Neural Belief-Propagation Decoders for Quantum Error-Correcting Codes

Yehua Liu
Institut Quantique
In collaboration with Prof. David Poulin
INTRIQ 2018
Outline

• Classical error correction and graphical representation
• Belief-propagation decoders
• Quantum codes and the failure of BP
• Neural belief propagation (NBP)
• Results and conclusions
Error correction stabilizes information science.
Error Correction

MacKay, “Information Theory, Inference, and Learning Algorithms”
Repetition Code and Its Tanner Graph

Encoding 1 *logical* bit into 3 *physical* bits:

- \(0 \rightarrow 000\)
- \(1 \rightarrow 111\)

Code space \{000, 111\} is defined by parity checks in *Tanner graph*:
Code Space
Error & Syndrome
The Decoding Problem

To find the *most probable* error that is *consistent* with the syndrome.
Quantifying “Probable” by Factor Graph

\[
\begin{align*}
    p(0) &= 1 - p_{\text{err}} \\
    p(1) &= p_{\text{err}} \\
    f(x, y; s) &= \begin{cases} 
                    1, & \text{if } s = x + y, \\
                    0, & \text{otherwise.}
                  \end{cases} \\
    \rho(x, y, z; s, t) &= Z^{-1} p(x)p(y)p(z)f(x, y; s)f(y, z; t)
\end{align*}
\]

Statistical \textit{inference} in a graph:

\begin{itemize}
    \item Fix part of the graph and find the most probable pattern of the \textit{missing} part.
    \item Important paradigm in many research areas:
        Information science, machine Learning, physics, economics...
\end{itemize}
Statistical Physics Perspective

\[ \rho (\tilde{e}; \tilde{s}) = Z^{-1} e^{-E(\tilde{e}; \tilde{s})/T} \]

- The lower the \( p_{err} \), the lower the temperature \( T \).
- Decoding: finding the ground state (NP-complete).
- If the ground state is \textit{not degenerate}, thermal average in low \( T \) recovers it:

\[ \tilde{e}_{GS} \approx \langle \tilde{e} \rangle_T. \]

Mean-field theory for thermal average?
Belief Propagation

Iteration $t = 1, 2, \ldots, T - 1$

\[
\mu_{v \rightarrow c}^{(t+1)} = l_v + \sum_{c_i \in \mathcal{N}(v)/c} \mu_{c_i \rightarrow v}^{(t)}
\]

\[
\mu_{c \rightarrow v}^{(t+1)} = (-1)^s c 2 \tanh^{-1} \prod_{v_i \in \mathcal{N}(c)/v} \tanh \frac{\mu_{v_i \rightarrow c}^{(t)}}{2}
\]

Output at $t = T$

\[
\mu_v = l_v + \sum_{c \in \mathcal{N}(v)} \mu_{c \rightarrow v}^{(T)}
\]

J. Pearl, 1982
M. Mezard, G. Parisi, and M. Virasoro, 1986
\[ \mu_{v \rightarrow c}^{(t+1)} = l_v + \sum_{c' \in N(v)/c} \mu_{c' \rightarrow v}^{(t)} \]

\[ l_v = \log\left(\frac{1 - p_{err}}{p_{err}}\right) \text{ is the prior.} \]
\[
\mu_{c \rightarrow v}^{(t+1)} = (-1)^s c \cdot 2 \tanh^{-1} \prod_{v \in N(c) \setminus v} \tanh \frac{\mu_{v \rightarrow c}^{(t)}}{2}
\]
\[ \mu_v = l_v + \sum_{c \in \mathcal{N}(v)} \mu_{c \rightarrow v}^{(T)} \]

\( e_v \) is inferred to be \( \Theta(-\mu_v) \).
Low-Density Parity-Check Codes

• Random Tanner graph with low variable and check degrees
  • Edges are sparse.
• BP decoders are both efficient and accurate
  • BP is exact for tree graphs.
  • BP also works well for graphs with (long) loops.
• Cornerstone of (classical) modern coding theory
• Widely used in wireless communication
## Classical vs. Quantum Codes

<table>
<thead>
<tr>
<th></th>
<th>Classical Codes</th>
<th>Quantum Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Error</strong></td>
<td>( \hat{e} = {e_1, e_2, e_3} )</td>
<td>( \hat{e} = X_1^{e_1}X_2^{e_2}X_3^{e_3}Z_1^{e_4}Z_2^{e_5}Z_3^{e_6} )</td>
</tr>
<tr>
<td><strong>Check</strong></td>
<td>( \vec{H}_1 = {1,1,0}, \vec{H}_2 = {0,1,1} )</td>
<td>( \hat{S}_1 = ZZI, \hat{S}_2 = IZZ ).</td>
</tr>
<tr>
<td><strong>Syndrome</strong></td>
<td>( s_i = \vec{H}_i \cdot \hat{e} )</td>
<td>( s_i = [\hat{S}_i, \hat{e}] = \hat{S}_i \cdot \hat{M} \hat{e} )</td>
</tr>
<tr>
<td><strong>Successful decoding</strong></td>
<td>( \hat{e}_{inf} = \hat{e} )</td>
<td>( [\hat{S}<em>i, \hat{e}\hat{e}</em>{inf}] = [\hat{L}<em>i, \hat{e}\hat{e}</em>{inf}] = 0 )</td>
</tr>
<tr>
<td><strong>Constraint</strong></td>
<td>None</td>
<td><strong>All checks commute</strong></td>
</tr>
</tbody>
</table>
Difficulties

For quantum codes:

1. Extensive short loops in the Tanner graph
2. Extensive degeneracy in the ground state

*BP decoders break down.* (D. Poulin & Y. Chung, 2008)

Adapting BP to the quantum regime.

See review paper: Z. Babar *et al.*, 2015
New Perspective on BP Equations

Iteration

\[
\mu_{v \rightarrow c}^{(t+1)} = l_v b_v^{(t)} + \sum_{c' \in \mathcal{N}(v)/c} \mu_{c' \rightarrow v}^{(t)} w_{c'v,vc}^{(t)}
\]

\[
\log \tanh \frac{\mu_{c \rightarrow v}^{(t+1)}}{2} = i \pi s_c b_c^{(t)} + \sum_{v' \in \mathcal{N}(c)/v} \log \tanh \frac{\mu_{v' \rightarrow c}^{(t)}}{2} w_{v'c, cv}^{(t)}
\]

Output

\[
\mu_v = l_v b_v^{(T)} + \sum_{c \in \mathcal{N}(v)} \mu_{c \rightarrow v}^{(T)} w_{cv,v}^{(T)}
\]

*Weighted sum, bias & nonlinearity: feed-forward neural network!*

E. Nachmani, Y. Be'ery, and D. Burshtein, 2016
Adapting BP by Machine Learning

- **Initialize** a neural network to the classical BP.
- **Train** the neural network with “quantum data”.
- **BP learns** to adapt to the quantum regime.

*YHL & D. Poulin, to submit, 2018*
Neural Belief Propagation (NBP)

\[ t \xrightarrow{\mu_{c \rightarrow v}} CVVC \xrightarrow{\mu_{v \rightarrow c}} VCCV \xrightarrow{\mu_{c \rightarrow v}} t+1 \]

\[ \times N_c \]

\textbf{Marginalization & Multi-Loss}

\textbf{Residual Connection}

\textbf{Prior}

\textbf{Syndrome}

\textbf{YHL & D. Poulin, to submit, 2018}
Guiding the Learning with Loss Function

• Gradient-based machine learning:
\[ \Delta w \propto -\frac{\partial \mathcal{L}}{\partial w}, \Delta b \propto -\frac{\partial \mathcal{L}}{\partial b}. \]

• \( \mathcal{L}(w, b) \) is the **loss function** that guides the learning.

• A suitable choice for decoding quantum codes?

\[
\mathcal{L} = \sum_{i=1}^{N_S} f \left( \hat{e}_{tot} \cdot M \tilde{S}_i \right) + \sum_{i=1}^{N_L} f \left( \hat{e}_{tot} \cdot M \tilde{L}_i \right) \\
\hat{e}_{tot} = \sigma(\mu) + \hat{e} \\
f(x) = |\sin(\pi x/2)|
\]

YHL & D. Poulin, to submit, 2018
Toric Code: Z Checks
Toric Code: X Checks
Toric Code

Rate $= \frac{1}{L^2}$

YHL & D. Poulin, to submit, 2018
Toric Code: Development of Threshold

YHL & D. Poulin, to submit, 2018
Quantum LDPC Codes with High Rates

Rate = 0.125

Rate ~ 0.2

YHL & D. Poulin, to submit, 2018
Conclusions

• The belief propagation algorithm, widely used in many research areas, has a neural-network representation.

• Training greatly enhances the performance of BP decoders for quantum error-correcting codes.

• Training BP with “quantum data” might solve other intricate problems in quantum many-body physics.

Thank you for your attention!

YHL & D. Poulin, to submit, 2018