Policy Gradient Approach

Of Variational Quantum Circuits

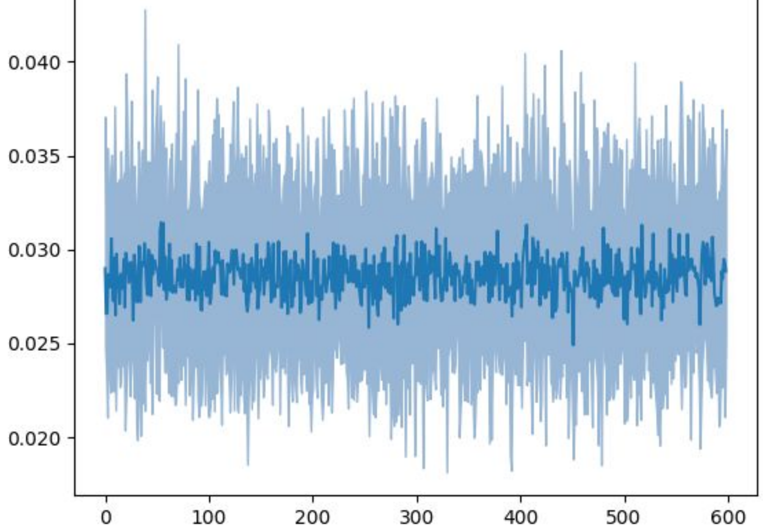
Ross Mojgani, Owen Guo, Kobe Ng, Akhil Veeraghanta

n (max points: 17)

	0	1	2	3
	Content is not explained at a level appropriate for the class	Some content is explained at a level appropriate for the class; many complex topics are included and not made accessible	Most content is explained at an appropriate level; some complex topics are included but are not made accessible, or only minimal effort is made	Content is clearly explained at an appropriate level; significant effort is made to make complex topics accessible to classmates.
	No visualizations are used to explain concepts	Some basic visualizations are used to explain concepts	Good-quality graphics are included to explain concepts and enhance the presentation experience.	Graphics used to explain concepts are innovative and/or of high quality and significantly enhance the presentation experience.
	No live demo is included	A basic live demo is included but is not used to demonstrate reproduction of the work	A live demo is performed that reproduces key elements of the work; some audience input may be used	A live demo is performed that reproduces key elements of work in an innovative way, or that uses substantial audience input
	Software implementation is not discussed at all	Some aspects of the software are discussed	Most aspects of the software are discussed	A detailed overview of the software is provided, with key design decisions outlined.
ty	Ease-of-reproducibility of the work is not discussed	Ease-of-reproducibility or issues encountered are briefly discussed	Ease-of-reproducibility and issues encountered are clearly discussed	
	Presentation is dry and not engaging, and/or hard to follow.	Some parts of the presentation are engaging and easy to follow.	Most of the presentation is engaging and easy to follow.	Presentation is extremely engaging and/or includes active learning or audience participation mechanisms.

lity will be graded by your classmates, as well as myself, according to this scale. You will receive int value.

that some aspect of the algorithm prohibits a full live demo. This will not result in a grade of 0 dled on a case-by-case basis.



Outline

- Quantum Compilation
- Variational quantum algorithms (VQA)
- Problems with traditional VQA optimization
- Anzatz Circuit Creation
- Policy Gradient Reinforcement Learning Approach
- Policy Gradient Descent
- Software Overview
- Live Demo
- Results
- Assumptions/Challenges
- Q&A

High Level Idea of Paper

- Use variational quantum circuits for quantum compilation
- Traditional optimization suffers from barren plateaus
- Use policy gradient methods
  - Optimize in the distribution parameter space rather than variational parameters

Policy Gradient Approach to Compilation of Variational Quantum Circuits

D. A. Herrera-Martí[1]

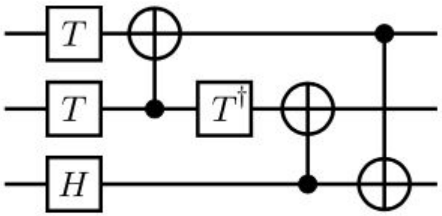
[1] *Université Grenoble Alpes, CEA List, 38000 Grenoble, France*

(Dated: November 22, 2021)

We propose a method for finding approximate compilations of quantum circuits, based on techniques from policy gradient reinforcement learning. The choice of a stochastic policy allows us to rephrase the optimization problem in terms of probability distributions, rather than variational parameters. This implies that searching for the optimal configuration is done by optimizing over the distribution parameters, rather than over the circuit free angles. The upshot of this is that we can

Quantum Compilation

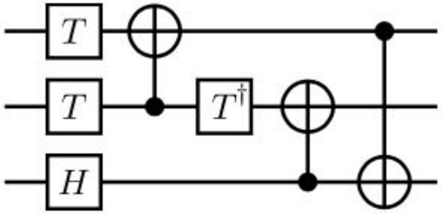
- Goal: approximate a unitary transform as a sequence of gates selected from a universal gate set.



Synthesis

Quantum Compilation

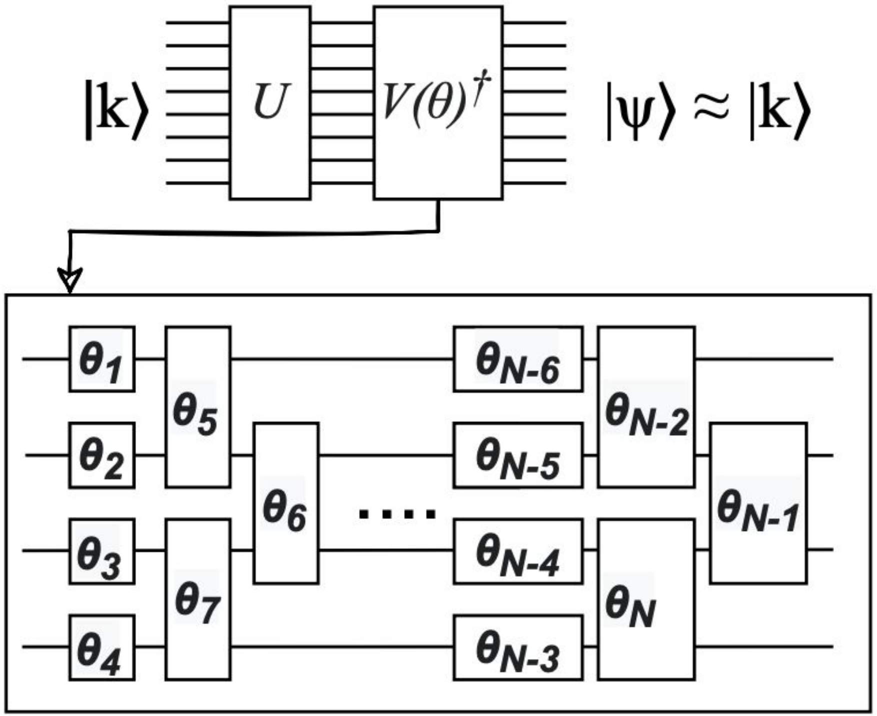
- Goal: approximate a unitary transform as a sequence of gates selected from a universal gate set
- Solovay-Kitaev theorem shows that it's possible
  - But doesn't explain how to find optimal solution



Synthesis

Quantum Compilation w/ VQA

- Create an Ansatz circuit
- Find θ that maximizes overlap between final state and the initial state
- Ansatz circuit consists of single-qubit Ry and two-qubit Rzz rotations



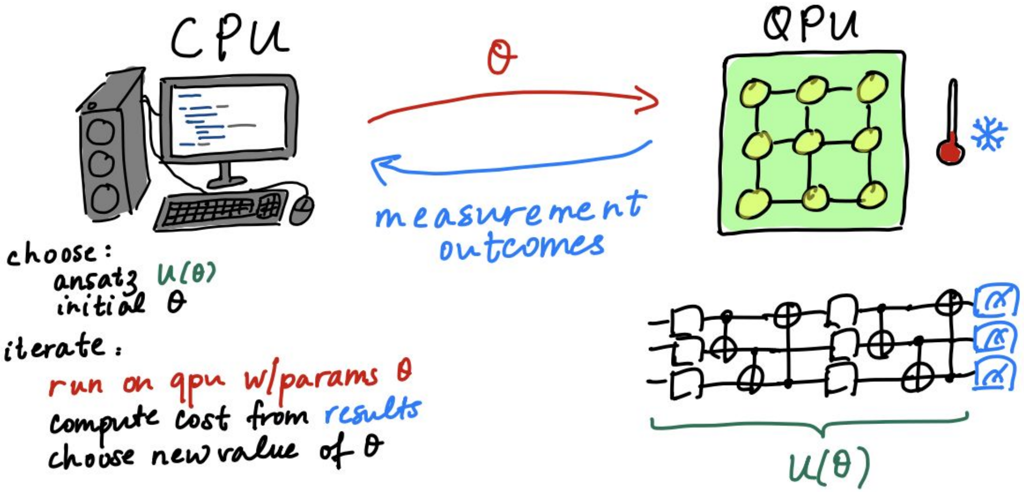
Fidelity Metric

- Metric to assess performance of compilation

$$\hat{F}(\theta) = \frac{1}{m}\sum_{k}^{m} |\langle k|V(\theta)^{\dagger}U|k\rangle|^{2} \qquad (1)$$

Variational Quantum Algorithms (VQAs)

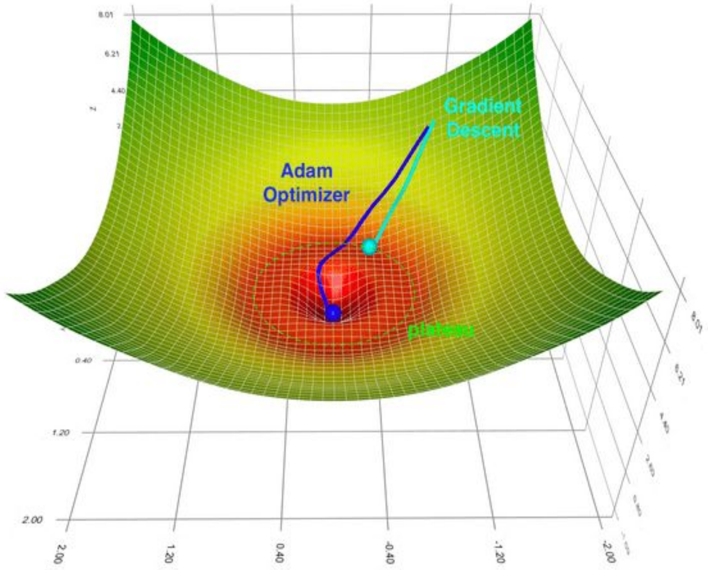
- Feature an iterative exchange between classical and quantum devices. (Sometimes called "hybrid" quantum-classical algorithms)



Source CPEN400Q

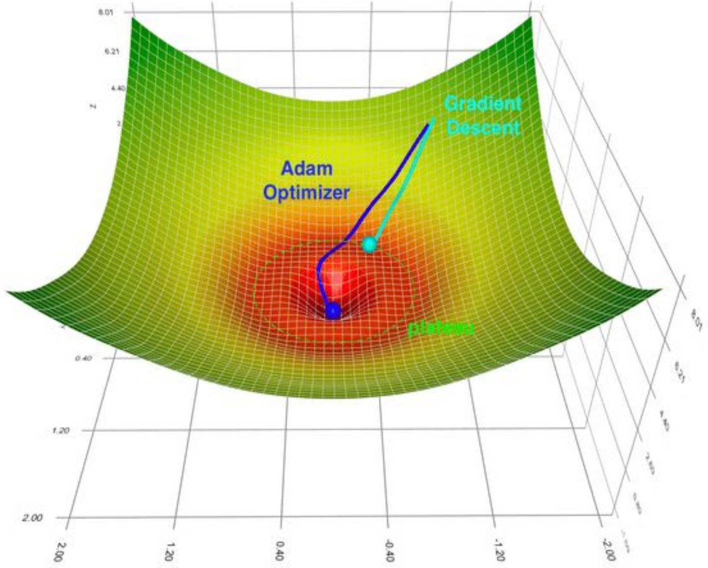
Problems with traditional VQA optimization

- Vanishing gradient significantly makes training harder
- Barren Plateaus in the parameter space
- Even gradient-free optimizers don't solve this problem



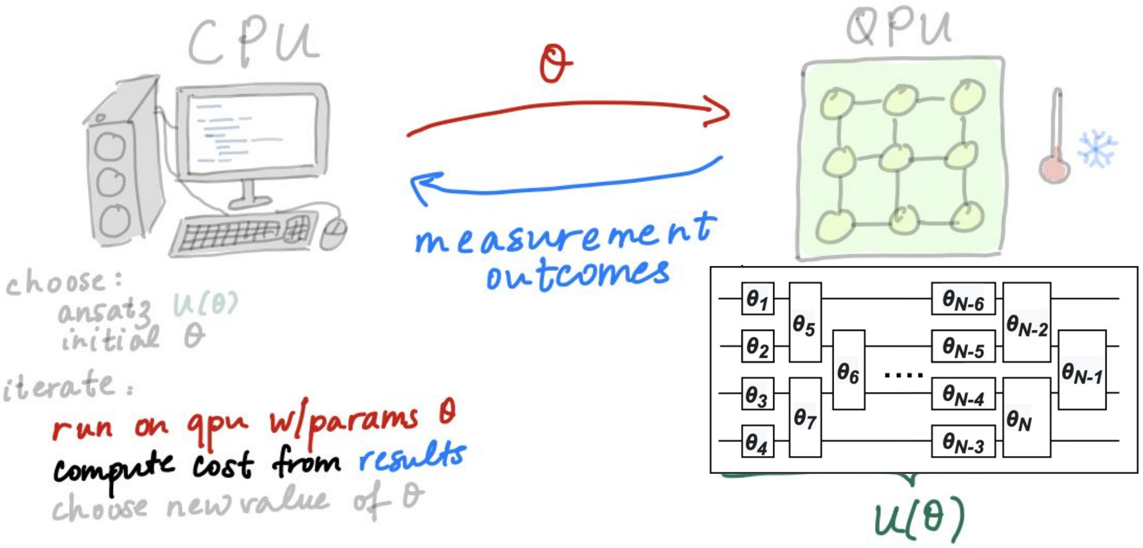
Solution to traditional VQA optimization

- Policy Gradient Reinforcement Learning
  - Mitigates effects of barren plateaus by optimizing in the distribution parameter space
- In particular policy gradient methods

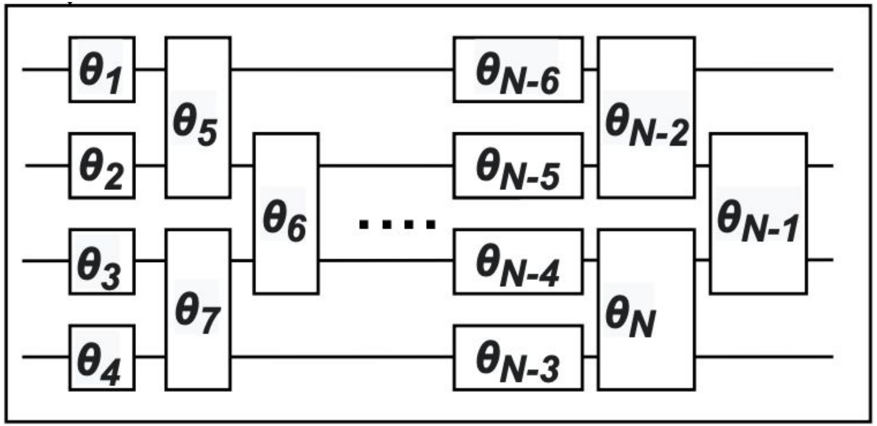


Parameterized quantum circuits

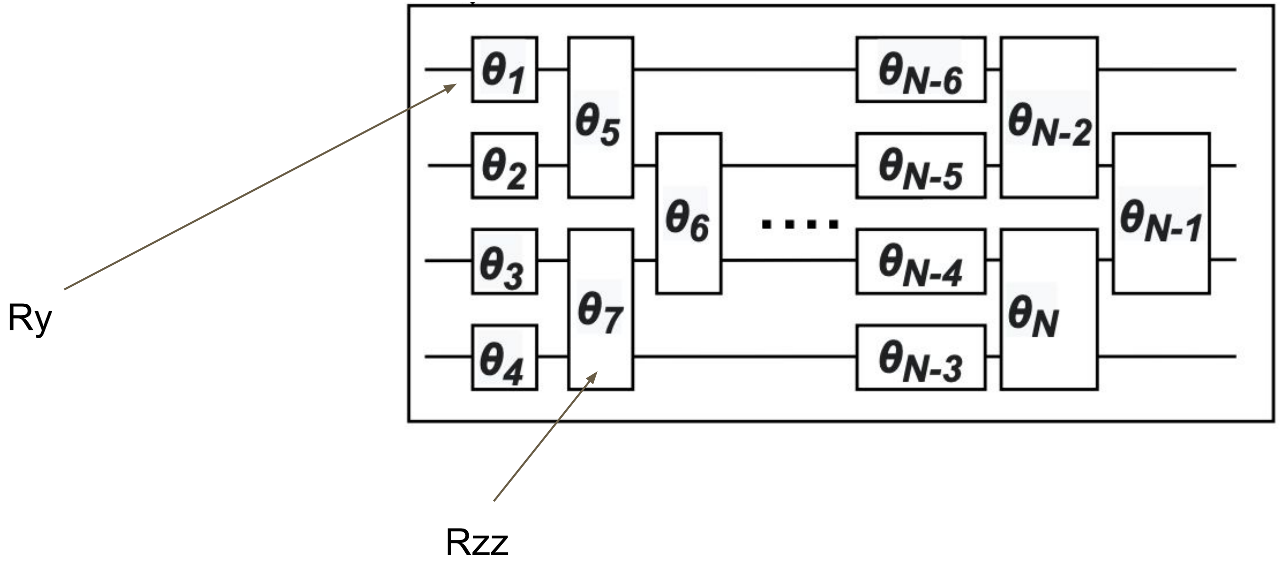
Parametrized circuits are used to assist in evaluation of a cost function which represents a particular problem.



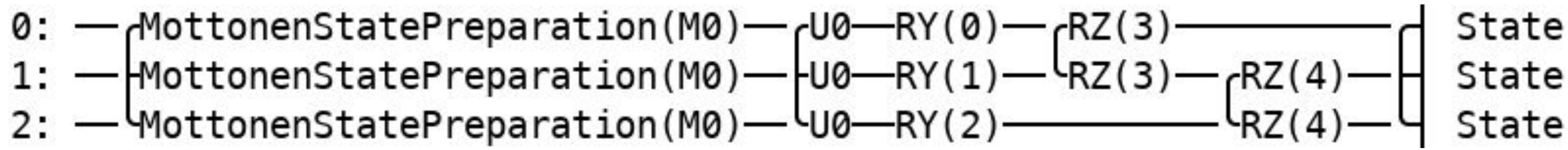
Anatz Circuit Creation



Anatz Circuit Creation

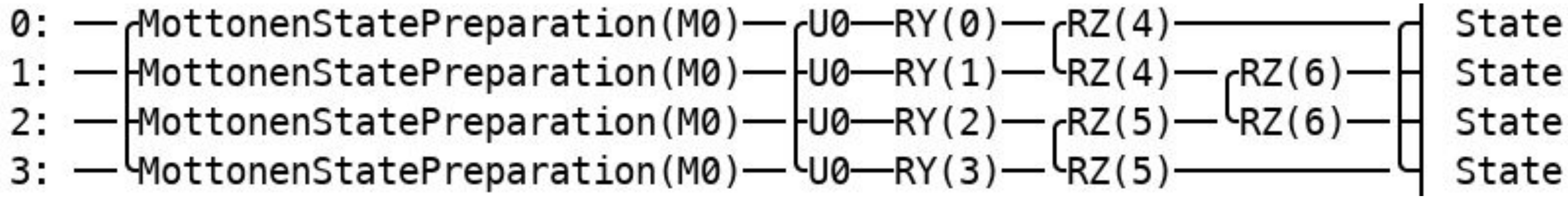


Anatz Circuit Creation



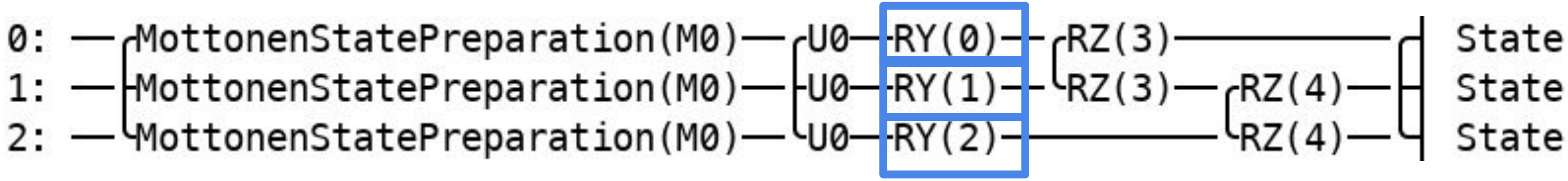
Ry Rzz

3 Qubit - 1 Layer



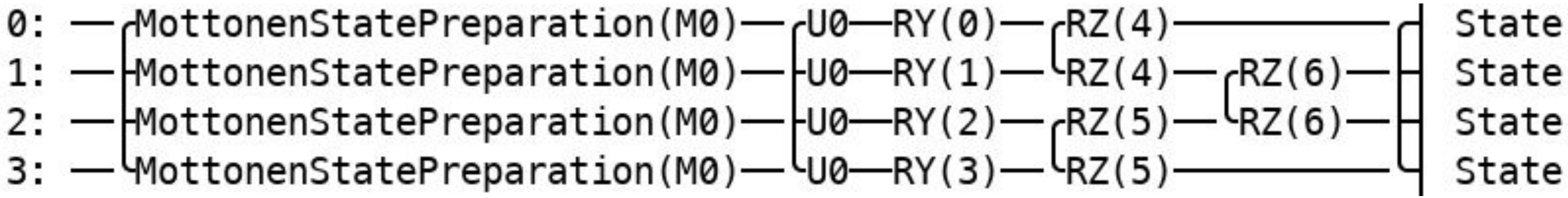
4 Qubit - 1 Layer

Anatz Circuit Creation



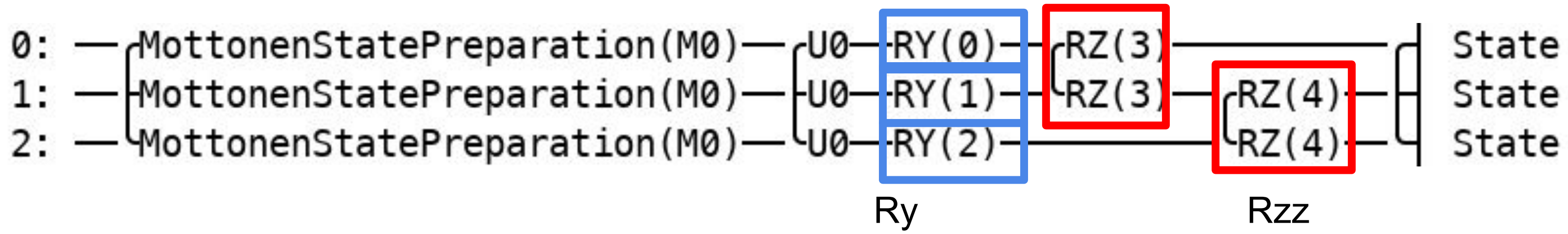
Ry

3 Qubit - 1 Layer

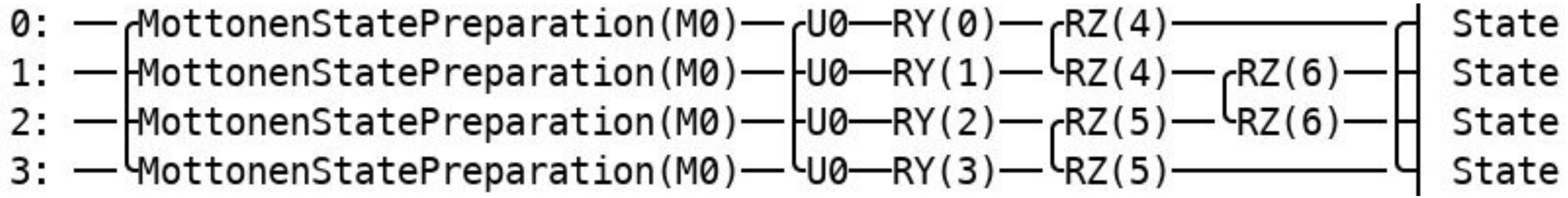


4 Qubit - 1 Layer

Anatz Circuit Creation



3 Qubit - 1 Layer



4 Qubit - 1 Layer

Policy Gradient Reinforcement Learning (PGRL) Approach

- Optimizing parametrized policies with respect to the expected return by gradient descent
- Mitigates the effect of barren plateaus in the training of VQAs
- Therefore it is possible to classically estimate the gradient at all times

$$\theta \sim \pi(x; \mu, \Sigma) = \frac{1}{\sqrt{2\pi|\Sigma|}} e^{-(x-\mu)\Sigma^{-1}(x-\mu)^T}$$

Objective function

- Corresponds to an average of the end reward
- i.e the asymptotic fidelity, over initial states (each sampled with probability pk)
- All possible actions (given by the current policy π(θ | μ, Σ))

$$J = \mathbb{E}_{\pi_\mu}[F] = \sum_k^m p_k \sum_\theta \pi(\theta|\mu, \Sigma)|\langle k|V(\theta)^\dagger U|k\rangle|^2 \quad (3)$$

Gradient of objective function

- Have used the so-called “policy gradient theorem”, which amounts to applying the chain rule to the policy function
- Allows to write the gradient of an expectation value as the expectation of a log likelihood times a cost function
- Only calculate gradient of mean vector

$$\nabla_\mu J = \sum_k^m p_k \sum_\theta \pi(\theta|\mu,\Sigma)\nabla_\mu \log \pi(\theta|\mu,\Sigma)|\langle k|V(\theta)^\dagger U|k\rangle|^2$$

$$\Sigma(t) = (1 - t/T)\Sigma_i + t/T\Sigma_f$$

RMSProp

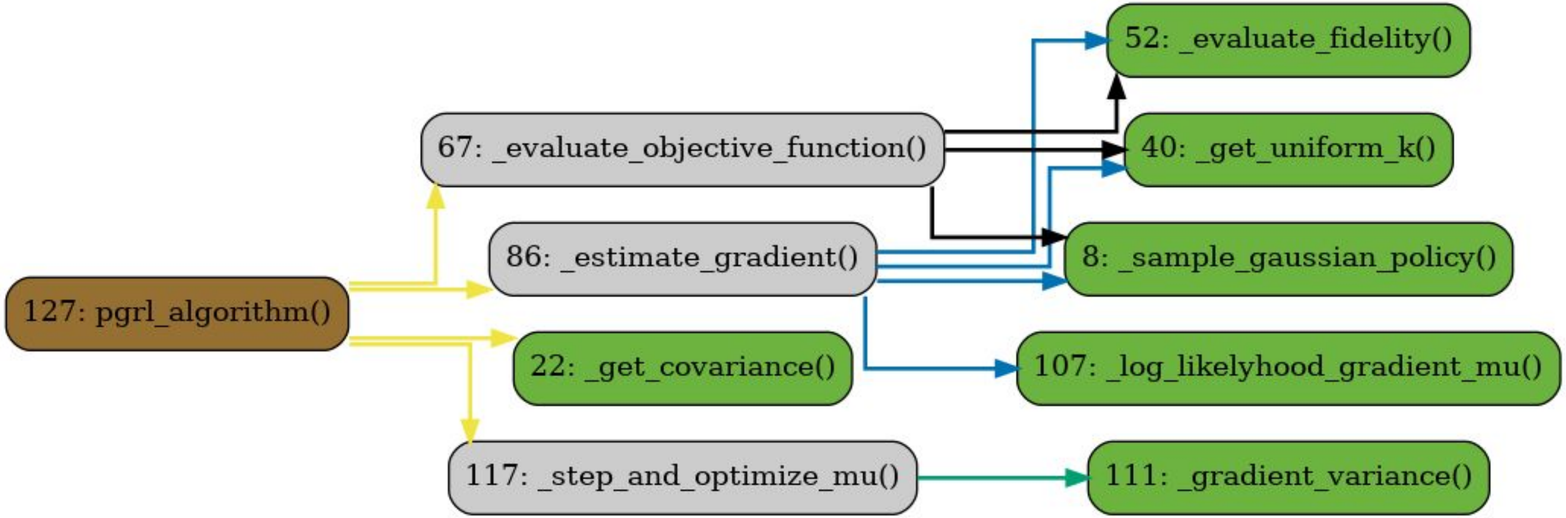
- Once a gradient has been estimated, RMSProp is applied as the update rule
- RMSprop is an adaptive learning rate method with better convergence properties than simple gradient ascent methods

$$\sigma_g^{(t)} = \gamma \sigma_g^{(t-1)} + (1-\gamma)(\nabla_\xi J|_g)^2$$

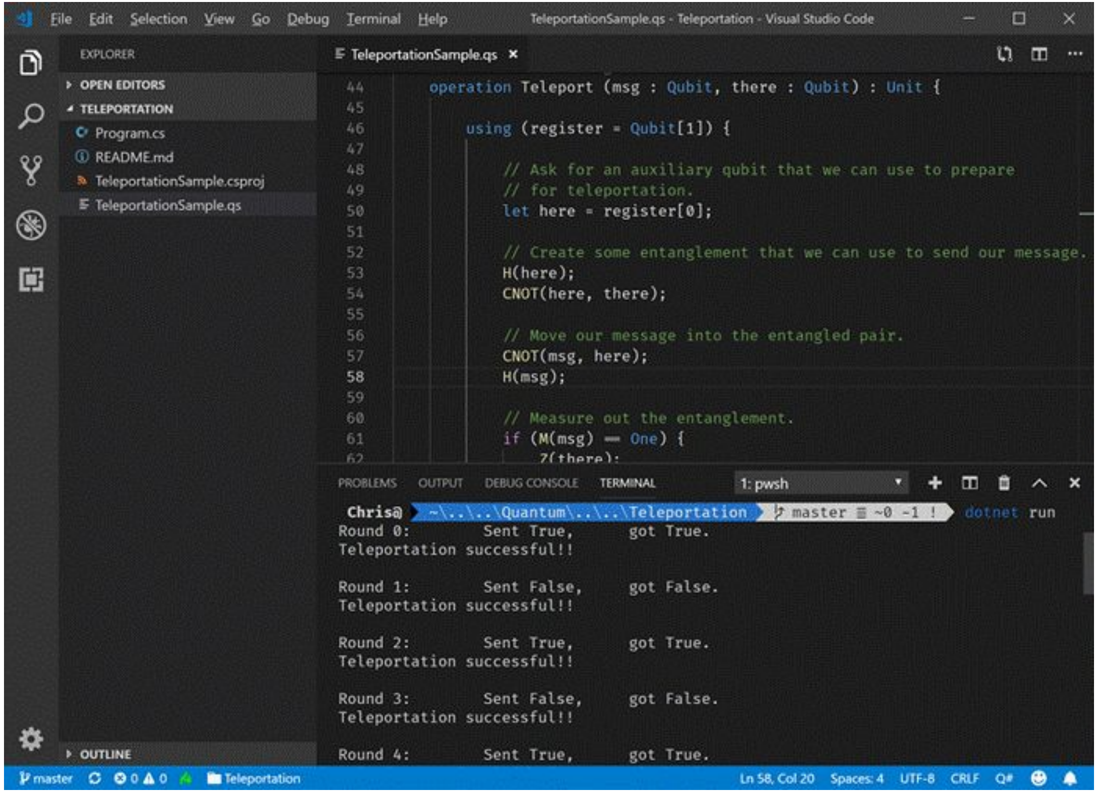
$$\xi \leftarrow \xi + \eta \frac{\nabla_\xi J|_g}{\sqrt{\sigma_g^{(t)}} + \varepsilon}$$

(i) computing a moving discounted average of the gradient variances and
(ii) dividing the update step by the discounted variance

Software Overview

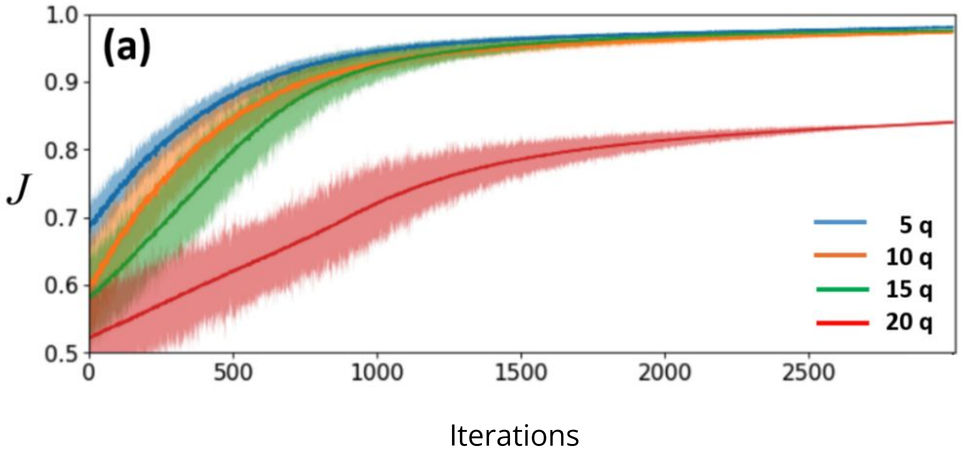


Live Demo



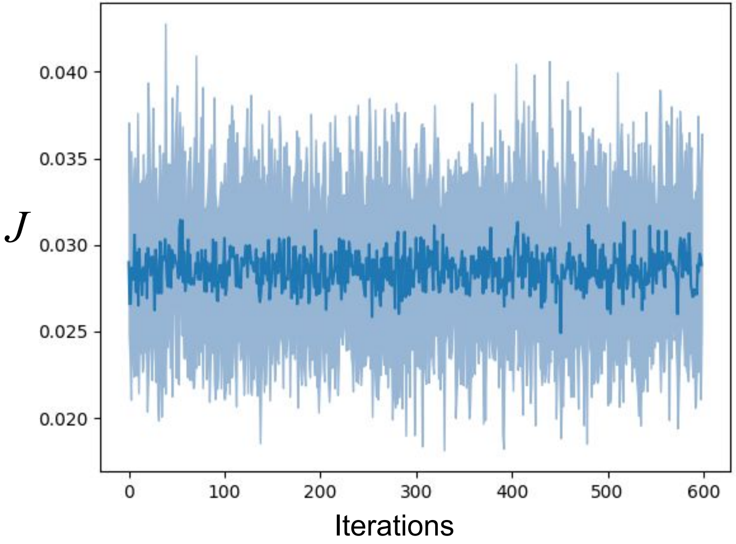
File Edit Selection View Go Debug Terminal Help
TeleportationSample.qs - Teleportation - Visual Studio Code
EXPLORER
OPEN EDITORS
TELEPORTATION
Program.cs
README.md
TeleportationSample.csproj
TeleportationSample.qs
operation Teleport (msg : Qubit, there : Qubit) : Unit {
using (register = Qubit[1]) {
// Ask for an auxiliary qubit that we can use to prepare
// for teleportation.
let here = register[0];
// Create some entanglement that we can use to send our message.
H(here);
CNOT(here, there);
// Move our message into the entangled pair.
CNOT(msg, here);
H(msg);
// Measure out the entanglement.
if (M(msg) == One) {
Z(there);
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL
1: pwsh
Chris@ ~\..\..\Quantum\..\..\Teleportation master dotnet run
Round 0: Sent True, got True.
Teleportation successful!!
Round 1: Sent False, got False.
Teleportation successful!!
Round 2: Sent True, got True.
Teleportation successful!!
Round 3: Sent False, got False.
Teleportation successful!!
Round 4: Sent True, got True.
OUTLINE
master
Teleportation
Ln 58, Col 20 Spaces: 4 UTF-8 CRLF Q#

Results



Results from paper

Reward for a 5 qubit test with 600 iterations.



Results from our implementation on 5 qubits

Assumptions and challenges (1)

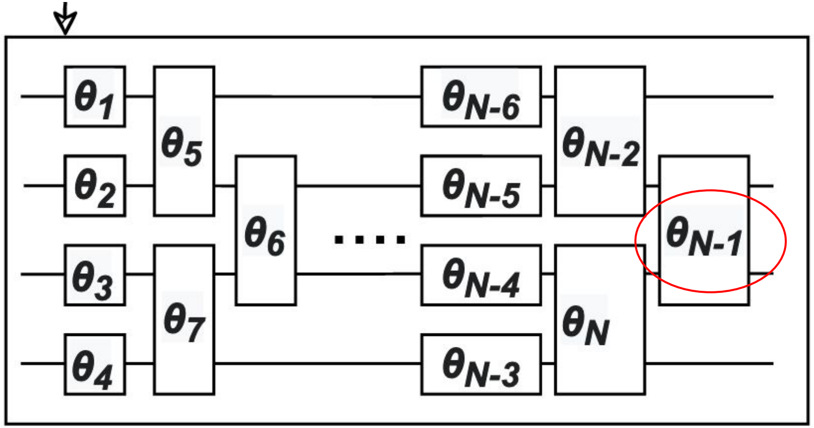
- Ambiguity of key equations
  - These equations describe the reward function
  - Several reasonable interpretations
  - None of them worked

$$J = \mathbb{E}_{\pi_\mu}[F] = \sum_k^m p_k \boxed{\sum_\theta \pi(\theta|\mu,\Sigma)} |\langle k|V(\theta)^\dagger U|k\rangle|^2 \quad (3)$$

$$\nabla_\mu J = \sum_k^m p_k \boxed{\sum_\theta \pi(\theta|\mu,\Sigma)} \nabla_\mu \log \pi(\theta|\mu,\Sigma) |\langle k|V(\theta)^\dagger U|k\rangle|^2$$

Assumptions/Challenges (2)

- State |k> and unitary matrices come from a smaller subset
  - The size of the smaller subset is undescribed
- Core parameters are unlisted in the paper
  - Number of layers used



Conclusion

- Paper suggests that reinforcement learning policies can help solve barren plateau problems in VQAs
- The paper used this to show successful quantum compilation
- Unable to reproduce results from paper due to lack of details in method

Sources

- CPEN 400q lecture slides (https://github.com/glassnotes/CPEN-400Q)
- https://www.researchgate.net/figure/color-online-One-iteration-of-the-Solovay-Kitaev-algorithm-applied-to-finding-a-braid_fig7_243437978
- Policy Gradient Approach to Compilation of Variational Quantum Circuits https://arxiv.org/abs/2111.10227