Quantum Information, Machine Learning and Knot Theory

Onkar Parrikar

Department of Physics & Astronomy
University of Pennsylvania.

Math/Strings Regional Meeting at Duke University, 04/20/2019.
Based on

Part 1: Quantum Information
Entanglement Entropy in Qubits: Brief Overview

- The basic example of an entangled state between two qubits is

$$\left| \psi_{\text{Bell}} \right\rangle = \frac{1}{\sqrt{2}} \left( \left| 0 \right\rangle \otimes \left| 0 \right\rangle + \left| 1 \right\rangle \otimes \left| 1 \right\rangle \right).$$

If we trace over one of the qubits, we obtain a mixed state

$$\rho = \text{Tr}_2 \left| \psi_{\text{Bell}} \right\rangle \langle \psi_{\text{Bell}} \right| = \frac{1}{2} \left( \left| 0 \right\rangle \langle 0 \right| + \left| 1 \right\rangle \langle 1 \right| .$$

We can associate an entropy to it, namely the Von-Neumann entropy, often called the entanglement entropy

$$S(\rho) = -\text{Tr}(\rho \ln \rho) = \ln(2).$$

This is to be contrasted against unentangled product states like

$$\left| 0 \right\rangle \otimes \left| 0 \right\rangle, \left| 0 \right\rangle \otimes \left| 1 \right\rangle, \left| + \right\rangle \otimes \left| + \right\rangle,$$
Entanglement Entropy in Qubits: Brief Overview

- The basic example of an entangled state between two qubits is

\[ |\psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \right). \]

- If we trace over one of the qubits, we obtain a mixed state

\[ \rho = \text{Tr}_2 |\psi_{\text{Bell}}\rangle \langle \psi_{\text{Bell}}| = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|. \]
Entanglement Entropy in Qubits: Brief Overview

- The basic example of an entangled state between two qubits is

\[ |\psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \right). \]

- If we trace over one of the qubits, we obtain a mixed state

\[ \rho = \text{Tr}_2 |\psi_{\text{Bell}}\rangle \langle \psi_{\text{Bell}}| = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|. \]

- We can associate an entropy to it, namely the Von-Neumann entropy, often called the entanglement entropy

\[ S(\rho) = -\text{Tr}_1 \left( \rho \ln \rho \right) = \ln(2). \]
Entanglement Entropy in Qubits: Brief Overview

- The basic example of an entangled state between two qubits is
  \[ |\psi_{\text{Bell}}\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle \otimes |0\rangle + |1\rangle \otimes |1\rangle \right). \]

- If we trace over one of the qubits, we obtain a mixed state
  \[ \rho = \text{Tr}_2 |\psi_{\text{Bell}}\rangle \langle \psi_{\text{Bell}}| = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|. \]

- We can associate an entropy to it, namely the Von-Neumann entropy, often called the entanglement entropy
  \[ S(\rho) = -\text{Tr}_1 \left( \rho \ln \rho \right) = \ln(2). \]

- This is to be contrasted against unentangled product states like
  \[ |0\rangle \otimes |0\rangle, \ |0\rangle \otimes |1\rangle, \ |+\rangle \otimes |+\rangle \text{ etc.} \]
Entanglement Entropy in Qubits: Brief Overview

- With more qubits, one can construct more interesting entangled states. For example, with three qubits we have [Dur et al ’00]

\[ |GHZ\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle). \]

\[ |W\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle). \]
Entanglement Entropy in Qubits: Brief Overview

- With more qubits, one can construct more interesting entangled states. For example, with three qubits we have [Dur et al ’00]

\[ |GHZ\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle). \]

\[ |W\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle). \]

- The GHZ state has the property that if we trace over one qubit, then the reduced state is separable, i.e., it is a classical mixture of product states:

\[ \text{Tr}_3 |GHZ\rangle\langle GHZ| = \frac{1}{2} |00\rangle\langle 00| + \frac{1}{2} |11\rangle\langle 11|. \]
With more qubits, one can construct more interesting entangled states. For example, with three qubits we have [Dur et al ’00]

\[ |GHZ\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle). \]

\[ |W\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle). \]

The GHZ state has the property that if we trace over one qubit, then the reduced state is separable, i.e., it is a classical mixture of product states:

\[ \text{Tr}_3 |GHZ\rangle\langle GHZ| = \frac{1}{2} |00\rangle\langle 00| + \frac{1}{2} |11\rangle\langle 11|. \]

On the contrary, the W-state is not separable:

\[ \text{Tr}_3 |W\rangle\langle W| = \frac{1}{3} |00\rangle\langle 00| + \frac{2}{3} |\Psi^+\rangle\langle \Psi^+|, \quad |\Psi^+\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}. \]
Entanglement in Topological Quantum Field Theory

- We will study entanglement structure of a certain class of states in Chern-Simons theory.
We will study entanglement structure of a certain class of states in Chern-Simons theory.

The action for \( d = 3 \) Chern-Simons gauge theory at level \( k \) is given by [Witten '89]

\[
S_{CS}[A] = \frac{k}{4\pi} \int_M \text{Tr} \left( A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)
\]
Entanglement in Topological Quantum Field Theory

- We will study entanglement structure of a certain class of states in Chern-Simons theory.
- The action for $d = 3$ Chern-Simons gauge theory at level $k$ is given by [Witten ’89]

$$S_{CS}[A] = \frac{k}{4\pi} \int_M \text{Tr} \left( A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right)$$

- We will consider the theory for gauge groups $U(1)$ and $SU(2)$. 
Which states?

- The states we will consider are created by performing the Euclidean path integral of Chern-Simons theory on 3-manifolds $M_n$ with boundary consisting of $n$ copies of $T^2$. 

![Diagram of $M_3$ with $T^2$ boundary]
**Which states?**

- The states we will consider are created by performing the Euclidean path integral of Chern-Simons theory on 3-manifolds $M_n$ with boundary consisting of $n$ copies of $T^2$.

- For a given $M_n$ of this form, the path-integral of Chern-Simons theory on $M_n$ defines a state

$$|\Psi\rangle \in \mathcal{H}(T^2) \otimes \mathcal{H}(T^2) \otimes ... \otimes \mathcal{H}(T^2)$$

$$\Psi[A_{(0)}] = \int_{A|\Sigma=A_{(0)}} [DA] e^{iS_{CS}[A]}$$
Link-Complements

- Clearly, the choice of $M_n$ is far from unique. But there is a simple way to construct such manifolds.
Link-Complements

- Clearly, the choice of $M_n$ is far from unique. But there is a simple way to construct such manifolds.
- We start with a closed 3-manifold (i.e., a compact 3-manifold without boundary) $X$, and an $n$-component link in $X$

$$L^n = L_1 \cup L_2 \cup \cdots L_n$$

![Link Diagram]
Link-Complements

- Clearly, the choice of $M_n$ is far from unique. But there is a simple way to construct such manifolds.
- We start with a closed 3-manifold (i.e., a compact 3-manifold without boundary) $X$, and an $n$-component link in $X$

$$\mathcal{L}^n = L_1 \cup L_2 \cup \cdots L_n$$

- Let us take $X$ to be the 3-sphere $S^3$ for simplicity.
We then remove a tubular neighbourhood $N(L^n)$ of $L^n$ from $S^3$. 
We then remove a tubular neighbourhood $N(\mathcal{L}^n)$ of $\mathcal{L}^n$ from $S^3$. The manifold $M_n = S^3 - N(\mathcal{L}^n)$ is called the link complement of $\mathcal{L}^n$. 
Link-Complements

- We then remove a tubular neighbourhood $N(\mathcal{L}^n)$ of $\mathcal{L}^n$ from $S^3$.
- The manifold $M_n = S^3 - N(\mathcal{L}^n)$ is called the link complement of $\mathcal{L}^n$.
- It has the desired property, namely that

$$\partial M_n = T^2 \cup T^2 \cup \cdots \cup T^2.$$
We then remove a tubular neighbourhood $N(\mathcal{L}^n)$ of $\mathcal{L}^n$ from $S^3$.

The manifold $M_n = S^3 - N(\mathcal{L}^n)$ is called the link complement of $\mathcal{L}^n$.

It has the desired property, namely that

$$\partial M_n = T^2 \cup T^2 \cup \cdots \cup T^2.$$ 

The path-integral of Chern-Simons theory on the link-complement assigns to a link $\mathcal{L}^n$ in $S^3$ a state $|\mathcal{L}^n\rangle \in \mathcal{H}(T^2)^\otimes n$. 
The Hilbert space on a Torus

• Before moving further, let us recall some details about the Hilbert space of CS theory on a torus [Witten ’89].
The Hilbert space on a Torus

- Before moving further, let us recall some details about the Hilbert space of CS theory on a torus [Witten ’89].
- To construct a basis, we perform the path-integral on the “interior” solid torus, with a Wilson line in an integrable representation $R_j$ placed along the non-contractible cycle in the bulk. We call this state $|j\rangle$. 
The Hilbert space on a Torus

- Before moving further, let us recall some details about the Hilbert space of CS theory on a torus [Witten ’89].

- To construct a basis, we perform the path-integral on the “interior” solid torus, with a Wilson line in an integrable representation $R_j$ placed along the non-contractible cycle in the bulk. We call this state $|j\rangle$.

- The Hilbert space is finite dimensional for compact groups. (For $SU(2)$, the basis is labelled by spins $j = 0, \frac{1}{2}, \cdots \frac{k}{2}$.)
Now we can write the state prepared by path integration on the link complement $S^3 - \mathcal{L}^n$ in this basis as:

$$|\mathcal{L}^n\rangle = \sum_{j_1, \ldots, j_n} C_{\mathcal{L}^n}(j_1, j_2, \ldots, j_n) |j_1\rangle \otimes |j_2\rangle \cdots \otimes |j_n\rangle$$

These are called colored link invariants. (For $G = SU(2)$ they are called colored Jones polynomials.)
Back to Link complements

- Now we can write the state prepared by path integration on the link complement $S^3 - \mathcal{L}^n$ in this basis as:

$$|\mathcal{L}^n\rangle = \sum_{j_1,\ldots,j_n} C_{\mathcal{L}^n}(j_1, j_2, \cdots j_n) |j_1\rangle \otimes |j_2\rangle \cdots \otimes |j_n\rangle$$

- A little bit of thought shows that

$$C_{\mathcal{L}^n}(j_1, \cdots, j_n) = \left\langle \text{Tr}_{R_{j_1}^*} (e^{\oint L_1 A}) \cdots \text{Tr}_{R_{j_n}^*} (e^{\oint L_n A}) \right\rangle_{S^3}$$

These are called colored link invariants. (For $G = SU(2)$ they are called colored Jones polynomials.)
Back to Link complements

- Now we can write the state prepared by path integration on the link complement $S^3 - \mathcal{L}^n$ in this basis as:

$$|\mathcal{L}^n\rangle = \sum_{j_1, \ldots, j_n} C_{\mathcal{L}^n}(j_1, j_2, \cdots, j_n) |j_1\rangle \otimes |j_2\rangle \cdots \otimes |j_n\rangle$$

- A little bit of thought shows that

$$C_{\mathcal{L}^n}(j_1, \cdots, j_n) = \left\langle \text{Tr}_{R^*_{j_1}} (e^{\oint L_1 A}) \cdots \text{Tr}_{R^*_{j_n}} (e^{\oint L_n A}) \right\rangle_{S^3}$$

- These are called colored link invariants. (For $G = SU(2)$ they are called colored Jones polynomials.)
Entanglement Entropy

- We wish to study the entanglement structure of these states.
Entanglement Entropy

- We wish to study the entanglement structure of these states.
- So we partition the $n$-component link into an $m$-component sub-link $\mathcal{L}_A^m$ and the rest $\mathcal{L}_{\bar{A}}^{n-m}$.

\[
\mathcal{L}^n = \mathcal{L}_A^m \cup \mathcal{L}_{\bar{A}}^{n-m}
\]
Entanglement Entropy

- We wish to study the entanglement structure of these states.
- So we partition the $n$-component link into an $m$-component sub-link $\mathcal{L}_A^m$ and the rest $\mathcal{L}_{\bar{A}}^{n-m}$.

\[ \mathcal{L}^n = \mathcal{L}_A^m \cup \mathcal{L}_{\bar{A}}^{n-m} \]

- The reduced density matrix is obtained by tracing out $\bar{A}$:

\[ \rho_A = \frac{1}{\langle \mathcal{L}_n | \mathcal{L}_n \rangle} \text{Tr}_{\bar{A}} |\mathcal{L}_n\rangle \langle \mathcal{L}_n| \]
We wish to study the entanglement structure of these states.

So we partition the $n$-component link into an $m$-component sub-link $\mathcal{L}_A^m$ and the rest $\mathcal{L}_{\tilde{A}}^{n-m}$.

The reduced density matrix is obtained by tracing out $\tilde{A}$:

$$\rho_A = \frac{1}{\langle \mathcal{L}_n | \mathcal{L}_n \rangle} \text{Tr}_{\tilde{A}} |\mathcal{L}_n\rangle \langle \mathcal{L}_n|$$

The entanglement entropy is given by the Von Neumann entropy of this density matrix:

$$S_{EE} = -\text{Tr}_{\mathcal{L}_A} (\rho_A \ln \rho_A)$$
Example 0: The Unlink

- To see why these entropies are potentially interesting, we consider the simple but illuminating example of the unlink.
Example 0: The Unlink

- To see why these entropies are potentially interesting, we consider the simple but illuminating example of the unlink.
- So take $\mathcal{L}^n$ to be $n$ un-linked knots.
Example 0: The Unlink

- To see why these entropies are potentially interesting, we consider the simple but illuminating example of the unlink.
- So take $\mathcal{L}^n$ to be $n$ un-linked knots.

\[
|\mathcal{L}^n\rangle \propto |L_1\rangle \otimes |L_2\rangle \cdots \otimes |L_n\rangle.
\]

- It is well-known that the colored link-invariant of the unlink factorizes (up to an overall constant) [Witten '89]

Remark: The entanglement entropies are all framing independent.
**Example 0: The Unlink**

- To see why these entropies are potentially interesting, we consider the simple but illuminating example of the unlink.
- So take $\mathcal{L}^n$ to be $n$ un-linked knots.

![Unlinked knots](image)

- It is well-known that the colored link-invariant of the unlink factorizes (up to an overall constant) [Witten '89]

$$|\mathcal{L}^n\rangle \propto |L_1\rangle \otimes |L_2\rangle \cdots \otimes |L_n\rangle.$$  

- Consequently all the entanglement entropies vanish. This is our first hint that quantum entanglement is tied in with topological linking.

Remark: The entanglement entropies are all framing independent.
Example 0: The Unlink

- To see why these entropies are potentially interesting, we consider the simple but illuminating example of the unlink.
- So take \( \mathcal{L}^n \) to be \( n \) un-linked knots.

\[
|\mathcal{L}^n\rangle \propto |L_1\rangle \otimes |L_2\rangle \cdots \otimes |L_n\rangle.
\]

- It is well-known that the colored link-invariant of the unlink factorizes (up to an overall constant) [Witten '89]

- Consequently all the entanglement entropies vanish. This is our first hint that quantum entanglement is tied in with topological linking.
- **Remark:** The entanglement entropies are all framing independent.
Example 1: $G = U(1)_k$

- For $G = U(1)$, we can give a completely general formula for the entropy of a bi-partition of a general $n$-link $\mathcal{L}^n$:

$$\mathcal{L}_A^m = L_1 \cup L_2 \cup \cdots \cup L_m, \quad \mathcal{L}_{\overline{A}}^{n-m} = L_{m+1} \cup L_{m+2} \cup \cdots \cup L_n$$
Example 1: $G = U(1)_k$

- For $G = U(1)$, we can give a completely general formula for the entropy of a bi-partition of a general $n$-link $\mathcal{L}^n$:

$$\mathcal{L}^m_A = L_1 \cup L_2 \cup \cdots \cup L_m, \quad \mathcal{L}^{n-m}_{\bar{A}} = L_{m+1} \cup L_{m+2} \cup \cdots \cup L_n$$

- To state the answer for the entropy, we first define the linking matrix between the two sublinks:

$$G_{A,\bar{A}} = \begin{pmatrix} \ell_{1,m+1} & \ell_{2,m+1} & \cdots & \ell_{m,m+1} \\ \ell_{1,m+2} & \ell_{2,m+2} & \cdots & \ell_{m,m+2} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{1,n} & \ell_{2,n} & \cdots & \ell_{m,n} \end{pmatrix}$$
Example 1: $G = U(1)_k$

- For $G = U(1)$, we can give a completely general formula for the entropy of a bi-partition of a general $n$-link $\mathcal{L}^n$:

$$\mathcal{L}^m_A = L_1 \cup L_2 \cup \cdots \cup L_m, \quad \mathcal{L}^{n-m}_{\overline{A}} = L_{m+1} \cup L_{m+2} \cup \cdots \cup L_n$$

- To state the answer for the entropy, we first define the linking matrix between the two sublinks

$$G_{A,\overline{A}} = \begin{pmatrix}
\ell_{1,m+1} & \ell_{2,m+1} & \cdots & \ell_{m,m+1} \\
\ell_{1,m+2} & \ell_{2,m+2} & \cdots & \ell_{m,m+2} \\
\vdots & \vdots & \ddots & \vdots \\
\ell_{1,n} & \ell_{2,n} & \cdots & \ell_{m,n}
\end{pmatrix}$$

- Then, the entanglement entropy is given by

$$S_{EE} = \ln(k^m | \ker G_{A,\overline{A}}|)$$
Example 1: $G = U(1)_k$

- For $G = U(1)$, we can give a completely general formula for the entropy of a bi-partition of a general $n$-link $\mathcal{L}^n$:

$$\mathcal{L}_A^m = L_1 \cup L_2 \cup \cdots \cup L_m, \quad \mathcal{L}_{A}^{n-m} = L_{m+1} \cup L_{m+2} \cup \cdots \cup L_n$$

- To state the answer for the entropy, we first define the linking matrix between the two sublinks

$$G_{A,\bar{A}} = \begin{pmatrix}
\ell_{1,m+1} & \ell_{2,m+1} & \cdots & \ell_{m,m+1} \\
\ell_{1,m+2} & \ell_{2,m+2} & \cdots & \ell_{m,m+2} \\
\vdots & \vdots & \ddots & \vdots \\
\ell_{1,n} & \ell_{2,n} & \cdots & \ell_{m,n}
\end{pmatrix}$$

- Then, the entanglement entropy is given by

Claim

$$S_{EE} = \ln \left( \frac{k^m}{|\ker G_{A,\bar{A}}|} \right)$$
Minimal Genus Bound

- The entanglement entropy measures the obstruction to the splitting of a link between its sublink.
Minimal Genus Bound

- The entanglement entropy measures the obstruction to the splitting of a link between its sublink.
- Given an $n$-component link $\mathcal{L}^n \subset S^3$ and a bi-parition $\mathcal{L}^n = \mathcal{L}_A^m \cup \mathcal{L}_{\bar{A}}^{n-m}$, a separating surface $\Sigma$ is a connected, compact, oriented two-dimensional surface-without-boundary such that $\mathcal{L}_A^m$ is contained inside $\Sigma$, and $\mathcal{L}_{\bar{A}}^{n-m}$ is contained outside $\Sigma$.
The entanglement entropy measures the obstruction to the splitting of a link between its sublink.

Given an \( n \)-component link \( L^n \subset S^3 \) and a bi-parition \( L^n = L^m_A \cup L^{n-m}_{\overline{A}} \), a separating surface \( \Sigma \) is a connected, compact, oriented two-dimensional surface-without-boundary such that \( L^m_A \) in contained inside \( \Sigma \), and \( L^{n-m}_{\overline{A}} \) is contained outside \( \Sigma \).

The separating surface is not unique, but there is a unique such surface of \textit{minimal-genus}.
We claim the following general bound:

Claim

$$\min (g_\Sigma) \geq c_k S_{EE},$$

where $c_k$ is a positive constant which depends on the level $k$. 
We claim the following general bound:

\[
\min (g_\Sigma) \geq c_k S_{EE},
\]

where \(c_k\) is a positive constant which depends on the level \(k\).

This can be proved by cutting open the path-integral along separating surfaces:
We claim the following general bound:

\[ \min (g_\Sigma) \geq c_k S_{EE}, \]

where \(c_k\) is a positive constant which depends on the level \(k\).

This can be proved by cutting open the path-integral along separating surfaces:

This is reminiscent of the area-law bounds in tensor network descriptions of critical states [Nozaki et al ’12, Pastawski et al ’15].
Classifying Entanglement Structure of Links

- We begin with some definitions:

\[ |\text{GHZ}\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle) \]

\[ |\text{W}\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle) \]
We begin with some definitions:

- A link will be called **GHZ-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed (i.e., has a non-trivial entropy) but is separable (i.e., a convex combination of product states) on all the remaining sub-factors.

- A link will be called **W-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed, but is not always separable (i.e., a convex combination of product states) on all the remaining sub-factors.

\[ |\text{GHZ}\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle) \]

\[ |\text{W}\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle) \]
We begin with some definitions:

- A link will be called **GHZ-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed (i.e., has a non-trivial entropy) but is separable (i.e., a convex combination of product states) on all the remaining sub-factors.

- E.g.,

\[ |GHZ\rangle = \frac{1}{\sqrt{2}} (|000\rangle + |111\rangle) . \]

- A link will be called **W-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed, but is not always separable (i.e., a convex combination of product states) on all the remaining sub-factors.

- E.g.,

\[ |W\rangle = \frac{1}{\sqrt{3}} (|001\rangle + |010\rangle + |100\rangle) . \]
We begin with some definitions:

- A link will be called **GHZ-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed (i.e., has a non-trivial entropy) but is separable (i.e., a convex combination of product states) on all the remaining sub-factors.

- E.g.,
  \[
  \left|GHZ\rightangle = \frac{1}{\sqrt{2}} \left( \left|000\rightangle + \left|111\rightangle \right).
  \]

- A link will be called **W-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed, but is not always separable on the remaining sub-factors.
Classifying Entanglement Structure of Links

- We begin with some definitions:
- A link will be called **GHZ-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed (i.e., has a non-trivial entropy) but is separable (i.e., a convex combination of product states) on all the remaining sub-factors.
- E.g.,
  \[ |GHZ\rangle = \frac{1}{\sqrt{2}} \left( |000\rangle + |111\rangle \right). \]
- A link will be called **W-like** if the reduced density matrix obtained by tracing out any sub-factor is mixed, but is not always separable on the remaining sub-factors.
- E.g.,
  \[ |W\rangle = \frac{1}{\sqrt{3}} \left( |001\rangle + |010\rangle + |100\rangle \right). \]
Classifying Entanglement Structure of Links

- From the knot theory side, we will focus on two important topological classes of links, namely torus links and hyperbolic links.
Classifying Entanglement Structure of Links

- From the knot theory side, we will focus on two important topological classes of links, namely torus links and hyperbolic links.
- In fact, all non-split, alternating, prime links are either torus or hyperbolic [Menasco ’84].
Torus links

- Torus links are links which can be drawn on the surface of a torus without self-intersections.

\[ 2_1^2 \quad 4_1^2 \quad 6_3^3 \]
Torus links

- Torus links are links which can be drawn on the surface of a torus without self-intersections.

  \[
  2_1^2 \quad 4_1^2 \quad 6_3^3
  \]

- The following general result is true:

  Claim
  All torus links (with three or more components) have a GHZ-like entanglement structure. This can be proved by using the special structure of the colored link invariants of torus links [Labadista et al.'00].
Torus links

- Torus links are links which can be drawn on the surface of a torus without self-intersections.

\[ 2_1^2 \quad 4_1^2 \quad 6_3^3 \]

- The following general result is true:

**Claim**

All torus links (with three of more components) have a GHZ-like entanglement structure.
Torus links

- Torus links are links which can be drawn on the surface of a torus without self-intersections.

- The following general result is true:

Claim

All torus links (with three of more components) have a GHZ-like entanglement structure.

- This can be proved by using the special structure of the colored link invariants of torus links [Labadista et al’ 00].
Hyperbolic Links

- Hyperbolic links are links whose link-complements admit a hyperbolic structure.

\[ \begin{align*}
\text{Hyperbolic Links} & \\
\text{Hyperbolic links are links whose link-complements admit a hyperbolic structure.}
\end{align*} \]
Hyperbolic Links

- Hyperbolic links are links whose link-complements admit a hyperbolic structure.

- In contrast with torus links, we cannot study the entanglement structure of hyperbolic links in complete generality.
Hyperbolic Links

- Hyperbolic links are links whose link-complements admit a hyperbolic structure.

\[ 5_1^2 \quad 6_2^3 \]

- In contrast with torus links, we cannot study the entanglement structure of hyperbolic links in complete generality.

Conjecture

Hyperbolic links (with three or more components) have a W-like entanglement structure.
In order to test this, we need to use entanglement negativity [Peres ’96, Vidal & Werner ’02, Rangamani & Rota ’15].
Entanglement Negativity

- In order to test this, we need to use entanglement negativity \cite{Peres96,Vidal02,Rangamani15}.

- For a given (possibly mixed) density matrix $\rho$ on a bi-partite system, we define the partial transpose $\rho^\Gamma$:

  $$\langle j_1, j_2 | \rho^\Gamma | \tilde{j}_1, \tilde{j}_2 \rangle = \langle \tilde{j}_1, \tilde{j}_2 | \rho | j_1, j_2 \rangle.$$
Entanglement Negativity

• In order to test this, we need to use entanglement negativity [Peres ’96, Vidal & Werner ’02, Rangamani & Rota ’15].

• For a given (possibly mixed) density matrix \( \rho \) on a bi-partite system, we define the partial transpose \( \rho^\Gamma \):

\[
\langle j_1, j_2 | \rho^\Gamma | \tilde{j}_1, \tilde{j}_2 \rangle = \langle \tilde{j}_1, \tilde{j}_2 | \rho | j_1, j_2 \rangle.
\]

• Then the negativity is defined as

\[
\mathcal{N} = \frac{||\rho^\Gamma|| - 1}{2},
\]

where \( ||A|| = \text{Tr} \left( \sqrt{A^\dagger A} \right) \) is the trace norm.
A non-zero value of $\mathcal{N}$ is a sufficient (but not necessary) condition for the reduced density matrix to be non-separable.
A non-zero value of $\mathcal{N}$ is a sufficient (but not necessary) condition for the reduced density matrix to be non-separable.

We numerically computed the entanglement negativities for 20 3-component hyperbolic links.
We found in all the cases that the links had W-like entanglement. This provides some evidence that hyperbolic links generically have W-like entanglement.
Part 2: Machine Learning
The Volume conjecture

- For a knot $K$, let $J_{K,N}(q)$ be the colored Jones polynomial, where $N = 2j$ is the color and

$$q = e^{\frac{2\pi i}{k+2}}.$$
The Volume conjecture

- For a knot $K$, let $J_{K,N}(q)$ be the colored Jones polynomial, where $N = 2j$ is the color and
  \[ q = e^{\frac{2\pi i}{k+2}}. \]
- The volume conjecture states that: [Kashaev’97, Murakami’01, Gukov’05]
The Volume conjecture

- For a knot $K$, let $J_{K,N}(q)$ be the colored Jones polynomial, where $N = 2j$ is the color and
  \[ q = e^{\frac{2\pi i}{k+2}}. \]

- The volume conjecture states that: [Kashaev’97, Murakami’01, Gukov’05]

Volume conjecture

\[
\lim_{N \to \infty} \frac{2\pi \log |J_{K,N}(e^{\frac{2\pi i}{N}})|}{N} = \text{Vol}(K).
\]
The Volume conjecture

- For a knot $K$, let $J_{K,N}(q)$ be the colored Jones polynomial, where $N = 2j$ is the color and
  \[ q = e^{\frac{2\pi i}{k+2}}. \]

- The volume conjecture states that: [Kashaev’97, Murakami’01, Gukov’05]

Volume conjecture

\[
\lim_{N \to \infty} \frac{2\pi \log |J_{K,N}(e^{\frac{2\pi i}{N}})|}{N} = \text{Vol}(K) .
\]

- Note that the double-scaling limit $k \to \infty$, $N \to \infty$ with $N/k = 1$ is a weak-coupling but strong back-reaction limit.
The Volume conjecture

- For a knot $K$, let $J_{K,N}(q)$ be the colored Jones polynomial, where $N = 2j$ is the color and
  \[ q = e^{\frac{2\pi i}{k+2}}. \]

- The volume conjecture states that: [Kashaev’97, Murakami’01, Gukov’05]

Volume conjecture

\[
\lim_{N \to \infty} \frac{2\pi \log |J_{K,N}(e^{\frac{2\pi i}{N}})|}{N} = \text{Vol}(K).
\]

- Note that the double-scaling limit $k \to \infty$, $N \to \infty$ with $N/k = 1$ is a weak-coupling but strong back-reaction limit.

- In this limit, the colored Jones polynomial knows about the hyperbolic volume.
Generalized Volume conjecture?

- This begs the question: does the ordinary Jones polynomial \((N = 1)\) also satisfy some version of the volume conjecture?
Generalized Volume conjecture?

- This begs the question: does the ordinary Jones polynomial ($N = 1$) also satisfy some version of the volume conjecture?
- This is motivated by an observation due to Nathan Dunfield: if one plots the volume $v$ vs. $\log|J_K(-1)|$, there seems to be an approximately linear dependence [Figure taken from Dunfield’s webpage]
Generalized Volume conjecture?

- This begs the question: does the ordinary Jones polynomial \((N = 1)\) also satisfy some version of the volume conjecture?
- This is motivated by an observation due to Nathan Dunfield: if one plots the volume \(v\) vs. \(\log |J_K(-1)|\), there seems to be an approximately linear dependence [Figure taken from Dunfield’s webpage]

- But this only seems to work for alternating knots, and fails badly for non-alternating knots.
Generalized Volume conjecture?

- Another hint is the “volumish” bound [Dasbach, Lin ’04], which bounds the volume in terms of coefficients of the Jones polynomial:
Generalized Volume conjecture?

- Another hint is the “volumish” bound [Dasbach, Lin ’04], which bounds the volume in terms of coefficients of the Jones polynomial:

\[ J_K(q) = a_n q^n + a_{n+1} q^{n+1} + \cdots + a_{m-1} q^{m-1} + a_m q^m \]

\[ 2v_0 \left( \max(|a_{m-1}|, |a_{n+1}|) - 1 \right) < \text{Vol} < 10v_0 \left( |a_{m-1}| + |a_{n+1}| - 1 \right) \]
Generalized Volume conjecture?

• Another hint is the “volumish” bound [Dasbach, Lin ’04], which bounds the volume in terms of coefficients of the Jones polynomial:

\[ J_K(q) = a_n q^n + a_{n+1} q^{n+1} + \cdots a_{m-1} q^{m-1} + a_m q^m \]

\[ 2v_0 \left( \max(|a_{m-1}|, |a_{n+1}|) - 1 \right) < \text{Vol} < 10v_0 \left( |a_{m-1}| + |a_{n+1}| - 1 \right) \]

• But this bound is not very tight:

Further, the bounds are only proven for alternating knots.
Machine Learning

- Machine learning is the perfectly suited to this type of pattern-recognition problem.
Machine Learning

- Machine learning is the perfectly suited to this type of pattern-recognition problem.
- A neural network is a function which is constructed by training on several examples.
Machine Learning

- Machine learning is the perfectly suited to this type of pattern-recognition problem.
- A neural network is a function which is constructed by training on several examples.
- Suppose that we have a dataset $\mathcal{D} = \{J_1, J_2, \ldots, J_m\}$, and to every element of $\mathcal{D}$, there is an associated element in another set $\mathcal{V}$:

$$A : \{J_1, J_2, \ldots, J_m\} \mapsto \{v_1, v_2, \ldots, v_m\} \subset \mathcal{V}.$$
Machine Learning

- Machine learning is the perfectly suited to this type of pattern-recognition problem.
- A neural network is a function which is constructed by training on several examples.
- Suppose that we have a dataset $\mathcal{D} = \{J_1, J_2, \ldots, J_m\}$, and to every element of $\mathcal{D}$, there is an associated element in another set $\mathcal{V}$:

$$A : \{J_1, J_2, \ldots, J_m\} \mapsto \{v_1, v_2, \ldots, v_m\} \subset \mathcal{V}.$$  

- In our case, the $J_i$ are the Jones polynomials of knots, and the $v_i$ are the volumes of those knots.
Machine Learning

- Machine learning is the perfectly suited to this type of pattern-recognition problem.
- A neural network is a function which is constructed by training on several examples.
- Suppose that we have a dataset \( \mathcal{D} = \{J_1, J_2, \ldots, J_m\} \), and to every element of \( \mathcal{D} \), there is an associated element in another set \( \mathcal{V} \):
  \[
  A : \{J_1, J_2, \ldots, J_m\} \mapsto \{v_1, v_2, \ldots, v_m\} \subset \mathcal{V}.
  \]
- In our case, the \( J_i \) are the Jones polynomials of knots, and the \( v_i \) are the volumes of those knots.
- A neural network \( f_\theta \) is a function (with an \textit{a priori} chosen architecture) which is designed to approximate the associations \( A \) efficiently.
Neural Net architecture

- The architecture of the neural net looks as follows:

\[ f^\theta(\vec{J}_K) = \sum_i \sigma(W^2_{\theta} \sigma(W^1_{\theta} \cdot \vec{J}_K + \vec{b}_1_{\theta}) + \vec{b}_2_{\theta})_i, \]

where \( W^j_{\theta} \) and \( \vec{b}^j_{\theta} \) are the weight matrices and bias vectors, and \( \sigma \) is a non-linear activation function.

The intermediate vectors are taken to be 100-dimensional.

The non-linear function is the logistic sigmoid:

\[ \sigma(x) = \frac{1}{1 + e^{-x}}. \]
Neural Net architecture

- The architecture of the neural net looks as follows:

  ![Neural Net Diagram]

- We encode the Jones polynomial in a vector $\vec{J}_K = (a_n, \cdots, a_m)$, and feed it to the network:

  $$f_\theta(\vec{J}_K) = \sum_i \sigma \left( W^2_\theta \cdot \sigma (W^1_\theta \cdot \vec{J}_K + \vec{b}^1_\theta) + \vec{b}^2_\theta \right)^i,$$

  where $W^j_\theta$ and $\vec{b}^j_\theta$ are the weight matrices and bias vectors, and $\sigma$ is a non-linear activation function.
**Neural Net architecture**

- The architecture of the neural net looks as follows:

![Neural Net Diagram]

- We encode the Jones polynomial in a vector $\vec{J}_K = (a_n, \cdots, a_m)$, and feed it to the network:

$$f_\theta(\vec{J}_K) = \sum_i \sigma \left( W_2^i \cdot \sigma (W_1^i \cdot \vec{J}_K + b_1^i) + b_2^i \right)$$

where $W_2^i$ and $b_2^i$ are the weight matrices and bias vectors, and $\sigma$ is a non-linear activation function.

- The intermediate vectors are taken to be 100-dimensional.
Neural Net architecture

- The architecture of the neural net looks as follows:

- We encode the Jones polynomial in a vector $\vec{J}_K = (a_n, \cdots, a_m)$, and feed it to the network:

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

where $W_{\theta}^j$ and $\vec{b}_{\theta}^j$ are the weight matrices and bias vectors, and $\sigma$ is a non-linear activation function.

- The intermediate vectors are taken to be 100-dimensional.

- The non-linear function is the logistic sigmoid: $\sigma(x) = \frac{1}{1+e^{-x}}$. 

Onkar Parrikar (UPenn)  Math/Strings Regional Meeting
The neural net is taught the associations on the training set by tuning the internal parameters $\theta$ to approximate $A$ as closely as possible on $T$, by minimizing a suitable loss function:

$$h(\theta) = \sum_{i \in T} ||f_\theta(J_i) - v_i||_2^2.$$
\textbf{N} = \text{NetChain}\{\text{DotPlusLayer}[100],
\text{ElementwiseLayer}[\text{LogisticSigmoid}], \text{DotPlusLayer}[100],
\text{ElementwiseLayer}[\text{LogisticSigmoid}], \text{SummationLayer}[]\},
"Input" \rightarrow \{17\};

- For the network to learn \(A\), we divide the dataset \(\mathcal{D}\) into two parts: a training set, \(T = \{J_1, J_2, \ldots, J_n\}\) chosen at random from \(\mathcal{D}\), and its complement, \(T^c = \{J'_1, J'_2, \ldots, J'_{m-n}\}\).
\[
N = \text{NetChain}\{\text{DotPlusLayer}[100], \\
\text{ElementwiseLayer}[\text{LogisticSigmoid}], \text{DotPlusLayer}[100], \\
\text{ElementwiseLayer}[\text{LogisticSigmoid}], \text{SummationLayer}[]\}, \\
"\text{Input}" \to \{17\};
\]

- For the network to learn \( A \), we divide the dataset \( \mathcal{D} \) into two parts: a training set, \( T = \{J_1, J_2, \ldots, J_n\} \) chosen at random from \( \mathcal{D} \), and its complement, \( T^c = \{J'_1, J'_2, \ldots, J'_{m-n}\} \).
- The neural net is taught the associations on the training set by tuning the internal parameters \( \theta \) to approximate \( A \) as closely as possible on \( T \), by minimizing a suitable loss function:

\[
h(\theta) = \sum_{i \in T} ||f_\theta(J_i) - v_i||^2.
\]
Comparing with the true volumes

- Finally, we assess the performance of the trained network by applying it to the unseen inputs $J'_i \in T^c$ and comparing $f_\theta(J'_i)$ to the true answers $v'_i = A(J'_i)$. 

By training on as little as 10% of data, the network can predict the volume with an accuracy of 97.5%, for both alternating and non-alternating knots.
Comparing with the true volumes

- Finally, we assess the performance of the trained network by applying it to the unseen inputs $J'_i \in T^c$ and comparing $f_\theta(J'_i)$ to the true answers $v'_i = A(J'_i)$.

- By training on as little as 10% of data, the network can predict the volume with an accuracy of 97.5%, for both alternating and non-alternating knots.
The robustness of the network suggests that there might be a generalized volume conjecture which relates the hyperbolic volume to the Jones polynomial, i.e., the weak-backreaction but possibly strong-coupling regime.
Summary

- The robustness of the network suggests that there might be a generalized volume conjecture which relates the hyperbolic volume to the Jones polynomial, i.e., the weak-backreaction but possibly strong-coupling regime.
- Neural networks might provide a novel and useful technique to search for mathematical relationships between topological invariants.