Towards Extending Ising Model Coherence Times with Deep Recurrent Neural Networks

Authors: Yash Anand, Mark Gouzoulis, Linfred Kingston, Jander Kugelman, Jonathan Li, Duy Pham, Eashan Siddalingaiah, William Taylor, George Witt
Mentor: Mohammed Hafezi || Librarian: Jordan Sly

Introduction
For predicting decoherence of transverse Ising Models, we evaluate:
- Deep Recurrent Neural Networks (RNNs) and Transformers
- 2 - 10 spin setups with up to 500 timesteps
- Different 'program state' representations for the NNs

We further provide, with the goal of extending quantum computer coherence times:
- Sample predictions for the models
- Comparison of efficacy of different approaches for prediction

Mathematical Formulation
\[ H = \sum_i^{-N} \left( \frac{1}{2} J_{ij} \sigma_i^z \sigma_j^z \right) + \sum_{\sigma=\uparrow,\downarrow} \left( \sum_i \left( \frac{1}{2} J_{ij} \sigma_i^\sigma \sigma_{j+1}^\sigma \right) \right) \]
\[ \psi = \bigotimes_{i=1}^{N} \psi_i \]

Transformers

LSTM Proof of Concept
LSTM: 2 Spins

Oscillation around local minima suggests better tuning needed for learning rate schedules.

Initial Analysis

C 2, MSE: 0.000029651432789030334

Quick local convergence within only 20 minutes of training with full state information.

Scaling Transformers

Transformer Mean Square Error

Loss catapult due to tuning observed!

Short training transients demonstrate excellent convergence!

Echo State Net Tuning

Progress:
- Implemented ESN baseline, LSTM proof of concept, and scaled up to SOTA transformer models
- Implemented sweeps across parameters for different models
- Implementing encoder/decoder to strengthen LSTM prediction

Takeaways

Overcome Issues:
- Improved models by generating full information simulated data
- Previous work was less accurate for small scale systems, improved with RNNs
- New work shows more lower MSE predictions after hyperparameter tuning

Research Hypothesis
Attention-based SOTA models should be able to predict the decoherence of a transverse interacting Ising Model with 10 spins over 500 discrete steps to achieve a 1E-3 MSE.

Summary of Data Collection
Simulation data was generated using Qutip to evolve N spins over 500 discrete time steps. Tests were performed on both simple and transverse Hamiltonians. The mean square error between predicted states and Qutip modeled data is used to compare efficacy of models over small time-clusters.

Overcome Issues:
- Improved models by generating full information simulated data
- Previous work was less accurate for small scale systems, improved with RNNs
- New work shows more lower MSE predictions after hyperparameter tuning

Future Research Goals
- Extend model prediction to 2D lattice
- Ablation testing across LSTM / ESN / Transformer models for cost efficiency
- Observable data tests for practical application

References:

Acknowledgments: We’d like to extend a big thank you to our Gemstone program leaders, Dr. Lovell and Dr. Lansverk, our mentor Dr. Hafezi and Librarian Jordan Sly. This project would not be possible without them!