Learning Knot Invariants from Latent Representations

Edison Vázquez[1,3*†], Radmila Sazdanovic[2*†], Ernesto Lupercio[1,3*†], Carlos Ruiz[3†], Aldo Guzmán-Sáenz[4†], Pedro Olivares[3, 5†]

[1*]Department of Mathematics, Center for Research and Advanced Studies of the National Polytechnic Institute, Av. Instituto Politécnico Nacional , Mexico City, 07360, Mexico City, Mexico.

[2]Department of Mathematics, North Carolina State University, 2311 Stinson Dr, Raleigh, 27695, North Carolina, USA.

[3]Colegio de Matemáticas Bourbaki, Privada San Antonio Cucul, Merida, 97119, Yucatan, Mexico.

[4]T. J. Watson Research Center, IBM Research, 1101 Kitchawan Rd, Yorktown Heights, 10598, NY, USA.

[5]Instituto Tecnológico Autónomo de México, ITAM, Río Hondo 1, Mexico City, 01080, Mexico.

*Corresponding author(s). E-mail(s): evazquez@math.cinvestav.mx; rsazdan@ncsu.edu; lupercio@math.cinvestav.mx;

Contributing authors: alfonso@escuela-bourbaki.com; aldo@ibm.com; polivar4@itam.mx;

†These authors contributed equally to this work.

Abstract

The abstract serves both as a general introduction to the topic and as a brief, non-technical summary of the main results and their implications. Authors are advised to check the author instructions for the journal they are submitting to

for word limits and if structural elements like subheadings, citations, or equations are permitted.

Keywords: keyword1, Keyword2, Keyword3, Keyword4

1 Introduction

Knot invariants such as the Alexander and Jones polynomials offer partial yet algebraically rich characterizations of knot structure, while the HOMFLYPT polynomial is a more powerful invariant that can distinguish knots which simpler invariants fail to separate. However, computing the HOMFLYPT polynomial is #P-hard in general [**JaegerVertiganWelsh1990**], [RS: The Jones and HOMFLY polynomials are both #P-hard] and remains intractable for complex diagrams even with recent advances in fixed-parameter tractability [**Burton2018**]. This motivates the use of data-driven approximations to infer HOMFLYPT representations from more accessible invariants like Alexander and Jones.

Most knot invariants can be represented as Laurent polynomials:

$$\Delta_K(t) = \sum_{k=-d}^{d} a_k \, t^k,$$

where K denotes the knot, d is the maximal degree, and each $a_k \in \mathbb{Z}$ is an integer coefficient. For the Alexander polynomial, these coefficients arise from a Seifert matrix V of a spanning surface of the knot, defined as:

$$\Delta_K(t) = \det(V - tV^T).$$

This algebraic form makes polynomial invariants naturally suitable for machine learning. By aligning exponents across the dataset, we embed each polynomial as a fixed-length vector $[a_{-d}, \ldots, a_d]$, allowing their use as numerical inputs in classification and regression tasks [**dlotko2025data**; **DlotkoGurnariSazdanovic2024**].

In this work, we investigate whether a neural network can approximate a function $f(A, J) \approx \text{HOMFLY}$, where A and J are the Alexander and Jones polynomial vectors, respectively. We consider two regimes: one using raw polynomial coefficients as input, and one using latent representations extracted via autoencoders. The latter consistently yields lower reconstruction and prediction error, suggesting that latent spaces capture more meaningful structure.

Analyzing reconstruction error across a dataset of over 300,000 knots, we observe a consistent set of examples that are poorly reconstructed despite being present in the training set. We refer to these as *hard knots*—instances whose structure is topologically complex or sparsely represented in the training distribution. Remarkably, these

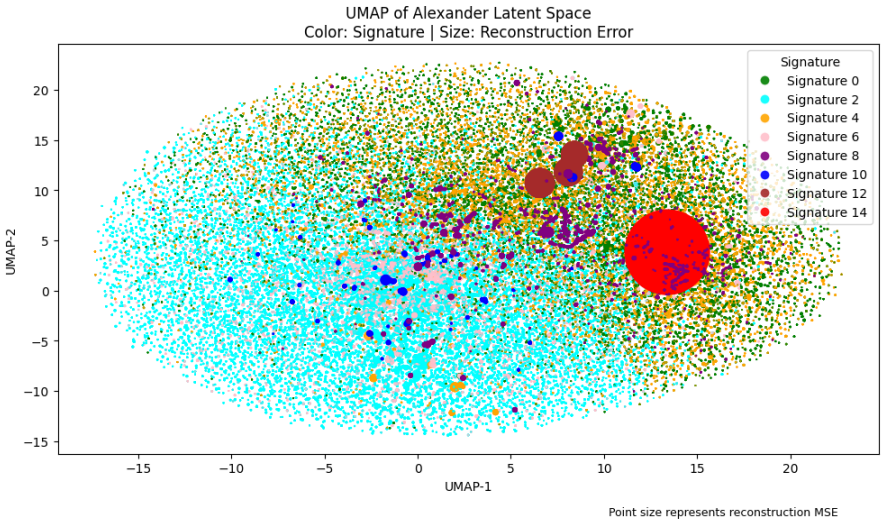


Fig. 1: **Latent UMAP of the Alexander autoencoder.** Each point is coloured by topological signature and scaled by $\log_{10}(\text{MSE} + 1)$. The largest red point (index 145,200, signature 14) dominates error in all three invariants.

hard knots correlate with known topological quantities such as the knot signature. Reconstruction error increases with signature, revealing a measurable limitation in the encoding capacity of the autoencoder.

Motivated by this relationship, we then train a multi-class classifier to predict knot signature directly from the latent representations of the Alexander, Jones, and HOMFLYPT polynomials. The model achieves high accuracy and recalls rare signature classes, reinforcing that these representations preserve significant topological information. Our findings suggest that latent representations of polynomial invariants not only provide accurate reconstructions but also reflect intrinsic complexity, opening a pathway toward using machine learning to probe and quantify topological structure in knot theory.

2 Results

2.1 Hard knots revealed by latent–space reconstruction error

After training an autoencoder on each polynomial representation (Alexander, Jones and HOMFLY–PT), we quantified reconstruction error over the entire data set.

To inspect how difficulty is organised, we projected the Alexander latent vectors of all 313 k knots with UMAP. **Figure 1** colours each knot by signature and scales point size with $\log_{10}(\text{MSE} + 1)$. Most knots form a dense, low-error manifold, yet a handful of outliers deviate markedly.

The distribution of errors is extremely skewed. Across the full data set the median MSE is below 10^{-2}, whereas the single worst knot (index 145,200, signature 14) reaches 3.6×10^{4} for the Alexander invariant and similarly dominates the Jones (6.2×10^{3}) and HOMFLY–PT (1.1×10^{4}) models. Removing this knot and retraining lowers the *mean* error by nearly two orders of magnitude for Alexander (from 1.18×10^{-1} to 1.1×10^{-3}) and reduces the *maximum* error from 3.6×10^{4} to 1.4; comparable improvements are observed for Jones and HOMFLY–PT (details in Table A1). Only eleven knots lie in the 99.9999^{th} percentile of error; all possess high signatures (10, 12 or 14) and cluster tightly in latent space (pastel markers in Fig. 1). Their coherent position suggests a common failure mode: highly-signed knots aggregate at the periphery of the learned manifold and are therefore reconstructed poorly.

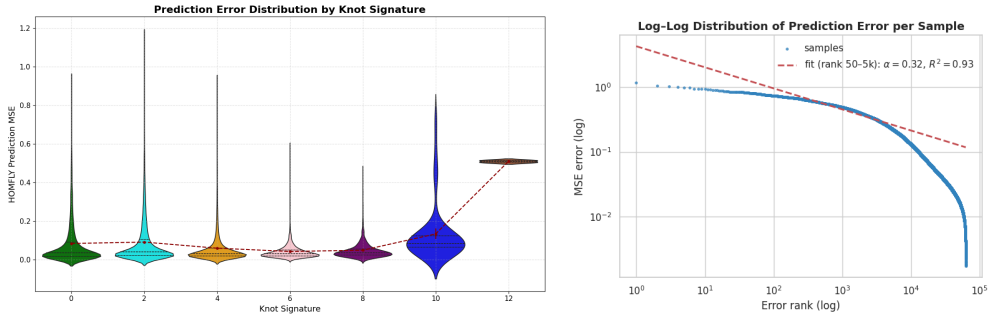
2.2 Predictability of the HOMFLYPT polynomial

Using the filtered data set (knot 145,200 removed), we trained six multilayer perceptrons (MLPs): three on raw polynomial coefficients and three on the corresponding autoencoder embeddings. On raw inputs the Alexander and Jones models achieve comparable performance (MSE = 0.168 and 0.178; $R^2 = 0.63$ and 0.64, respectively), while concatenating both polynomials gives a modest gain (MSE = 0.160, $R^2 = 0.65$). Training on the latent vectors yields slightly lower error for the individual invariants (Alexander: 0.164, 0.64; Jones: 0.177, 0.64), but the improvement becomes striking when the two latent spaces are combined: the MLP's mean-squared error drops by more than half, from 0.160 to 0.078, and its explained variance rises from 0.65 to 0.91. The result indicates that the embeddings capture complementary structure present in the Alexander and Jones invariants yet only weakly accessible in the raw coefficient space.

The predictive performance of MLPs trained on latent representations was comparable to or slightly better than models trained on raw polynomial coefficients. Notably, when combining the embeddings of Alexander and Jones, the prediction error dropped substantially, suggesting that the latent space captures complementary information from both invariants.

2.2.1 Error complexity and topological correlation

Having shown that latent representations markedly reduce the mean squared error, we next asked *how* whether the remaining error is distributed across knots and whether it correlates with topological complexity. A rank-frequency plot (Figure **??**a) of the MSE per knot on log-log axes follows a power law with exponent $\alpha = 0.99$ and goodness-of-fit $R^2 = 0.88$. Thus, a vanishingly small minority of knots accounts for a disproportionate share of the total loss, mirroring the long-tailed behavior seen in natural-image and linguistic datasets. To localize the source of this tail, we grouped residuals by the topological signature of each knot. Figure **??**b shows violin plots of $\log_{10}$(MSE) for every even signature from 0 to 14. The median error grows monotonically with signature, and the interquartile range widens sharply beyond 8, indicating increasing instability of the predictor. Signatures 10, 12 and 14 dominate the upper 0.0001 % of the error distribution, consistent with the single hard knot



(a) **(b)**

Fig. 2: **Residual-error structure and power-law behaviour. (a)** Violin plots of $\log_{10}(\text{MSE})$ grouped by knot signature. Medians rise steadily with signature, and high-signature knots (10–12) exhibit broader tails. **(b)** Log–log scatter of ranked prediction errors with a power-law fit over the central range (rank 50–5 000). The fitted line yields $\alpha' \approx 0.99$ and $R^2 \approx 0.93$, indicating a markedly linear regime in log–log space, whereas the extreme head and tail deviate from strict power-law behaviour.

(index 145,200, signature 14) that distorted the autoencoder in Sec . 2.1. Together, these results confirm that topological complexity, captured here by signature, strongly limits predictive capacity even after latent compression.

2.3 Predicting Signature

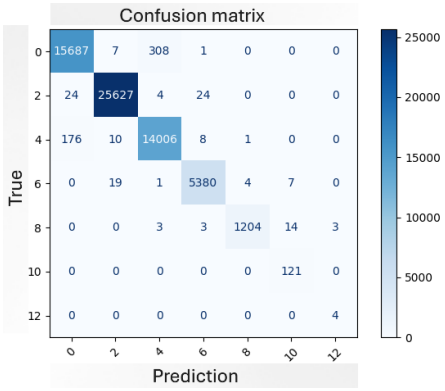
Having established that latent representations markedly reduce reconstruction error, we next asked whether they also encode sufficient information to infer a knot's topological signature. A seven-way MLP classifier trained on the combined Alexander + Jones embeddings attains **99 % overall accuracy** on the held-out validation set (Fig. 3b). Precision, recall and F_1 exceed 0.98 for signatures 0–8, which together comprise ∼99 % of the data. Performance deteriorates for the two rarest classes, signature 10 (121 examples) and signature 12 (4 examples), yet recall remains perfect, indicating that the minority knots are detected consistently even at the cost of a few false positives. Macro-averaged F_1 is 0.94, confirming that class imbalance has only a modest impact. The confusion matrix (Fig. 3a) reveals that the few misclassifications are confined to adjacent signatures (e.g. occasional confusion between 0 and 4, or 6 and 8). Taken together, these results demonstrate that the latent space captures signature-specific features with high fidelity, enabling accurate, end-to-end prediction of topological complexity.

3 Discussion

We set out to investigate whether neural latent spaces can capture - and even exploit - the structure of classical knot invariants. Three main insights emerge. First, autoencoders trained on Alexander and Jones polynomials compress more than 3.1×10^5

Signature	Precision	Recall	F_1
0	0.99	0.98	0.98
2	1.00	1.00	1.00
4	0.98	0.99	0.98
6	0.99	0.99	0.99
8	1.00	0.98	0.99
10	0.85	1.00	0.92
12	0.57	1.00	0.73
Acc.		0.99	
Macro	0.91	0.99	0.94
Wtd.	0.99	0.99	0.99

(a)



(b)

Fig. 3: **Prediction of knot signature from latent embeddings. (a)** Per-class precision, recall and F_1. **(b)** Confusion matrix; misclassifications are limited to adjacent signatures, and minority classes (10, 12) are retrieved with perfect recall.

knots into compact, information-rich embeddings. Regressors operating in this latent space predict the HOMFLY–PT polynomial better than equivalent models trained on raw coefficients, and the advantage nearly doubles when the two invariants are combined (MSE $= 7.8 \times 10^{-2}$ versus 1.6×10^{-1} on the same data; Sec. 2.1). Thus the latent geometry preserves complementary structure that is obscured when the polynomials are treated in isolation.

Second, the residual error distribution is heavy-tailed. After removing a single hard knot (index 145,200, signature 14), mean reconstruction error falls by almost two orders of magnitude and the power-law exponent of the tail settles at $\alpha \approx 1.02$ ($R^2 = 0.89$; Fig. 3a). This scale-free behaviour suggests that representation learning on knots inherits the same long-range complexity observed in other discrete physical systems (e.g. protein folds, spin glasses).

Third, error is strongly modulated by topological complexity. Median prediction error rises monotonically with signature (Fig. 3b), and every knot in the 99.9999^{th} percentile of error carries signature ≥ 10. In other words, the mapping

$$f_{\text{MLP+AE}} \colon (A,\, J) \longrightarrow \widetilde{\text{HOMFLYPT}}$$

is well approximated only in the low-signature regime where the two input invariants share sufficient mutual information. Extending this domain-either by regularising the latent manifold or by introducing signature-aware loss functions, remains an open challenge.

Latent embeddings are not merely predictive surrogates for HOMFLYPT; they also encode the signature itself. A seven-way classifier trained on the combined embeddings

attains **99 % overall accuracy**, with precision, recall and F_1 exceeding 0.98 for signatures 0–8 and macro-averaged $F_1 = 0.94$ (Fig. 3a). The scarce classes (10 and 12) are retrieved with perfect recall at the expense of several false positives, consistent with their extreme positions in latent space. Misclassifications are confined to adjacent signatures (Fig. 3b), indicating that most errors arise from genuine topological proximity rather than random noise. Taken together, these results show that neural embeddings not only compress classical invariants but also stratify knots by signature, enabling accurate end-to-end prediction of both quantitative (HOMFLYPT coefficients) and categorical (signature) properties. The remaining failure cases, high-signature knots that dominate the long-tail of error, highlight a frontier for future work: architectures and loss functions able to reconcile multiple, increasingly difficult invariants without discarding the rare but topologically rich examples that drive discovery.

4 Methods

4.1 Data

Our dataset consists of **313,231 knots**, each represented by three polynomial invariants: the **Alexander polynomial** (17 coefficients), the **Jones polynomial** (51 coefficients), and the **HOMFLYPT polynomial** (152 coefficients). In addition to these invariants, we also have access to knot-level topological descriptors such as **signature**, **number of crossings**, and **alternation class**. This uniform structure enables direct comparison and mapping between invariants across the same set of knots.

All polynomial vectors were standardized before training. Coefficient ranges vary significantly, motivating the use of autoencoders for normalization and dimensionality reduction

4.2 Neural Network baseline

To evaluate whether latent representations improve the learnability of the HOMFLYPT polynomial, we design a comparative study based on two neural network models:

- **Raw Polynomials (Baseline)**: We train a multilayer perceptron (MLP) to predict the HOMFLYPT polynomial directly from the concatenated coefficient vectors of the Alexander and Jones polynomials. This serves as a baseline to evaluate the predictive capacity of the original invariants in their raw form.
- **Latent Representations*:** We train autoencoders on the Alexander, Jones, and HOMFLYPT polynomials independently to obtain compressed embeddings. A second MLP is then trained to map the concatenated embeddings of Alexander and Jones to the latent space of HOMFLYPT. This setup tests whether the learned representations capture a more structured or learnable mapping.

The goal is to learn a meaningful latent space for each invariant that retains the essential topological structure while reducing dimensionality.

Each model is trained under the same hyperparameter settings and evaluated using mean squared error (MSE) on a held-out test set. This comparison allows us to isolate the value added by the autoencoder representations in terms of predictive accuracy and topological expressiveness.

4.3 Model Architecture and Training

To evaluate the mapping from knot invariants to the HOMFLYPT polynomial, we use a multilayer perceptron (MLP) as a function approximator. The model consists of four fully connected layers with ReLU activation functions between them, progressively expanding and contracting the hidden dimension:

- Input layer $\rightarrow$ 64 units $\rightarrow$ ReLU
- 64 $\rightarrow$ 128 $\rightarrow$ ReLU
- 64 $\rightarrow$ 128 $\rightarrow$ ReLU
- 128 $\rightarrow$ 64 $\rightarrow$ ReLU
- 64 $\rightarrow$ Output layer (size of HOMFLY vector)

This architecture is flexible enough to capture nonlinear interactions while maintaining efficiency.

We use the mean squared error (MSE) loss between the predicted output and the true HOMFLYPT vector. The model is trained using the Adam optimizer with a learning rate of $1x10^{-3}$ for 100 epochs. During training, we monitor both training and validation loss using mini-batch gradient descent with fixed batch size (via PyTorch's DataLoader). Validation is performed every 10 epochs to ensure generalization and early signal of overfitting.

This setup is used for both the raw input experiment (Alexander and Jones polynomials concatenated directly) and the latent representation experiment (embeddings learned through autoencoders).

4.3.1 Latent Representation via Autoencoders

To obtain compressed representations of knot invariants, we trained three independent autoencoders — one for each polynomial: Alexander, Jones, and HOMFLYPT.

Each autoencoder consists of a fully connected encoder and decoder with the following architecture:

- **Encoder:** A stack of three dense layers with ReLU activations and batch normalization, followed by a bottleneck (latent) layer using a tanh activation.
- **Decoder:** A symmetric structure that mirrors the encoder, with a final linear output layer.

The architecture parameters vary by polynomial due to their input dimensionality:

Polynomial	Input Size	Latent Size	Layer Configuration (Encoder & Decoder)
Jones	51	25	[128, 64, 64]
Alexander	17	9	[64, 32, 32]
HOMFLY	152	75	[256, 128, 128]

Each model was trained using the mean squared error (MSE) loss function and the Adam optimizer with a learning rate of $5x10^{-4}$. Training was conducted for up to 30 epochs with a batch size of 128, and early stopping was applied with a patience of 5 epochs to prevent overfitting.

To assess reconstruction quality, we compute the per-sample MSE between the original polynomial vectors and their autoencoder reconstructions.These latent embeddings are then used as inputs for the subsequent mapping model , enabling a comparison between raw and learned representations.

4.3.2 Predicting Signature

We aim to predict the signature of a knot from the concatenated embeddings of the Alexander, Jones, and HOMFLYPT polynomial invariants. To this end, we train a Multi-Layer Perceptron (MLP) as a multi-class classification model.

However, before detailing the model architecture, it is important to highlight a significant issue: the dataset is highly imbalanced in terms of signature distribution, as shown below:

Signature	Count	%
0	80,016	25.5
2	128,393	40.9
4	71,003	22.9
6	27,058	8.6
8	6,135	1.9
10	605	0.19
12	20	0.01

To mitigate this imbalance, we applied a class weighting strategy during training: we assigned higher weights to underrepresented classes and lower weights to frequent ones, so that the model is not biased toward predicting only the most common signatures.

The model architecture is the following:

- **Input layer:** Receives the concatenated embeddings (dimension = 25 + 9 + 32 = 66).
- **Hidden layers:**
  - Dense layer with 128 units and ReLU activation
  - Dropout (rate = 0.3) for regularization
  - Dense layer with 64 units and ReLU activation
- **Output layer:**

  - Final dense layer with 7 units (one per signature class: 0, 2, 4, 6, 8, 10, 12)
  - Softmax activation for multi-class classification

- **Loss function:** Categorical Cross-Entropy with **class weighting** to address label imbalance.
- **Optimizer:** Adam with learning rate 5×10^{-4}.
- **Batch size:** 128.
- **Epochs:** Up to 50, with **early stopping** (patience = 5).
- **Metrics:** Accuracy and per-class performance.

Supplementary information. If your article has accompanying supplementary file/s please state so here.

Authors reporting data from electrophoretic gels and blots should supply the full unprocessed scans for key as part of their Supplementary information. This may be requested by the editorial team/s if it is missing.

Please refer to Journal-level guidance for any specific requirements.

Acknowledgements. Acknowledgements are not compulsory. Where included they should be brief. Grant or contribution numbers may be acknowledged.

Please refer to Journal-level guidance for any specific requirements.

Declarations

Some journals require declarations to be submitted in a standardised format. Please check the Instructions for Authors of the journal to which you are submitting to see if you need to complete this section. If yes, your manuscript must contain the following sections under the heading 'Declarations':

- Funding
- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use)
- Ethics approval and consent to participate
- Consent for publication
- Data availability
- Materials availability
- Code availability
- Author contribution

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Editorial Policies for:

Springer journals and proceedings: https://www.springer.com/gp/editorial-policies

Nature Portfolio journals: https://www.nature.com/nature-research/editorial-policies

Scientific Reports: https://www.nature.com/srep/journal-policies/editorial-policies

BMC journals: https://www.biomedcentral.com/getpublished/editorial-policies

Appendix A Results

Polynomial	**Mean MSE**	**Δ Mean**	**Max MSE**	**Δ Max**	**Min MSE**
Jones	0.02279	*-0.01262*	181.87	*-5972.29*	$5\text{x}10^{-5}$
Alexander	0.00110	*-0.11719*	1.3869	*-36375.47*	$2\text{x}10^{-5}$
HOMFLY	0.01227	*-0.03235*	160.71	*-10365.06*	$2.5\text{x}10^{-4}$

Table A1: Comparison of reconstruction error statistics for the autoencoders trained with and without the hard knot (index 145,200). The table reports the mean, maximum, and minimum mean squared error (MSE) per polynomial type. The Δ columns indicate the change in error after removing the hard knot from the training set.