Reinforcement Learning Based Quantum Circuit Optimization via ZX-Calculus

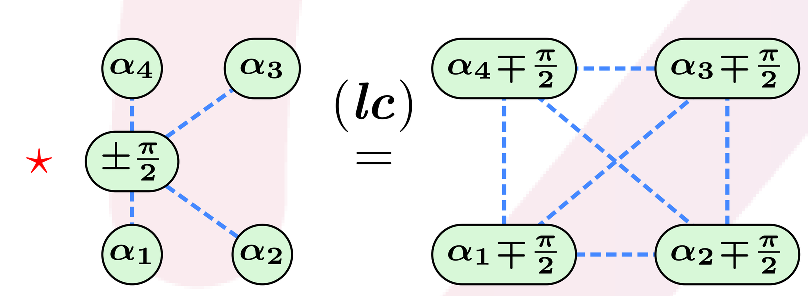
Jordi Riu, Jan Nogué, Gerard Vilaplana, Artur Garcia-Saez and Marta P. Estarellas
jan.nogue@qilimanjaro.tech   Qilimanjaro Quantum Tech, Barcelona, Spain

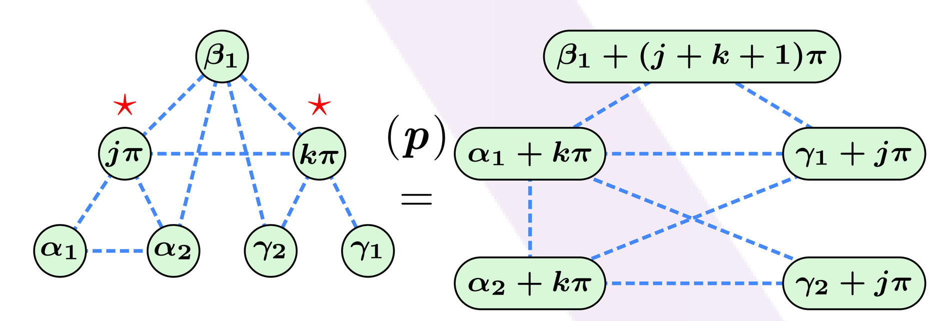
The issue

- **Problem 1:** Current Reinforcement Learning (RL) techniques for Quantum Circuit Optimization (QCO) suffer from a **high-dimensional action space**.
- **Problem 2: ZX-approaches for QCO.** Transform the Quantum Circuit into a ZX-diagram and apply graph-theoretic rules to simplify it. A circuit extraction process needed to obtain the final circuit from the ZX-diagram.
  - PROs: Small action space.
  - CONs: **Optimal order of rules application is hard** to find. Often increase the two qubit gate count.
- **Motivation**: Implement a RL agent that learns the optimal sequence of actions in the ZX-diagram such that the gate count is reduced after circuit extraction.
- **Ultimate goal**: Achieve a circuit optimization algorithm useful for target circuits of the NISQ era.

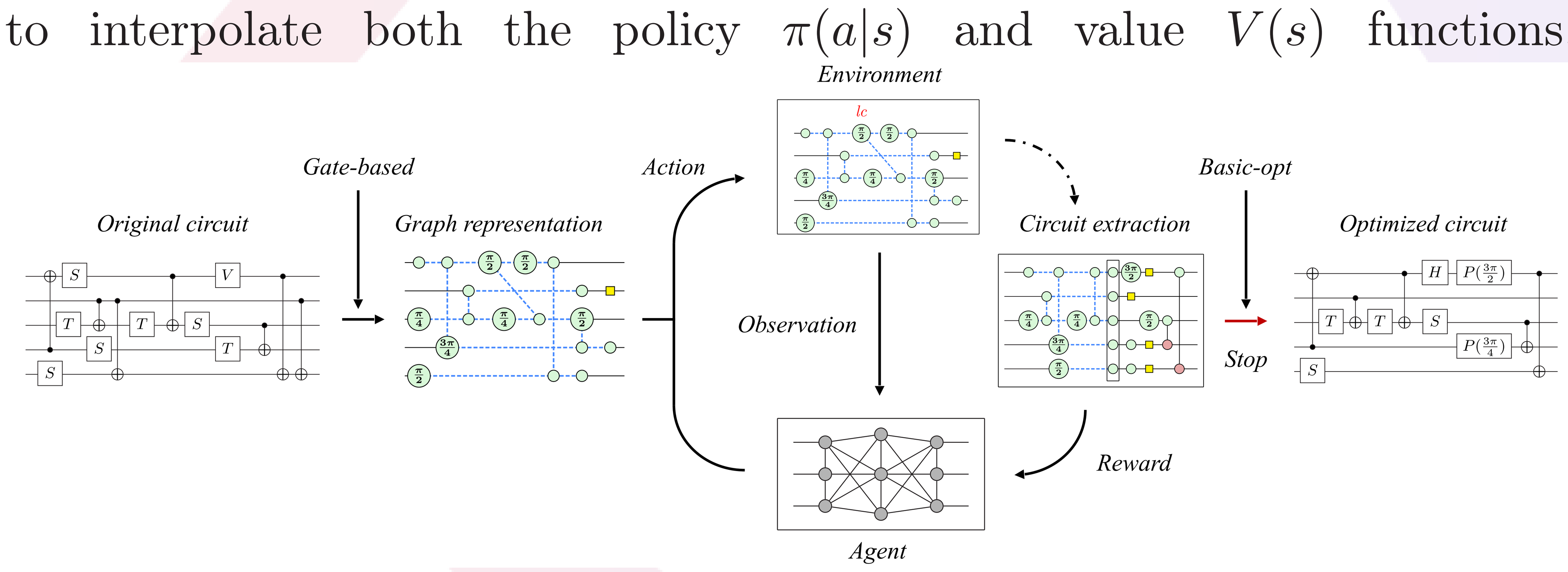
The approach

Actions to simplify a ZX-diagram. Two principal rules to simplify a ZX-diagram: Local complementation ($\boldsymbol{lc}$) and pivoting ($\boldsymbol{p}$).





Overview of the rl-zx approach. We use the Proximal Policy Optimization (PPO) to train the agent and a Graph Neural Network to interpolate both the policy $\pi(a|s)$ and value $V(s)$ functions.



The **reward function** is defined as the difference in gates after an action is applied, allowing for single and two qubit gate target reduction.

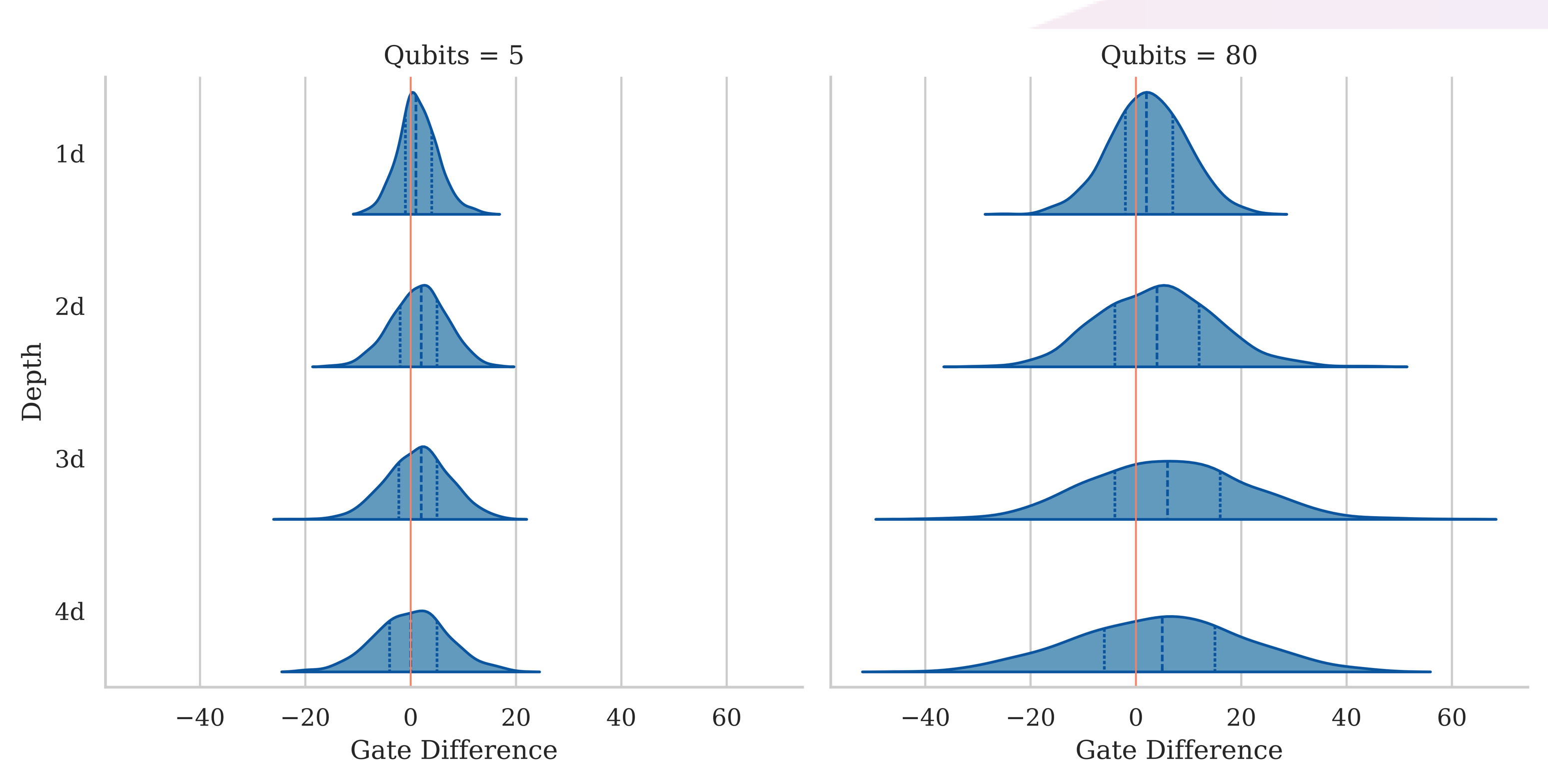
Impact [1]

- Showcased its applicability for relevant NISQ algorithms in **relativistic quantum chemistry**.
- Our approach obtained a **60% two qubit gate reduction** in a 12-qubit UCC ansatz w.r.t the best circuit obtained with traditional methods.
- Obtained a **1% quality result** w.r.t to simulation on IonQ Forte-I hardware.

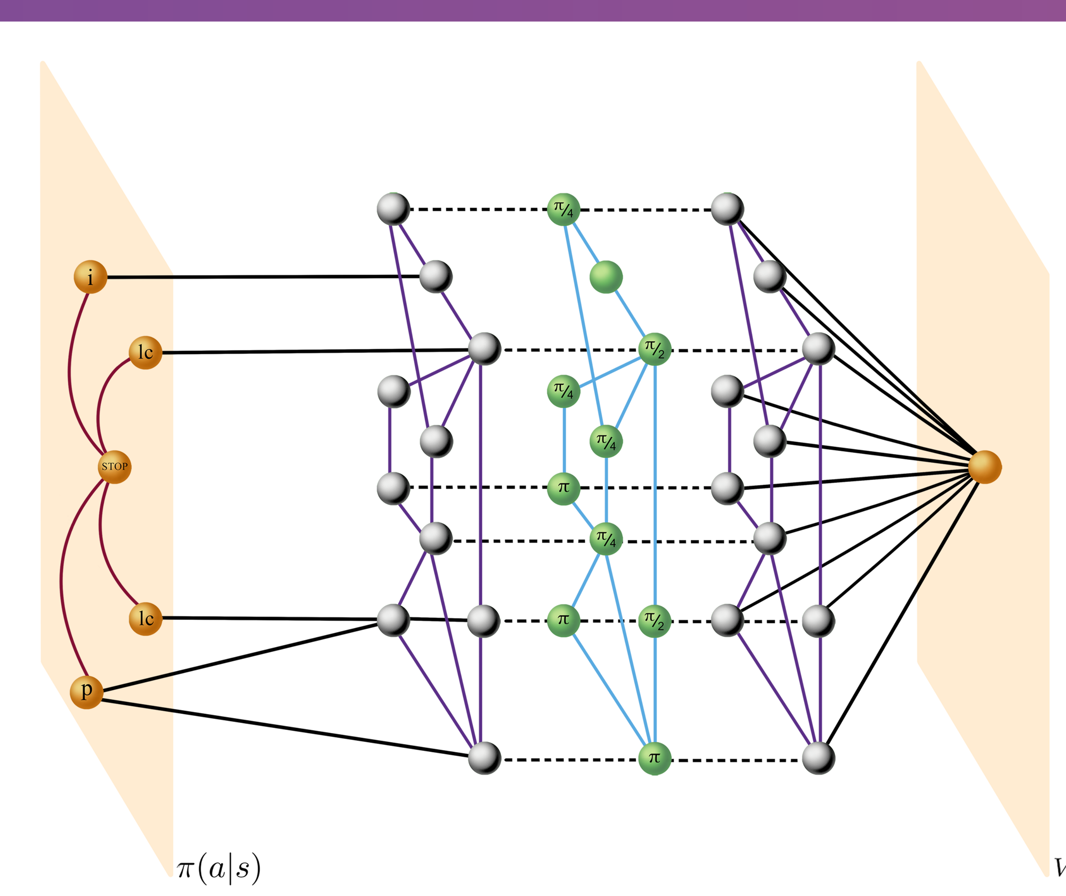
References

[1] Palak Chawla et al. Relativistic vqe calculations of molecular electric dipole moments on trapped ion quantum hardware, 2024.

Additional details



(a) Distribution of the difference in two qubit gates between our trained agent and the state-of-the-art ZX-based algorithm for QCO.



(b) Schematic overview of the actor and policy networks.