Applied topology methods in knot theory

Radmila Sazdanovic

NC State University

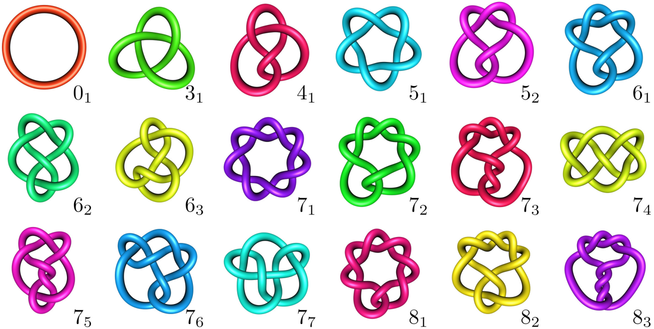
Optimal Transport, Topological Data Analysis and Applications to Shape and Machine Learning//Mathematical Bisciences Institute OSU

29 July 2020

Joint work with P. Dlotko, J. Levitt. M. Hajij

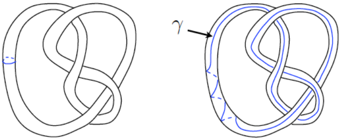
- How to vectorize data, how to get a point cloud out of the shapes?
  Idea: Use a vector consisting of numerical descriptors
- Methods:
  - Machine learning: How to use ML with infinite data that is hard to sample in a reasonable way. a
  - Topological data analysis
  - Hybrid between TDA and statistics: PCA combined with appropriate filtration
- Goals: improving identification and circumventing obstacles from computational complexity of shape descriptors.

- Input: point clouds obtained from knot invariants
  - Machine learning
  - PCA Principal Component Analysis with different filtrations
  - Topological data analysis
- Goals:
  - characterizing discriminative power of knot invariants for knot detection
  - comparing knot invariants
  - experimental evidence for conjectures



Knot is an equivalence class of smooth embeddings $f : S^1 \to \mathbb{R}^3$ up to ambient isotopy.

- Problems are easy to state but a remarkable breath of techniques are employed in answering these questions
  - combinatorial
  - algebraic
  - geometric



- Knots are interesting on their own but they also provide information about 3- and 4-dimensional manifolds

Theorem (Lickorish-Wallace)

Every closed, connected, oriented 3-manifold can be obtained by doing surgery on a link in S^3.

#C	0	3	4	5	6	7	8	9	10
#PK	1	1	1	2	3	7	21	49	165

#C	11	12	13	14	15
#PK	552	2,176	9,988	46,972	253,293

#C	16	17	18	19
#PK	1,388,705	8,053,393	48,266,466	294,130,458

TABLE: Number of prime knots for a given crossing number.
Legend: Number of Crossings denoted by #C and number of Prime Knots dentoted by #PK

- Rolfsen's table
- The first 1,701, 936 knots (Hoste, Thistlethwaite, Weeks)
- Knot Atlas by Bar Natan and KnotInfo by Cha, Livingston
- Over 350×10^6 knots under 20 crossings (Burton's Regina)
- Tuzun and Sikora found no counter examples to the Jones unknot conjecture among knots with up to 23 crossing using 10,718,938,763,889 knot diagrams.
- Ernst and Sumners proved in 1987 that the number of distinct knots grows exponentially with the crossing number
- We consider the 9,755,329 knots with up to 17 crossings

- Knot invariants for knot classification and tabulation
- **Different viewpoint: knot invariants and their relations.**
- Dimension reduction:
  - Principal component analysis PCA by J. Levitt, M. Hajij, R.S.
  - non-linear technique based on an isometric mapping ISOMAP
- Machine learning - classification
  - A neural network approach to predicting and computing knot invariants by Mark C. Hughes
  - Deep Learning the Hyperbolic Volume of a Knot by V. Jejjalaa, A.Karb, O.Parrikarb
- Topological data analysis

DEF (KNOT INVARIANT)

is a function from the set of links to the set values which depend only on the equivalence class of knots and links.

INVARIANTS ~50

CLASSICAL numerical: components, linkings, colorings
polynomial: Alexander, Jones, Kauffman
algebraic: coloring and knot group

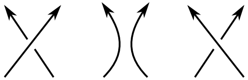
MODERN Homology theories: Khovanov link, Heegaard-Floer homology and numerical invariants defined based on their properties

Knot identification is hard.

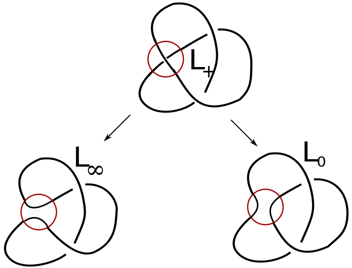
- $J_{2,q}(K)$ is the Jones polynomial and $J_{2,q}(\text{unknot}) = 1$.

$$(q^{1/2} - q^{-1/2})J_{L_0}(q) = q^{-1}J_{L_+}(q) - qJ_{L_-}(q)$$

- Alexander polynomial
$(x^{1/2} - x^{-1/2})\Delta_{L_0}(x) = \Delta_{L_+}(x) - \Delta_{L_-}(x)$
- Bar Natan and Van der Veen: Z_0-polynomial
- HOMFLY-PT polynomial



From left to right:
L_+, L_0 and L_-.



- Crossing number
- The signature $\sigma(K)$. It is computable from
  - Seifert surface S of genus g whose boundary is the knot by considering a Seifert form V a 2g by 2g matrix whose entreies are the linking numbers of pushoffs of generators of $H_1(S)$. Then the signature of a knot is a signature of $V + V^T$
  - the Alexander module.
- Combinatorial f-la for alternating knots: $\sigma(K) = s_A(D) - n_+(D) - 1 = -s_B(D) + n_-(D) + 1$ [Traczyk]
- Rasmussen s-invariant: coming from special grading of Khovanov homology that categorifies the Jones polynomial.

We associate to $J_{2,q}(K)$ a point in coefficient space.
Consider the knots of up to 6 crossings:

$$J_{2,q}(0_1) = 1$$
$$J_{2,q}(3_1) = -q^{-4} + q^{-3} + q^{-1}$$
$$J_{2,q}(4_1) = q^{-2} - q^{-1} + 1 - q + q^2$$
$$J_{2,q}(5_1) = -q^{-7} + q^{-6} - q^{-5} + q^{-4} + q^{-2}$$
$$J_{2,q}(5_2) = -q^{-6} + q^{-5} - q^{-4} + 2q^{-3} - q^{-2} + q^{-1}$$
$$J_{2,q}(6_1) = q^{-4} - q^{-3} + q^{-2} - 2q^{-1} + 2 - q + q^2$$
$$J_{2,q}(6_2) = q^{-5} - 2q^{-4} + 2q^{-3} - 2q^{-2} + 2q^{-1} - 1 + q$$
$$J_{2,q}(6_3) = -q^{-3} + 2q^{-2} - 2q^{-1} + 3 - 2q + 2q^2 - q^3$$

Mirror so that the extreme power is positive.
(Forces signature to always be positive)

$$
\begin{aligned}
J_{2,q}(0_1) &= 1 \\
J_{2,q}(\mathrm{mir}(3_1)) &= q + q^3 - q^4 \\
J_{2,q}(4_1) &= q^{-2} - q^{-1} + 1 - q + q^2 \\
J_{2,q}(\mathrm{mir}(5_1)) &= q^2 + q^4 - q^5 + q^6 - q^7 \\
J_{2,q}(\mathrm{mir}(5_2)) &= q - q^2 + 2q^3 - q^4 + q^5 - q^6 \\
J_{2,q}(\mathrm{mir}(6_1)) &= q^{-2} - q^{-1} + 2 - 2q + q^2 - q^3 + q^4 \\
J_{2,q}(\mathrm{mir}(6_2)) &= q^{-1} - 1 + 2q - 2q^2 + 2q^3 - 2q^4 + q^5 \\
J_{2,q}(6_3) &= -q^{-3} + 2q^{-2} - 2q^{-1} + 3 - 2q + 2q^2 - q^3
\end{aligned}
$$

Write as vectors so q^0 is in the same position in every vector

	q^{-3}	q^{-2}	q^{-1}	q^0	q^1	q^2	q^3	q^4	q^5	q^6	q^7
$J_{2,q}(0_1)$	0	0	0	1	0	0	0	0	0	0	0
$J_{2,q}(\mathrm{mir}(3_1))$	0	0	0	0	1	0	1	-1	0	0	0
$J_{2,q}(4_1)$	0	1	-1	1	-1	1	0	0	0	0	0
$J_{2,q}(\mathrm{mir}(5_1))$	0	0	0	0	0	1	0	1	-1	1	-1
$J_{2,q}(\mathrm{mir}(5_2))$	0	0	0	0	1	-1	2	-1	1	-1	0
$J_{2,q}(\mathrm{mir}(6_1))$	0	1	-1	2	-2	1	-1	1	0	0	0
$J_{2,q}(\mathrm{mir}(6_2))$	0	0	1	-1	2	-2	2	-2	1	0	0
$J_{2,q}(6_3)$	-1	2	-2	3	-2	2	-1	0	0	0	0

Theorem (Garoufalidis ’03)

For all simple knots up to 8 crossings and for all torus knots, the colored Jones polynomial determines the signature of the knot.

- Originally conjectured for all simple knots.
- A knot K is simple, if all the roots $\alpha \in \Delta(K)$ of the Alexander polynomial where $|\alpha| = 1$ have multiplicity 1.
- Garoufalidis conjectured that the the colored Jones polynomial $J_{N,q}(K)$ for simple knots determines the signature.

- They can be small
  $\{4_1, \mathrm{mir}(11n19)\}$ $0, 4$ $\sqrt{5}$
  $\{\mathrm{mir}(7_2), 11n88\}$ $2, 6$ $\sqrt{17}$
- There can be multiple in a grouping
  $\{\mathrm{mir}(5_2), 11n57, 13n3082\}$ $2, 6$ 3
  $\{13n137, \mathbf{13n627}, \mathbf{13n716}, 13n1539, \mathrm{mir}(13n1560), \ldots\}$ with signature either $0, 4$ and $L2$-norm equal to $\sqrt{509}$.[1]
- They can have the same number of crossings
  $\{11n28, 11n64\}$ $0, 4$
  $\{12n107, 12n171\}$ $2, 6$
- They can both be alternating
  $\{12a802, 12a1242\}$ $2, 6$ $\sqrt{215}$

[1] **13n627**, **13n716** have identical signature, determinat, Alexander, HOMFLYPT and Kauffman polynomials but different DT-codes

There are two major styles of machine learning:

- supervised learning gives classifiers or predictors
- unsupervised learning
  - clustering or dimension reduction
  - could improve knot type identification.

Challenges

- Sampling?
- Depending on the knots the data will belong to Euclidean spaces of different dimensions. Which distance to use?
- How to encode 2-variable polynomials such as HOMFLYPT and Khovanov polynomial to agree with the Jones vectors?

Definition (classifiers)

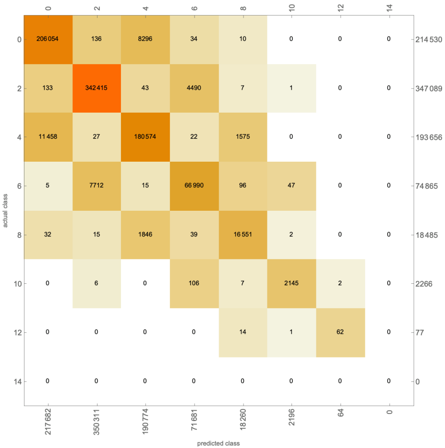
functions from a given input space to a finite number of known outputs. Rules for new inputs come from a set of known cases.

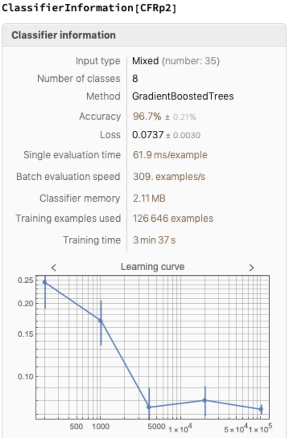
- Given a geometric space of inputs, the classification of inputs into bins can be defined by dividing hyperplanes.
- These can be defined globally, or pairwise.
- The input space can be covered by probability distributions.
- Unknown points are determined by polling their k nearest neighbors.

Definition (predictors)

are functions from a given input space to a known output space. Predictors are sort of a collection of regression methods.

- Ordinary Least Squares Regression (OLSR)
- Linear Regression
- Stepwise Regression
- Multivariate Adaptive Regression Splines (MARS)
- Locally Estimated Scatterplot Smoothing (LOESS)





Yes & No. Accurate 98%+ on training and 94% on test data.

Theorem (Hajij-Levitt-S.)

- *Let K be a knot with the span of the Jones polynomial equal to s,*
- *K_{nn} the knot with Jones span less than s that is the nearest neighbor to K under the L_2-norm.*

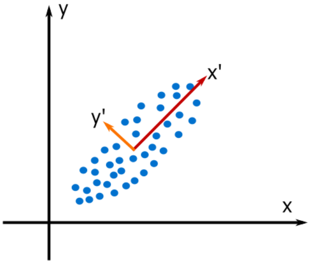
Then the probability $P(\sigma(K) = \sigma(K_{nn})) > 95\%$

Understanding the discriminative power of the Jones polynomial, its relations to other classical invariants of knots and links, as well as the information encoded in its coefficients, conjectured to be related to the hyperbolic volume of the knot, are important problems in low-dimensional topology.

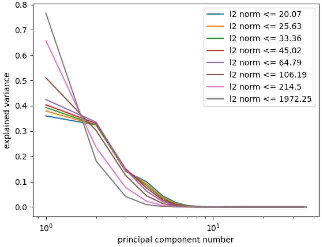
Principal Component Analysis (PCA)

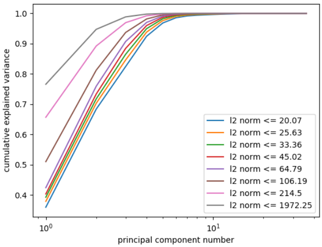
- Orthogonal linear transformation to a new orthonormal basis. (Via Eigendecomposition of the covariance matrix)

- Ranks directions based on maximal variance.
- Eigenvalue λ_i: the *explained variance* associated with the corresponding eigenvector.
- PCA of a single data set describes its dimensionality.



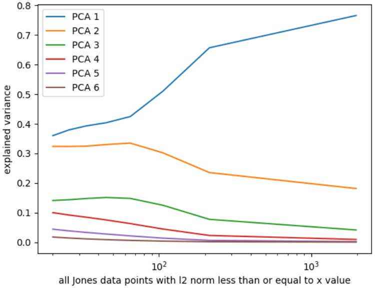
- Nested sequences can be obtained by filtering by any well-ordered index set.
- PCA of a nested sequence of a family of point cloud provides insights on how the eigenvectors and the variances of the evolve as the size of the point cloud grows.

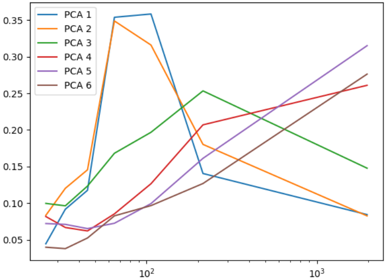




- Left: the explained variance is plotted against the index of the principal vector.
- Right: the incremental summation S_k is plotted against the same.

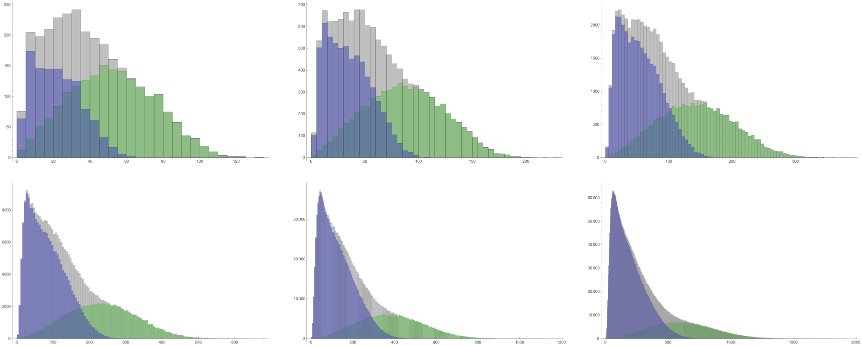
Explained Variance Filtered by Norm: Jones





- Left: Explained variance plotted against the size of the l_2 norm-filtered point cloud. $\overline{\lambda_1}$ grows slowly as we increase the size of this knot family, while the explained variances $\overline{\lambda_i}$ for $i \geq 2$ decrease.
- Right: The PCA bases obtained from the filtration stabilize as we pass through filtration.

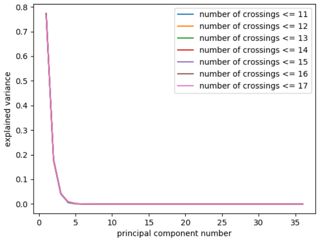
Jones Vector Norm Distribution

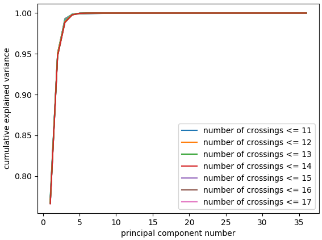


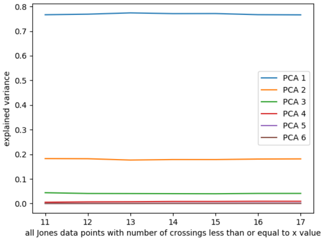
The distribution of the l_2-norms (total count vs. norm) for the

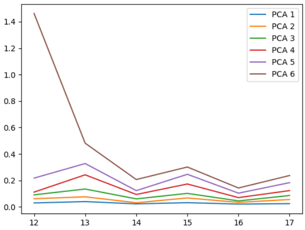
- Alternating (Green), NonAlternating (Blue), and Combined (Grey) knots
- up to 12, 13, 14, 15, 16, and then 17 crossings when taken left to right, top to bottom.

EV Filtered by Crossing: Jones

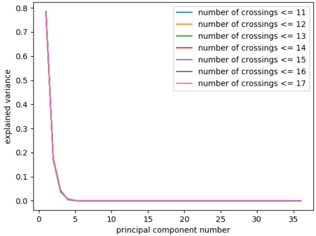


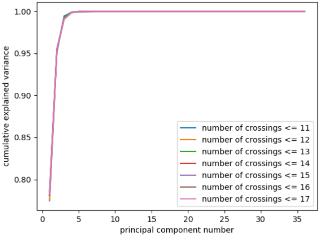


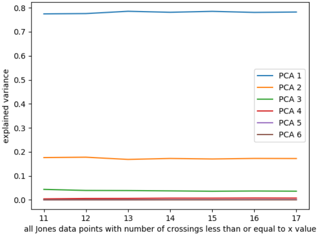


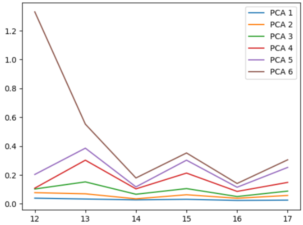


EV Filtered by Crossing: Alternating Jones

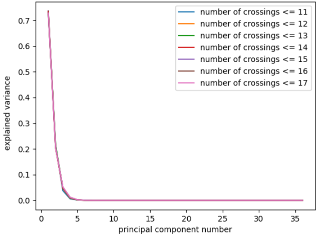


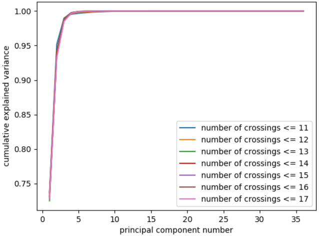


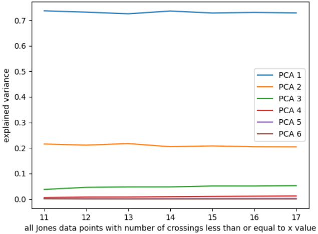


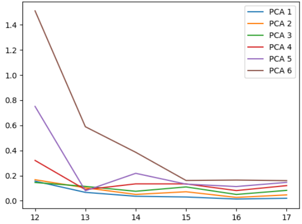


EV Filtered by Crossing: NonAlternating

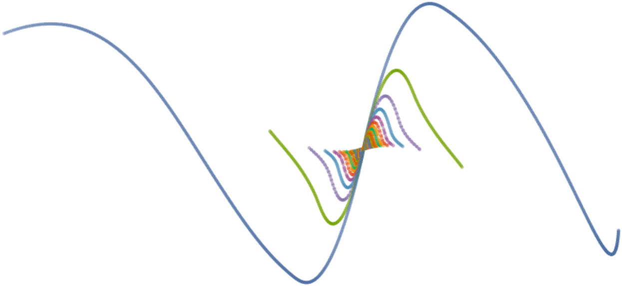


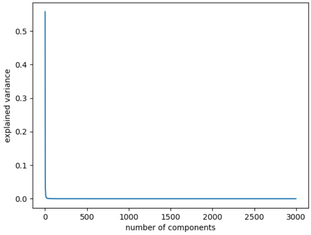


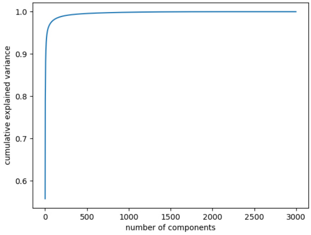




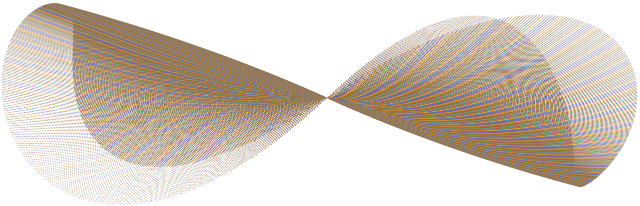
Explained Variance for Torus Knots

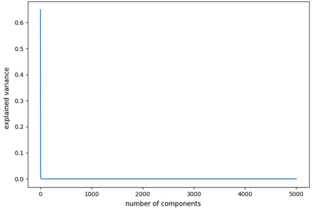


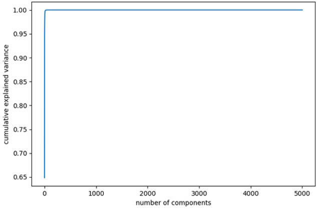




EV for Double Twist Link Knots

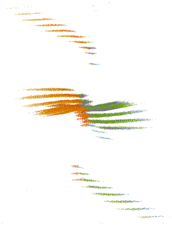




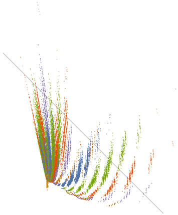


PCA Results: $\Delta \approx 1D$, $Z_0 \approx 2D$, $J \approx 3D$

(A) Jones is ≈ 3 dimensional



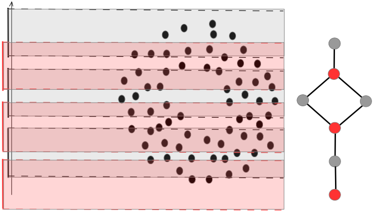
(B) Alexander is ≈ 1 dimensional



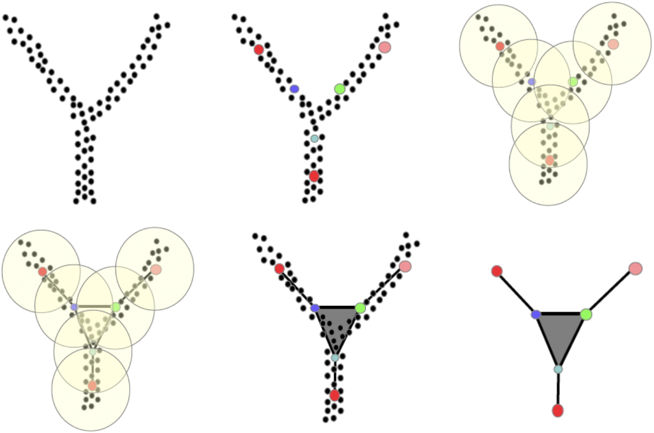
(C) Z_0 is ≈ 2 dimensional

Mapper:

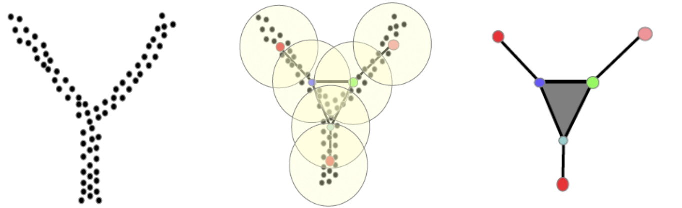
- Cover of the line (number of cover elements, percentage of overlap).
- **Lens function.**
- Clustering algorithm.



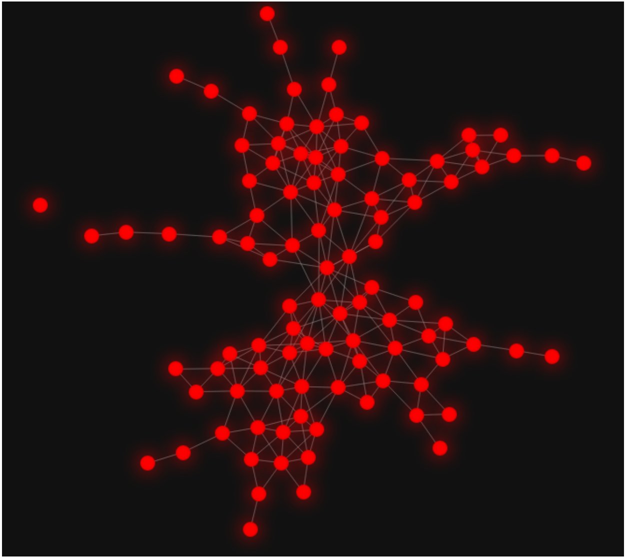
1. The choice of a lens fucntion affects the connectivity of the Mapper graphs
2. Analysis is often complicated.
3. Result may not stable with respect to the cover of the line.



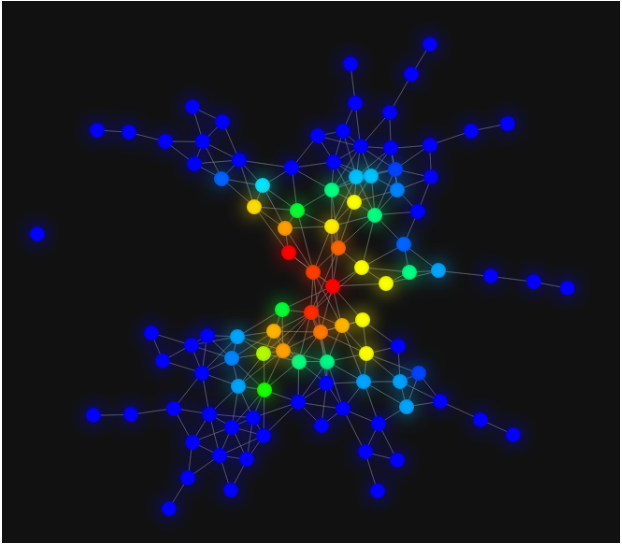
Ball mapper by P. Dłotko built on a simple, geometry–based idea might be used as an alternative.



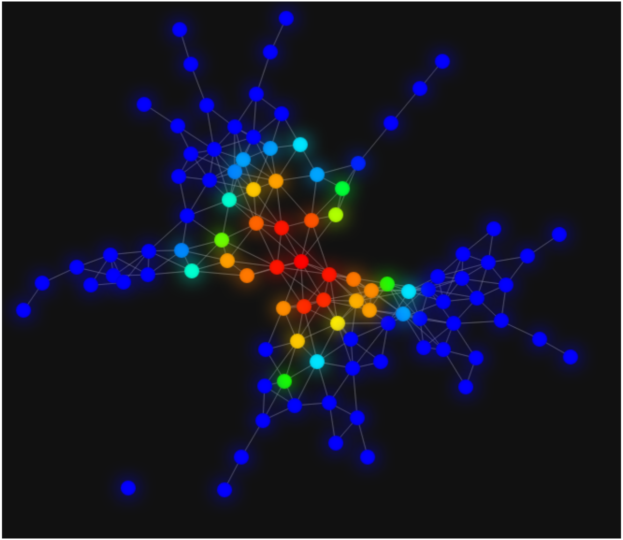
- Geometrically stable, easy-to-use exploratory data analysis tool
- Based on the nerve complex construction
- Requires only one parameter–radius.
- In some cases it provides a more accurate information than Mapper but it might be less amenable to interpretations.
- Compatible with machine learning



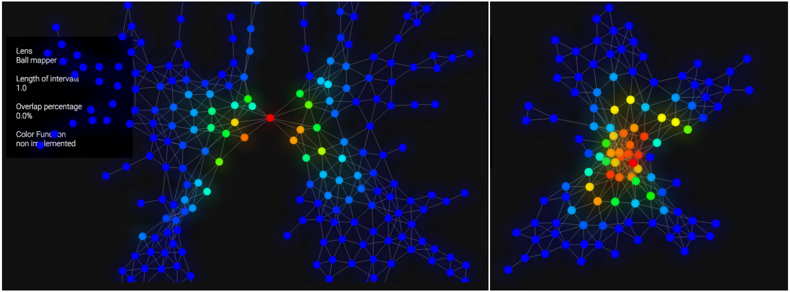
JONES DATA COLORED BY THE CROSSING NUMBER



JONES DATA: ALTERNATING VS. NON–ALTERNATING

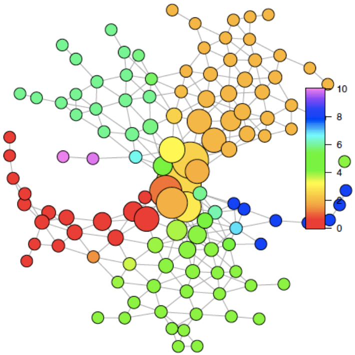


JONES DATA: ALTERNATING VS. NON–ALTERNATING

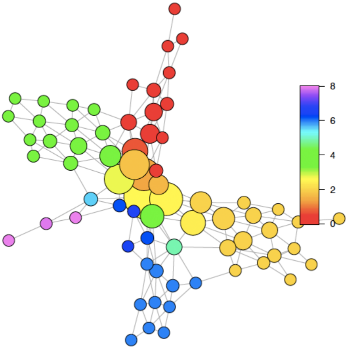


(A) Alternating (B) Non-alternating

FIGURE: Knots up to 17 crossings, colored by crossing number, $L2$ norm.

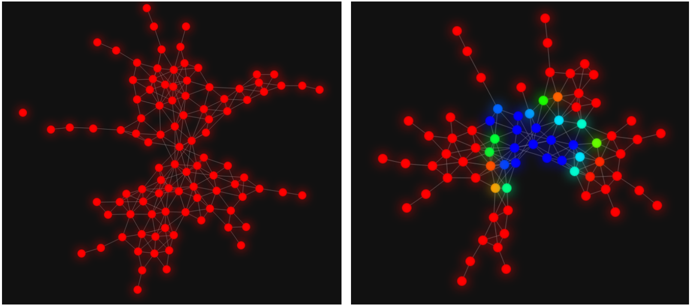


(A) $L2$



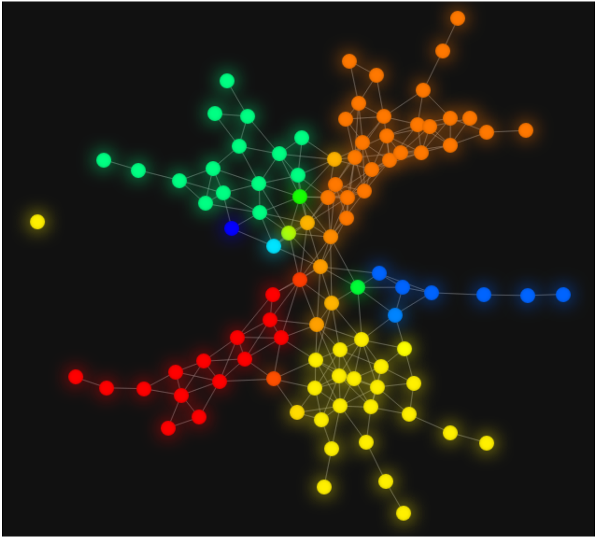
(B) $L1$

FIGURE: Knots up to 17 crossings



Space of Jones Polynomials for knots up to 17 crossings.

JONES DATA VS. KNOT SIGNATURE



SIGNATURE: ALTERNATING KNOTS

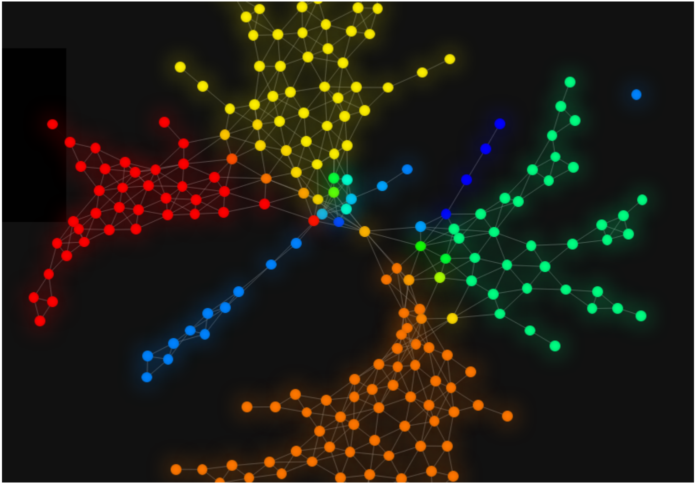


FIGURE: Alternating knots up to 17 crossings, r=120, colored by signature values, $L1$ norm.

JONES DATA: SIGNATURE

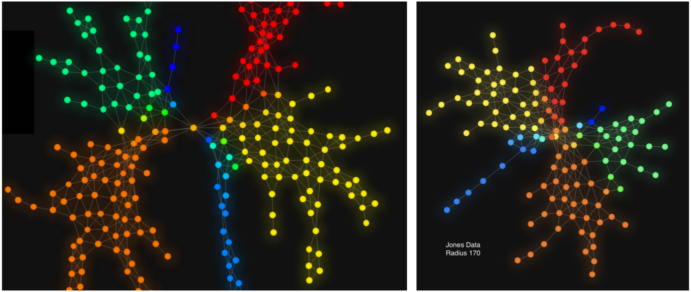
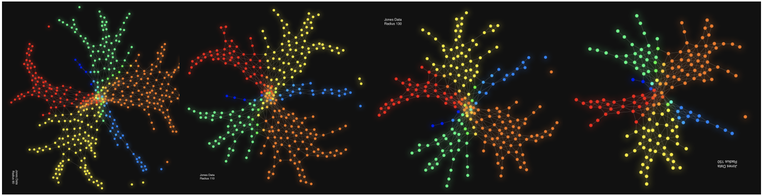


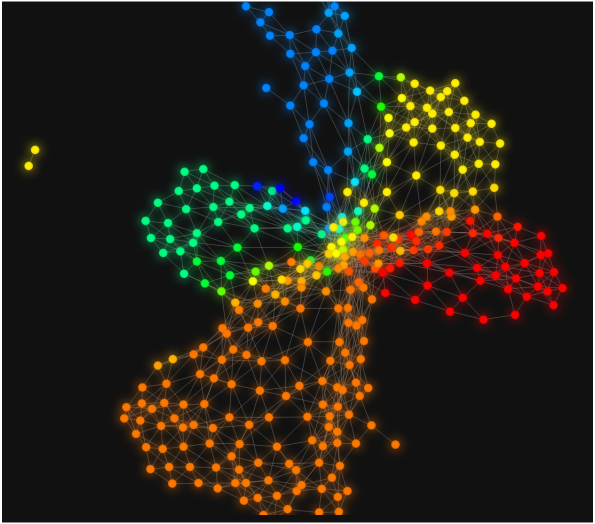
FIGURE: Signature on alternating vs. non-alternating knots Jones data cloud

JONES POLY VS. SIGNATURE

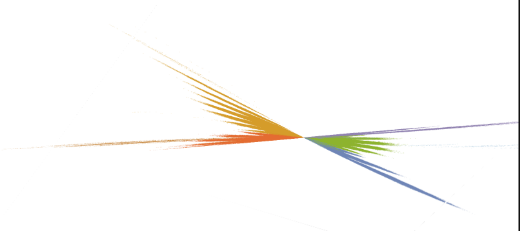


(A) r=90 (B) r=110 (C) r=130 (D) r=150

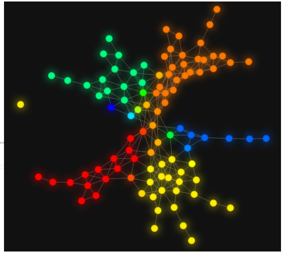
FIGURE: Knots up to 17 crossings, colored by signature values, $L2$ norm.



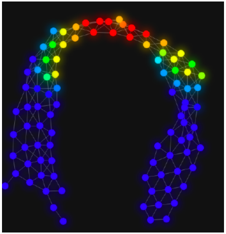
Signature in the center, $r = 20$.



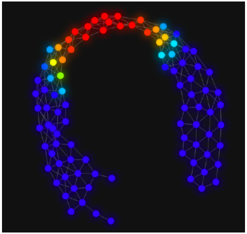
(A) 2d planes determine signature away from the center



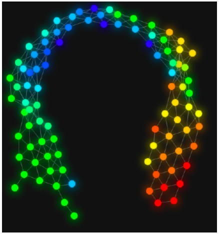
(B) "flares" determine signature



(A) Crossing number



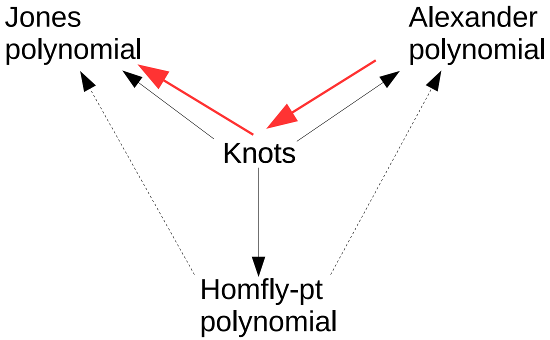
(B) Alt vs. non-alt

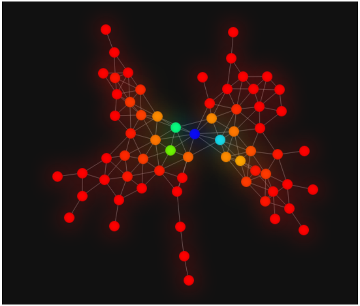


(C) signature

FIGURE: Knots up to 17 crossings, radius 45.

HOW DOES ALEXANDER DATA COMPARE TO JONES?

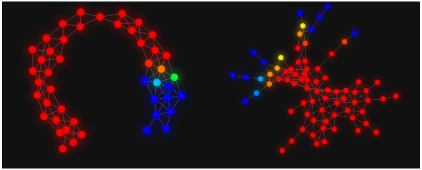
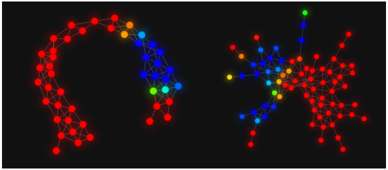
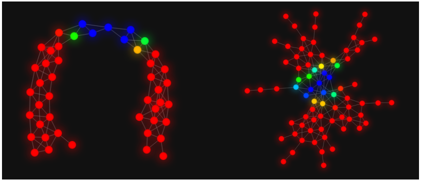
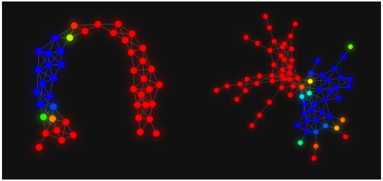




(A) Jones data: Center node 130k, peripheral 1-10k.

(B) Alexander data: Center node 70k, peripheral 3-10k.

HOW DOES ALEXANDER DATA COMPARE TO JONES?

Both PCA and Ball Mapper approaches have/provide

- stability under crossing number filtration both PCA and Ball Mapper
- similar structural properties of Jones data
- tools for analyzing and comparing knot invariants

Theorem (Lee; Shumakovitch)

Khovanov homology of alternating knots is determined by the Jones polynomial and signature.

Future projects

Formalize above observations, analyze and compare other knot invariants and other data sets.

Thank you!