Critical Points in Quantum Generative Models

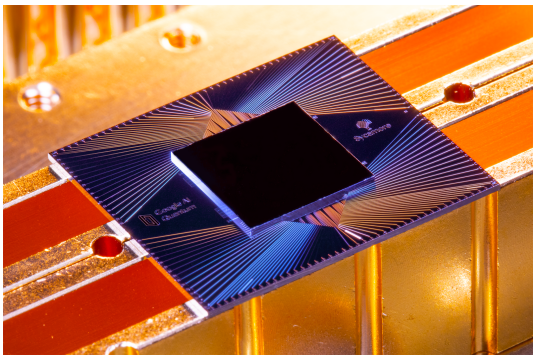
Eric R. Anschuetz

MIT Center for Theoretical Physics

ICLR 2022

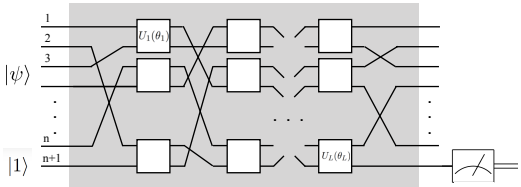
Quantum Computing

- Quantum computers believed to be superpolynomially more efficient than "classical" computers in solving certain tasks
- Currently in an era of noisy, intermediate-scale quantum devices
- What can they do that's useful?



Quantum Generative Models

- Quantum neural network[1]—generalization of "classical" neural networks

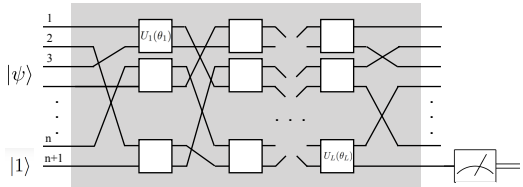


- Provably more expressive than classical counterparts[2]

[1] Farhi and Neven 2018.
[2] Gao et al. 2021.

Quantum Generative Models

- Quantum neural network[1]—generalization of "classical" neural networks

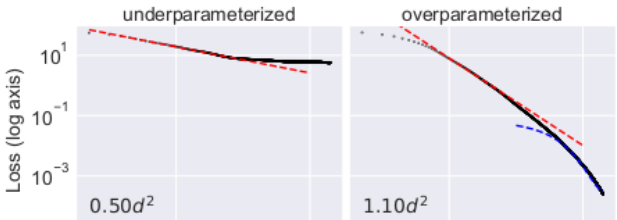


- Provably more expressive than classical counterparts[2]
- Are these models trainable?

[1] Farhi and Neven 2018.
[2] Gao et al. 2021.

Untrainability Results

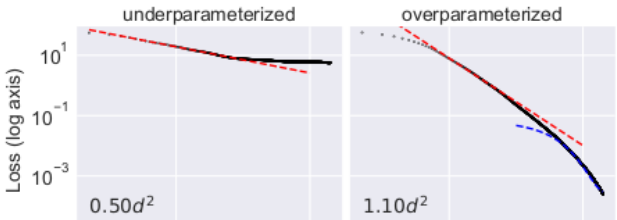
- Analytic results: gradient vanishes exponentially in model size for deep models[3]
- Numerical results: poor local minima quality in shallow models[4], in contrast with typical behavior of local minima in neural networks[5]



[3] McClean et al. 2018; Cerezo et al. 2021.
[4] Kiani, Lloyd, and Maity 2020.
[5] Choromanska et al. 2015.

Untrainability Results

- Analytic results: gradient vanishes exponentially in model size for deep models[3]
- Numerical results: poor local minima quality in shallow models[4], in contrast with typical behavior of local minima in neural networks[5]



- Can these numerical results be proven?

[3] McClean et al. 2018; Cerezo et al. 2021.
[4] Kiani, Lloyd, and Maity 2020.
[5] Choromanska et al. 2015.

Overview

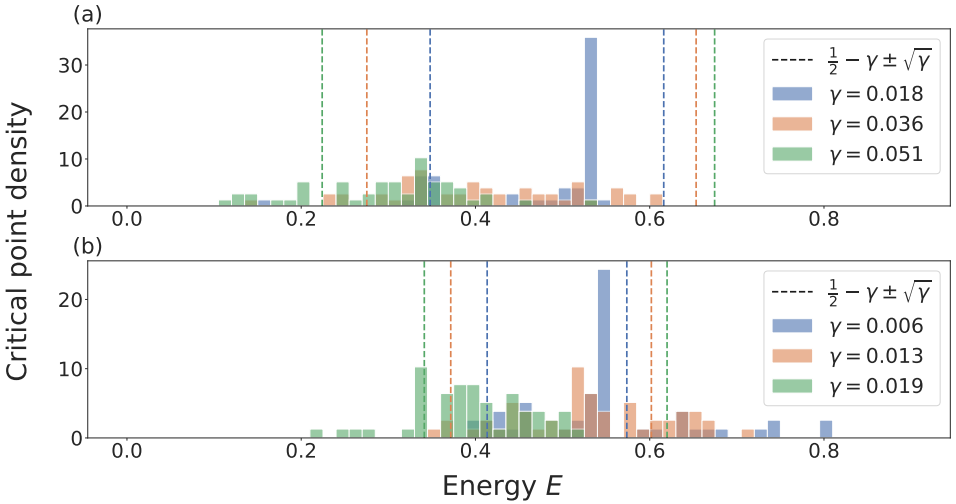
We show analytically:

- The existence of this trainability phase transition at $\gamma = 1$, governed by "order parameter"

$$\gamma \sim \frac{p}{2^{n+1}}$$

- The asymptotic distribution of local minima in these models
- Heuristic reasons why certain classes of models may not experience this poor quality of local minima

Numerical Confirmation



Conclusion

- Even shallow quantum models can be difficult to train!
- Is there a way to show similar results for more general quantum models?

Thank you!