Curriculum reinforcement learning for optimization of variational quantum circuit architectures

Mateusz Ostaszewski[1] , Wojciech Masarczyk[2], Lea M. Trenkwalder[3], Eleanor Scerri[4], Vedran Dunjko[4]

[1] Institute of Theoretical and Applied Informatics, Polish Academy of Sciences, Gliwice, Poland, [2] Warsaw University of Technology, Warsaw, Poland, [3] Institute for Theoretical Physics, University of Innsbruck, Innsbruck, Austria, [4] Leiden University, Leiden, The Netherlands.

Introduction

As we are entering the so called Noisy Intermediate Scale Quantum (NISQ) [1] technology era, the search for more suitable algorithms under NISQ restrictions is becoming ever important. Perhaps the most promising classes of such algorithms are based on variational circuit methods, applied to problems in quantum chemistry. This problem is believed to be intractable in general, yet the quantum Variational Quantum Eigensolver (VQE) [2] algorithm can provide solutions in regimes which beyond the reach of classical algorithms, while maintaining NISQ-friendly properties.

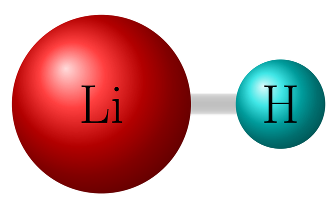
Variational Quantum Eigensolver

The VQE objective is to prepare the state $|\psi(\vec{\theta})\rangle$ which can be used to approximate the ground state of a given Hamiltonian H by the variational principle

$$E_{\min} \leq \min_{\vec{\theta}} \langle 0|U^\dagger(\vec{\theta})HU(\vec{\theta})|0\rangle = \min_{\vec{\theta}} \langle \psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle,$$

where $E_{\min}$ is a ground state energy of H.

Hamiltonian H models electronic structure of chemical molecule

$H \sim$ 

Chemical accuracy

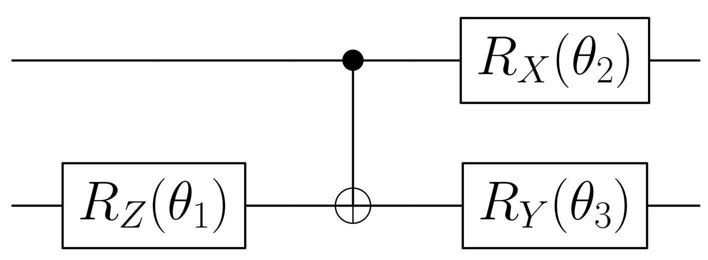
Find such $\vec{\theta}$ that

$$|\langle\psi(\vec{\theta})|H|\psi(\vec{\theta})\rangle - E_{\min}| < 0.001 \ \ Hartree$$

where $E_{\min}$ is minimal eigenvalue of H.

Architecture design

The parametrized state is prepared by applying $U(\vec{\theta})$ which can be decomposed into quantum circuit

$U(\vec{\theta}) \sim$ 

The architecture itself can also be optimized for the constraints of NISQ:

- reduce number of gates with high fidelity error – two-qubit gates,
- reduce depth of the circuit – decoherence noise.

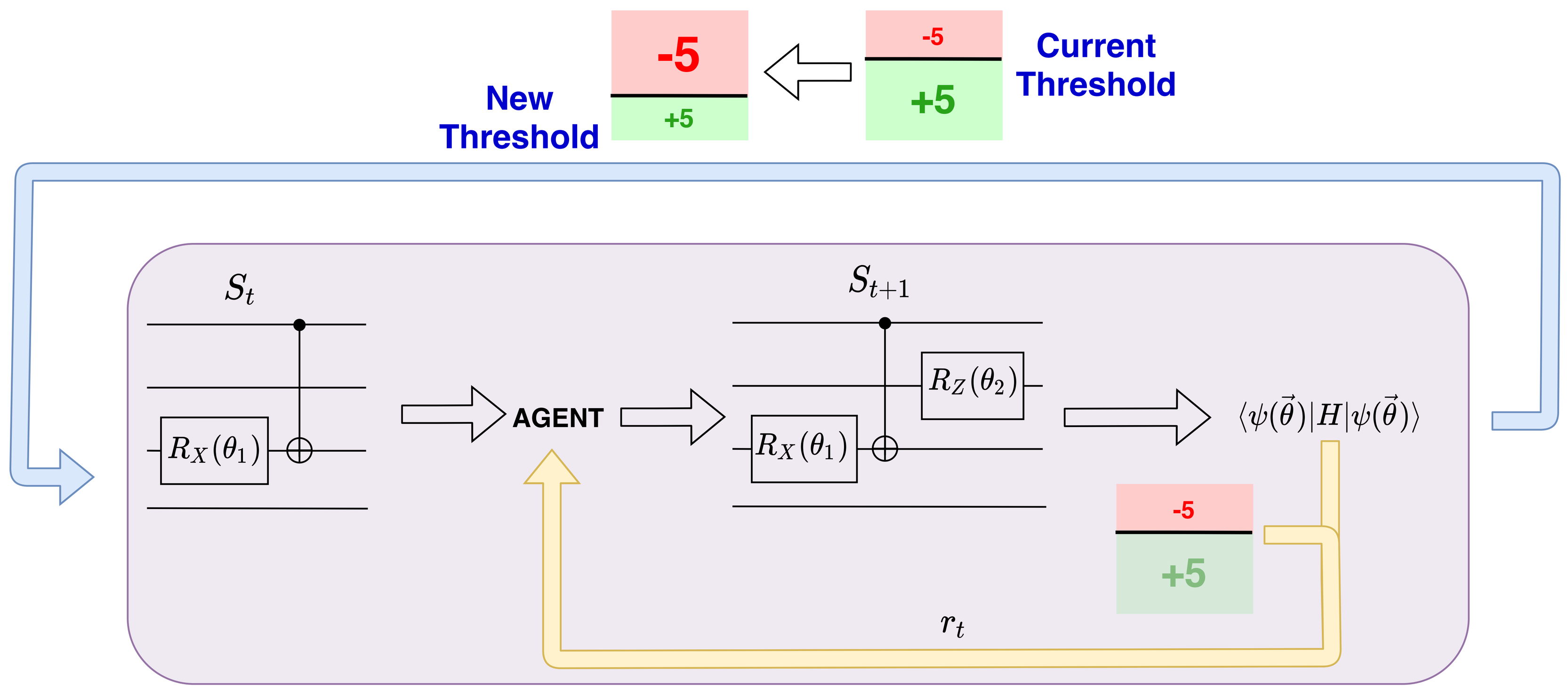
Curriculum agent

Reinforcement learning setup

- **RL state** – representation of the current quantum circuit, with energy,
- **RL action** – all the possible placings of a one– and two–qubit gates,
- **RL reward** – proportional to the difference between the previous and the current energy, or when threshold it reached +5, or when end of episode and threshold not reached -5.

Curriculum learning:

- agent is trained in the same environment in multiple rounds with increasing complexity – **in order to improve learning process**
- **threshold** i.e. distance to exact ground energy after which agent receives positive reward, **is lowered**, increasing the difficulty of task.



Experiment design

- **Problem**: Creating concise circuits surpassing chemical precision for LiH problem
- **Baselines**: Random Agent (RA), Tabula Rasa Agent (TR)
- **Solution Quality Indicators**: min energy obtained, min number of gates, min depth of circuit

Results after one one-step optimization

	#(0.1%)	avg #gates	min #gates	avg depth	min depth	avg #CX	min #CX
RA	5	29.4	26	17.4	13	13.8	12
TR	10.5	30.3	23	21.38	13	22.72	13
CA	**24.28**	**20.21**	**13**	**11.21**	**8**	**10.41**	**6**

Results after multi-step optimization

- **Baselines**: Random Agent (RA), Tabula Rasa Agent (TR)
- **Standard VQE approaches**: Hardware Efficient (HE) and UCCSD architectures

	avg distance	min (dist)	#(0.1%)	avg #gates	min #gates	min depth
RA	0.00041	0.00009	5	29.4	26	13
HE	0.00239	0.00230	N.A.	33	33	12
UCCSD	**0.00038**	0.00038	N.A.	430	430	430
TR	0.00049	0.00013	129.71	30.68	23	13
CA	0.00043	**0.00007**	**846.29**	**16.21**	**13**	**6**

Bibliography

[1] J. Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 79 (2), 2018.

[2] A. Peruzzo, J. McClean, P. Shadbolt, M. Yung, X. Zhou, P. J. Love,A. Aspuru-Guzik, J.L. O'brien A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 4213 (5), 2014.

Acknowledgements

MO acknowledge the support of the Foundation for Polish Science (FNP) under grant number POIR.04.04.00-00-17C1/18-00. LMT acknowledges support by the Austrian Science Fund FWF within the DK-ALM (W1259-N27). This work was also supported by the Dutch Research Council(NWO/OCW), as part of the Quantum Software Consortium programme (project number 024.003.037). VD and ES acknowledges the support of SURFsara through the QC4QC project. This research was partially funded by the Grant of Priority Research Domain at Warsaw University of Technology - Artificial Intelligence and Robotics.