Quantum Walk Neural Networks

Stefan Dernbach*
Arman Mohseni-Kabir*
Don Towsley*
Siddharth Pal**

*University of Massachusetts Amherst
**Raytheon BBN

This document does not contain technology or technical data controlled under either the U.S. International Traffic in Arms Regulations or the U.S. Export Administration Regulations
Motivating Task: Social Networks

Social networks:
- Extremely Feature Rich
- Shape of the graph plays obvious role

Tasks:
- Community Detection
- Link Prediction
- Node Classification/Regression
Motivating Task: Molecules

Molecular Graphs:
- Node/Edge Features
- Structure important to behavior

Task:
- Graph Labeling
Contributions

2 Different Paradigms to view our work:

- A new neural network architecture based off of a quantum algorithm
- A classical method to learn a quantum operator

Directly learn a diffusion matrix for a graph specific to a task.

Outperforms state of the art graph signal processing methods for a temperature prediction task.
Graph Notation

\[ G = \{ \mathcal{V}, \mathcal{E} \} \]

\[ A_{i,j} = \begin{cases} 
1 & \text{if } (v_i, v_j) \in \mathcal{E} \\
0 & \text{otherwise}
\end{cases} \]

\[ D_{i,i} = \sum_j A_{i,j} \]
Random Walk

\[ x(t) = W^\top x(t-1) \]

\[ W = \begin{cases} 
  D^{-1}A & \text{Classical Random Walk} \\
  (D + \alpha I)^{-1}(A + \alpha I) & \text{Lazy Random Walk} \\
  (D + NaI)^{-1}(A + \alpha 1) & \text{Random Walk with Teleports} 
\end{cases} \]
Diffusion Convolutional Neural Networks

\[ y = h \left( \sum_{i} \Theta_i (W_i^i)^\top X + b \right) \]
Quantum Walks

A particle on a graph can be described by a product state of:

1. A position state

\[ \{ |v\rangle : v \in V \} \]

2. A spin state

\[ \{ |i\rangle : i \in 1...d_{max} \} \]
Quantum Walks on a lattice

We Define Two Operators

1. A Coin Operator

\[ C|v, i\rangle \]

2. A Shift Operator

\[ S = |\uparrow\rangle\langle\uparrow| \otimes \sum_i |v + 1\rangle\langle v| \]

\[ + |\downarrow\rangle\langle\downarrow| \otimes \sum_i |v - 1\rangle\langle v| \]

Quantum Walk

\[ U = S(C_v \otimes I_N) \]
Quantum Walks on a 2D Torus
Quantum Walks on a 2D Torus
Quantum Walks on Arbitrary Graphs

A particle on a graph can be described by a product state of:

1. A position state

\[ \{ |v\rangle : v \in \mathcal{V} \} \]

2. A spin state

\[ \{ |i\rangle : i \in 1 \ldots d_{\text{max}} \} \]
Quantum Walks on a lattice

We Define Two Operators

1. A Coin Operator

\[ C|v, i\rangle \]

2. A Swap Operator

\[ S|u, v, A_{uv}\rangle = \begin{cases} |u, v, A_{uv}\rangle & A_{uv} = 0 \\ |v, u, A_{uv}\rangle & A_{uv} = 1 \end{cases} \]

Quantum Walk

\[ U = S(C_v \otimes I_N) \]
A Quantum Walk: Initial State
A Quantum Walk: Coin Operator
A Quantum Walk: Swap Operator
Transforming Quantum Walks into a Classical Algorithm
Neural Network Layout

Initial Amplitudes: $\Phi^{(0)}$

Feature Matrix: $X$

Graph:

Output: $Y$
QWNNs

Quantum Step Layer

Input: $\Phi^{(t)}$

For each node $v_i$:

$\hat{\Phi}_i^{(t+1)} \leftarrow \Phi_i^{(t)} \cdot C_i$

$\Phi^{(t+1)} \leftarrow \hat{\Phi}^{(t+1)} \cdot S$

Diffusion Layer

Input: $\Phi^{(T)}$, $X$

$W = \sum_s \Phi_s^{(T)} \odot \Phi_s^{(T)}$

$y = h(W^T X + b)$
Comparison to Diffusion Convolution Layer

**DCNN:** \[ y = h \left( \sum_i \Theta_i (W^i)^\top X + b \right) \]

**QWNN:** \[ y = h (W^\top X + b) \]
Experiment

409 Weather Stations

Graph Formed for K-Nearst Neighbors with K=8

High Temperature Readings Daily for 1 year

Task: Predict the Following Days Temperature
# Test Results

<table>
<thead>
<tr>
<th></th>
<th>Walk Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Spectral</td>
<td>7.32</td>
</tr>
<tr>
<td>Diffusion Conv.</td>
<td>7.85</td>
</tr>
<tr>
<td><strong>QW Amp+Coin</strong></td>
<td>11.93</td>
</tr>
<tr>
<td>QW Amp</td>
<td>12.01</td>
</tr>
<tr>
<td>QW Coin</td>
<td>14.33</td>
</tr>
</tbody>
</table>
Discussion and Questions

Machine Learning Advantages from Quantum Walks:

○ Can tune performance to different sections of the graph
○ Can handle data from specific nodes differently

Quantum Advantages from NN

○ Can be implemented today (albeit slowly)
QTML 2017 (Nov 6-8)

Quantum Techniques in Machine Learning

qtml2017.di.univr.it
- Program
- Slides
- Attendees
Quantum Reinforcement Learning

• “Progress in Quantum Reinforcement Learning.” Vedran Dunjko (Invited Talk).

Quantum Random Walks

• “Quantum State Engineering Using One-Dimensional Discrete-Time Quantum Walks.” Luca Innocenti et al.
Quantum Neural Networks

- “Quantum Autoencoders for Efficient Compression of Quantum Data.” Jonathan Romero, Jonathan Olson, and Alan Aspuru-Guzik
- “Neural Networks Quantum States, String-Bond States and Chiral Topological States.” Ivan Glasser.
- “Quantum Error Correction with Recurrent Neural Networks.” P.Baireuther, et. al.
Quantum Tensor Networks

- “Towards Quantum Machine Learning with Tensor Networks.” William Huggins and Miles Stoudenmire.
- “Quantum Entanglement Simulators Inspired by Tensor Networks.” Shi-Ju Ran.
Others

Looking Towards the Future

- “Prospects in Quantum Machine Learning.” Seth Lloyd (Invited Talk)
- “Towards Quantum-Assisted Artificial Intelligence.” Peter Wirrek (Invited Talk).
Opportunities and Challenges for Quantum-Assisted Machine Learning in Near-Term Quantum Computers

- Potential impact across social and natural sciences, engineering, and more
- Bayesian inference
- Deep learning
- Probabilistic programming
- Others...

Hypothesis: intractable sampling problems enhanced by quantum sampling
Best Search Directions for Near-Term Realizable Quantum ML

i. Focus on problems that are currently hard and intractable for the ML community

ii. Focus on datasets with potentially intrinsic quantum-like correlations, making quantum computers indispensable

iii. Focus on hybrid quantum algorithms that can be easily integrated in the intractable step of the ML algorithmic pipeline
Opportunities in QAML

Sampling Applications

- Exact Inference, learning, and model selection are intractable in all but the most trivial topologies
  - RBMs vs. Bms
- MCMC suffers from slow-mixing in complex distributions
  - Multi-modal distributions
- Quantum Devices and quantum gibbs sampling have the potential to sample efficiently from complex probability distributions
Opportunities in QAML

Quantum-Like Correlations

- Interference
- Entanglement
- Contextuality
- Etc.
Quantum-Like Correlations: An Example

Experiment Set Up

- Contestants play a game with equal probability to win $200 or lose $100
- After the first round they are asked if they want to gamble again.
- If the player is told the result of the first gamble:
  - $P(G|W)=0.69$
  - $P(G|L)=0.59$
- If the player is not told the result:
  - $P(G)=0.39$
- The results violate the law of total probability:
  - $P(G) = P(G|W) \ P(W) + P(G|L) \ P(L)$
- Can be explained as an interference phenomenon allowing a much better fit with a quantum model.

Challenges in QAML

- Classical and Quantum Model Compatibility
  - Specifying temperature for Gibbs Distribution

- Green Line: Classical Contrastive Divergence

- Blue line: $T$, the instance dependent effective temperature, is estimated each learning iteration

- Red line: Physical temperature of the device is used for $T$. 
Limited Qubit Connectivity

- Required qubit-to-qubit interactions not available in the device leads to additional overhead.
- Quantum Annealer: Produce an embedding in the physical layout increasing the required number of qubits.
- Gate model: Overhead comes from additional number of swaps.
- Similar Issues for some solutions to continuous variables
  - May lose any advantage due to readout time
Quantum Assisted Helmholtz Machine
Quantum Assisted Helmholtz Machine

Generative Model:

\[ P(v) = \sum_u P(v|u)P_QC(u) \]

Samples from the quantum device are described by the diagonal elements of a parametrized density matrix:

\[ P_QC(u) = \langle u|\hat{\rho}|u \rangle \]

Example: Density as quantum Gibbs distribution

\[ \hat{\rho} = e^{-\beta H} / Z \]
Quantum Assisted Helmholtz Machine
Quantum Assisted Helmholtz Machine