(K)NOT MACHINE LEARNING

Vishnu Jejjala
University of the Witwatersrand

15 December 2020
string_data 2020

Collaborators

Jessica Craven

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

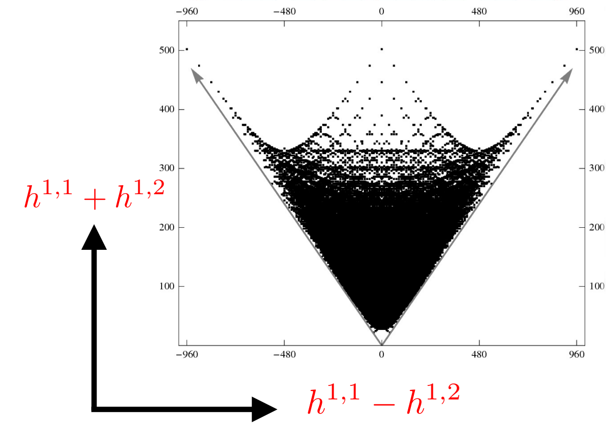
Arjun Kar

UBC

"Disentangling a Deep Learned Volume Formula"

arXiv:2012.03955

Mathematical Phenomenology

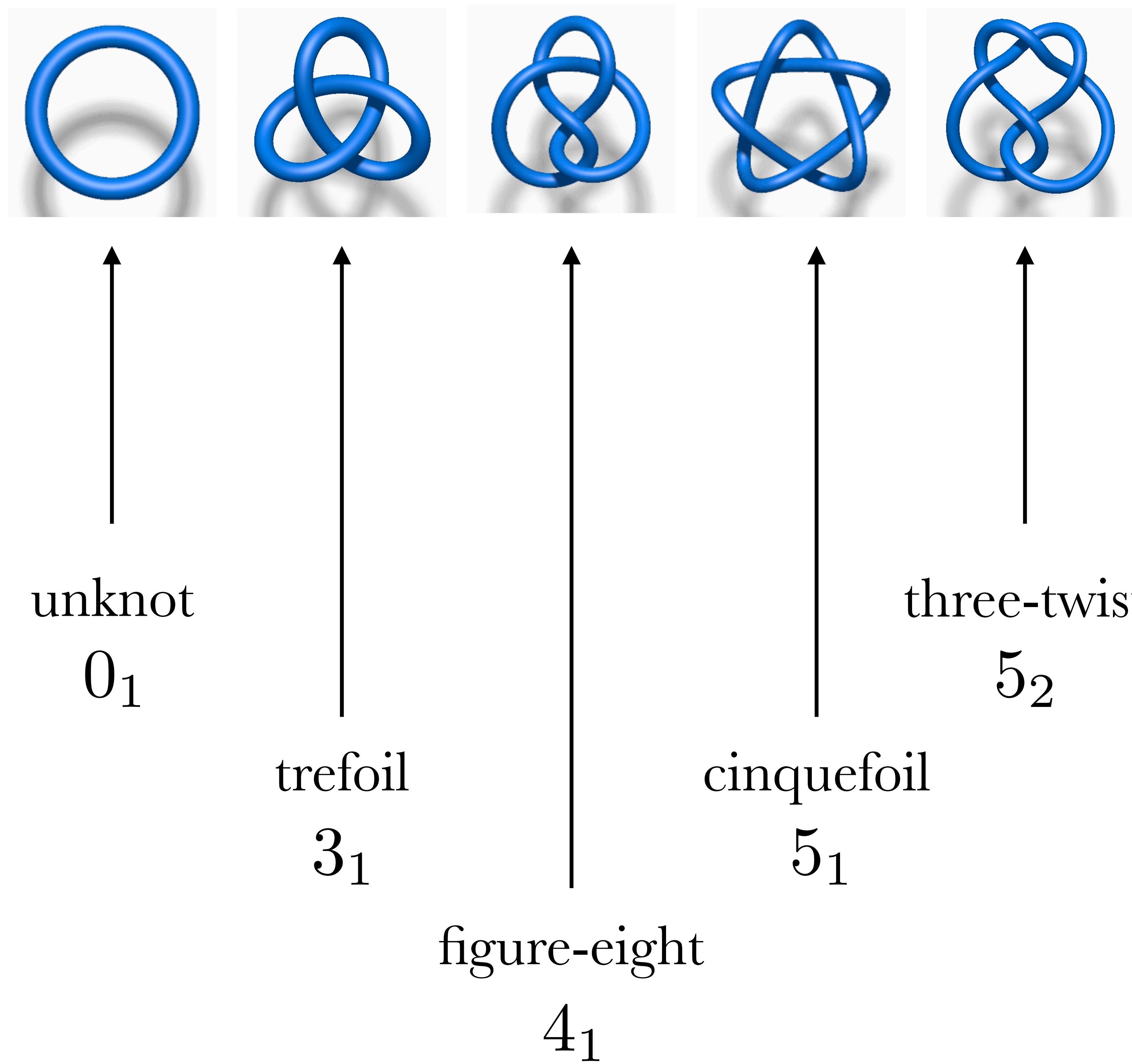


- Note patterns then look for an explanation
  - — Mirror symmetry is prototype example
  - — Knot theory provides another case study

- Use machine learning to train a computer to calculate in **`hep-th`**, **`math`**
  - — Black box gives **probably approximately correct** answers
  - — So far, we have mainly used ML to identify associations
  - — Want to bridge this success to new analytic results and methods

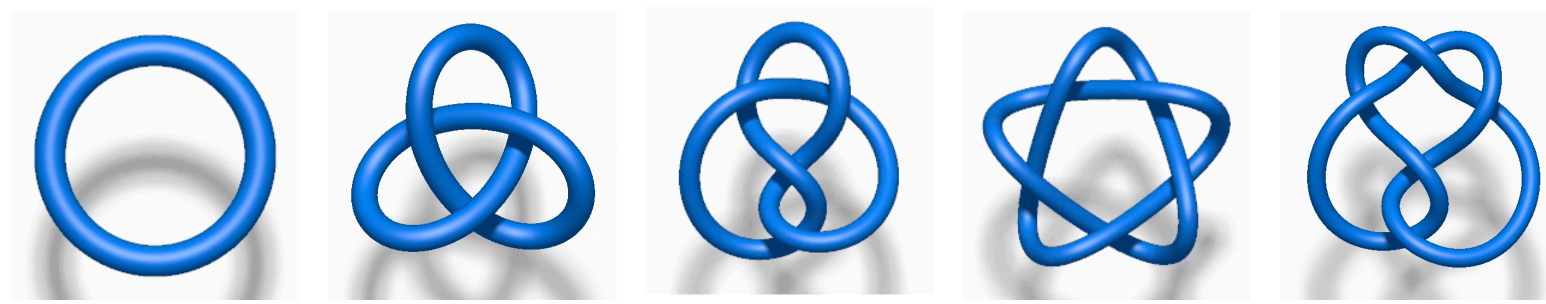
Dramatis Personae

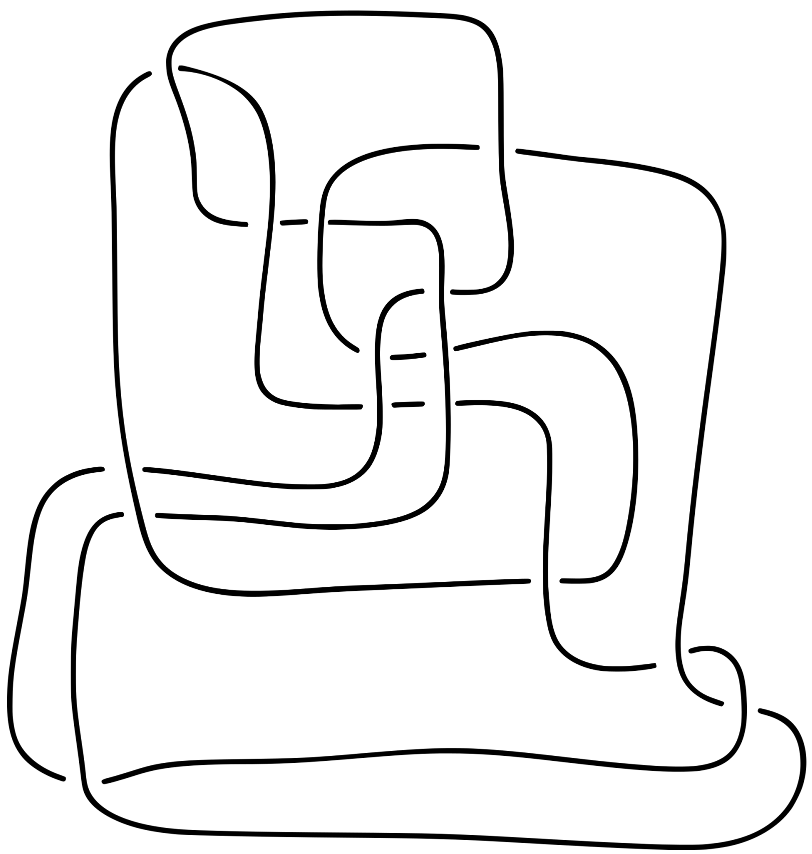
Knot: $S^1 \subset S^3$; e.g.,



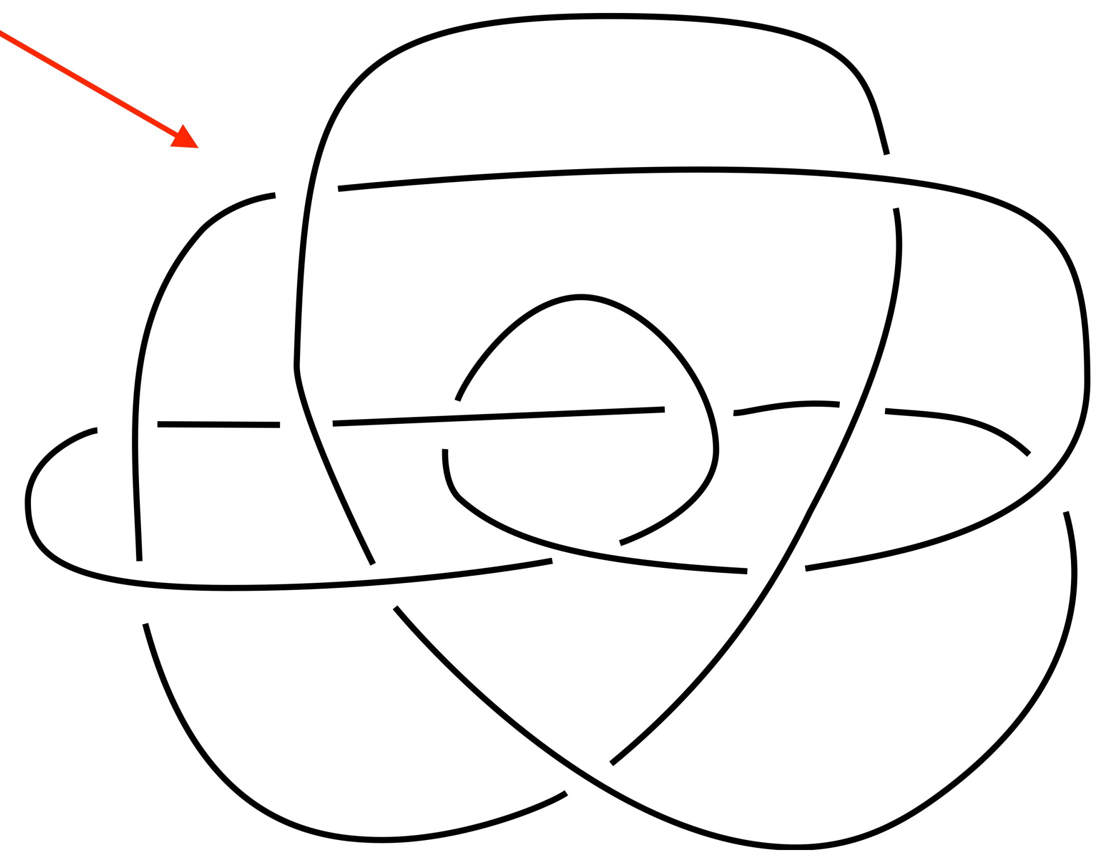
Dramatis Personae

Knot: $S^1 \subset S^3$; e.g.,





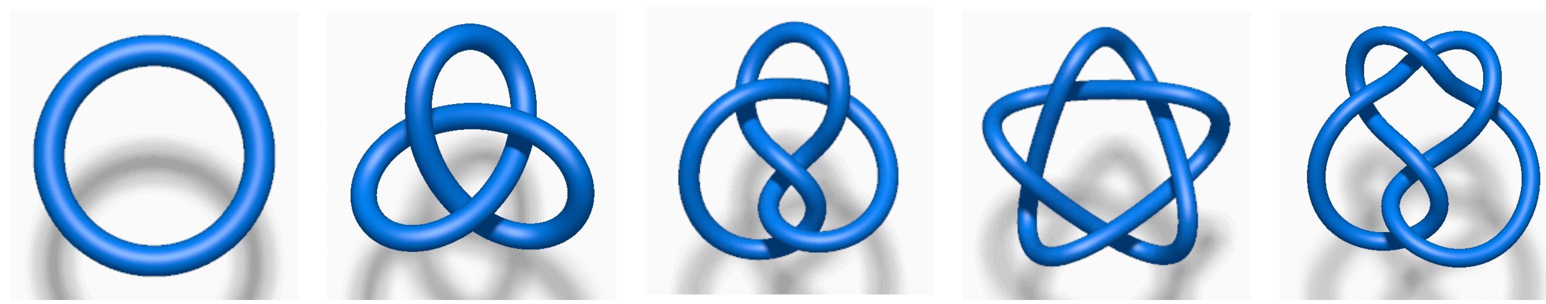
Thistlethwaite unknot



Ochiai unknot

Dramatis Personae

Knot: $S^1 \subset S^3$; e.g.,



Jones polynomial: $J(K;q) = (-q^{\frac{3}{4}})^{w(K)} \frac{\langle K \rangle}{\langle \bigcirc \rangle}$

$\langle \times \rangle = q^{\frac{1}{4}} \langle \asymp \rangle + \frac{1}{q^{\frac{1}{4}}} \langle)(\rangle$

$w(K) = \text{overhand} - \text{underhand}$

Jones (1985)

topological invariant: independent of how the knot is drawn

Question: how to calculate these?

Answer: quantum field theory!

Topological Invariants

- On a manifold $\mathcal{M}$ with metric $g_{\mu\nu}$, a topological invariant enjoys:

$$\frac{\delta}{\delta g_{\mu\nu}} \langle \mathcal{O}_1 \dots \mathcal{O}_n \rangle = 0$$

- In Chern–Simons theory, the operators are Wilson loops

$$U_R(\gamma) = \mathrm{tr}_R \, \mathcal{P} \exp \left(i \oint_\gamma A \right)$$

- The colored Jones polynomial is a knot invariant:

$$J_n(K; q = e^{2\pi i/(k+2)}) = \frac{\int_{\mathcal{U}} [DA] \; U_n(K) e^{iS_{\mathrm{CS}}(A)}}{\int_{\mathcal{U}} [DA] \; U_n(0_1) e^{iS_{\mathrm{CS}}(A)}} = \frac{\langle U_n(K) \rangle}{\langle U_n(0_1) \rangle}$$

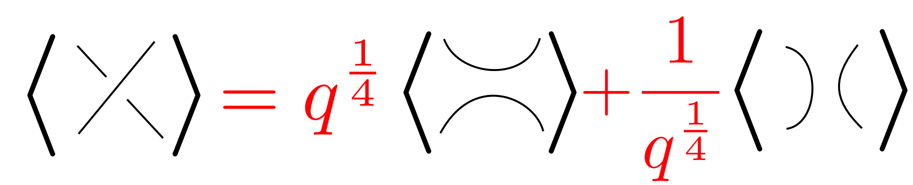
$$S_{\mathrm{CS}}(A) = \frac{k}{4\pi} \int_{\mathcal{M}} \mathrm{tr} \left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right) , \quad Z(\mathcal{M}) = \int_{\mathcal{U}} [DA] \; e^{iS_{\mathrm{CS}}(A)}$$

Witten (1989)

Dramatis Personae

Knot: $S^1 \subset S^3$; e.g.,

Jones polynomial: $J(K;q) = (-q^{\frac{3}{4}})^{w(K)} \dfrac{\langle K \rangle}{\langle \bigcirc \rangle}$



$w(K) = \text{overhand} - \text{underhand}$

vev of Wilson loop operator along K in

$\square$ for $SU(2)$ Chern–Simons on S^3

Jones (1985)
Witten (1989)

$$J_2(4_1;q) = q^{-2} - q^{-1} + 1 - q + q^2 \ , \quad q = e^{\frac{2\pi i}{k+2}}$$

Hyperbolic volume: volume of $S^3 \backslash K$ is another knot invariant

computed from tetrahedral decomposition of knot complement

Thurston (1978)

Dramatis Personae

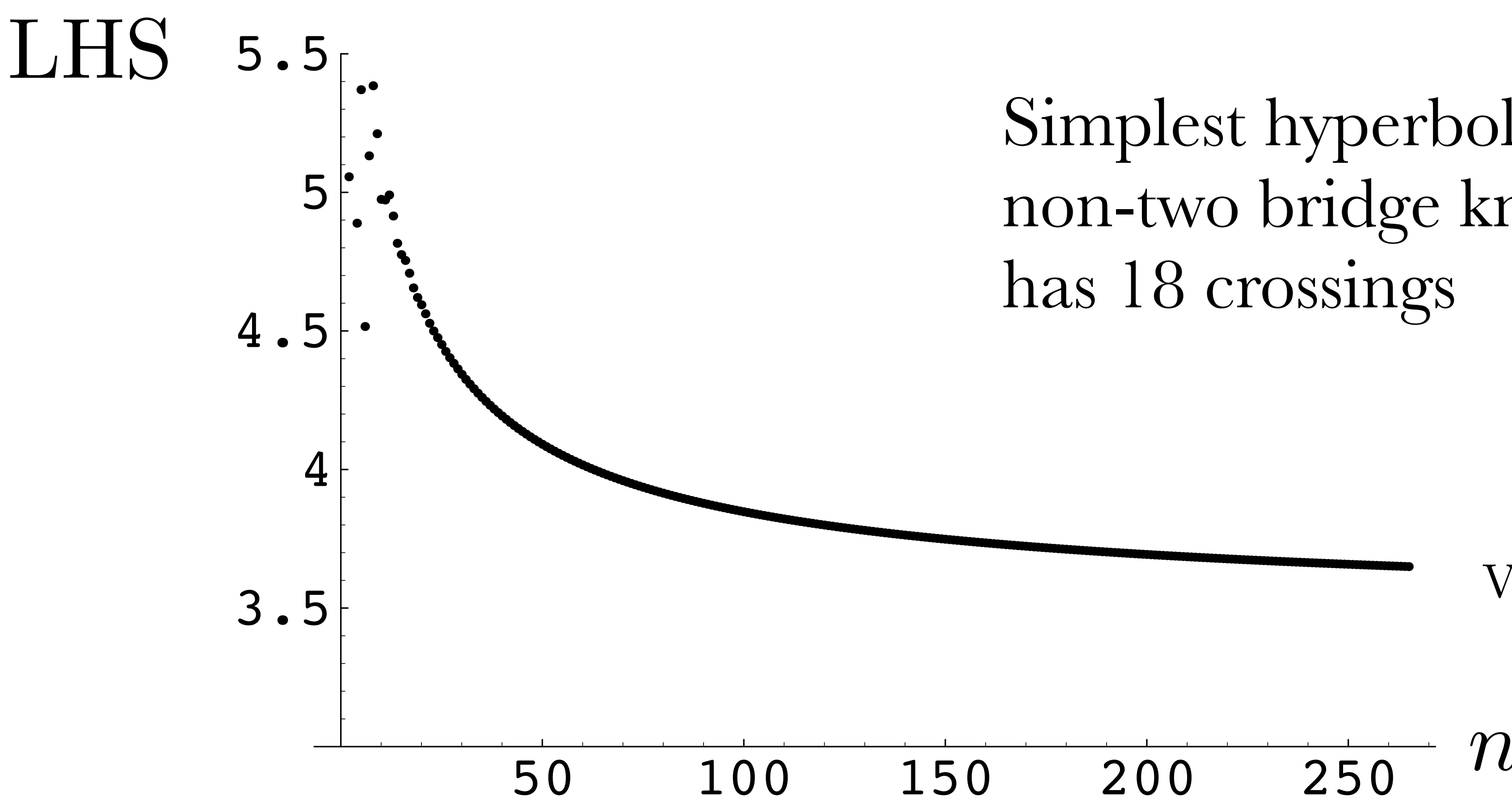
<u>Volume conjecture:</u>

$$\lim_{n\to\infty} \frac{2\pi \log |J_n(K;\omega_n)|}{n} = \mathrm{Vol}(S^3 \setminus K)$$

Kashaev (1997)
Murakami x 2 (2001)
Gukov (2005)

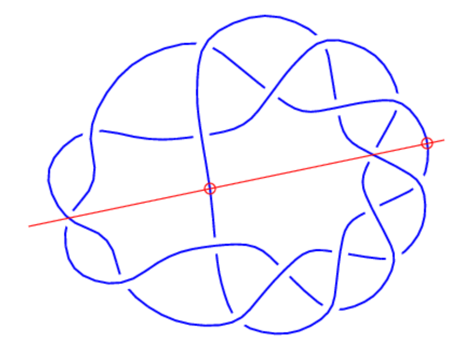
$$\omega_n = e^{\frac{2\pi i}{n}}$$

In fact, we take $n, k \to \infty$



$\mathrm{Vol}(K_0) = 3.474247\ldots$

Simplest hyperbolic non-two bridge knot, has 18 crossings



$\sigma_1^2(\sigma_1\sigma_2)^8$

Behavior is not monotonic!

Garoufalidis, Lan (2004)

Dramatis Personae

Volume conjecture:

$$\lim_{n\to\infty} \frac{2\pi \log |J_n(K;\omega_n)|}{n} = \mathrm{Vol}(S^3 \setminus K)$$

$$\omega_n = e^{\frac{2\pi i}{n}}$$

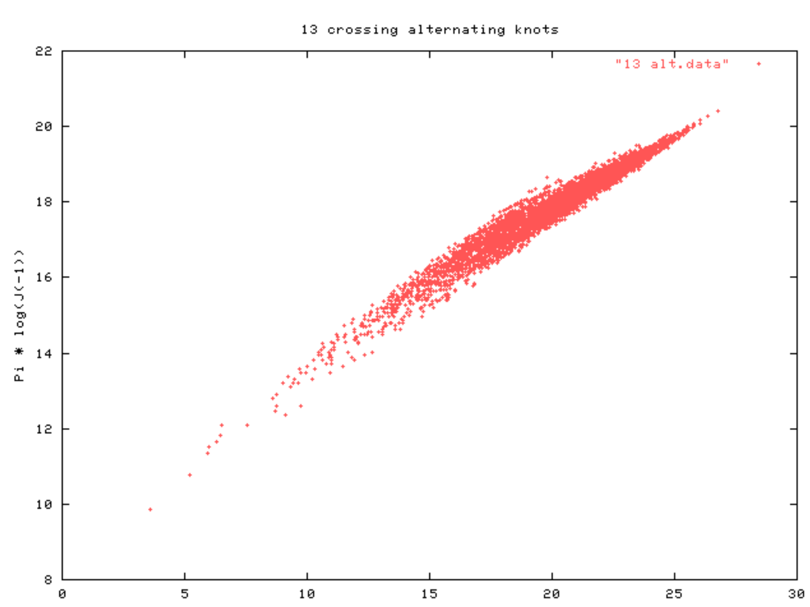
Kashaev (1997)
Murakami x 2 (2001)
Gukov (2005)

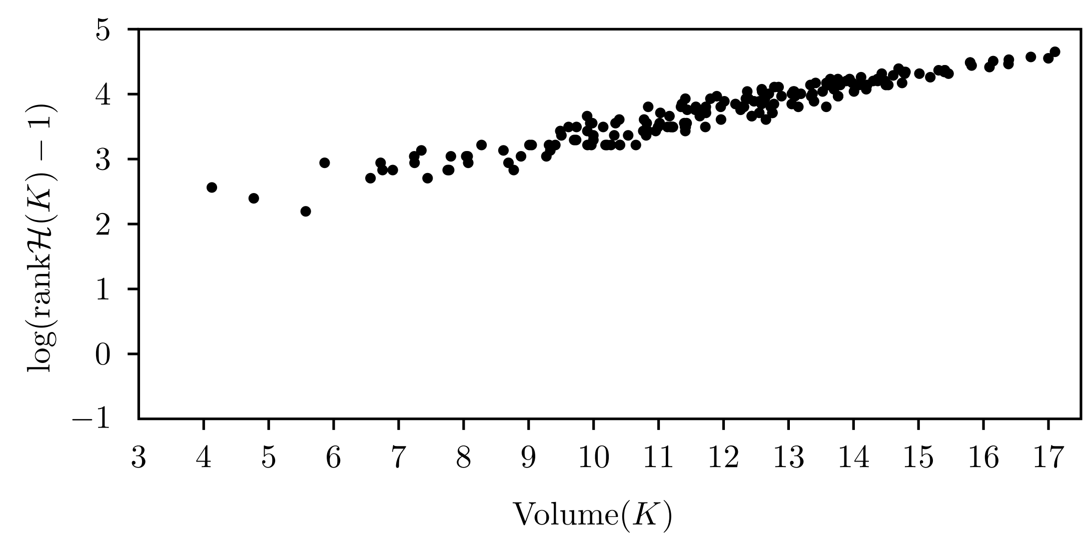
Khovanov homology: a homology theory $\mathcal{H}_K$ whose graded Euler characteristic is $J_2(K;q)$; explains why coefficients are integers

Khovanov (2000)
Bar-Natan (2002)

$$\log|J_2(K;-1)| \ , \ \ \log(\mathrm{rank}(\mathcal{H}_K)-1) \propto \mathrm{Vol}(S^3 \setminus K)$$

Dunfield (2000)
Khovanov (2002)





Dramatis Personae

Volume conjecture: $$\lim_{n\to\infty} \frac{2\pi \log |J_n(K;\omega_n)|}{n} = \mathrm{Vol}(S^3 \setminus K)$$

$$\omega_n = e^{\frac{2\pi i}{n}}$$

Kashaev (1997)
Murakami x 2 (2001)
Gukov (2005)

Khovanov homology: a homology theory $\mathcal{H}_K$ whose graded Euler characteristic is $J_2(K;q)$; explains why coefficients are integers

Khovanov (2000)
Bar-Natan (2002)

Volume-ish theorem: $J_2(K;q) = a_n q^n + \ldots + a_m q^m$

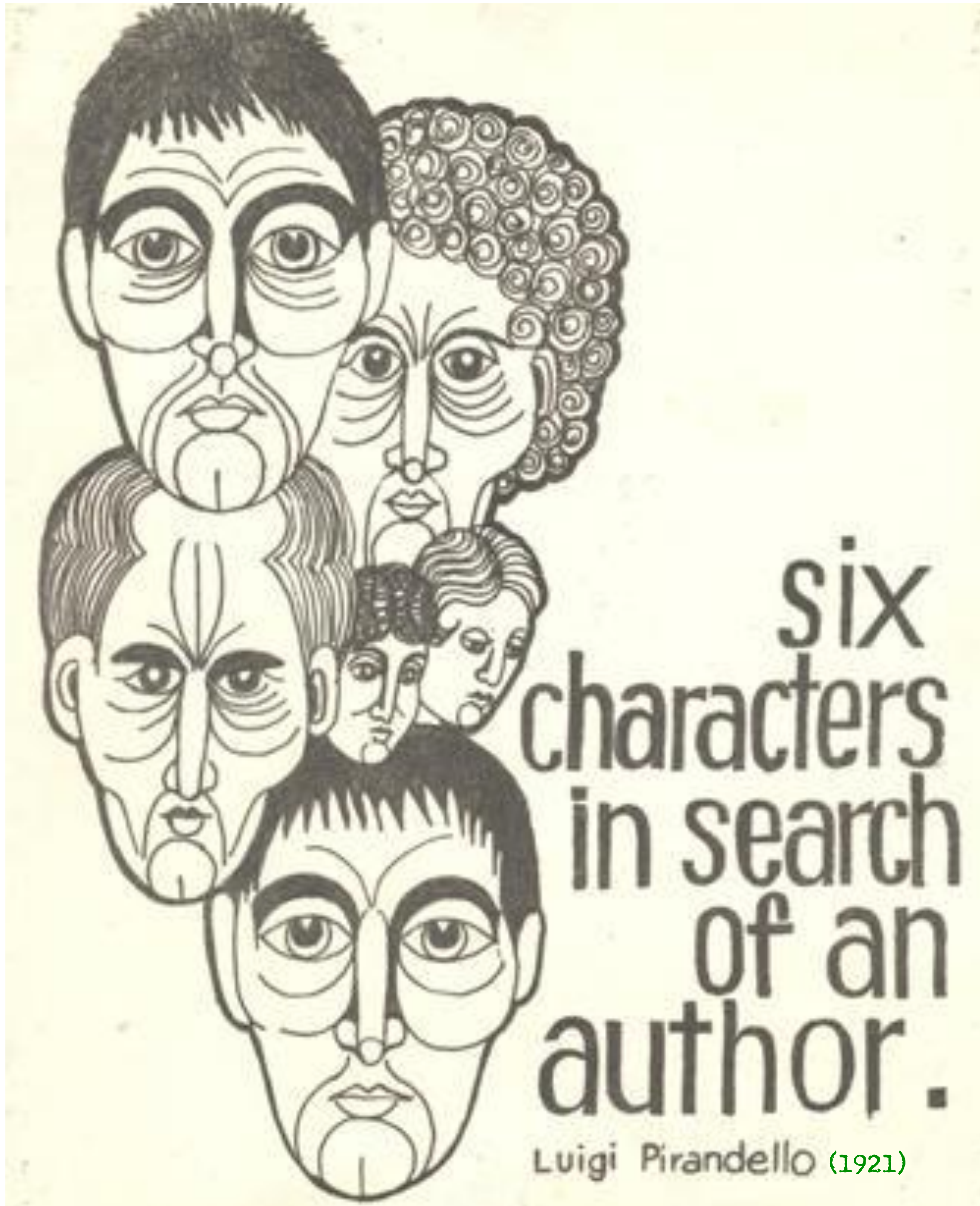
$$2v_0(\max(|a_{m-1}|, |a_{n+1}|) - 1) \leq \mathrm{Vol}(K) \leq 10v_0(|a_{m-1}| + |a_{n+1}| - 1)$$

for alternating knots; $v_0 = -\int_0^{\pi/3} dt\ \log(2\sin\frac{t}{2})$, volume of regular ideal tetrahedron

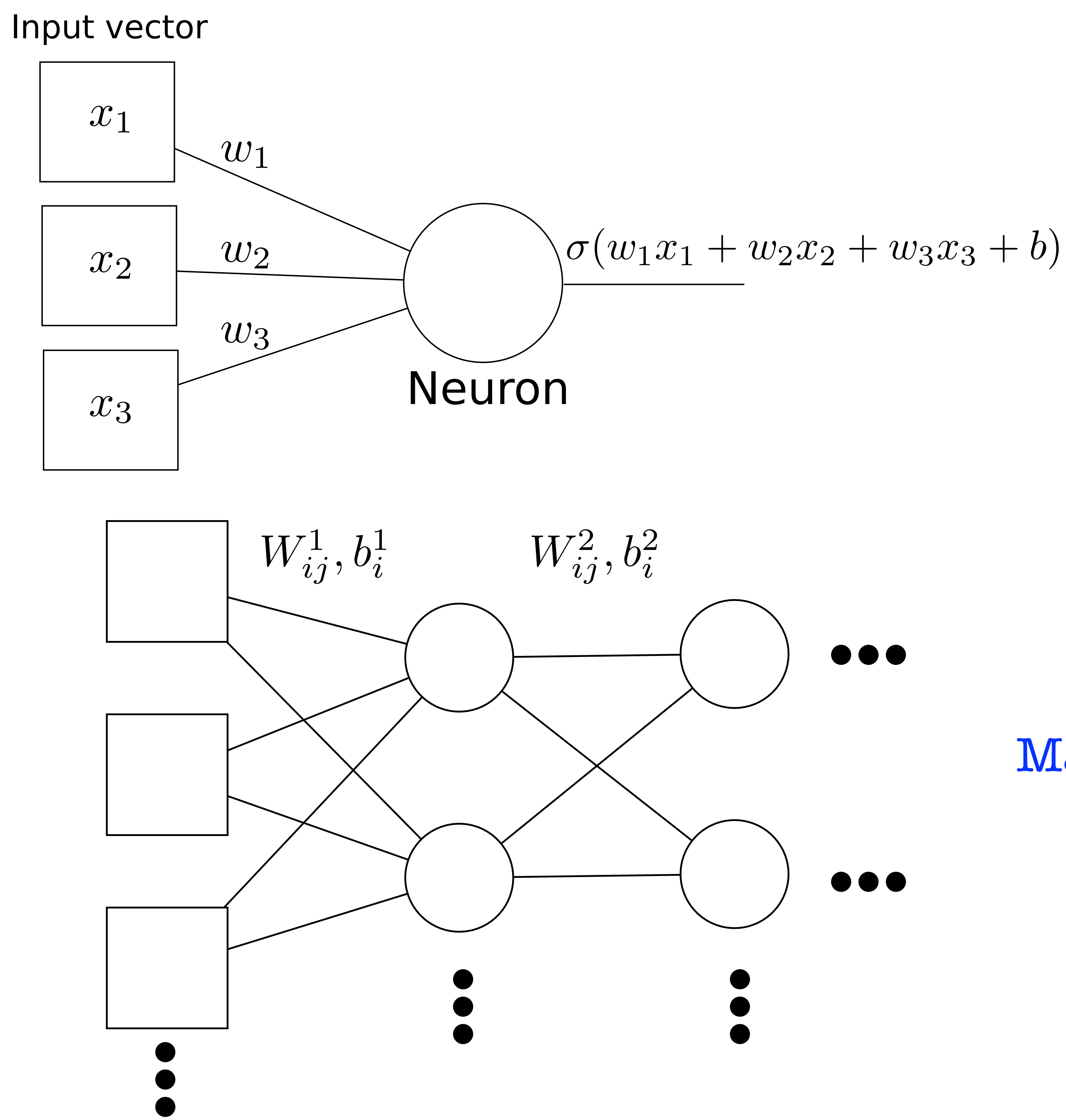
≈ 1.0149416

Dasbach, Lin (2007)

Dramatis Personae



Feedforward Neural Networks

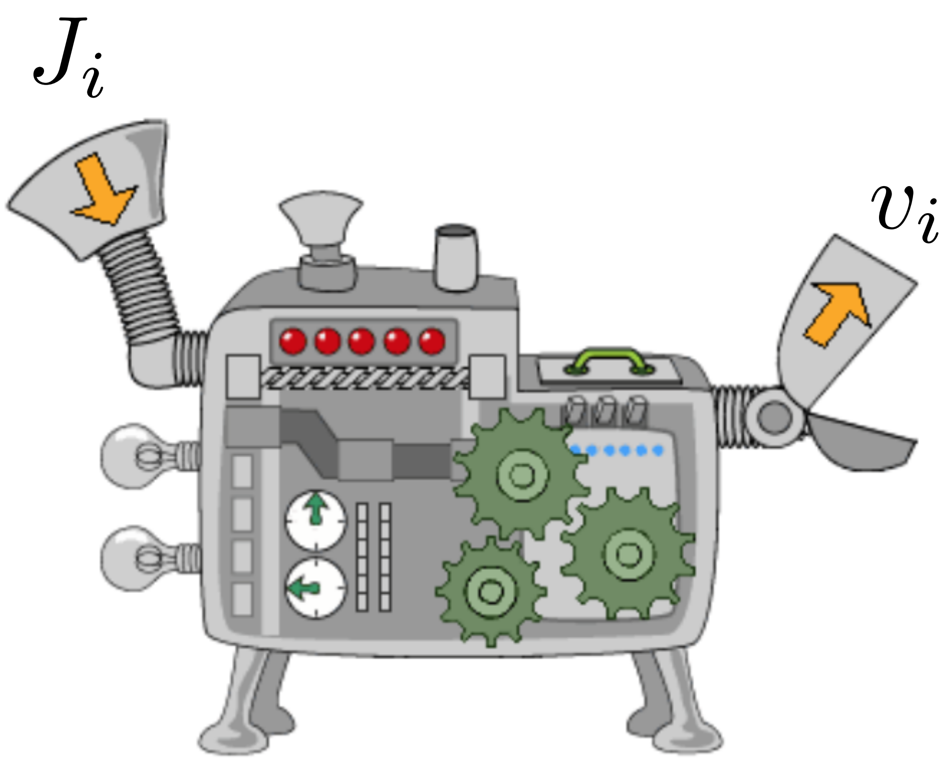


Rosenblatt (1957)

Mathematica 10+

Schematic representation of feedforward neural network. The top figure denotes the perceptron (a single neuron), the bottom, the multiple neurons and multiple layers of the neural network.

Neural Network

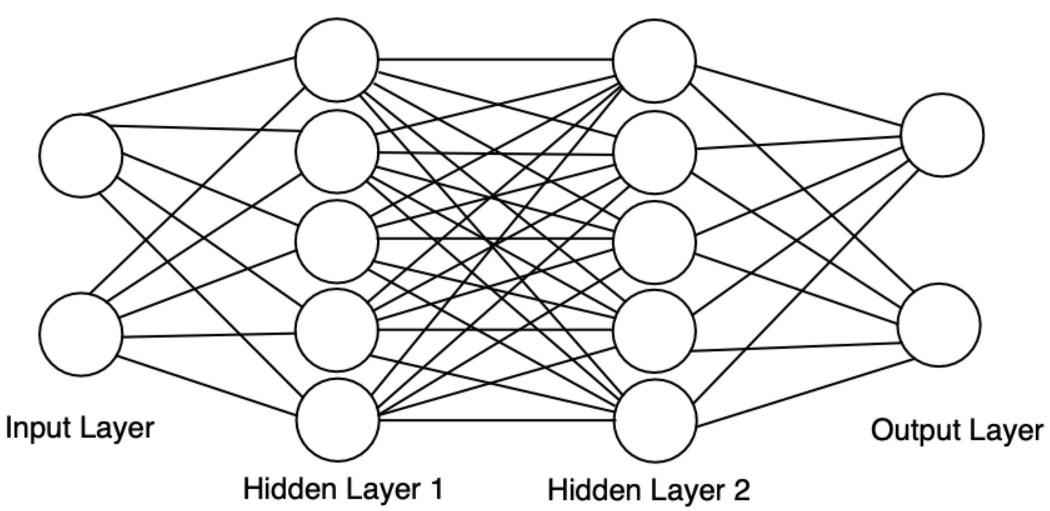


$$\{J_1, \ldots, J_n\} \longrightarrow \{v_1, \ldots, v_n\}$$

$$J_i \in T$$

$$\{J'_1, \ldots, J'_m\} \longrightarrow ???$$

$$J'_i \in T^c$$



$\vec{J}$ $\quad 100 \times 18 \quad 100 \times 100 \quad \sum_{a=1}^{100}$

12000 hyperparameters

Jones polynomials are represented as 18-vectors

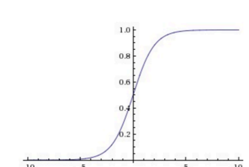
$$\vec{J}_K = (\text{min}, \text{max}, \text{coeffs}, 0, \ldots, 0)$$

Two layer neural network in **Mathematica**

$$f_\theta(\vec{J}_K) = \sum_a \sigma\left(W_\theta^2 \cdot \sigma(W_\theta^1 \cdot \vec{J}_K + \vec{b}_\theta^1) + \vec{b}_\theta^2\right)^a$$

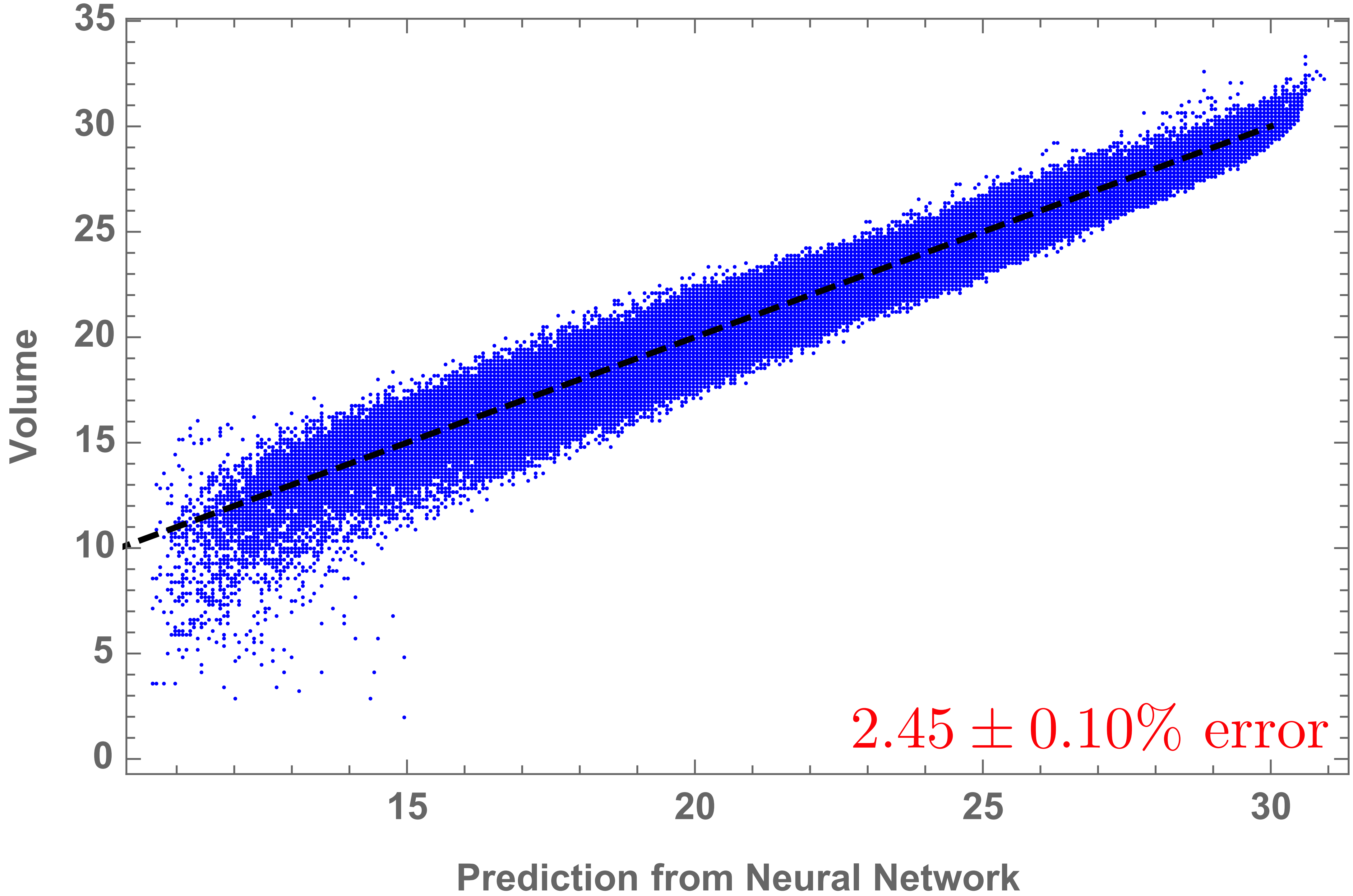
Logistic sigmoids for the hidden layers

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$



VJ, Kar, Parrikar (2019)

Neural Network



trained on 10% of the $313,209$ knots up to 15 crossings

VJ, Kar, Parrikar (2019)

Result

$$v_i = f(J_i) + \text{small corrections}$$

- J_i does not uniquely identify a knot

  e.g., 4_1 and K11n19 have same Jones polynomial, different volumes

- $174,619$ unique Jones polynomials

  2.83% average spread in volumes for a Jones polynomial
  intrinsic mitigation against overfitting

- Same applies to 1,701,913 hyperbolic knots up to 16 crossings
  (database compiled from **Knot Atlas** and SnapPy)

VJ, Kar, Parrikar (2019)

Entr'acte

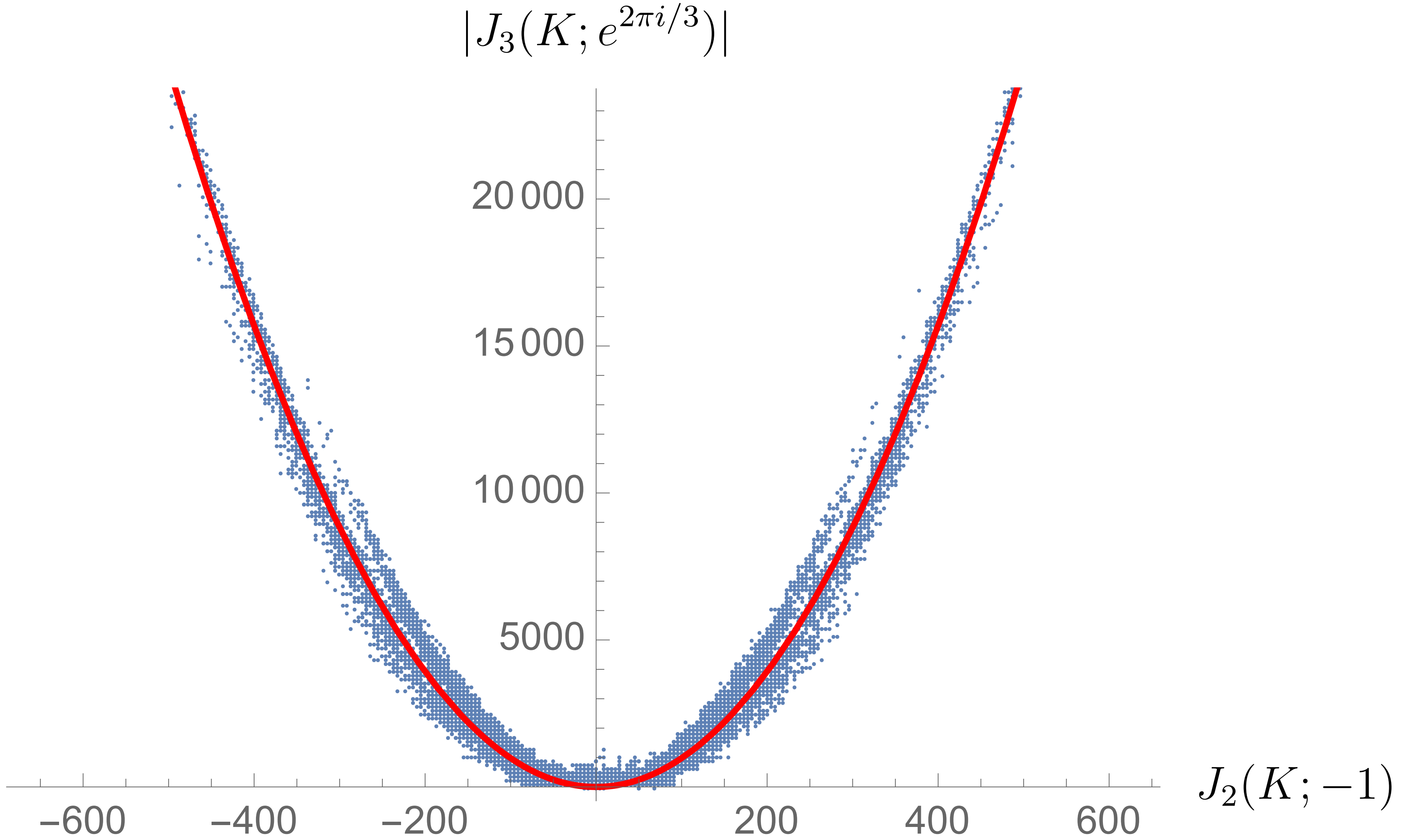
$$v_i = f(J_i) + \text{small corrections}$$

We seek to reverse engineer the neural network
to obtain an analytic expression for
the volume as a function of the Jones polynomial

To interpret the formula, we use machinery of
analytically continued Chern–Simons theory

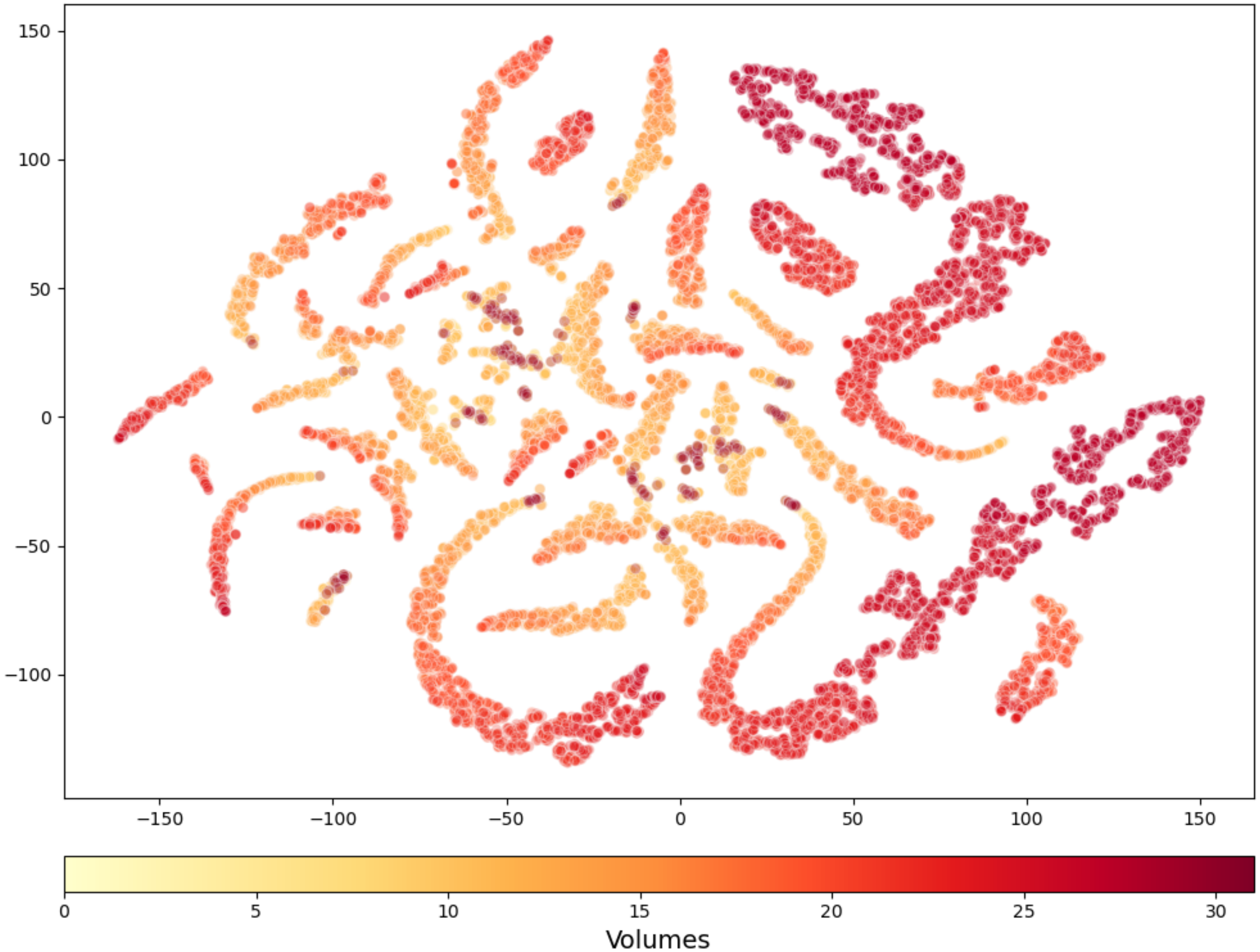
Towards the Volume Conjecture

- The volume conjecture: $$\lim_{n\to\infty} \frac{2\pi \log |J_n(K;\omega_n)|}{n} = \mathrm{Vol}(S^3 \setminus K)$$

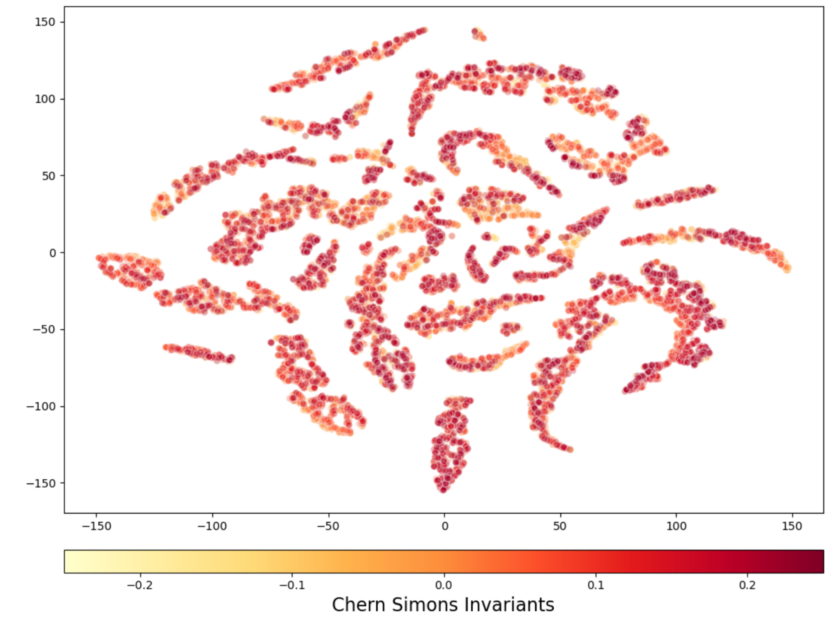


- 11,921 colored Jones polynomials at $n = 3$

t-SNE



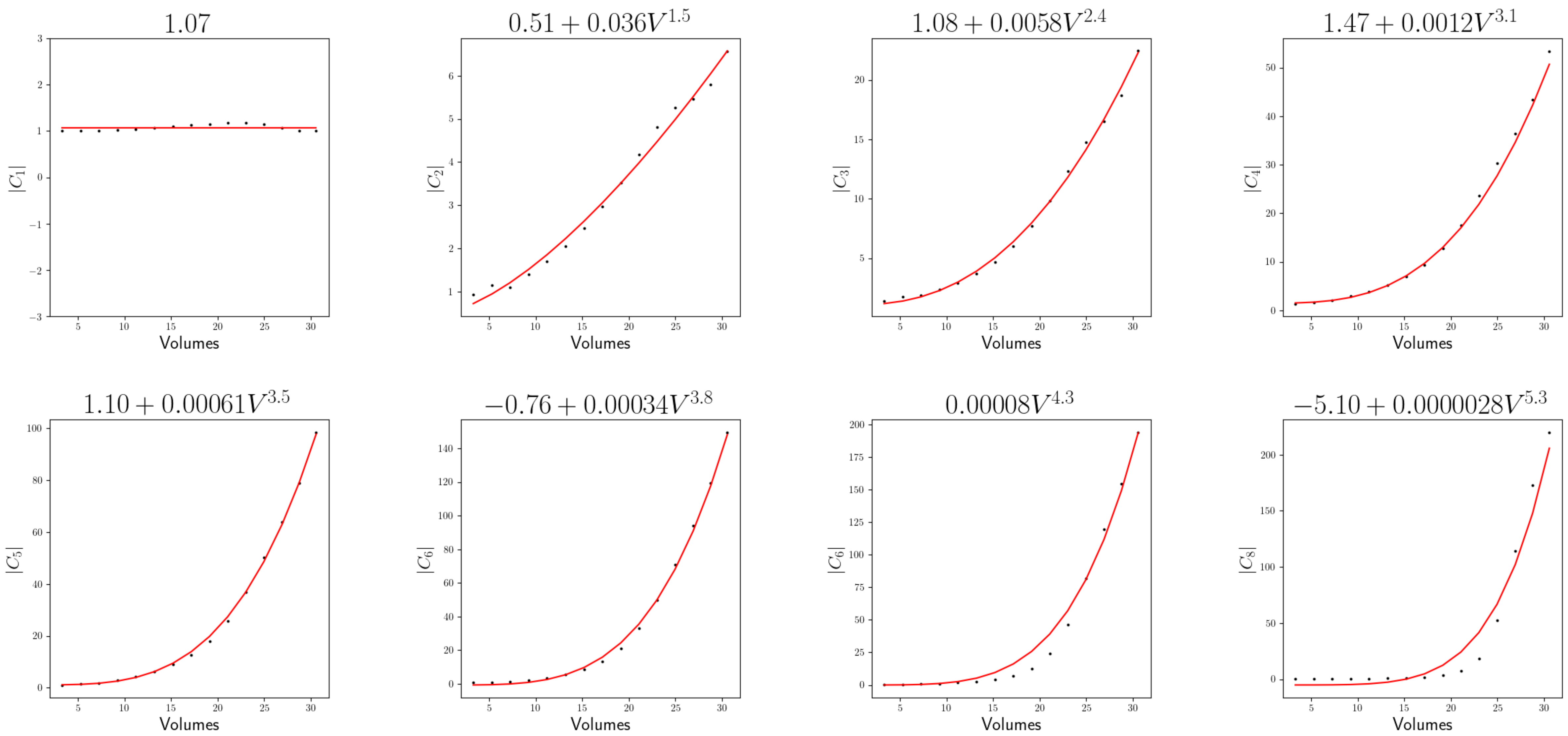
$$\lim_{n\to\infty} \frac{2\pi \log J_n(K;\omega_n)}{n}$$
$$= \mathrm{Vol}(S^3 \setminus K) + 2\pi^2 i\mathrm{CS}(K)$$



Volume is learnable from coefficients

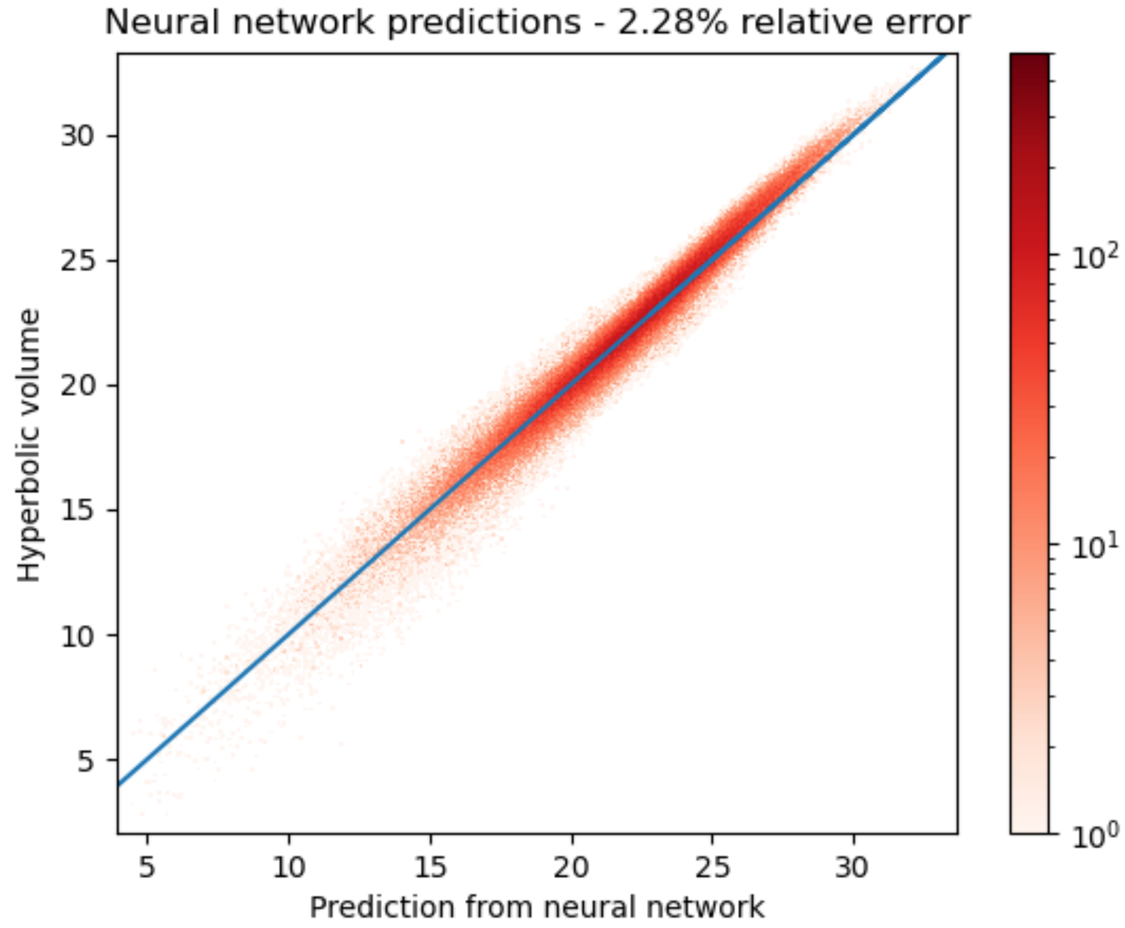
CS invariant probably is not

Coefficients



No Degrees Needed

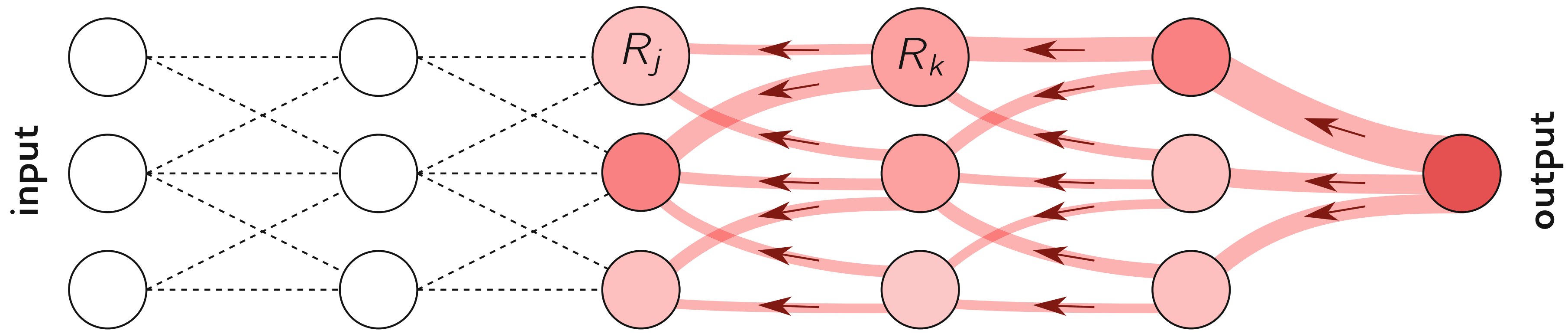
- Suppose we drop the degrees and provide only the coefficients; Jones polynomial is no longer recoverable from the input vector
- Results are unchanged!



N.B.: we have switched to `Python 3` using `GPU-Tensorflow` with `Keras` wrapper
two hidden layers, 100 neurons/layer, ReLu activation, mean squared loss, `Adam` optimizer

Layer-wise Relevance Propagation

- To determine which inputs carry the most weight, propagate backward starting from output layer employing a conservation property



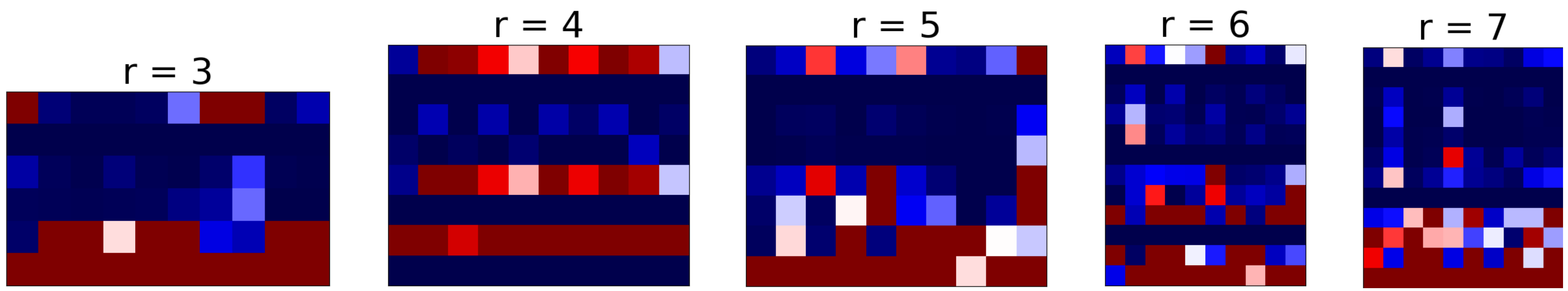
Montavon et al. (2019)

- Compute relevance score for a neuron using activations, weights, and biases

$$R_j^{m-1} = \sum_k \frac{a_j^{m-1} W_{jk}^m + b_k^m}{\sum_l a_l^{m-1} W_{lk}^m + b_k^m} R_k^m \quad , \quad \sum_k R_k^m = 1$$

j^{th} neuron in layer $m-1$

Layer-wise Relevance Propagation



- Each column is a single input corresponding to evaluations of the Jones polynomial at phases $e^{\frac{2\pi i p}{r+2}}$, $0 \leq 2p \leq r+2$, $p \in \mathbb{Z}$

- Ten different knots

- We show the relevances (red is most relevant) and notice that the same input features light up

Relevant Phases

r	Error	Relevant roots	Fractional levels	Error (relevant roots)
3	3.48%	$e^{4\pi i/5}$	$\frac{1}{2}$	3.8%
4	6.66%	-1	0	6.78%
5	3.48%	$e^{6\pi i/7}$	$\frac{1}{3}$	3.38%
6	2.94%	$e^{3\pi i/4}$, -1	$\frac{2}{3}$, 0	3%
7	5.37%	$e^{8\pi i/9}$	$\frac{1}{4}$	5.32%
8	2.50%	$e^{3\pi i/5}$, $e^{4\pi i/5}$, -1	$\frac{4}{3}$, $\frac{1}{2}$, 0	2.5%
9	2.74%	$e^{8\pi i/11}$, $e^{10\pi i/11}$	$\frac{3}{4}$, $\frac{1}{5}$	2.85%
10	3.51%	$e^{2\pi i/3}$,$e^{5\pi i/6}$, -1	1, $\frac{2}{5}$, 0	4.39%
11	2.51%	$e^{8\pi i/13}$, $e^{10\pi i/13}$, $e^{12\pi i/13}$	$\frac{5}{4}$, $\frac{3}{5}$, $\frac{1}{6}$	2.44%
12	2.39%	$e^{5\pi i/7}$, $e^{6\pi i/7}$, -1	$\frac{4}{5}$, $\frac{1}{3}$, 0	2.75%
13	2.52%	$e^{2\pi i/3}$, $e^{4\pi i/5}$, $e^{14\pi i/15}$	1, $\frac{1}{2}$, $\frac{1}{7}$	2.43%
14	2.58%	$e^{3\pi i/4}$, $e^{7\pi i/8}$, -1	$\frac{2}{3}$, $\frac{2}{7}$, 0	2.55%
15	2.38%	$e^{12\pi i/17}$, $e^{14\pi i/17}$, $e^{16\pi i/17}$	$\frac{5}{6}$, $\frac{3}{7}$, $\frac{1}{8}$	2.4%
16	2.57%	$e^{2\pi i/3}$, $e^{7\pi i/9}$, $e^{8\pi i/9}$, -1	1, $\frac{4}{7}$, $\frac{1}{4}$, 0	2.45%
17	2.65%	$e^{14\pi i/19}$, $e^{16\pi i/19}$, $e^{18\pi i/19}$,	$\frac{5}{7}$, $\frac{3}{8}$, $\frac{1}{9}$	2.46%
18	2.49%	$e^{4\pi i/5}$, $e^{9\pi i/10}$, -1	$\frac{1}{2}$, $\frac{2}{9}$, 0	2.52%
19	2.45%	$e^{2\pi i/3}$, $e^{16\pi i/21}$, $e^{6\pi i/7}$, $e^{20\pi i/21}$	1, $\frac{5}{8}$, $\frac{1}{3}$, $\frac{1}{10}$	2.43%
20	2.79%	$e^{8\pi i/11}$, $e^{9\pi i/11}$, $e^{10\pi i/11}$, -1	$\frac{3}{4}$, $\frac{4}{9}$, $\frac{1}{5}$, 0	2.4%

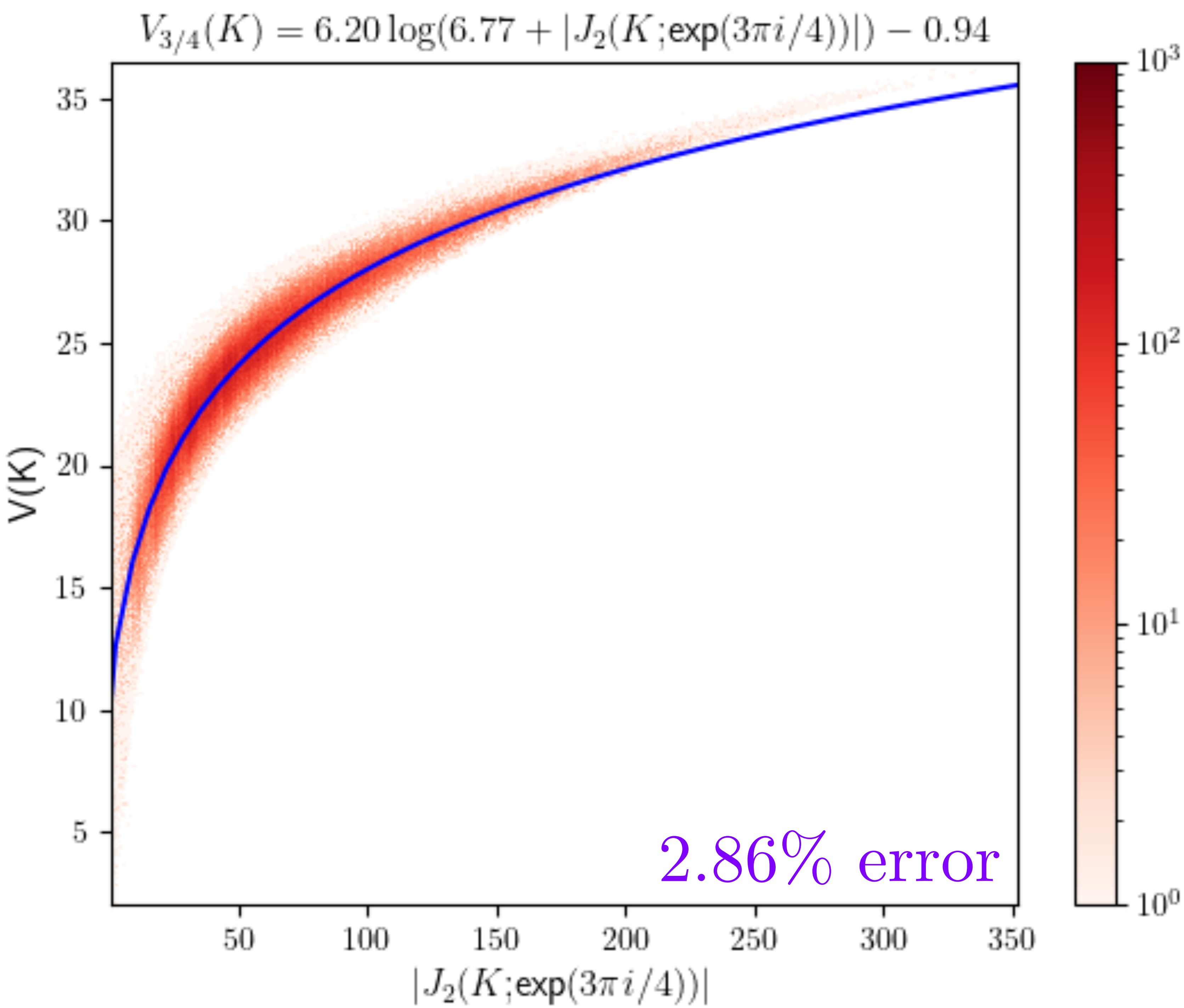
$$e^{ix} = e^{\frac{2\pi i}{k+2}}$$

Phenomenological Function

$$\omega = e^{\frac{3\pi i}{4}}$$

$$\Updownarrow$$

$$k = \frac{2}{3}$$

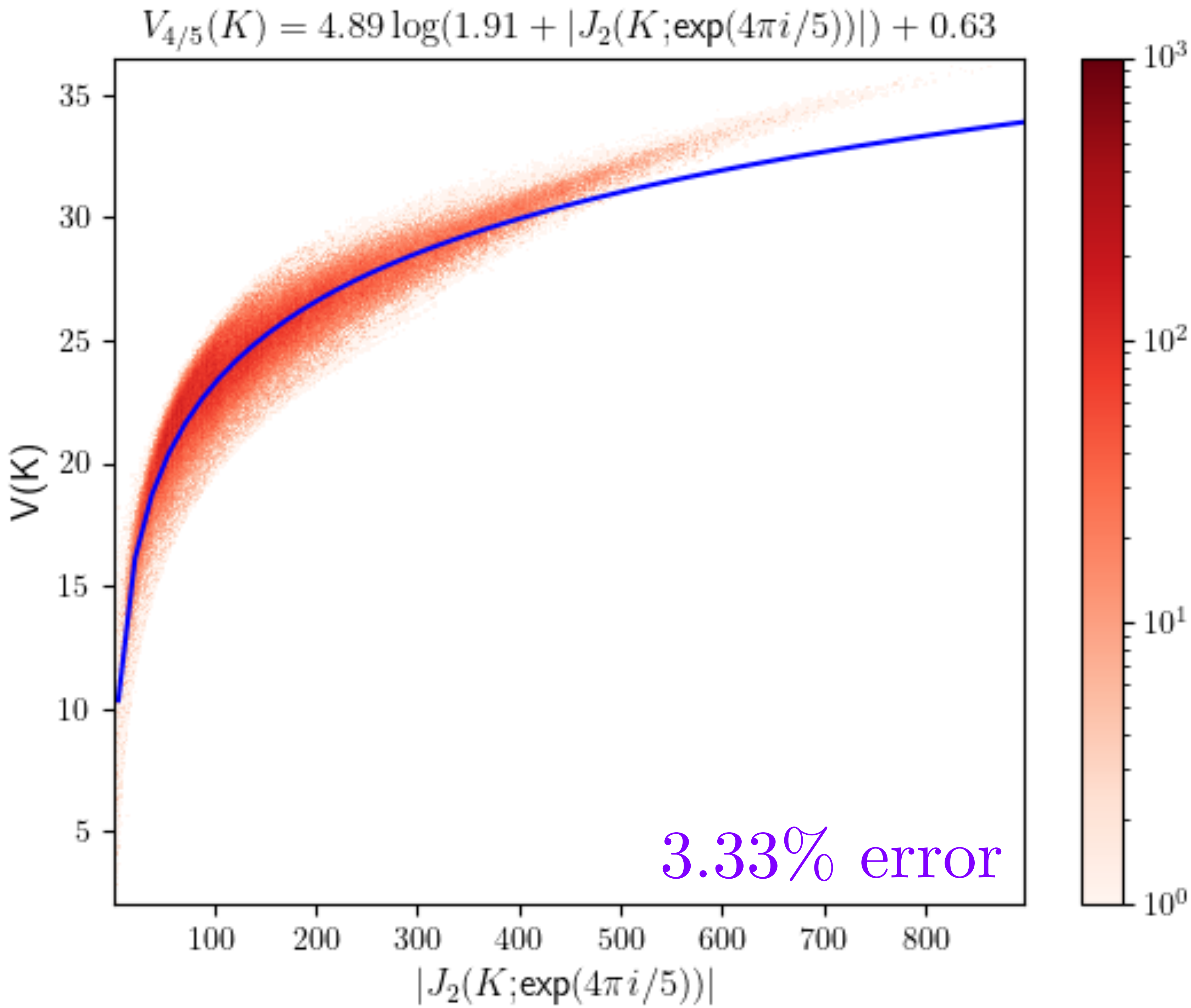


- Parameters fixed via curve fitting routines in Mathematica

Phenomenological Function

$\omega = e^{\frac{4\pi i}{5}}$

$k = \frac{1}{2}$



- Parameters fixed via curve fitting routines in Mathematica

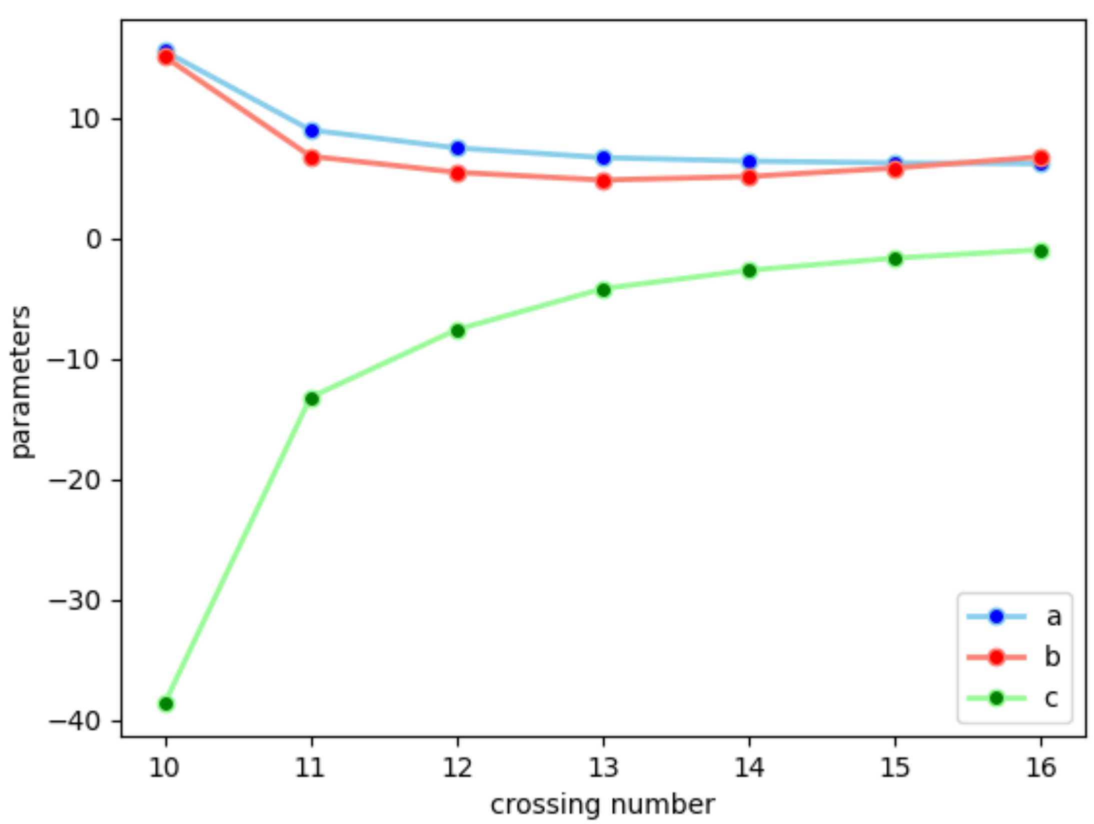
Phenomenological Function

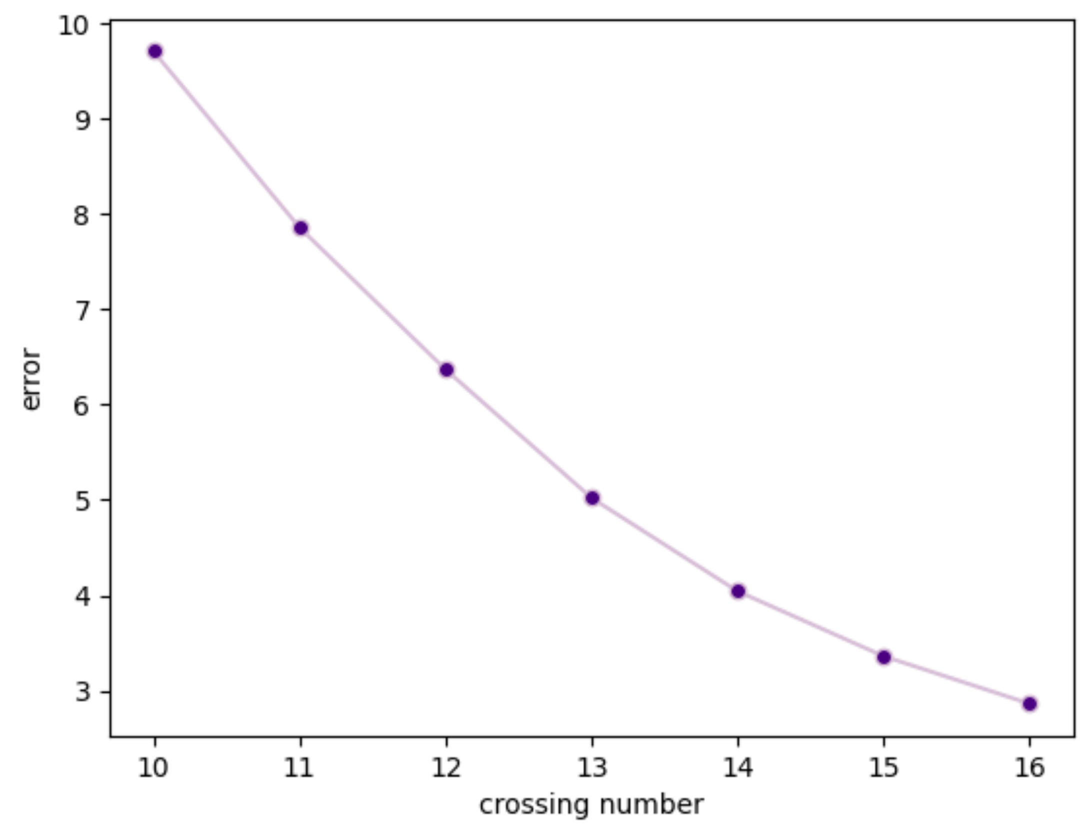
$$V_{3/4}(S^3 \setminus K) = 6.20 \log(|J_2(K; e^{\frac{3\pi i}{4}})| + 6.77) - 0.94$$

2.86% error compared to 2.28% error for neural network

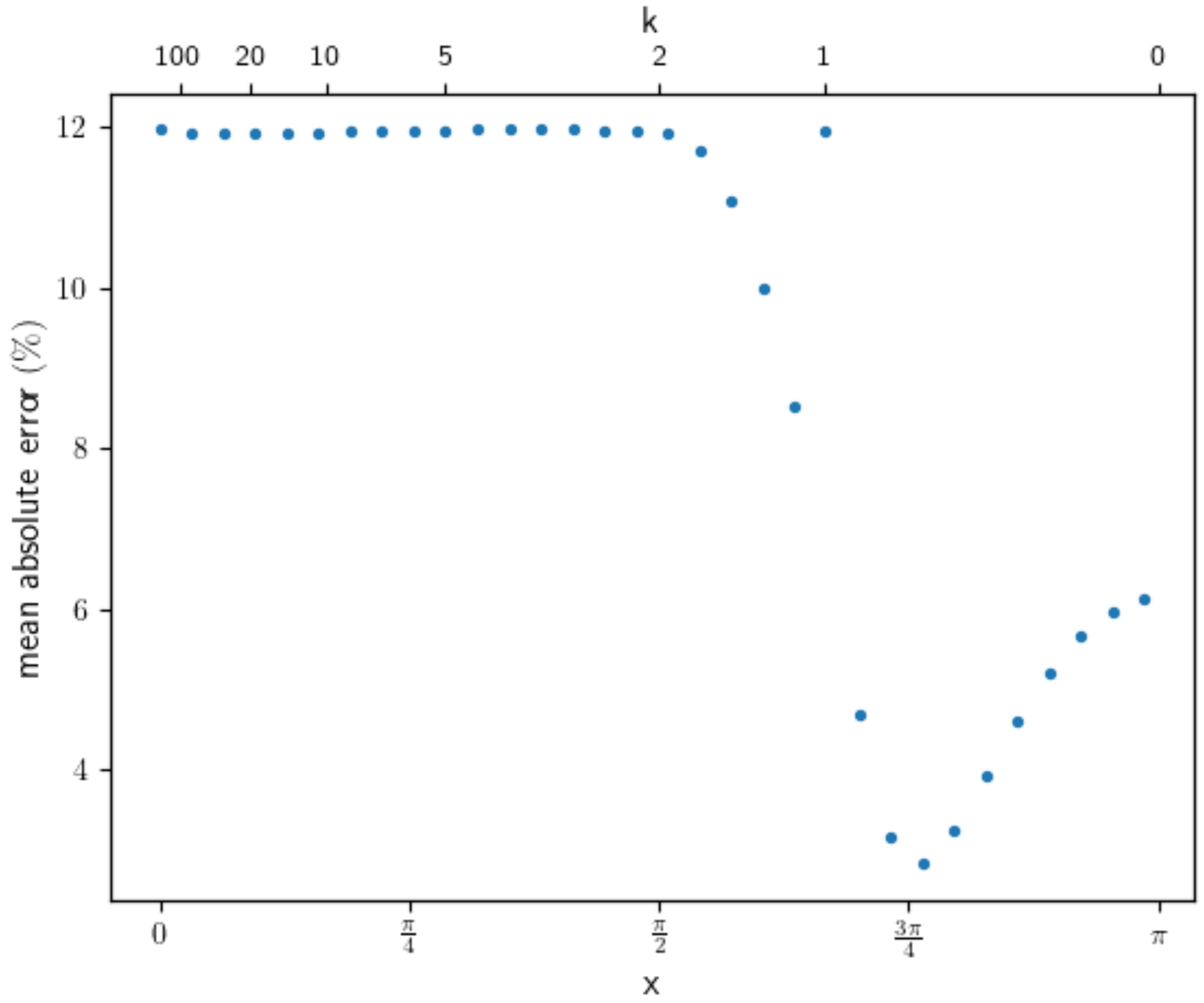
corresponds to Chern–Simons level $k = \frac{2}{3}$

- Parameters of fit robust as a function of crossing number



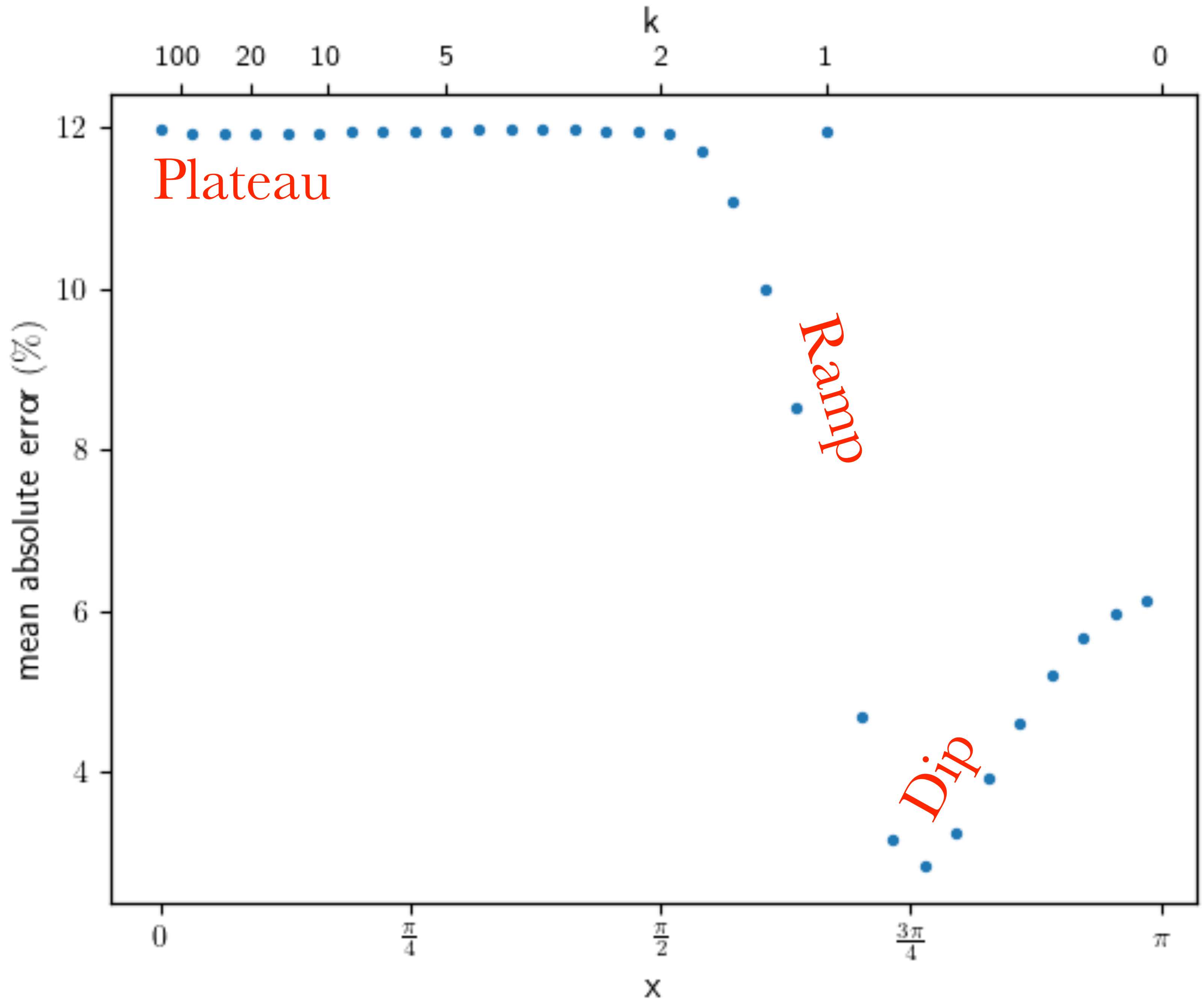


Best Phase



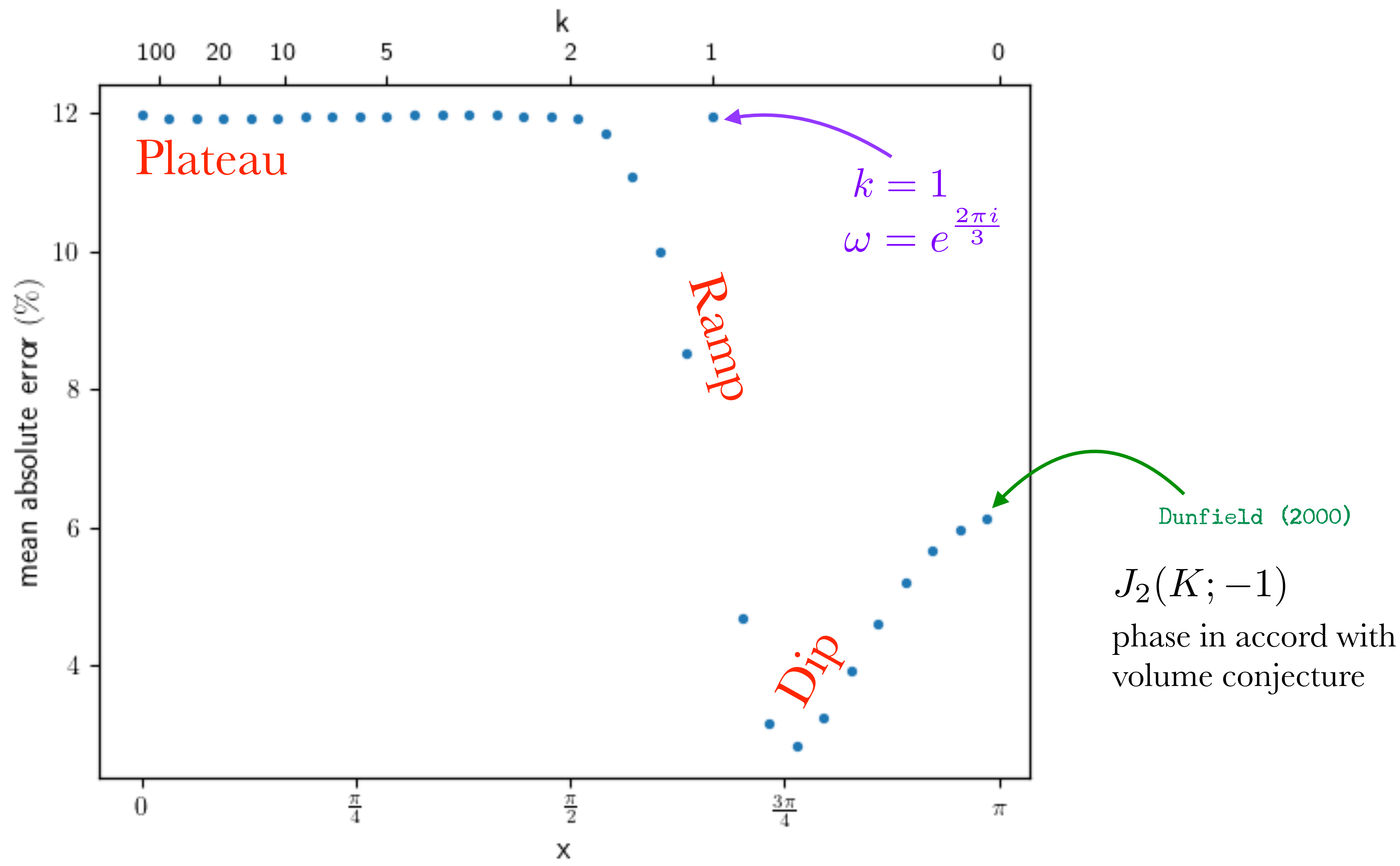
- Best fit to $V(x) = a \log(|J_2(K; e^{ix})| + b) + c$ at $x = 2.3$

Best Phase



- Best fit to $V(x) = a \log(|J_2(K; e^{ix})| + b) + c$ at $x = 2.3$

Best Phase



- Best fit to $V(x) = a \log(|J_2(K; e^{ix})| + b) + c$ at $x = 2.3$

Chern–Simons Theory

- Recall that $S_{\rm CS} = \dfrac{k}{4\pi}\displaystyle\int_{\mathcal{M}} {\rm tr}\left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A\right)$

- Under the gauge transformation $A_\mu \mapsto g^{-1} A_\mu g + g^{-1}\partial_\mu g$,

$$\Delta S_{\rm CS} = \frac{k}{4\pi}\int d^3x\ \epsilon^{\mu\nu\rho}\left(\partial_\mu {\rm tr}\left((\partial_\nu g)g^{-1}A_\rho\right) + \frac{1}{3}{\rm tr}\left(g^{-1}\partial_\mu g g^{-1}\partial_\nu g g^{-1}\partial_\rho g\right)\right)$$

- Associated to large gauge transformations, we recognize the *winding*

$$w(g) = \frac{1}{24\pi^2}\int d^3x\ \epsilon^{\mu\nu\rho}{\rm tr}\left(g^{-1}\partial_\mu g g^{-1}\partial_\nu g g^{-1}\partial_\rho g\right) \in \mathbb{Z}$$

- This implies that the level $k \in \mathbb{Z}$

- So what does fractional level mean?

- In Abelian Chern–Simons theory, can make sense of $k = \dfrac{1}{2}$

Moore (2019)

Analytic Continuation

- We can analytically continue the level
- Appeal to Morse theory; the integration cycle $\mathcal{C}$ used to compute path integral is decomposed in terms of Lefschetz thimbles
- We encounter Stokes phenomena and the decomposition changes
- The Stokes phenomena are determined by $\gamma := \dfrac{n-1}{k}$

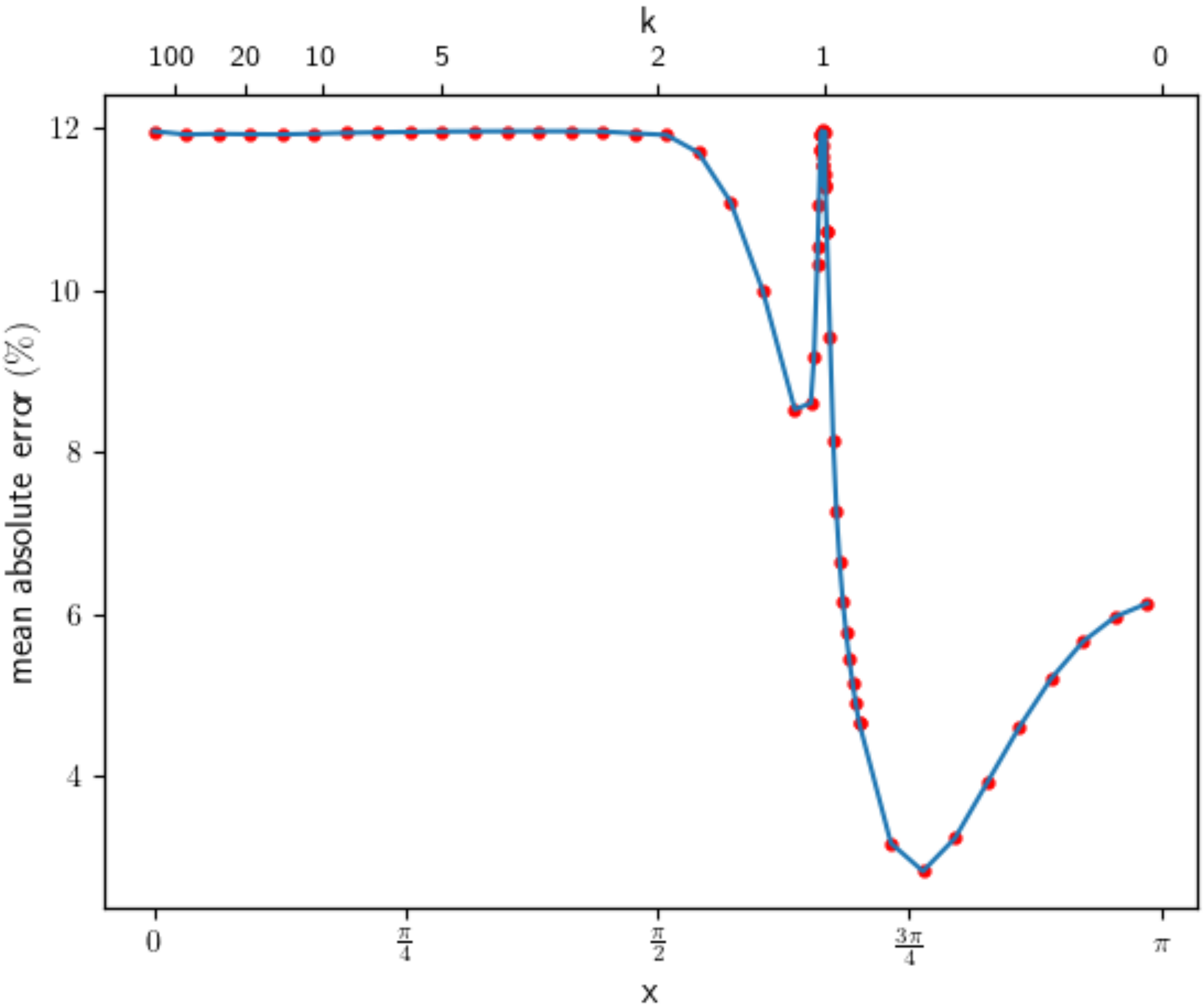
  N.B.: for Jones polynomial $n=2$ and $\gamma = k^{-1}$
- In semiclassical limit, $\gamma \to 1$, $k \to \infty$ and there can be a contribution

  $e^{iS_{\mathrm{CS}}}(1-e^{2\pi i k})$ to path integral from pair of $SL(2,\mathbb{C})$ critical points

  after analytic continuation
- Volume conjecture essentially states that this behavior occurs for geometric conjugate connection for every knot at $n = k+2$

Witten (2010)

Volume Functions



The Shape of Things

Plateau: $k > 2$

$V(K) = \overline{V}$

this gives 11.97% error for knots up to 16 crossings

corresponds to latent correlations in the dataset

Minimum: near $k = \frac{2}{3}$ or $\gamma = \frac{3}{2}$

Lefschetz thimbles contain geometric conjugate $SL(2, \mathbb{C})$ connection we expect in semiclassical limit

Ramp: $\frac{2}{3} < k < 2$

knots begin to lose access to geometric conjugate connection

at $\gamma = \frac{2}{3}$ or $k = \frac{3}{2}$, the geometric conjugate connection of 4_1 enters

Witten (2010)

The Shape of Things

Dip: $0 < k < \frac{2}{3}$

knots retain geometric conjugate connection even as $\gamma \gg 1$

Spike: near $k = 1$

at integer values of level with $k + 1 \geq n$, the path integral receives contributions only from $SU(2)$-valued critical points

i.e., no analytic continuation is necessary

because we lose knowledge of the geometric conjugate connection, the error becomes high

Conclusion: geometric conjugate connection is crucial to success of approximation formula

Prospectus

- Inequalities à la volume-ish theorem using analytically continued Chern–Simons theory
- Monotonic version of the volume conjecture
- Better understanding of what problems are machine learnable in mathematics and physics — failed experiments may teach us something!
- Why are the volume and CS invariant different?
- Reverse engineer other machine learned results

THANK YOU!