

A deep learning approach to noise prediction and circuit optimization for near-term quantum devices



Alexander Zlokapa, Alexandru Gheorghiu (advisor)

Introduction

Rationale

Noisy Intermediate Scale Quantum (NISQ) devices face critical challenges in achieving high fidelity computations.¹

Decoherence, crosstalk between qubits, and hardware-specific phenomena can make realistic noise modeling difficult.

Heuristics such as gate count minimization are often used to attempt to reduce noise.^{2,3} Although effective, such heuristics are a proxy for a true noise model of quantum hardware.⁴

Overview

Equivalent circuit generation. A dictionary of equivalent circuit patterns is generated to rewrite an arbitrary circuit in equivalent form, enabling dataset generation and circuit optimization.

Noise prediction model. Deep learning may be used to estimate the noise difference between a pair of equivalent quantum circuits.

Greedy circuit optimization. A greedy optimization algorithm enables rapid traversal of the space of equivalent circuits to minimize expected noise according to the noise prediction model.

Simulated results. Results of the optimization scheme are shown for a simple noise model simulating the IBM Q Melbourne device.

Equivalent Circuit Generation

An arbitrary quantum computation on any number of qubits can be generated by a finite set of gates called a *universal gate set*. We consider the universal set of X, Y, Z, H, S, T and CNOT gates.

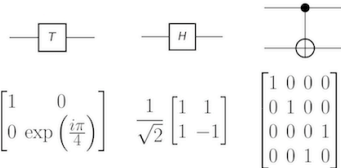


Figure 1. Example gates (T gate, Hadamard gate, CNOT gate) and their corresponding unitary matrix operators.

A brute-force search is performed over all possible circuits with 2 qubits and up to 4 gates. Each circuit is categorized into equivalence classes labeled by matrix form (up to a phase).

Each equality may be applied as a substitution in a region of a larger circuit, allowing a given circuit to be rewritten as an equivalent circuit during dataset generation and circuit optimization. We find that 2 qubits and 4 gates is sufficiently large to generate families of equivalent circuits.

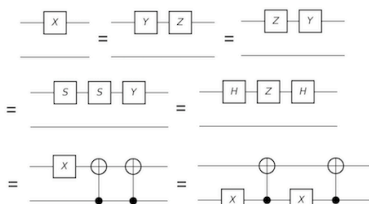


Figure 2. Equivalent circuits to a single X gate, showing only 7 out of 198 equivalent 2-qubit circuits with up to 4 gates.

Simulated Noise Model

Classical simulation

We simulate quantum circuits under a noise model consisting of three types of errors. At each time step, an error is applied with fixed probability as estimated from the IBM Q Melbourne hardware.⁵

- **Single-qubit gate errors:** a single qubit depolarizing error followed by a single qubit thermal relaxation error.
- **Two-qubit gate errors:** a two-qubit depolarizing error followed by single-qubit thermal relaxation errors on both qubits in the gate.
- **Single-qubit readout errors:** error on the classical bit value obtained during measurement on individual qubits.

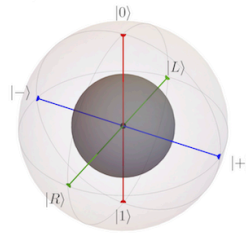


Figure 3. Depolarizing noise linearly combines the current state ρ with the maximally mixed state for an n qubit circuit according to the map $\rho \rightarrow E(\rho)$ given by:

$$E(\rho) = (1 - \lambda)\rho + \frac{\lambda}{2^n}I$$

where λ is the depolarizing error parameter. This effectively shrinks the Bloch sphere (shown above), moving the system out of a pure state.

Dataset

We generate 1.6 million training examples by selecting all possible pairs of 50 equivalent circuits for each of 668 random circuits. Each circuit has up to 8 qubits and 200 gates.

Each training example consists of a 16-channel image encoding 2 circuits ordered by a topological sort on a directed acyclic graph representation.

Noise is measured as the energy distance⁶ D between the resulting probability distributions of measurements of noisy and noiseless simulation.

$$D^2(\mathbf{P}_{\text{noisy}}) = \int_{-\infty}^{\infty} (\mathbf{P}_{\text{noisy}} - \mathbf{P}_{\text{clean}}) db$$

Deep Learning Noise Model

Convolutional neural networks (CNNs) have been shown to achieve state-of-the-art results for regression problems.⁷

ResNet⁸ avoids vanishing gradients associated with deep neural networks by introducing shortcut connections that skip layers.

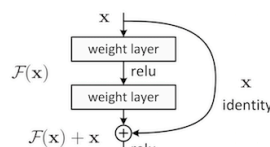
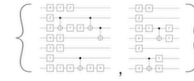


Figure 4. Building block of ResNet.⁸ A shortcut connection adds layer input x to the learned residual mapping $F(x)$.

Deep Learning Noise Model (cont.)



Equivalent circuit pair. Generated by greedy peephole optimization of sub-circuit patterns.

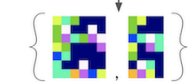
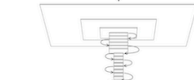


Image encoding. Topologically sorted DAG to 2 images concatenated to a 16-channel image.



Deep neural network. ResNet architecture modified for regression instead of classification. Due to its noisiness, the change in CNOT gate count is concatenated to the last layer.



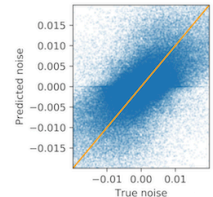
Noise prediction. Energy distance between clean and noisy readout probability distributions.

Figure 5. Overview of deep learning noise model.

Results & Conclusion

	Deep learning noise model	Gate count noise model
R^2	0.401 ± 0.003	0.314 ± 0.003
ΔD	-0.023 ± 0.001	-0.020 ± 0.001

A gate count linear model yields a significantly lower R^2 for noise prediction, showing that the deep learning noise model is generally expected to outperform gate count minimization heuristics.



Iterative peephole optimization is used to greedily replace equivalent 2-qubit sub-circuits of up to 4 gates, minimizing noise either by gate count D_{GC} or deep learning D_{DL} . An improvement of 15% in noise reduction over gate count minimization is found.

$$\frac{(D_{DL} - D_{GC})}{(D_{GC})} = 15\% \pm 3\%$$

Given the higher complexity of real-world hardware, we expect the advantage of deep learning noise mitigation over heuristics such as gate count minimization to be larger in practice. Nevertheless, we observe a significant improvement even with simple noise models.

Acknowledgments

We wish to thank Professor Thomas Vidick of Caltech, who provided helpful insight, comments and support for this project.

References

- [1] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018.
- [2] V. Kliuchnikov and D. Maslov, "Optimization of Clifford circuits," *Phys. Rev. A*, vol. 88, p. 052307, Nov 2013.
- [3] Y. Nam, N. J. Ross, Y. Su, A. M. Childs, and D. Maslov, "Automated optimization of large quantum circuits with continuous parameters," *npj Quantum Information*, vol. 4, no. 1, p. 23, 2018.
- [4] G. W. Dueck, A. Pathak, M. M. Rahman, A. Shukla, and A. Banerjee, "Optimization of circuits for IBM's five-qubit quantum computers," 2018 23rd EuroMicro Conference on Digital System Design (DSD), pp. 689-694, 2018.
- [5] G. Aleksandrowicz et al., "Qiskit: An open-source framework for quantum computing," 2019.
- [6] G. J. Székely, "E-statistics: The energy of statistical samples," *Bowling Green State University, Department of Mathematics and Statistics Technical Report*, vol. 3, no. 05, pp. 1-18, 2003.
- [7] S. Lathauillière, P. Mészáros, X. Alameda-Pineda, and R. Horaud, "A comprehensive analysis of deep regression," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-1, 2019.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, June 2016.