Efficient quantum circuit compression using Reinforcement Learning

Noah Berthusen[1,2], Peter P. Orth[1,3]

[1]Ames Laboratory, Ames, Iowa, USA
[2]Department of Electrical and Computer Engineering, Iowa State University, Ames, Iowa, USA
[3]Department of Physics and Astronomy, Iowa State University, Ames, Iowa, USA

Iowa State University
of Science and Technology

Abstract

Computations by the current generation of noisy intermediate scale quantum (NISQ) computers are often plagued by errors such as decoherence and cross talk. Such errors severely limit the depth of NISQ quantum circuits, yet many quantum algorithms that show promise of a quantum speedup require deep circuits and prolonged coherence times. In this work, we propose leveraging Reinforcement Learning (RL) to intelligently build quantum circuits that can recreate given target states, given no information about the circuit used to construct them. The RL agent learns about the hidden system by receiving rewards based on local observables, calculated using the target state, and the fidelity of the final state. By constraining the depth of the circuits built by the agent, we hypothesize that this approach allows us to compress the depth of quantum circuits necessary to create the target state. One important application of our method is dynamic quantum simulation, where the target state is a time-evolved state using a given Hamiltonian and a Trotterized quantum circuit. Our method promises quantum simulations out to longer final times than are currently feasible on NISQ devices.

Background

- Doing long time evolutions $e^{-iHt}|\Psi_0\rangle$ is infeasible due to the many number of Trotter steps

$$e^{-iHt} = \lim_{n\to\infty}\left(\prod_j e^{-iH_jt/n}\right)^n$$

- Compress a variable length Trotter evolution into a fixed length circuit
- Variational methods such as the variational quantum eigensolver have been used to do this [1]
- Variational methods have also been controlled (optimized) by RL [2]
- Goal is to use RL to intelligently build circuits to recreate time-evolved state opposed to optimizing parameters in fixed circuit

$$|\Psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$

- Single qubit state is completely described by two angles θ, ϕ

Reinforcement learning

- Reinforcement learning is the logical extension of **Markov Decision Processes (MDPs)**.
- Formally, MDPs are represented as a four-tuple, $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$
- Goal of agent in a MDP is to maximize the rewards it receives or, equivalently, to maximize the *value function*.

$$V_\pi(s) = \sum_a \pi(a|s)\sum_{s',r}\mathcal{P}(s',r|s,a)\Big(\mathcal{R}(s',s,a) + \gamma V_\pi(s')\Big)$$

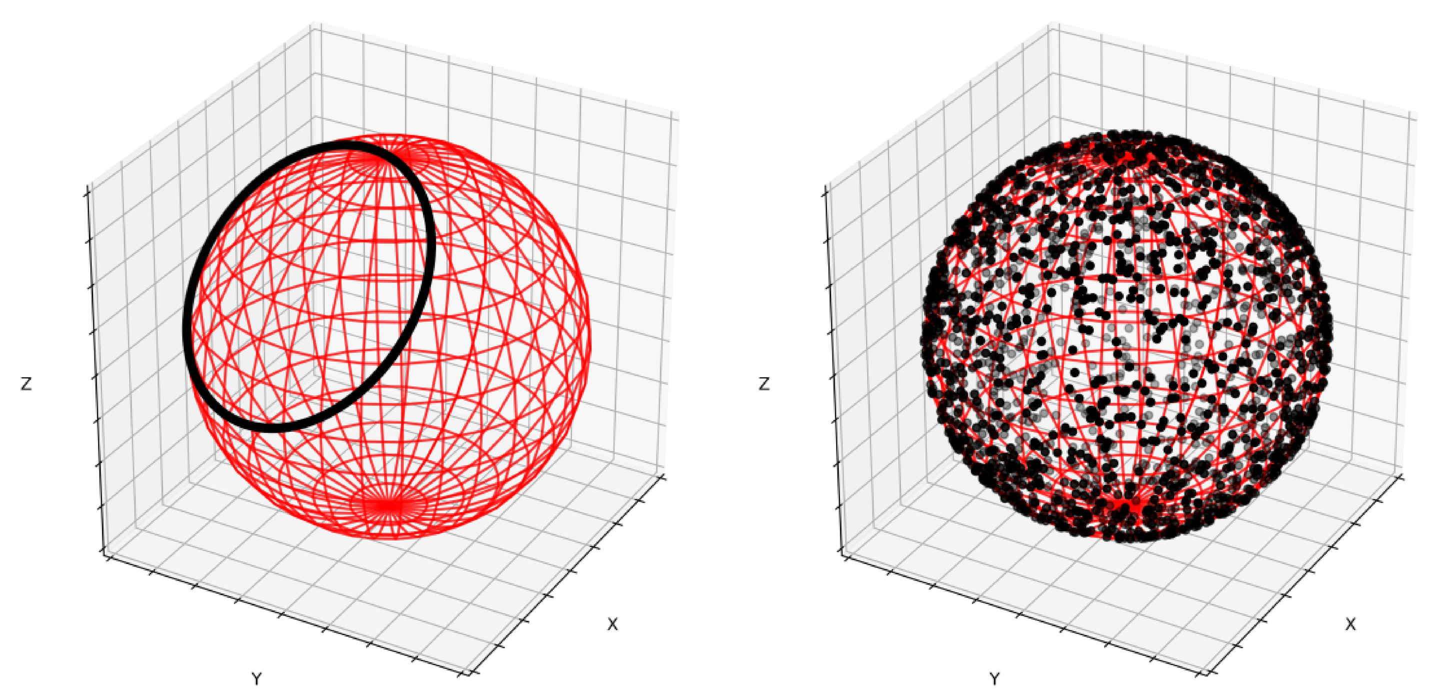
- Entire environment is known, so optimal policy can be found with *policy improvement*

$$\pi'(s) = \text{argmax}_a \sum_{s',r}\mathcal{P}(s',r|s,a)\Big(\mathcal{R}(s',s,a) + \gamma V_\pi(s')\Big)$$

- For more complex environments that are not completely described, approximate methods such as *proximal policy optimization* are needed.

Single qubit dynamics

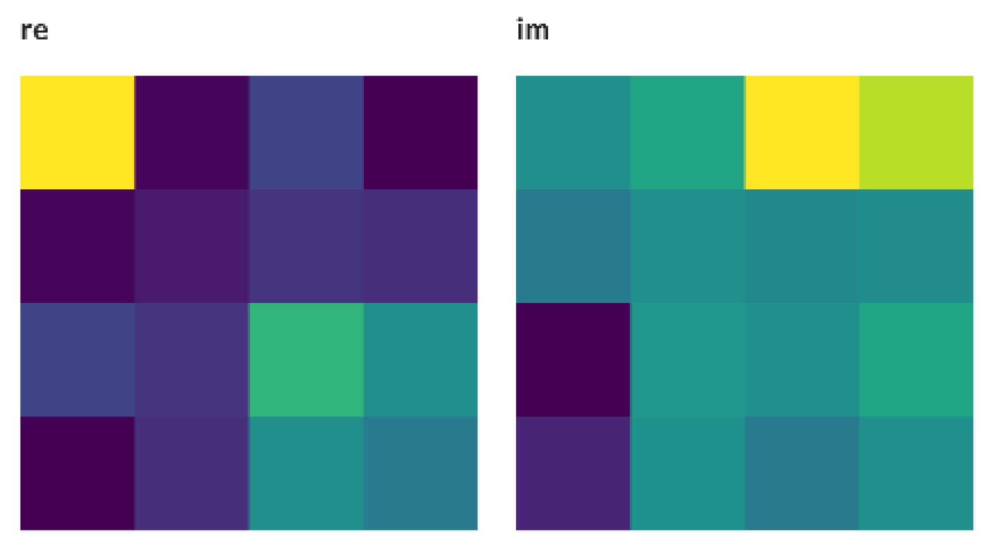
- Formulate single qubit dynamics as a MDP
  - ▷ $\mathcal{S}$ – Bloch sphere discretized by $\epsilon = \pi/k$
  - ▷ $\mathcal{A}$ – Discrete gate set {H, T, I}
  - ▷ $\mathcal{P}$ – Apply each gate in gate set to a random state in a MDP state and record distribution
  - ▷ $\mathcal{R}$ – $\mathcal{R}(s) = 0$ if s is the goal state, $\mathcal{R} = -1$ otherwise
- Investigate $(HT)^n|0\rangle$ (left) and random unitaries (right)



- Average fidelity $|\langle\psi_g|\psi_f\rangle|^2$ of 0.997 is able to be reached using a sequence of 8.64 gates from the gate set

Multi qubit dynamics

- Upper bound on the number of MDP states: k^{n^2-1} for n qubits
- State is instead real and imaginary part of density matrix

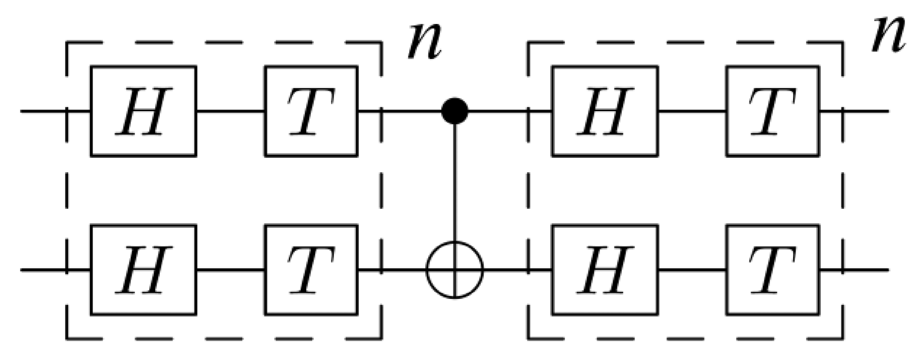


- Approximate value function with parameters $\theta \in \mathbb{R}^d$

$$\pi(a|s,\theta) = P\{a_t = a | s_t = s, \theta_t = \theta\}$$

- θ represents weights in a convolutional neural network (CNN)

- Bell states easily able to be found by RL agent – most likely due to the fact that only small number (~ 3) gates are needed
- $[(HT)^m]^{\otimes n}$ also feasible, finds fidelities $\geq$ 0.70



- Adding entanglement substantially increases difficult of the problem
- In general, it is computationally complex to approximate arbitrary quantum states with accuracy ϵ, $\Omega(2^n \log(1/\epsilon)/\log(n))$ [3]
- Agent needs more information in reward, entanglement entropy, closeness of density matrices, etc.
- Other variational methods such as VQE are more robust and practical for state compression
- RL still stands to potentially be useful in controlling hardware [4]

References

[1] Y. Yao, N. Gomes, F. Zhang, *et al*, arXiv quant-ph, 2011.00622 (2020).

[2] A. Bolens, M. Heyl. arxiv quant-ph 2006.16269 (2020).

[3] M. A. Nielsen, I. L. Chuang. Cambridge University Press (2000).

[4] M. Niu, S. Boxio, *et al*, npj Quantum Information 5 33 (2019)

Contact information: nfb1@iastate.edu and porth@iastate.edu