomial evaluations using points in the square with corners $\pm 1 \pm i$. The evaluation of the Jones polynomial at the point $t = -0.6 + 0.1i$ achieved 96.11% accuracy over five runs. Using only the real part of the evaluation, we obtain an accuracy of 87.22%. On the larger dataset of knots up to 15 crossings, the accuracies are 92.92% (full evaluation) and 81.83% (real part of the evaluation), while on the random SnapPy dataset the networks achieved accuracies of 73.37% (full evaluation) and 69.83% (real part of the evaluation). The results of the search for an optimal evaluation are shown in figure 2. This heatmap can be interpreted in terms of known

[14] In this and other experiments we compare the accuracy of our neural networks' predictions to a baseline accuracy we obtained by simply predicting the mean (for regression tasks) or mode (for classification tasks) of the target output invariant.

[15] A correlation of this sort, between the Jones polynomial and a homological integer invariant, was discovered for the s-invariant in [35].

[16] See [32] for a more quantitative explanation.

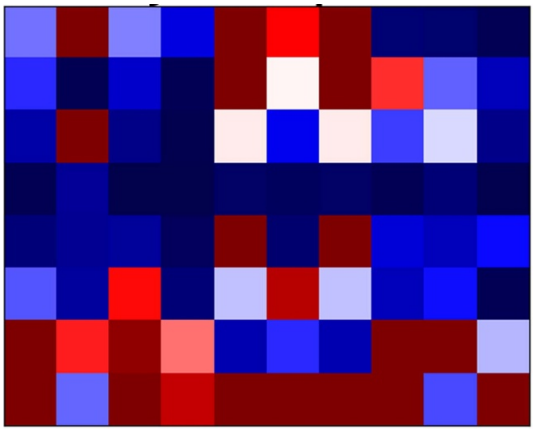

Figure 1. Layerwise relevance propagation results from a neural network trained to learn ε using evaluations of the Jones polynomial. Each column represents a randomly selected knot, and every pair of rows represent the real and imaginary parts of the Jones polynomial evaluations at a choice of point in the complex plane. Red pixels have higher relevance to the output prediction than blue pixels. Notice that certain evaluations are consistently more relevant than others. In this example, the final two rows are the most relevant. Although there are other rows that have high relevance for some knots, at least one of the final two rows (which correspond to a single evaluation) are highly relevant for every knot. These correspond to evaluations of the Jones polynomial at $t = -0.98 + 0.88i$. Similar images were obtained for other choices of input knots and evaluation points.

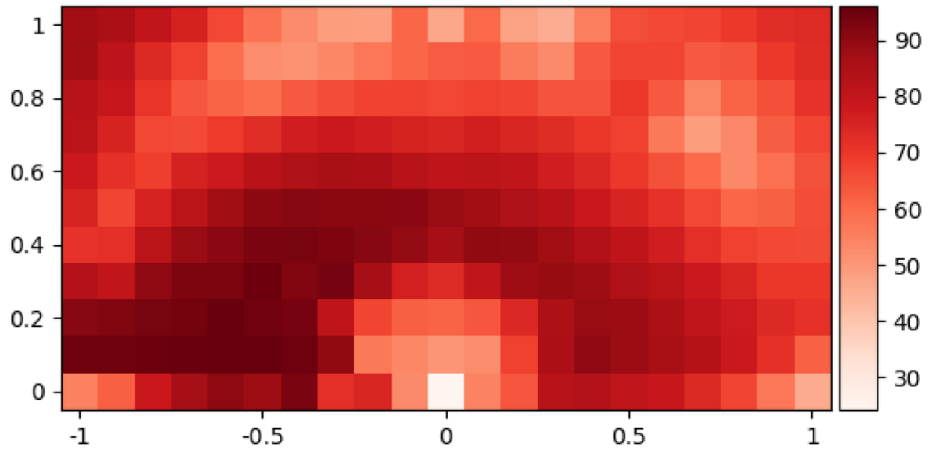

Figure 2. Heatmap showing accuracies of neural network predictions for ε from single evaluations of the Jones polynomial. The horizontal axis is the real part of the point where the polynomial is evaluated, and the vertical axis is the imaginary part. The colouring shows the accuracy of the neural network predictions. Each point is averaged over five training runs. Only the upper half plane is included, since the Jones polynomial is holomorphic. The best-performing point is at $t = -0.6 + 0.1i$.

evaluations of the Jones polynomial. For example, evaluating the Jones polynomial at $t = -1$ yields $\pm\det(K)$. The light coloring of the corresponding pixels in figure 2 indicates that the determinant is not particularly effective at predicting the value of ε, a fact which we observed by training a neural network to predict ε directly from $\det(K)$. Likewise, the evaluation of the Jones polynomial at $t = i$ depends only on the Arf invariant of K; the light coloration at the corresponding point of the heatmap indicates the Arf invariant does not yield accurate predictions of ε either. Similar conclusions can be drawn from the heatmap for τ in figure 3.

2.3.3. Jones/HOMFLY polynomials → Ozsváth–Szabó τ

The Jones polynomial predicts τ to 88.86% accuracy, while the accuracy from making predictions from the modal value of τ on our dataset is 29.21%. On the random SnapPy dataset the Jones polynomial predicted τ to an accuracy of 77.57%. The HOMFLY polynomial predicts τ to 88.44% accuracy. Scanning evaluations of the Jones polynomial, we find that the evaluation at $t = -0.7 + 0.1i$ predicts the Ozsváth–Szabó τ-invariant to 94.33% accuracy over five runs (though this decreased to an accuracy of only 54.18% on the random SnapPy dataset). This is surprising, since the evaluation at a single number outperforms the full Jones polynomial. One possible explanation for this fact is that the information required to accurately predict τ is more obscured in the full polynomial—and therefore the relationship is more difficult to learn—while it

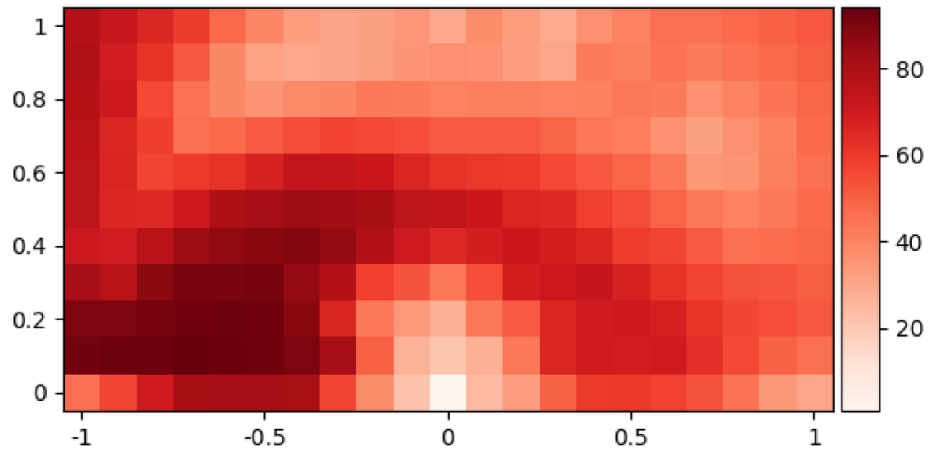

Figure 3. Heatmap showing accuracies of neural network predictions for τ from single evaluations of the Jones polynomial. The horizontal axis is the real part of the point where the polynomial is evaluated, and the vertical axis is the imaginary part. Each point is averaged over five training runs. Only the upper half plane is included, since the Jones polynomial is holomorphic. The best-performing point is at $t = -0.7 + 0.1i$.

becomes more direct after a single evaluation. A similar relationship between the Rasmussen s-invariant and the Khovanov polynomial of a knot was also observed in [35], for example. As well, evaluations of the Jones polynomial at the phase $e^{3\pi i/4}$ performed almost as well as the coefficients of the full Jones polynomial in predicting the volume of the knot complement of hyperbolic knots [32]. This is understood in the analytical continuation of Chern–Simons theory [33], for which the approximation formula works well for fractional Chern–Simons levels for which the geometric conjugate connection makes a contribution to the Chern–Simons path integral. A similar effect may explain the results quoted here.

Neither the real nor the imaginary parts alone are enough to predict τ accurately. The magnitude and phase do not give accurate predictions either. The results of the search for an optimal evaluation are shown in figure 3. All experiments for this case study used the KnotInfo dataset except where otherwise indicated.

2.3.4. Jones/Kauffman polynomials → Turaev genus g_T

The Jones polynomial is known to be related to the Turaev genus g_T via a bound coming from the span of the polynomial [53]. More precisely, if $\alpha_{\max}$ and $\alpha_{\min}$ are the maximum and minimum powers of t appearing in the Jones polynomial respectively, the Turaev genus is bounded by

$$g_T \leqslant c - |\alpha_{\max} - \alpha_{\min}| \,, \tag{2.1}$$

where c is the crossing number of the knot. In this case study we find a stronger relationship which uses either the full polynomial or the span in a nontrivial way that avoids using information about the crossing number.

The Jones polynomial predicts g_T to 91.41% accuracy, while the accuracy from predicting the modal value of g_T is 62.71%. In view of (2.1), this is not surprising, since (2.1) means the Jones polynomial contains some non-trivial information about the Turaev genus. The amount of information contained by the Jones polynomial is similar to estimating the crossing number itself, though the performance of the network when explicitly given the crossing data is better: we find that training a neural network on $c - |\alpha_{\max} - \alpha_{\min}|$ yields an accuracy of 99.89%[17]. Training the neural network on only the span of the Jones polynomial achieves an accuracy of 90.47%, which indicates that the span contains nearly as much information as the full Jones polynomial. One rather trivial explanation for this may be the following: the distribution of spans of the Jones polynomial correlates with crossing number in our dataset. This would explain the strong performance of the network but would not point to any deep underlying relationship like (2.1) which allows one to avoid the crossing number.

The Kauffman polynomial predicts g_T to 91.26% accuracy. As the Jones polynomial is a specialization of the Kauffman polynomial, this is to be expected. Most evaluations of the Jones polynomial that were tested do not perform better than the baseline 62% accuracy. The best performing evaluation we tested was at the point $t = -1 + 0.2i$, which predicts the Turaev genus to 78.89% accuracy over five runs.

[17] Simply predicting the upper bound provided by the inequality (2.1) for the Turaev genus on all knots only yields an accuracy of 65.96%. The improved performance of the neural network given $c - |\alpha_{\max} - \alpha_{\min}|$ may be explained as follows: on the KnotInfo dataset an accuracy of above 99% can be achieved by guessing 0 for the Turaev genus if $c - |\alpha_{\max} - \alpha_{\min}| = 0$, and guessing 1 otherwise.

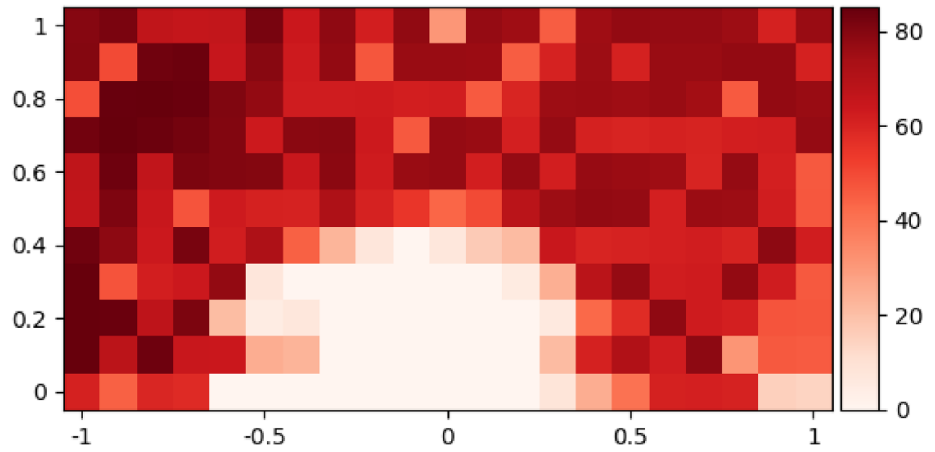

Figure 4. Heatmap showing accuracies of neural network predictions for longitude length from single evaluations of the Jones polynomial. The horizontal axis is the real part of the point where the polynomial is evaluated, and the vertical axis is the imaginary part. Each point is averaged over five training runs. Only the upper half plane is included, since the Jones polynomial is holomorphic. One of the best-performing points is at $t = -1 + 0.2i$.

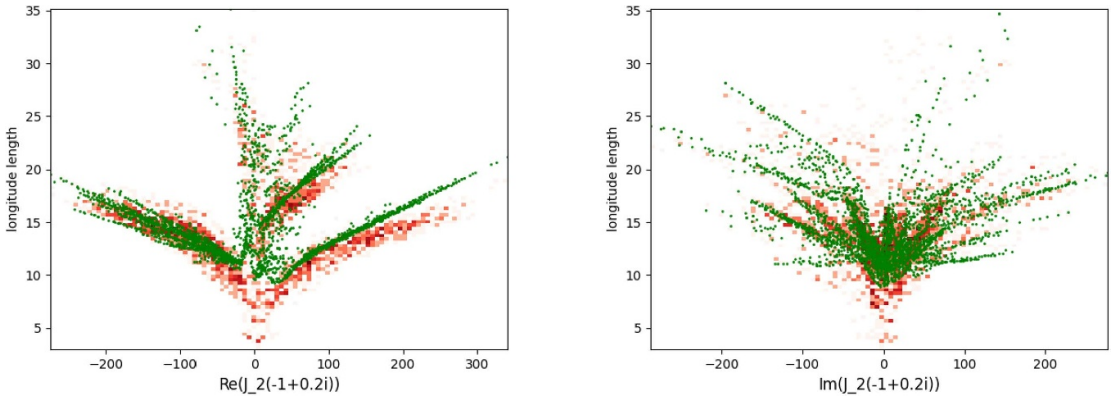

Figure 5. Predictions from the neural network trained on an evaluation of the Jones polynomial at $t = -1 + 0.2i$. (Left) The real part of the Jones polynomial plotted against the longitude length. (Right) The imaginary part of the Jones polynomial plotted against the longitude length. The red points show the actual longitude length (darker red indicating higher density of points) and the green points show the neural network predictions.

2.3.5. Jones/HOMFLY/Conway/Alexander polynomials → Longitude length ℓ

With the exception of the generalized volume conjecture, polynomial invariants are not known to be related to hyperbolic invariants. In this case study we find a very robust correlation between many different polynomial invariants and a particular hyperbolic invariant. We discuss implications of this result in section 3.

Using the Jones polynomial, the neural network predicted the longitude length to 89.62% accuracy (averaged over five runs). The accuracy from predicting the mean longitude length for every knot is 46.62%, so the Jones polynomial provides a significant improvement.

One can also investigate how well the neural network can perform when training on only an evaluation of the Jones polynomial at a single point. These experiments were performed using the KnotInfo dataset, and the results are displayed as a heatmap in figure 4. Many evaluations of those sampled achieve over 80% accuracy. The evaluation of the Jones polynomial at $t = -1 - 0.2i$ achieved one of the best accuracies: 87.51% over five runs. On the expanded SnapPy dataset the evaluation at this point achieved an accuracy of 79.57%. We also plot the longitude length against the real and imaginary parts of the Jones polynomial evaluation at $t = -1 - 0.2i$ (figure 5). Note that the evaluation at $t = -1$ also performs well, signalling a relation with the determinant $\det K$. The determinant is equal to the absolute value of the Jones polynomial evaluated at -1, and can similarly be extracted from the Alexander polynomial. Since the determinant is a common feature of both polynomials, we might suspect that it is really $\det K$ which serves as a predictor of the longitude length. As we will see in the next case study, $\det K$ does perform well, but at least a bit worse

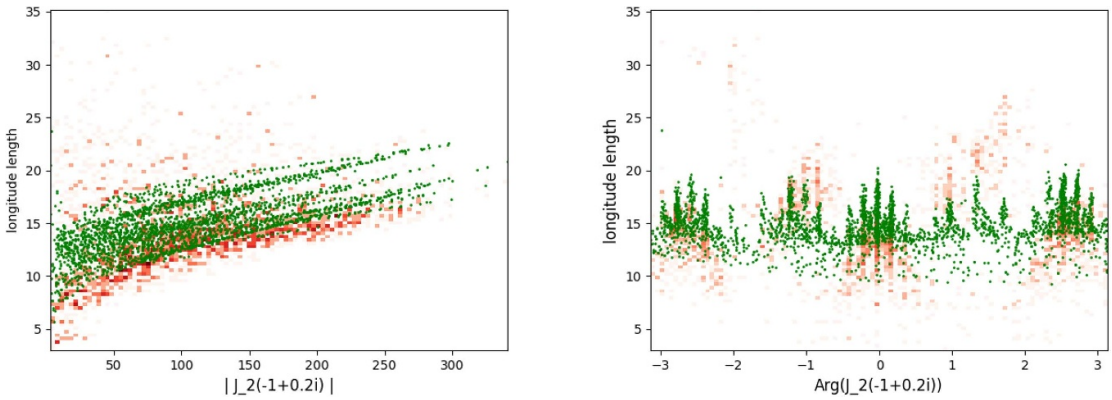

Figure 6. Predictions from the neural network trained on an evaluation of the Jones polynomial at $t = -1 + 0.2i$. (Left) The magnitude of the Jones polynomial plotted against the longitude length. (Right) The phase of the Jones polynomial plotted against the longitude length. The red points show the actual longitude length (darker red indicating higher density of points) and the green points show the neural network predictions. The plot of the magnitude is very similar in form to the $\det K$ correlation in figure 7.

than the full polynomial invariants. So the full polynomials contain some extra information about ℓ when compared to $\det K$.

We repeated the above experiments using the magnitude and phase of the evaluations as inputs, rather than the real and imaginary parts. The performance dropped to around 82.82%, averaged over five runs. Using this form of the evaluation may obscure some information about the longitude length. In figure 6, we plot the longitude length against the magnitude and phase of the Jones polynomial evaluation at $t = -1 - 0.2i$.

The Conway, HOMFLY, and Alexander polynomials all give good predictions of the longitude length as well (88.73%, 88.47%, and 89.36% accuracy, respectively).

2.3.6. Longitude length predictions

Initial experiments indicate that certain integer invariants, namely the three genus g_3, Ozsváth–Szabó τ-invariant, Rasmussen s-invariant, and determinant $\det K$, all predict the longitude length ℓ quite well. These invariants are not known to be related to hyperbolic invariants, as we discuss in section 3.

To explore these correspondences, we expand our experiments to include the larger dataset as well as knots with 16 crossings, for a total of over 1.7 million knots. Training on only 10% of the data for 100 epochs, we find that $\det K$, τ, and g_3 predict ℓ with 85.00%, 86.97%, and 84.51% accuracy, respectively (82.49%, 84.72%, and 83.55% respectively on the largest SnapPy dataset). Plots of these invariants against the longitude length are shown in figure 7. In table 3, we show the average longitude length for each value of τ and g_3 in the dataset. If we simply predict the mean value of ℓ for a particular τ or g_3, we get 86.89% and 84.68% accuracy, respectively. This recovers the accuracy of the neural network. Indeed, it explains the majority of the strong performance of the polynomial invariants in predicting ℓ as well, since $\det K$ may be extracted from those invariants by a simple evaluation.

2.3.7. Multipartite correlations

Thus far, all of the case studies we have discussed actually involved relations between individual invariants. In our experiments, we also trained neural networks to predict a single invariant from multiple invariants, and we filtered the experiments to isolate truly multipartite correlations. For instance, we generated a list of experiments with two input invariants which achieved greater than 80% accuracy where neither of the two inputs alone could predict the output with greater than 80% accuracy. Such a list contains tripartite correlation between the two inputs and one output. We performed similar experiments looking for invariants that could be predicted well from triples of input invariants, likewise filtering out those that could be explained by simpler correlations.

Perhaps surprisingly, we were not able to find many interesting examples of multipartite correlation. The few that did appear at the tripartite level with high accuracy were often of the form $(\text{genus}, \varepsilon) \to$ homological invariant, where (genus) could be the smooth or topological four genus, three genus, or others and (homological invariant) could be the s- or τ-invariants. These sorts of correlations are expected in the following sense. The homological invariants are related to the various genus quantities, and indeed this was one of the original motivations for their invention [11]. The only subtlety is that a genus is non-negative while the homological invariants can be negative. The ε-invariant, which is either ± 1 or zero,

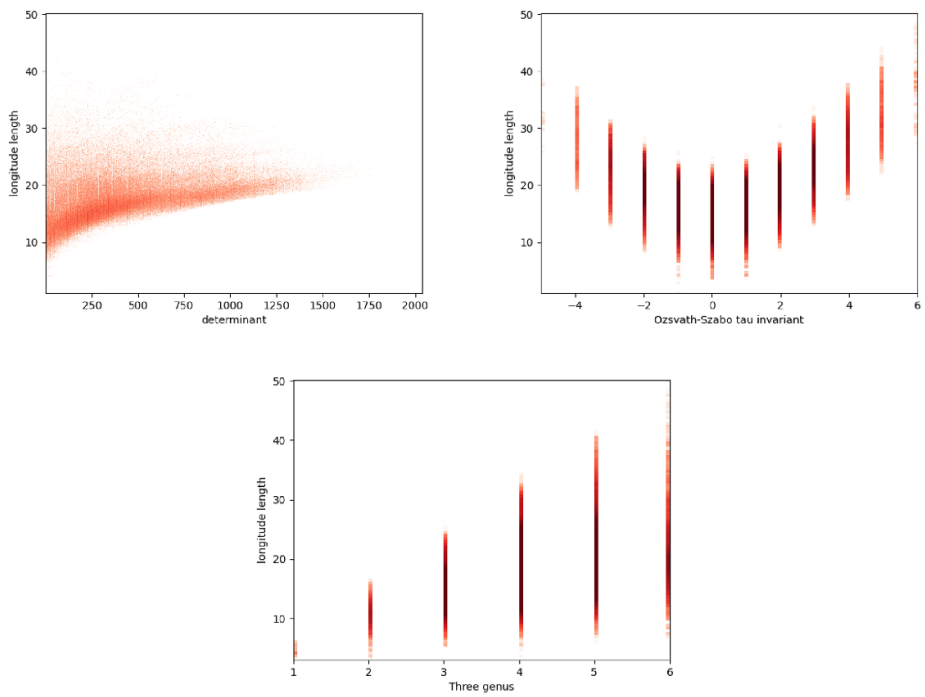

Figure 7. The determinant, Ozsváth–Szabó τ-invariant, and the three genus plotted against the longitude length. Darker regions correspond to a higher density of knots.

Table 3. Average longitude lengths for given values of τ and g_3. We notice that $|\tau|$ may be sufficient to predict ℓ to high accuracy.

τ	Mean ℓ	g_3	Mean ℓ
-5	32.14 ± 3.45	—	—
-4	27.54 ± 4.18	—	—
-3	23.08 ± 3.37	—	—
-2	19.31 ± 2.77	—	—
-1	16.33 ± 2.61	—	—
0	15.10 ± 2.56	—	—
1	16.55 ± 2.50	1	4.79 ± 0.76
2	19.52 ± 2.61	2	10.74 ± 1.76
3	23.21 ± 3.02	3	14.62 ± 2.62
4	27.25 ± 3.52	4	17.58 ± 3.43
5	31.92 ± 4.07	5	19.56 ± 4.00
6	37.68 ± 5.03	6	20.58 ± 4.85

seems to give just enough additional information to fix the sign, as in the datasets we used it is very often the case that the sign of ε matches the sign of the homological invariants τ or s.

Another tripartite correlation which is unusual at first sight involves tasks of the form $(\ell, \varepsilon) \to$ homological invariant, where the homological invariant is again τ or s. This correlation is similar in form to the one we just discussed: the longitude length gives a reasonable estimate of the magnitude of the homological invariant, and ε provides the sign. The theoretical underpinning of this correlation is much less clear, and is on the same footing as our experiments showing that ℓ may be predicted from τ. Given that empirical observation, these tripartite correlations are also not surprising, since we know that for knots in our datasets the distribution of ℓ is highly correlated with τ.

3. Discussion

We have automated a large number of machine learning experiments which search knot datasets for novel relationships between knot invariants. In addition to recovering earlier relationships we had either only discovered [35, 54] or both discovered [31] and explained [32], we have found new relationships between broad classes of knot invariants. We now discuss the mathematics and physics of these invariant classes in more detail and speculate on possible interpretations of our results.

3.1. Khovanov and Floer homology

The first type of novel relation involves two of the most important knot homology theories: Khovanov and Floer homology. A homology theory is an algebraic structure consisting of a sequence of vector spaces C_k and 'differentials' $d_k : C_k \to C_{k+1}$ obeying $d_{k+1} \circ d_k = 0$, which allows one to take the quotients $H_k \equiv \ker(d_{k+1})/\mathrm{im}(d_k)$.

Khovanov homology [8] is defined by using vector spaces C_k which are built from 'smoothings' of a knot diagram. Each crossing in the diagram can be uncrossed in two different ways, and when a choice is made at every crossing we are left with a set of disjoint circles called a smoothing. After weighting these circles by tensor products of vector spaces and combining them in a certain way, one arrives at Khovanov's C_k. The Khovanov differential is defined by noticing that the set of all smoothings forms a 'cube of resolutions,' where edges connect two smoothings that differ by a difference in choice at only a single crossing. Then the d_k are formed from maps which either 'fuse' the corresponding tensor factors (if the difference in smoothings corresponds to two circles merging into one) or 'split' a single tensor factor into two (if the difference in smoothings corresponds to a single circle splitting into two). As we mentioned in section 1, the graded Euler characteristic of the resulting homology theory Kh_k is the unnormalized Jones polynomial. There are also proposals [15, 23, 55, 56] for physical theories with Hilbert spaces that are supposed to be isomorphic (in certain sectors) to Khovanov homology.

Knot Floer homology [10, 57, 58] can be constructed rather similarly, via a cube of resolutions [59], but here the resolutions can include singular points where a crossing is made to overlap at a double point. The graded Euler characteristic of this homology theory KFH_k is the Alexander polynomial. The underlying physical theories are related to $\mathcal{N} = 2$ supersymmetric gauge theories [60–62]. To our knowledge, there are no known physical explanations for why the gauge and string theories relevant for Khovanov homology [23, 27] would be related to the variants of $\mathcal{N} = 2$ gauge theories in four dimensions [62] relevant for Floer homology.

While these two knot homology theories are not obviously related by physics, there are a few mathematical connections regarding the underlying Lie algebras [63, 64] and spectral sequences between the homology theories [16, 65–68]. In particular, the existence of a connection between Khovanov and Floer homology was suspected already in [65], and is cited as motivation for the development of the s-invariant [11]. A detailed understanding of the meaning of spectral sequences between the Hilbert spaces of supersymmetric gauge theories seems like a prerequisite for any physical explanation of the mathematical relationships between the two homology theories, and preliminary steps toward this goal have been explored [69]. However, we emphasize that the known mathematical relations also do not provide an explanation for the connection we have found between the Jones polynomial and Floer homology invariants like τ or ε.

Such specific relationships between Khovanov and Floer homologies may be different facets of the same underlying mysterious connection referred to as 'the FK correspondence' in [65]. With an eye toward understanding how our results extend outside of our chosen dataset, we note that [65] comments that this correspondence (explicitly between the rank of the homology theories and the s- and τ-invariants) is known to fail for sufficiently complicated knots. As such, it is possible that the predictions of τ and ε from the Jones polynomial that we have observed may break down when the complexity of the knot becomes great enough to violate the FK correspondence of [65] on average. However, if these correlations survive, they imply new robust entries in the dictionary of the FK correspondence.

3.2. Knot polynomials and hyperbolic invariants

One of the most well-known and influential conjectures in knot theory is the volume conjecture (1.1) [24–27]. From a coarse-grained perspective, the volume conjecture relates a quantum algebraic invariant (the colored Jones polynomial) to a classical geometric invariant (the hyperbolic volume of the knot complement). There is a detailed physical understanding of the mechanism which could give rise to something like the volume conjecture: it involves standard notions of analytic continuation and Picard–Lefschetz theory, but applied to a Feynman path integral instead of an ordinary integral [33, 70]. Indeed, this physical underpinning was used in [32] to give a quantitative explanation of the performance of a neural network [31] which predicted the volume from the Jones polynomial. Importantly, the classical volume invariant arises in the physics story as a saddle point value of the action evaluated on a certain 'geometric' gauge field configuration.

Here, we have found novel connections between quantum algebraic invariants and a hyperbolic invariant which is not known to appear as the saddle point value of some action integral: the hyperbolic longitude length ℓ. The algebraic invariants are the standard knot polynomials: Jones, HOMFLY-PT, Alexander, Conway, and Kauffman. It is curious that this relation with the longitude seems to persist between polynomials which serve as Euler characteristics of different homology theories. A similar statement cannot

be made for the hyperbolic volume. This may be a manifestation of the same mysterious connection between homology theories discussed above and in [65].

The most naïve way to interpret a strong correlation between the Jones polynomial and the longitude length is as a sort of generalization of the volume conjecture to other aspects of hyperbolic geometry. On one side of the relation, the necessary quantum algebraic ingredients are the colored Jones polynomials $J_n(t;K)$ with a choice of evaluation point t_n (or possibly multiple such points) and the semiclassical limit $n \to \infty$ with $\lim_{n\to\infty} t_n = t_\infty$ finite.

On the other side of the relation, we cannot simply use the classical action of $SL(2,\mathbb{C})$ Chern–Simons theory, because this already leads to the hyperbolic volume of the knot complement $S^3 \setminus K$. We need a new ingredient to land on the longitude length instead of the volume, and a natural candidate is the Wilson loop operator. The usual Wilson loop operator in gauge theories is interpreted roughly as the worldline of a charged particle coupled to the gauge field, but in gravitational theories the analogous object can measure properties of the spacetime like geodesic lengths. There is a well-known relation between $SL(2,\mathbb{C})$ Chern–Simons theory and gravity [71], so it may not be too much of a stretch to use the $SL(2,\mathbb{C})$ gauge theory Wilson loop operator to compute a geodesic length in the classical limit.

With these building blocks in hand, we can write a very rough conjectural relationship

$$\lim_{n\to\infty} \left(\log |J_n\left(e^{2\pi i/n}; K'(K)\right)| - \frac{n}{2\pi} \mathrm{Vol}(K) \right) \approx \ell \,, \tag{3.1}$$

where $K'(K)$ is a two-component link formed from K (a Wilson loop in the $SU(2)$ irreducible representation of dimension n) and an auxiliary Wilson loop which wraps the longitude of the cusp neighborhood in $S^3 \setminus K$ labeled by the fundamental representation of $SU(2)$. This auxiliary component should arise from a framing of K with zero self-linking; a self-linked auxiliary component would give rise to a winding geodesic in the cusp neighborhood associated with K in the complement $S^3 \setminus K$.

The intuition for this conjecture is that the knot K has a dimension scaling much faster than the auxiliary component, so its Wilson loop forms a background upon which the auxiliary component can be evaluated. In the large-n limit, the volume conjecture says that this background is the hyperbolic metric on the complement, so the classical limit of the auxiliary Wilson loop computes the exponential of the longitude length. This effect is a bit like the backreaction of D-branes which occurs in the AdS/CFT correspondence [72][18]. In the semiclassical limit, D-branes backreact and source a geometric background upon which strings and other objects such as probe branes propagate. In this situation, our intuition is that the Wilson loop with a dimension that scales with the Chern–Simons level will form a background, namely the hyperbolic geometry, upon which the lighter Wilson loop will measure a length. The total classical action is then the sum of the background volume term and the longitude length, so subtracting the volume and removing the classical divergence leaves behind the finite longitude length.

Unfortunately, this conjecture has several serious drawbacks. The fact that the longitude length is defined by enlarging the cusp neighborhood is not accounted for, but perhaps it emerges naturally as some kind of repulsion between the two loops. Perhaps more seriously, the proposed formula should imply a similar one for the meridian length, and our experiments have not revealed any such correlation[19]. Furthermore, the reason we have included the approximate symbol $\approx$ is because it will generally be difficult to disentangle the $O(n^0)$ terms coming from quantum corrections to the classical value of the action (the hyperbolic volume) from the classical contribution of the longitude length, which would also enter at $O(n^0)$ in $\log|J_n|$. We do not know of a way to isolate the effect of the auxiliary Wilson loop; this is an interesting issue which warrants further exploration.

3.3. Integer and hyperbolic invariants

There is at least one quantity which is known to arise from both Khovanov and Floer theory: the determinant $\det K$. This invariant is related to evaluations of both the Jones and Alexander polynomials:

$$\det(K) = |J(-1;K)| = |\Delta(-1;K)| \,. \tag{3.2}$$

In light of our discussion of several polynomial invariants predicting the longitude length ℓ, we might suspect that $\det K$ can also predict ℓ successfully. However, this is not quite true, and there is a good reason

[18] Of course, there are several formal differences between the two situations. In AdS/CFT, one has two descriptions of the D-branes which are equivalent. In the volume conjecture, the complex critical point yielding the volume is a required contribution to the analytically continued path integral, and no secondary description which avoids it exists. The intuition we are using is really more related to the idea of emergent geometry in a semiclassical limit, and this theme appears in both situations.

[19] This may be an artifact of our small dataset. The meridian length is distributed quite tightly compared to the longitude length in our dataset, so there is not much room for improvement in prediction by using a polynomial invariant. If our dataset had a wider range of meridian lengths, perhaps we would notice an improvement.

that it should not be true: $\det K$ is an integer-valued invariant while ℓ is a continuum quantity. In the volume conjecture, there is a scaling limit which allows a continuum quantity to emerge from a sequence of quantum invariants, but there is no such limit for the determinant. The polynomial invariants allow a sort of fictitious scaling limit by evaluation at a sequence of points in the complex plane. In fact, the region of the complex plane which we found was relevant for the Jones polynomial in predicting ℓ was quite close to $t = -1$, slightly shifted into the positive imaginary half plane.

Nevertheless, it is surprisingly quite possible to predict ℓ using $\det K$. We found strong performance for the ℓ prediction task given $\det K$, τ, or the three genus g_3. This bizarre result suggests yet another mysterious connection, but this time between the integer valued homology invariants and continuum hyperbolic invariants. Some portion of this correlation may be explained by the relationship we found before between the Jones polynomial and the s- or τ-invariants. However, since the knot polynomials contain a great deal of information that is logically distinct from the more subtle homological invariants like s or τ, we cannot explain the integer invariant connection to hyperbolic invariants with only that relation.

3.4. Unsupervised learning

A potentially fruitful direction for future work involves unsupervised learning. In this work, we automated the supervised learning of a large number of knot invariants. However, this automation required selection of a small number of invariants on which to train, and the selection of an invariant to predict. We might instead have set up an unsupervised learning problem where we feed the entirety of the dataset into the machine and ask for interesting intrinsic correlations like cluster formation. This may require considerably more computational power as well as an extended dataset to extract more robust correlations. We leave this to future work.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/JessRachel97/Knot_searcher_experiments.

Acknowledgments

We thank Jim Halverson, Jen Hom, Allison Moore, and Fabian Ruehle for discussions. JC acknowledges support from the South African Research Chairs Initiative of the Department of Science and Innovation and the National Research Foundation and the Rosenbaum-Faber Family Graduate Fellowship. MH is supported by a grant from the National Science Foundation (DMS-2213295), and thanks the Max Planck Institute for Mathematics in Bonn Germany for hosting him during much of the work contained here. VJ is supported by the South African Research Chairs Initiative of the Department of Science and Innovation and the National Research Foundation. AK is supported by the Simons Foundation through the It from Qubit Collaboration.

Appendix A. Running your own experiments and viewing results

The file '`learn.py`' in the linked repository [73] can be used to run your own experiments. It uses the KnotInfo dataset and the same neural network architecture that was used in the experiments described in this paper. To run an experiment, run the command `python learn.py <num inputs> <inputs> <output>`. For instance, to learn the hyperbolic volume from Jones polynomials, run the command `python learn.py 1 jones_polynomial_vector volume`.

The program `draw_output.py`, along with the results text files, can be used to produce LaTeX tables of desired results. Download the program and the results files and run the command `python draw_output.py`. The program will prompt you to answer questions about which results you want to view. The example below will produce a `.tex` and a `.pdf` file containing table 4.

The results files, in plain text format, is included in the Github repository. The results for all the experiments can be found in `results_{*}_fin.txt`, where {*} should be replaced by 'one', 'two', or 'three'. The `{*}_pruned.txt` files contain results which have been pruned to ensure true correlation between the full input and the output. In other words, a two-input experiment $A + B \to C$ will only survive pruning if its performance is improved when compared to the one-input experiments $A \to C$ and $B \to C$. The file `results_x_high.txt` are those pruned results whose accuracy is over 90%. Due to the small size of the dataset, and the fact that each experiment included in the results was only run once, you may find that the results of your own experiments differ from those in the results files. Training occasionally fails resulting in a low accuracy, when in reality the accuracy is high for the majority of runs.

Table 4. Example of output from the `draw_output` program.

Input 1	Output	Accuracy	Mean/Mode	Number
conway polynomial vector	volume	0.899 846 101 969 9341	0.769 207 335 176 7865	2970
jones polynomial vector	volume	0.957 915 108 777 278	0.769 207 335 176 7865	2970
kauffman polynomial vector	volume	0.884 362 335 650 6978	0.769 207 335 176 7865	2970
homfly polynomial vector	volume	0.862 161 362 295 2917	0.769 207 335 176 7865	2970
alexander polynomial vector	volume	0.918 092 352 173 835	0.769 207 335 176 7865	2970

```
user@path# python draw_output.py
Output file name?
my_test
How many inputs per experiment?  Experiments can have either one, two, or three inputs
1
Minimum accuracy?  From 0 to 1, minimum accuracy achieved by neural network.
0.5
Maximum accuracy?
1
Minimum number of knots? Minimum number of knots used in experiment. There are 2978 knots in total. Experiments with fewer that 100 knots are not run.
100
Maximum number of knots?
3000
Comma separated list ofinputs (i.e. volume,alternating,quasipositive) or type [ALL] or  [POLY]
POLY
Comma separated list of outputs or type [ALL]
volume
```

Appendix B. Invariants

As this paper's intended audience includes non-experts in knot theory, we briefly define the invariants used in our case studies.

Alexander polynomial: Historically the first knot polynomial, the Alexander polynomial is defined as

$$\Delta(t;K) = \det\left(V - tV^T\right), \tag{B.1}$$

where V is a Seifert matrix of the knot. The Alexander polynomial enjoys the property that $\Delta(t;K) = \Delta(t^{-1};K)$. Half the degree of the Alexander polynomial supplies a lower bound for the three genus, which is the minimal genus of a Seifert surface for a knot. (See the description of the three genus below for more details.) Knot Floer homology [10] categorifies the Alexander polynomial.

Conway polynomial: The Conway polynomial $\nabla(t;K)$ is a reparametrization of the Alexander polynomial:

$$\Delta\left(t^2;K\right) = \nabla\left(t - t^{-1};K\right), \tag{B.2}$$

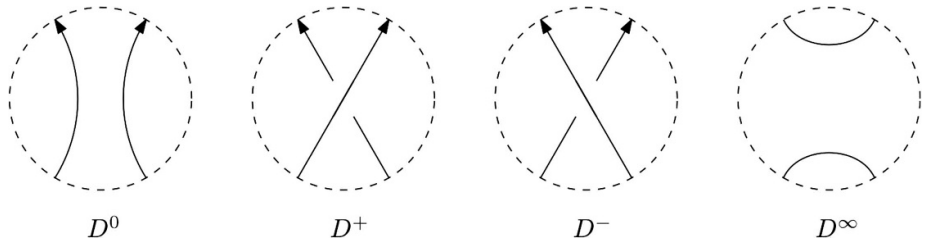

Figure 8. Local crossing configurations in D^0, D^+, D^-, and D^∞.

where this relation holds up to multiplication by an overall power of t. In terms of a Seifert matrix V of the knot K,

$$\nabla(t;K) = \det\left(t^{1/2}V - t^{-1/2}V^T\right) . \tag{B.3}$$

Determinant: The determinant of a knot is a three dimensional numerical invariant obtained from the Seifert matrix:

$$\det(K) = |\det\left(V + V^T\right)| . \tag{B.4}$$

As stated in (3.2), the determinant can also be computed from the absolute value of evaluations of the Jones or Alexander polynomials at -1.

Epsilon: As defined by Hom [12, 74], $\varepsilon(K)$ is an invariant of knots derived from comparing a pair of maps F_τ and G_τ on certain subquotient complexes of the knot Floer chain complex $CFK^\infty(K)$. The maps induced on homology by F_τ and G_τ cannot both be trivial, and $\varepsilon(K)$ takes values in $\{-1,0,1\}$ depending on which (or possibly both) of these two induced maps vanish.

If $\varepsilon(K)$ vanishes, then so does $\tau(K)$. Furthermore, for various classes of knots (e.g. Heegaard–Floer thin knots) the value of $\varepsilon(K)$ is equal to the sign of $\tau(K)$.

HOMFLY-PT polynomial: Consider the crossing and smoothings on a local region of a knot diagram as in figure 8. Denoting the unknot by $\bigcirc$, the HOMFLY-PT polynomial [4, 5] is determined by the skein relations:

I. $\alpha^{-1}P(\alpha,z;D^+) - \alpha P(\alpha,z;D^-) = zP(\alpha,z;D^0)$;
II. $P(\alpha,z;\bigcirc) = 1$.

The Jones and Alexander polynomials are specializations of the HOMFLY-PT polynomial:

$$J(t;K) = P\left(\alpha = t, z = t^{1/2} - t^{-1/2}; K\right) , \tag{B.5}$$

$$\Delta(t;K) = P\left(\alpha = 1, z = t^{1/2} - t^{-1/2}; K\right) , \tag{B.6}$$

where these relations hold up to an overall power of t. Analagous to the Jones polynomial discussed below, by taking $\alpha = t^N$, the HOMFLY-PT polynomial has an interpretation in $SU(N)$ Chern–Simons theory [3, 75].

Jones polynomial: The Jones polynomial has been defined in a number of equivalent ways, since its original definition by Jones in terms of von Neumann algebras. Here, we present a simple recursive definition due to Kauffman [2].

Let D^0, D^+, and D^- be three oriented diagrams which are identical except near a single crossing, where they are configured as shown in figure 8. Let $\bigcirc$ denote a knot diagram with no crossings, and let $D \sqcup \bigcirc$ denote a diagram consisting of the union of the diagram $\bigcirc$ and a disjoint knot diagram D.

The Jones polynomial then is a polynomial $J(t;D)$ in the variable t which satisfies the following properties:

I. $t^{-1}J(t;D^+) - tJ(t;D^-) = (t^{1/2} - t^{-1/2})J(t;D^0)$;
II. $J(t;D \sqcup \bigcirc) = -(t^{-1/2} + t^{1/2})J(t;D)$;
III. $J(t;\bigcirc) = 1$.

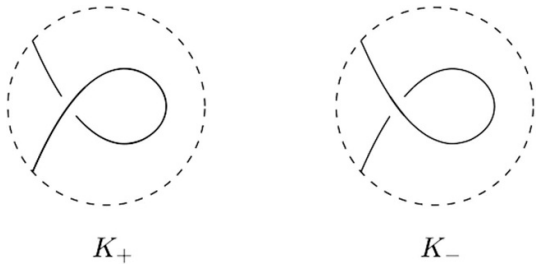

Figure 9. Adding a $+1$ kink (left) or a -1 kink (right) to obtain the knots K_+ and K_-, respectively.

Although defined in terms of a diagram D of K, the resulting polynomial $J(t;D)$ does not depend on the specific choice of D, and hence we may refer to the Jones polynomial $J(t;K)$ of the knot K. Khovanov homology [8] categorifies the Jones polynomial.

Witten [3] showed that the colored Jones polynomial is computed in Chern–Simons theory on a three manifold $\mathcal{M}$ as follows:

$$J_n(t;K) = \frac{\int [DA]\ e^{iS_{\rm CS}[A]}\ W_n(K)}{\int [DA]\ e^{iS_{\rm CS}[A]}\ W_n(\bigcirc)}\ , \tag{B.7}$$

where the Wilson loop operator computes the trace of the holonomy along a curve γ in the n-dimensional representation of $SU(2)$:

$$W_n(\gamma) = \mathrm{tr}_n\, \mathcal{P} \exp\left(i \oint_\gamma A \right)\ , \tag{B.8}$$

the Chern–Simons action is written in terms of a $\mathfrak{su}(2)$ valued gauge connection:

$$S_{\rm CS}[A] = \frac{k}{4\pi} \int_{\mathcal{M}} \mathrm{tr}\left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)\ , \quad k \in \mathbb{Z}^+\ , \tag{B.9}$$

and $t = e^{2\pi i/(k+2)}$. The theory is topological, i.e. independent of the metric on $\mathcal{M}$. Specializing to $n = 2$ recovers the Jones polynomial.

Kauffman polynomial: The Kauffman polynomial [76] is a two variable knot polynomial defined as,

$$F(a,z;K) = a^{-w(D)} L(a,z;D)\ . \tag{B.10}$$

Here $w(D)$ is the writhe of a knot diagram D, which is the difference between the number of positive crossings and negative crossings, and (referencing figure 8 with the orientations ignored) $L(a,z;D)$ is determined by skein relations:

I. $L(a,z;D_\pm) = a^{\pm 1} L(a,z;D)$;
II. $L(a,z;D^+) + L(a,z;D^-) = z\big(L(a,z;D^0) + L(a,z;D^\infty)\big)$;
III. $L(a,z;\bigcirc) = 1$.

Here, $D_\pm$ adds a ± 1 kink to the diagram D as in figure 9. The Kauffman polynomial is invariant under type 2 and type 3 Reidemeister moves, which, respecitvely, move one loop completely over another and move a string completely over or under a crossing. The Jones polynomial is a specialization of the Kauffman polynomial:

$$J(t;K) = F\left(-t^{3/4}, t^{-1/4} + t^{1/4}; K \right)\ . \tag{B.11}$$

The Kauffman polynomial has a definition in $SO(N)$ Chern–Simons theory [77].

Longitude length: A cusp of a hyperbolic knot K is a neighborhood of the knot intersected with the knot complement $S^3 \setminus K$. Such a neighborhood can be obtained as the image of a horoball H in the hyperbolic 3-space $\mathbb{H}^3$ under the universal cover $\rho : \mathbb{H}^3 \to S^3 \setminus K$. For such a horoball $H \subset \mathbb{H}^3$, the preimage $\rho^{-1}(\rho(H))$ will be a family of horoballs, and by increasing the size of H, we can arrange for these horoballs to be tangent to each other with disjoint interiors. In this case the image $\rho(H) \subset S^3 \setminus K$ is the maximal cusp of $S^3 \setminus K$.

The boundary of a cusp is the torus T^2, which has a meridian and a longitude. The lengths of geodesic representatives of these curves are invariants of K. The maximum longitude length is $5c(K) - 6$, where $c(K)$ is the crossing number of the knot [78].

Ozsváth–Szabó τ-invariant: Any knot $K \subset S^3$ defines a filtration on the Heegaard–Floer chain complex $\widehat{CF}(S^3)$ of S^3 (see [10] for the definition of $\widehat{CF}(S^3)$ and [58] or [79] for a definition of the filtration induced by K). If $\mathcal{F}_m(K)$ is the level-m subcomplex induced by the filtration, then the inclusion $\iota_m : \mathcal{F}_m(K) \hookrightarrow \widehat{CF}(S^3)$ induces a map $(\iota_m)_* : H(\mathcal{F}_m(K)) \to \widehat{HF}(S^3)$ on the homology of these chain complexes. The Ozsváth–Szabó τ-invariant $\tau(K)$ of K is defined to be the minimum integer m for which the map $(\iota_m)_*$ is nontrivial. The absolute value of τ gives a lower bound on the slice genus of a knot.

Rasmussen s-invariant: In [80], Lee defined a filtered chain complex associated to a knot diagram D, by perturbing the differential of the Khovanov chain complex of D. Lee proved that for a knot K the resulting homology always has rank 2, and Rasmussen proved that this homology is always supported in grading $s \pm 1$, where $s = s(K)$ is an invariant of K, called the Rasmussen s-invariant. The absolute value of s gives a lower bound on two times the slice genus of a knot.

Three genus: Every knot $K \subset S^3$ can be expressed as the boundary $K = \partial F$ for some compact, orientable, embedded surface $F \subset S^3$. Such a surface F is called a Seifert surface for K, and the minimal genus g_3 among all Seifert surfaces for K is called the three genus (or Seifert genus) of K.

Given a Seifert surface F for K of genus g, the corresponding Seifert matrix V is a $2g \times 2g$ matrix whose elements are the linking numbers between cycles a_i that are elements of a basis of $H_1(F)$ and their pushoffs a_i^+; i.e. the entries of V are given by $v_{ij} = \mathrm{lk}(a_i, a_j^+)$. This matrix can be used to define the Alexander and Conway polynomials as well as the determinant of a knot.

Turaev genus: Given a diagram D of a knot K, Turaev [53] described an algorithm for constructing a closed, orientable, unknotted surface $F(D) \subset S^3$ — called the Turaev surface of D — on which D sits as an alternating diagram. In particular, if D is an alternating diagram then $F(D)$ is a two-sphere in S^3. The minimal genus of $F(D)$ over all diagrams D of a knot K is called the Turaev genus of K, and is denoted by $g_T(K)$. A knot K is alternating if and only if its Turaev genus is zero.

If $J(t;K)$ is the Jones polynomial of K, let $\mathrm{span}(J(t;K))$ denote the difference between the largest and smallest exponents on nonzero monomials in $J(t;K)$, and let $c(K)$ denote the crossing number of K (i.e. the minimal number of crossings in any diagram of K). Then Turaev [53] proved that the $g_T(K)$ satisfies the following

$$\mathrm{span}\,(J(t;K)) \leqslant c(K) - g_T(K)\ . \tag{B.12}$$

Other invariants: The other 41 invariants we have investigated are tabulated below. The initial experiments used the data complied in KnotInfo [49] for knots up to 12 crossings. Invariants for knots up to 16 crossings in the expanded dataset were computed using SnapPy [50]. Real invariants are in bold. Boolean invariants are italicized.

arc index	braid index	braid length
bridge index	crosscap number	Morse–Novikov number
Nakanishi index	super bridge index	Thurston–Bennequin number
tunnel number	unknotting number	width
Arf invariant	3d clasp number	4d clasp number
smooth 4d crosscap number	topological 4d crosscap number	smooth concordance crosscap number
topological concordance crosscap number	smooth concordance order	algebraic concordance order
topological concordance order	smooth concordance genus	topological concordance genus
double slice genus	smooth four genus g	topological four genus
signature	*L-space*	**Chern–Simons invariant**
meridian length	**volume**	*alternating*
fibered	*almost alternating*	*adequate*
quasi-alternating	*positive braid*	*positive*
quasipositive	*strongly quasipositive*	

ORCID iDs

Mark Hughes https://orcid.org/0000-0001-6305-2949
Vishnu Jejjala https://orcid.org/0000-0003-2603-6717
Arjun Kar https://orcid.org/0000-0003-1943-4346

References

[1] Jones V F R 1985 A polynomial invariant for knots via von Neumann algebras *Bull. Am. Math. Soc.* **12** 103–11
[2] Kauffman L H 1987 State models and the Jones polynomial *Topology* **26** 395–407
[3] Witten E 1989 Quantum field theory and the Jones polynomial *Commun. Math. Phys.* **121** 351–99
[4] Freyd P, Yetter D, Hoste J, Lickorish W B R, Millett K and Ocneanu A 1985 A new polynomial invariant of knots and links *Bull. Am. Math. Soc.* **12** 239–46
[5] Przytycki J H and Traczyk P 2016 Invariants of links of Conway type (arXiv:1610.06679)
[6] Alexander J W 1923 A lemma on systems of knotted curves *Proc. Natl Acad. Sci.* **9** 93–95
[7] Alexander J W 1928 Topological invariants of knots and links *Trans. Am. Math. Soc.* **30** 275–306
[8] Khovanov M 2000 A categorification of the Jones polynomial *Duke Math. J.* **101** 359–426
[9] Bar-Natan D 2002 On Khovanov's categorification of the Jones polynomial *Algebr. Geom. Topol.* **2** 337–70
[10] Ozsváth P and Szabó Z 2004 Holomorphic disks and topological invariants for closed three-manifolds *Ann. Math.* **159** 1027–158
[11] Rasmussen J 2010 Khovanov homology and the slice genus *Invent. Math.* **182** 419–47
[12] Hom J 2014 Bordered Heegaard Floer homology and the tau-invariant of cable knots *J. Topol.* **7** 287–326
[13] Khovanov M and Rozansky L 2008 Matrix factorizations and link homology *Fundam. Math.* **199** 1–91
[14] Khovanov M and Rozansky L 2008 Matrix factorizations and link homology II *Geom. Topol.* **12** 1387–425
[15] Gukov S, Schwarz A S and Vafa C 2005 Khovanov-Rozansky homology and topological strings *Lett. Math. Phys.* **74** 53–74
[16] Dunfield N M, Gukov S and Rasmussen J 2005 The superpolynomial for knot homologies (arXiv:math/0505662)
[17] Gukov S, Putrov P and Vafa C 2017 Fivebranes and 3-manifold homology *J. High Energy Phys.* JHEP07(2017)071
[18] Gukov S, Pei D, Putrov P and Vafa C 2020 BPS spectra and 3-manifold invariants *J. Knot Theory Ramif.* **29** 2040003
[19] Gukov S and Manolescu C 2021 A two-variable series for knot complements *Quantum Topol.* **12** 1–109
[20] Chandler A and Gorsky E 2022 Structures in HOMFLY-PT homology (arXiv:2209.13058)
[21] Hoste J, Thistlethwaite M and Weeks J 1998 The first 1,701,936 knots *Math. Intell.* **20** 33–48
[22] Gaiotto D and Witten E 2012 Knot invariants from four-dimensional gauge theory *Adv. Theor. Math. Phys.* **16** 935–1086
[23] Witten E 2011 Fivebranes and knots (arXiv:1101.3216)
[24] Kashaev R M 1997 The hyperbolic volume of knots from quantum dilogarithm *Lett. Math. Phys.* **39** 269–75
[25] Murakami H and Murakami J 2001 The colored Jones polynomials and the simplicial volume of a knot *Acta Math.* **186** 85–104
[26] Murakami H, Murakami J, Okamoto M, Takata T and Yokota Y 2002 Kashaev's conjecture and the Chern-Simons invariants of knots and links *Exp. Math.* **11** 427–35
[27] Gukov S 2005 Three-dimensional quantum gravity, Chern-Simons theory and the A polynomial *Commun. Math. Phys.* **255** 577–627
[28] Hughes M C 2020 A neural network approach to predicting and computing knot invariants *J. Knot Theory Ramif.* **29** 2050005
[29] Dunfield N 2000 An interesting relationship between the Jones polynomial and hyperbolic volume (available at: https://faculty.math.illinois.edu/∼nmd/preprints/misc/dylan/index.html)
[30] Khovanov M 2003 Patterns in knot cohomology, I *Exp. Math.* **12** 365–74
[31] Jejjala V, Kar A and Parrikar O 2019 Deep learning the hyperbolic volume of a knot *Phys. Lett.* B **799** 135033
[32] Craven J, Jejjala V and Kar A 2021 Disentangling a deep learned volume formula *J. High Energy Phys.* JHEP06(2021)040
[33] Witten E 2011 Analytic continuation of Chern-Simons theory *Chern-Simons Gauge Theory: 20 Years After* (*AMS/IP Studies in Advanced Mathematics* vol 50) (American Mathematical Society) pp 347–446
[34] Montavon G, Binder A, Lapuschkin S, Samek W and Müller K-R 2019 Layer-wise relevance propagation: an overview *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning* (Springer) pp 193–209
[35] Craven J, Hughes M, Jejjala V and Kar A 2021 Learning knot invariants across dimensions (arXiv:2112.00016)
[36] Gukov S, Halverson J, Ruehle F and Sułkowski P 2020 Learning to unknot (arXiv:2010.16263)
[37] Gukov S, Halverson J, Manolescu C and Ruehle F 2023 Searching for ribbons with machine learning (arXiv:2304.09304)
[38] Levitt J S F, Hajij M and Sazdanovic R 2019 Big data approaches to knot theory: understanding the structure of the Jones polynomial (arXiv:1912.10086)
[39] Kauffman L, Russkikh N and Taimanov I 2020 Rectangular knot diagrams classification with deep learning (arXiv:2011.03498)
[40] Hajij M, Zamzmi G, Dawson M and Muller G 2020 Algebraically-informed deep networks (AIDN): a deep learning approach to represent algebraic structures (arXiv:2012.01141)
[41] Dłotko P, Gurnari D and Sazdanovic R 2021 Knot invariants and their relations: a topological perspective (arXiv:2109.00831)
[42] Hughes M, Eubanks A and Slone J 2021 Using generative adversarial networks to produce knots with specified invariants *AMS Special Session on Theoretical and Applied Perspectives in Machine Learning, I* (available at: https://meetings.ams.org/math/fall2021w/meetingapp.cgi/paper/11384)
[43] Hughes M, Eubanks A and Slone J 2021 Using deep learning to generate knots with prescribed invariants *AMS Special Session on Developments in Knot Theory and Low-Dimensional Topology, III* (available at: https://meetings.ams.org/math/fall2021c/meetingapp.cgi/paper/6594)
[44] Vernitski A *et al* 2022 Reinforcement learning algorithms for the untangling of braids *The Int. FLAIRS Conf. Proc.* vol 35
[45] Grünbaum D 2022 Narrowing the gap between combinatorial and hyperbolic knot invariants via deep learning *J. Knot Theory Ramif.* **31** 2250003
[46] Davies A *et al* 2021 Advancing mathematics by guiding human intuition with AI *Nature* **600** 70–74
[47] Davies A, Juhász A, Lackenby M and Tomasev N 2021 The signature and cusp geometry of hyperbolic knots (arXiv:2111.15323)
[48] Appel K and Haken W 1976 Every planar map is four colorable *Bull. Am. Math. Soc.* **82** 711–2
[49] Livingston C and Moore A H 2021 KnotInfo: table of knot invariants (available at: knotinfo.math.indiana.edu)
[50] Culler M, Dunfield N M, Goerner M and Weeks J R 2022 SnapPy, a computer program for studying the geometry and topology of 3-manifolds (available at: http://snappy.computop.org)
[51] Vaswani A 2017 Attention is all you need *Advances in Neural Information Processing Systems*
[52] Liu Z *et al* 2024 KAN: Kolmogorov-Arnold networks (arXiv:2404.19756)
[53] Turaev V G 1990 A simple proof of the Murasugi and Kauffman theorems on alternating links *New Developments In The Theory Of Knots* (World Scientific) pp 602–24
[54] Craven J, Hughes M, Jejjala V and Kar A 2022 (K)not machine learning *Nankai Symp. on Mathematical Dialogues: In Celebration of S.S.Chern's 110th Anniversary* (arXiv:2201.08846)
[55] Aganagic M 2020 Knot categorification from mirror symmetry, part I: coherent sheaves (arXiv:2004.14518)

[56] Aganagic M 2021 Knot categorification from mirror symmetry, part II: Lagrangians (arXiv:2105.06039)
[57] Floer A 1988 An instanton-invariant for 3-manifolds *Commun. Math. Phys.* **118** 215–40
[58] Rasmussen J A 2003 Floer homology and knot complements (arXiv:math/0306378)
[59] Ozsváth P and Szabó Z 2009 A cube of resolutions for knot Floer homology *J. Topol.* **2** 865–910
[60] Seiberg N and Witten E 1994 Monopoles, duality and chiral symmetry breaking in N = 2 supersymmetric QCD *Nucl. Phys.* B **431** 484–550
[61] Seiberg N and Witten E 1994 Electric-magnetic duality, monopole condensation and confinement in N = 2 supersymmetric Yang-Mills theory *Nucl. Phys.* B **426** 19–52
[62] Witten E 1994 Monopoles and four manifolds *Math. Res. Lett.* **1** 769–96
[63] Douglas C L and Manolescu C 2014 On the algebra of cornered Floer homology *J. Topol.* **7** 1–68
[64] Tian Y 2012 A categorification of U_q $sl(1,1)$ as an algebra (arXiv:1210.5680)
[65] Rasmussen J 2005 Knot polynomials and knot homologies (arXiv:math/0504045)
[66] Ozsváth P and Szabó Z 2005 On the Heegaard Floer homology of branched double-covers *Adv. Math.* **194** 1–33
[67] Dowlin N 2018 A spectral sequence from Khovanov homology to knot Floer homology (arXiv:1811.07848)
[68] Beliakova A, Putyra K K, Robert L-H and Wagner E 2022 A proof of Dunfield-Gukov-Rasmussen conjecture (arXiv:2210.00878)
[69] Gukov S, Nawata S, Saberi I, Stošić M and Sułkowski P 2016 Sequencing BPS spectra *J. High Energy Phys.* JHEP03(2016)004
[70] Witten E 2010 A new look at the path integral of quantum mechanics (arXiv:1009.6032)
[71] Witten E 1988 (2+1)-dimensional gravity as an exactly soluble system *Nucl. Phys.* B **311** 46
[72] Maldacena J M 1998 The large N limit of superconformal field theories and supergravity *Adv. Theor. Math. Phys.* **2** 231–52
[73] Craven J, Hughes M, Jejjala V and Kar A 2022 Accompanying code (available at: https://github.com/JessRachel97/Knot_searcher_experiments)
[74] Hom J 2014 The knot Floer complex and the smooth concordance group *Comment. Math. Helv.* **89** 537–70
[75] Rama Devi P, Govindarajan T R and Kaul R K 1993 Three-dimensional Chern-Simons theory as a theory of knots and links. 3. Compact semisimple group *Nucl. Phys.* B **402** 548–66
[76] Kauffman L H 1990 An invariant of regular isotopy *Trans. Am. Math. Soc.* **318** 417–71
[77] Horne J H 1990 Skein relations and Wilson loops in Chern-Simons gauge theory *Nucl. Phys.* B **334** 669–94
[78] Futer D, Kalfagianni E and Purcell J S 2016 A survey of hyperbolic knot theory *Int. Conf. on KNOTS* (Springer) pp 1–30
[79] Ozsváth P and Szabó Z 2003 Knot Floer homology and the four-ball genus *Geom. Topol.* **7** 615–39
[80] Lee E S 2005 An endomorphism of the Khovanov invariant *Adv. Math.* **197** 554–86