SAND2022-2450C

Sandia National Laboratories

How to benchmark a 100-qubit quantum computer using fewer than 100 circuits

PRESENTED BY

Timothy Proctor

Kenneth Rudinger, Stefan Seritan, Daniel Hothem, Jordan Hines, Thomas Catanach, Robin Blume-Kohout and Kevin Young

Quantum Performance Laboratory
@ Sandia National Laboratories
Livermore, CA and Albuquerque, NM, (and Chicago, Il)

Quantum Performance Laboratory

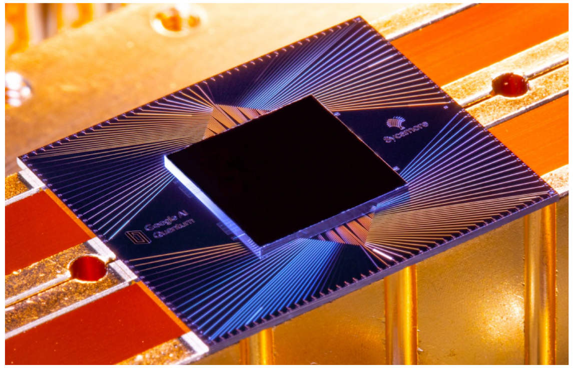
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

U.S. DEPARTMENT OF ENERGY NNSA

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Motivation

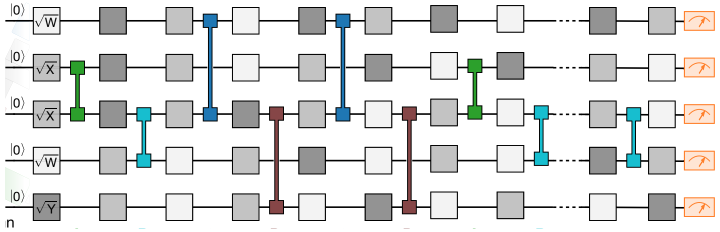
Want to buy my 100 qubit quantum computer?



Arute et al., Nature 574, 505 (2019)

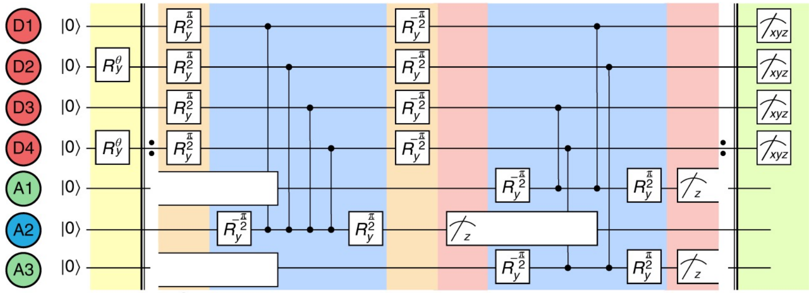
What applications or algorithms can it run?

Can it run random circuit sampling?



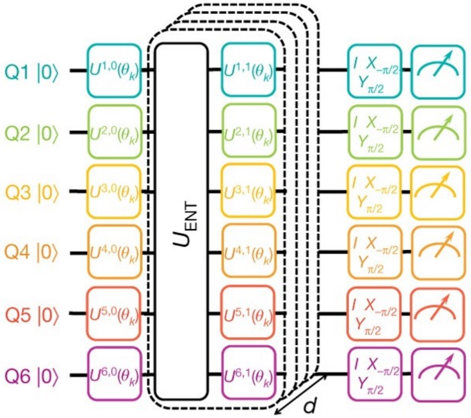
Arute et al., Nature 574, 505 (2019)

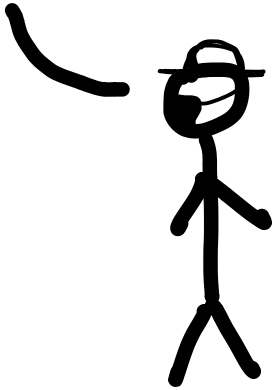
Quantum error correction?



Anderson et al., Nat. Phys. 16, 875 (2020)

What about QAOA?



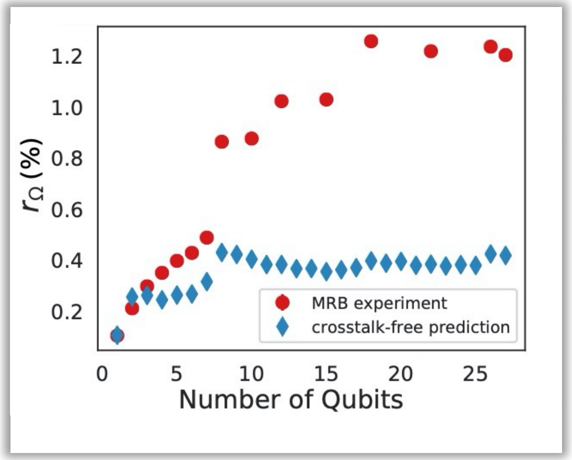


Kandala et al.,Nature 549, 242 (2017)

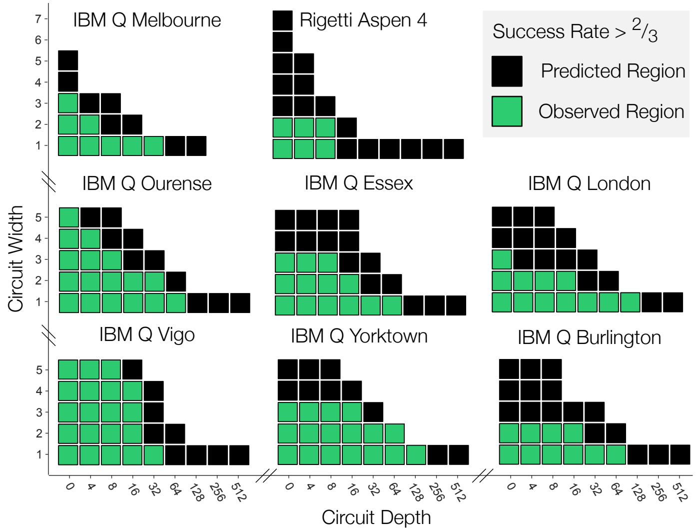
What about VQE? What about Phase Estimation? The QFT? Etc… etc…

We need holistic benchmarks

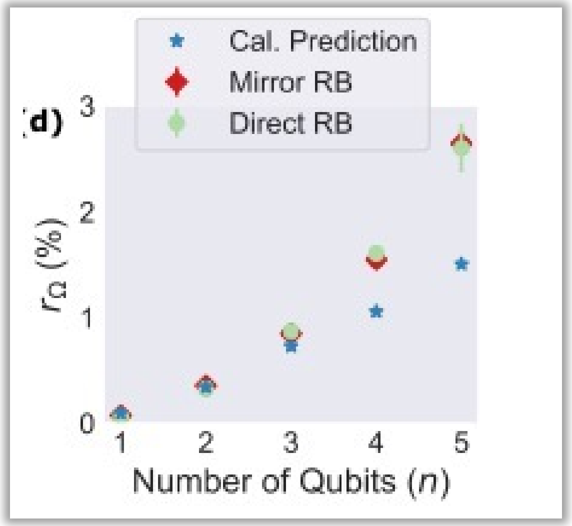
- Real-world quantum computers are subject to errors that only appear in many-qubit circuits.



See N38.00011 (Jordan Hines).



T. Proctor *et al.* Nature Physics 18, 75-79 (2022)



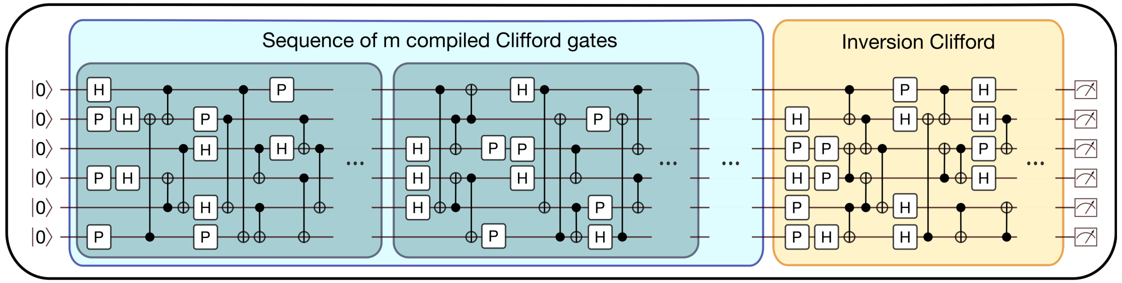
- So we cannot accurately predict the success rates of many-qubit circuits from one- and two-qubit gate error rates.

- To benchmark a 100-qubit quantum computer we're going to need some 100 qubit benchmarking circuits…

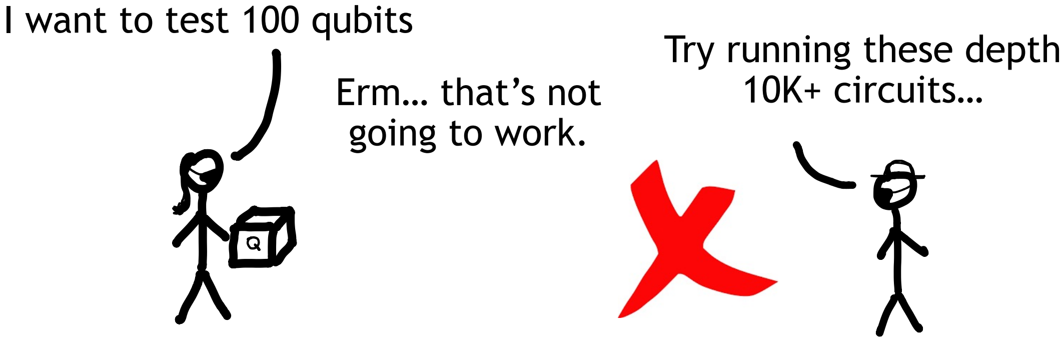
But most benchmarks aren't feasible on 50+ qubits

- Standard randomized benchmarking (RB) doesn't scale because it requires gate compilation.

Clifford Randomized Benchmarking

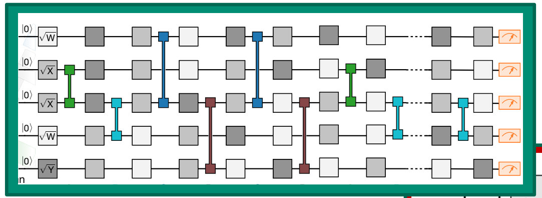


Standard RB: Magesan *et al*, PRL (2011). Figure from Proctor et al., PRL 100, 032328 (2019).



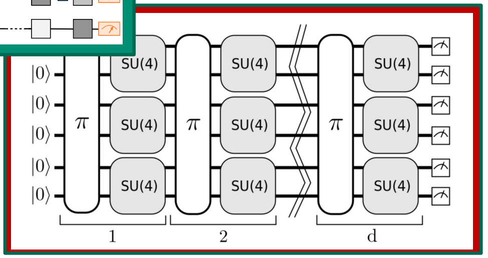
- Many other benchmarks require **exponentially expensive** classical computations.

Cross-entropy benchmarking



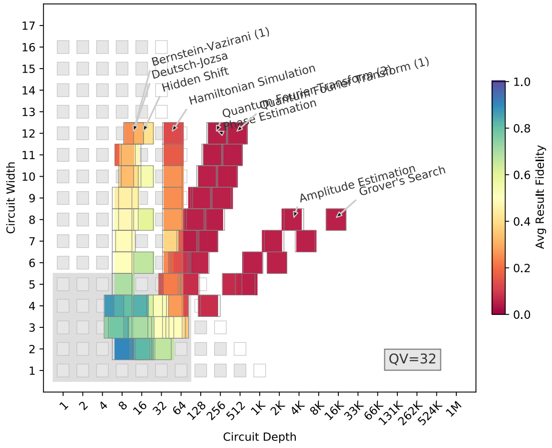
Arute et al., Nature 574, 505 (2019)

Quantum volume

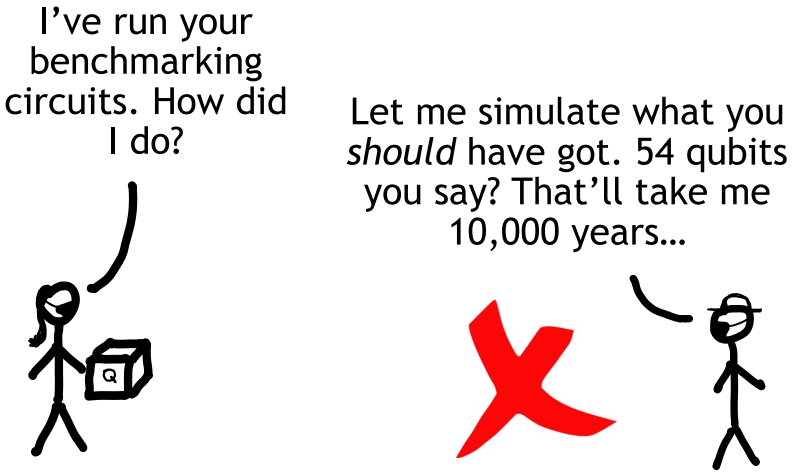


Cross et al., PRA 100, 032328 (2019)

Algorithmic benchmarks

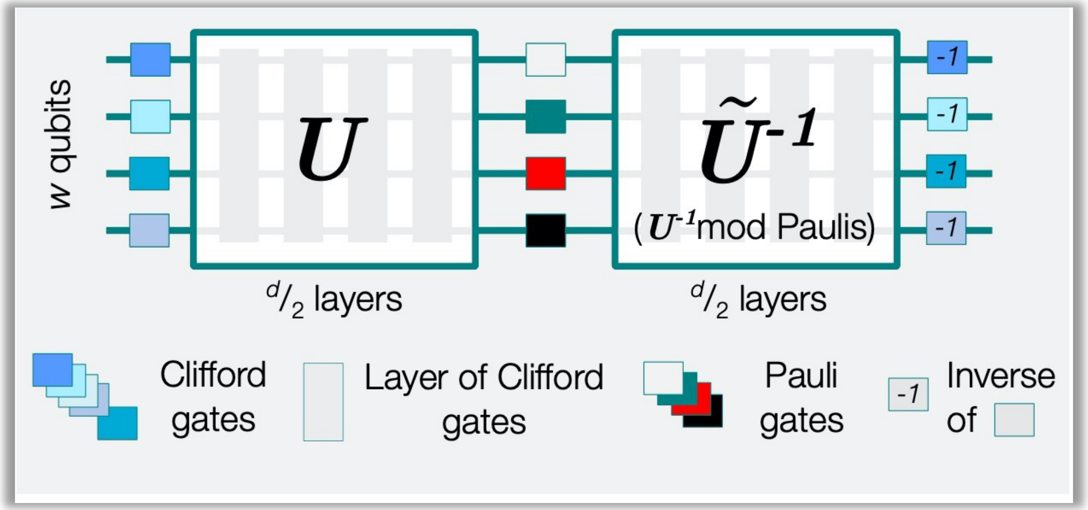


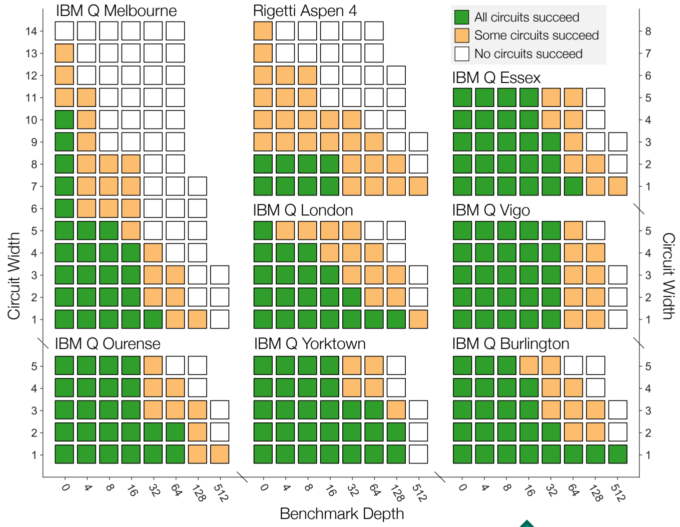
Lubinski et al., arXiv:2110.03137

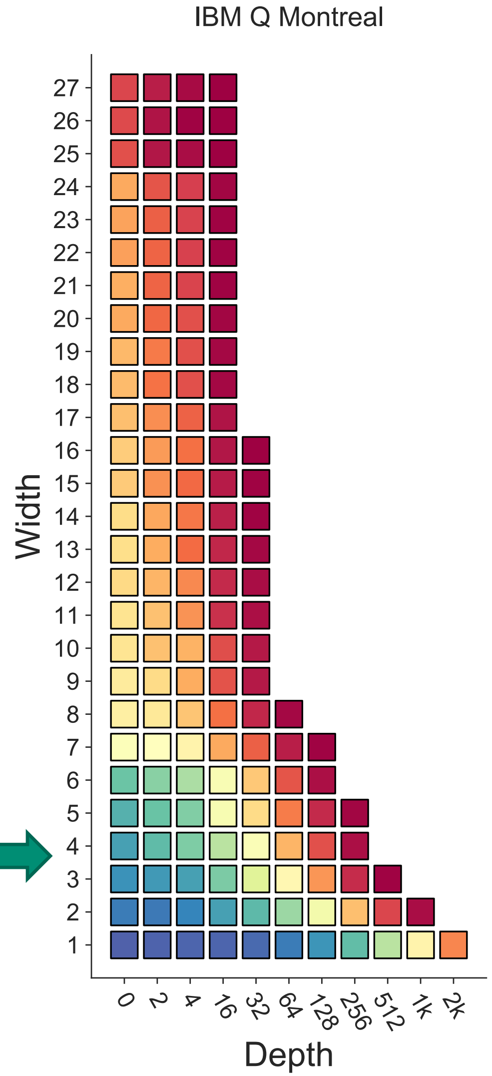


Scalable benchmarking using mirror circuits

- Mirror circuits[1] are a general technique for constructing *scalable* benchmarks







- They can be used to map out a device's performance as a function of circuit features, such as width and depth.[2]

- But generating these performance plots involves running a lot of circuits… **is all this data really necessary**?

Each square summarizes data from running 10s of circuits at that shape.

[1]T. Proctor *et al.* Nature Physics 18, 75-79 (2022), [2]Blume-Kohout and Young Quantum 4, 362 (2020)

Randomized benchmarking using randomized mirror circuits

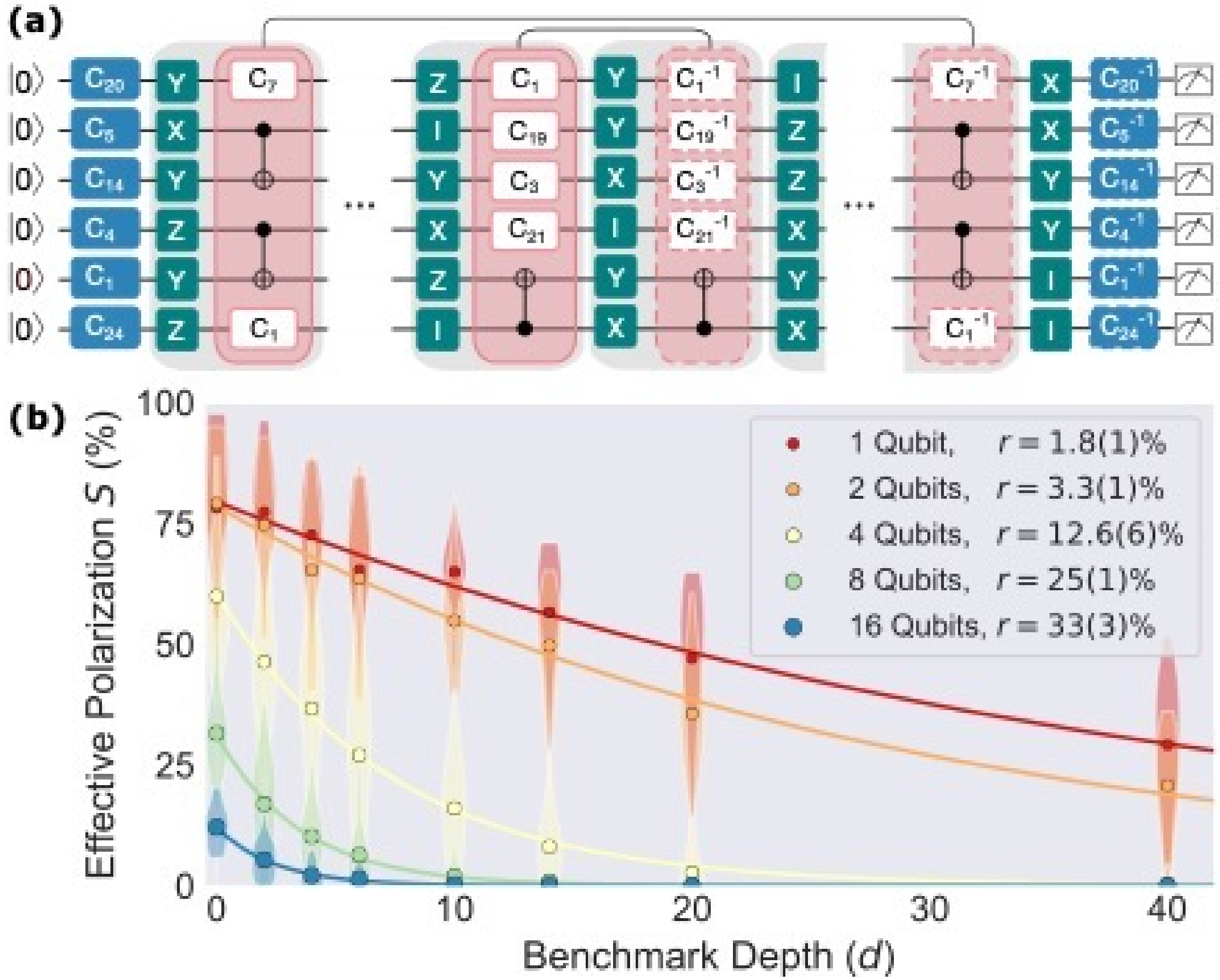
- Randomized mirror circuits[1] can be used to estimate average gate error rates, like traditional randomized benchmarking.[2]
- We compute each circuit's *effective polarization*:

$$S = \frac{4^n}{4^n - 1}\left[\sum_{k=0}^{n}\left(-\frac{1}{2}\right)^k h_k\right] - \frac{1}{4^n - 1},$$

where h_k is the rate that the output bit-string is a Hamming distance of k from the "target" bit-string.

- We fit the mean S as function of depth d to:

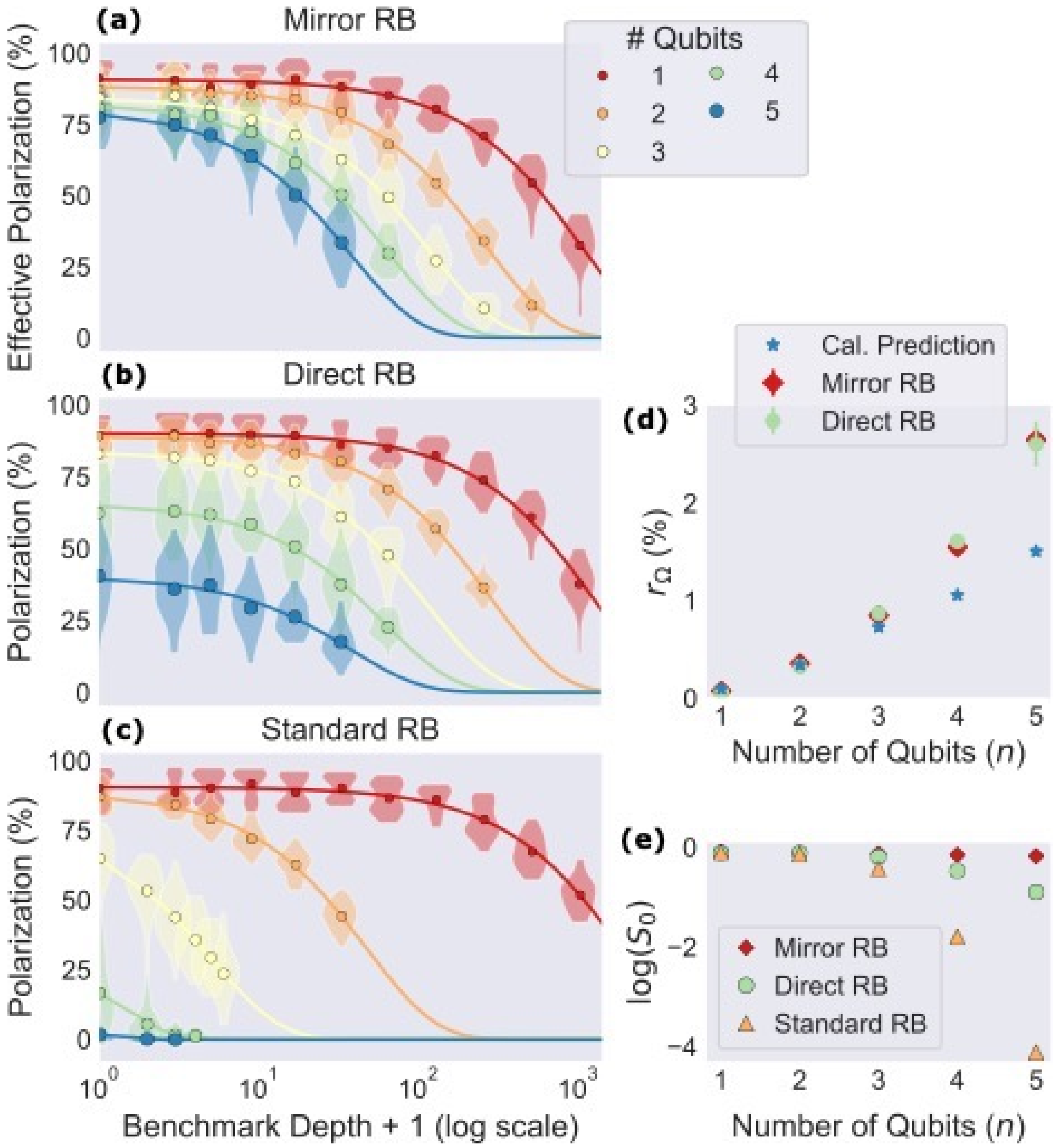
$$\bar{S}_d = Ap^d$$



[1]T. Proctor et al. arXiv:2112.09853 (2021), [2]Magesan *et al*, PRL (2011) [3]Proctor *et al*, PRL (2019)

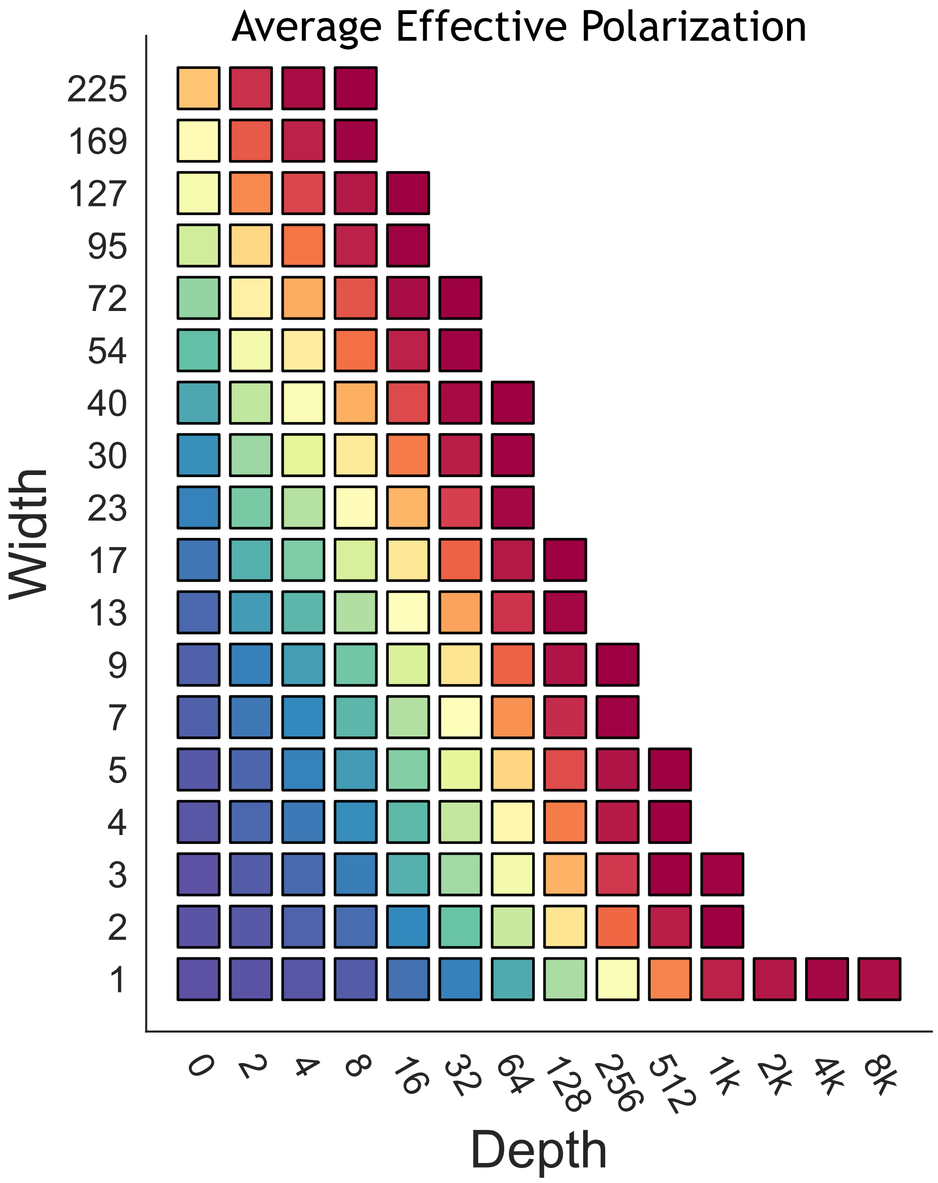
Randomized benchmarking using randomized mirror circuits

- Randomized mirror circuits[1] can be used to estimate average gate error rates, like traditional randomized benchmarking.[2]
- This method is a huge improvement on traditional RB.[2-3]
    - Traditional RB scales only to ~4-5 qubits.
    - RB with randomized mirror circuits scales to 100s or 1000s of qubits!



[2]Magesan *et al*, PRL (2011) [3]Proctor *et al*, PRL (2019)

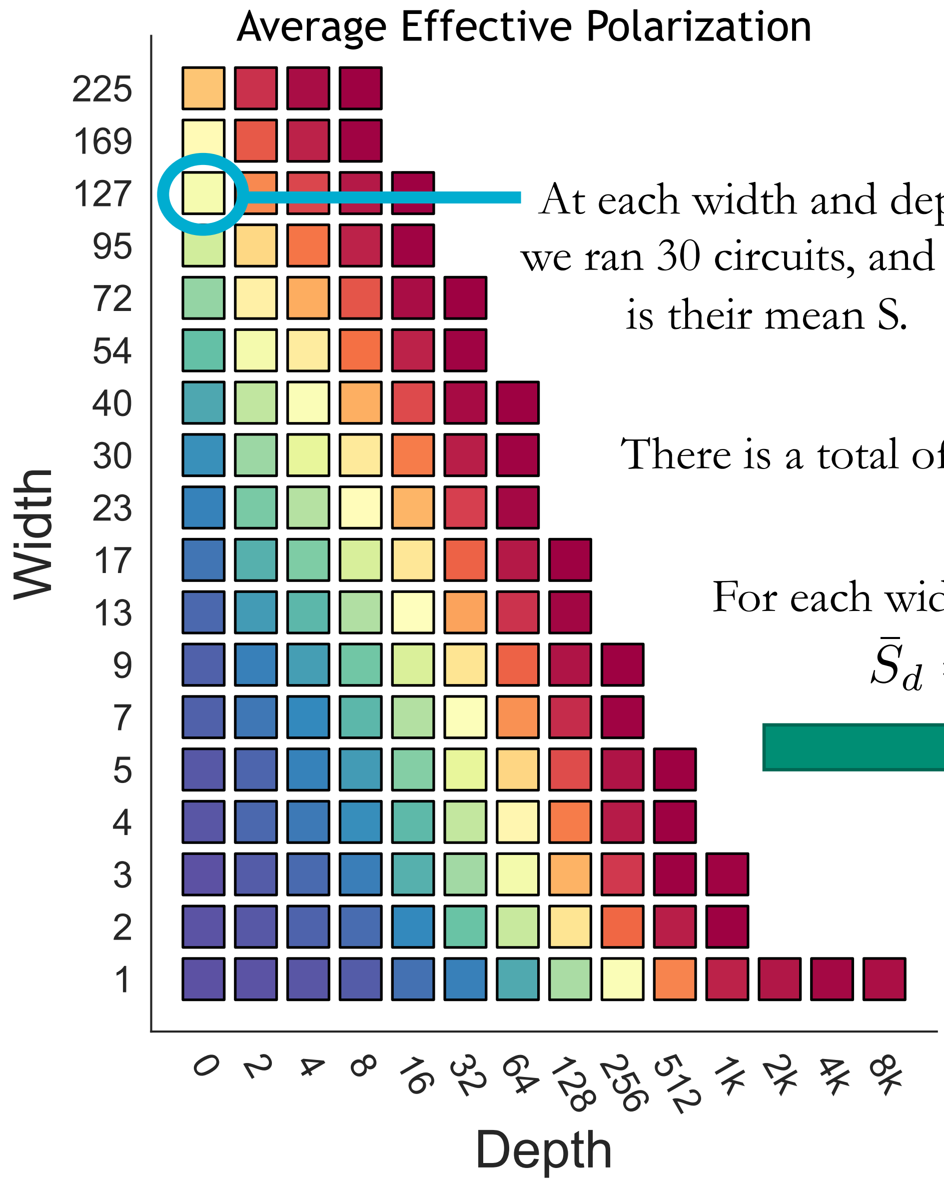
Benchmarking 225 qubits using 1000s of circuits (simulation)



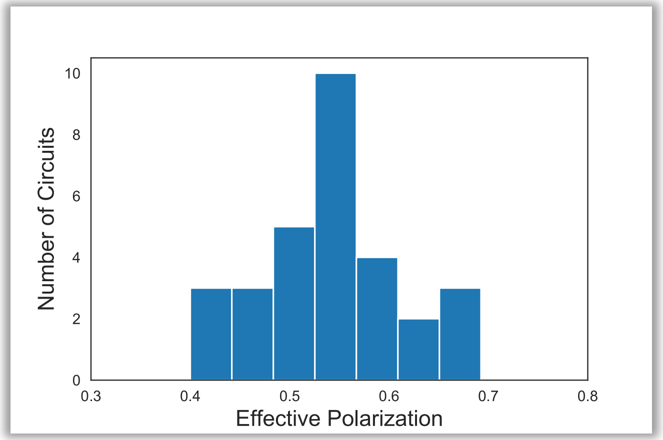
Simulation details

- 225 qubits connected in a 15 by 15 lattice.
- Clifford gates subject to Pauli stochastic errors.
  - ~0.1% error for 1-qubit gates.
  - ~1% error for 2-qubit gates.
  - ~0.5% readout error on each qubit.

Benchmarking 225 qubits using 1000s of circuits (simulation)



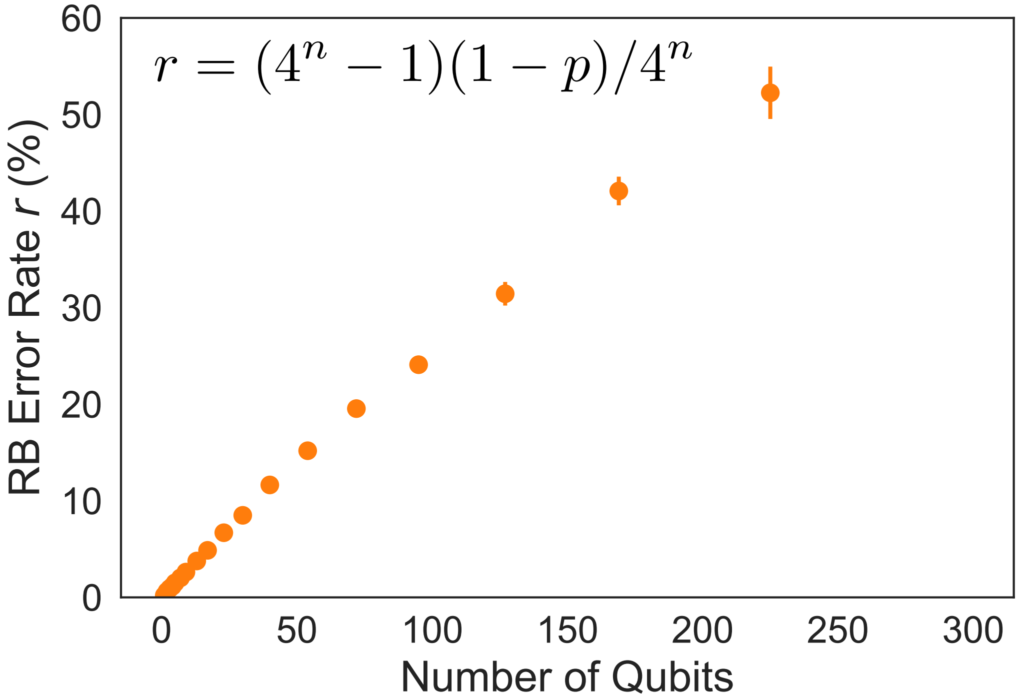
At each width and depth we ran 30 circuits, and this is their mean S.



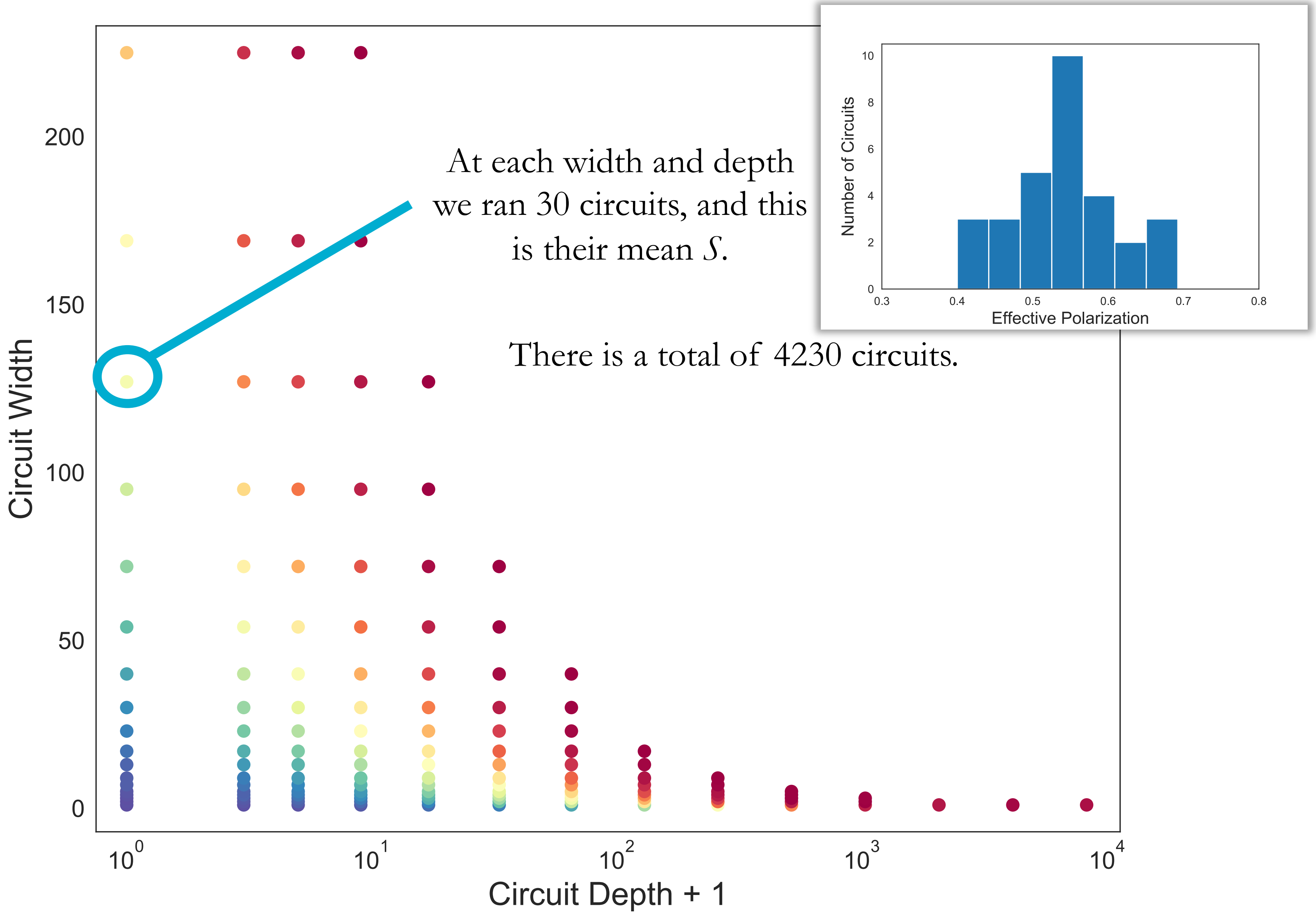
There is a total of 4230 circuits.

For each width, fit data to

$$\bar{S}_d = Ap^d$$

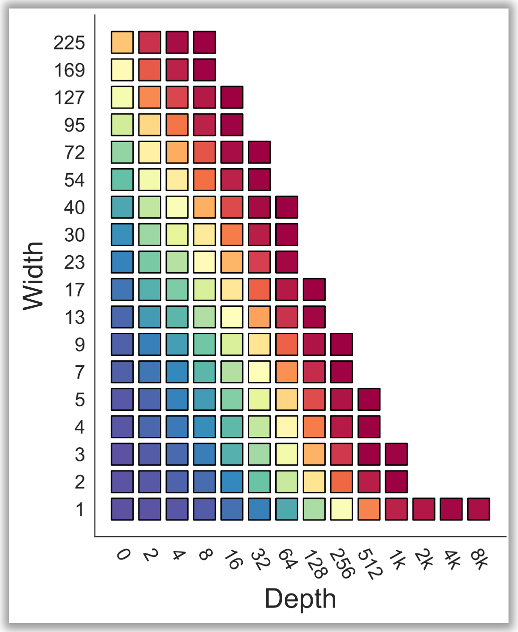


Benchmarking 225 qubits using 1000s of circuits (simulation)

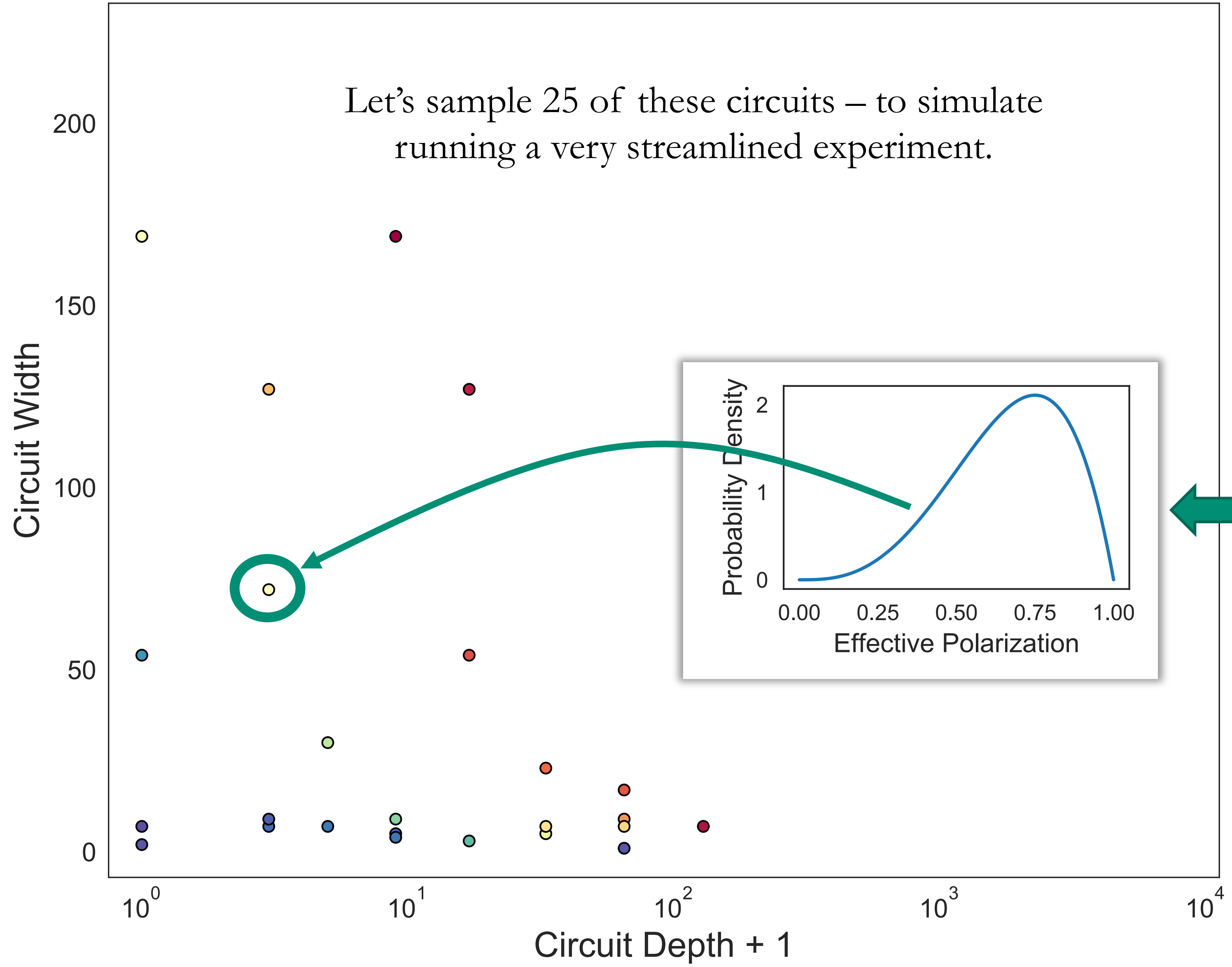


At each width and depth we ran 30 circuits, and this is their mean S.

There is a total of 4230 circuits.



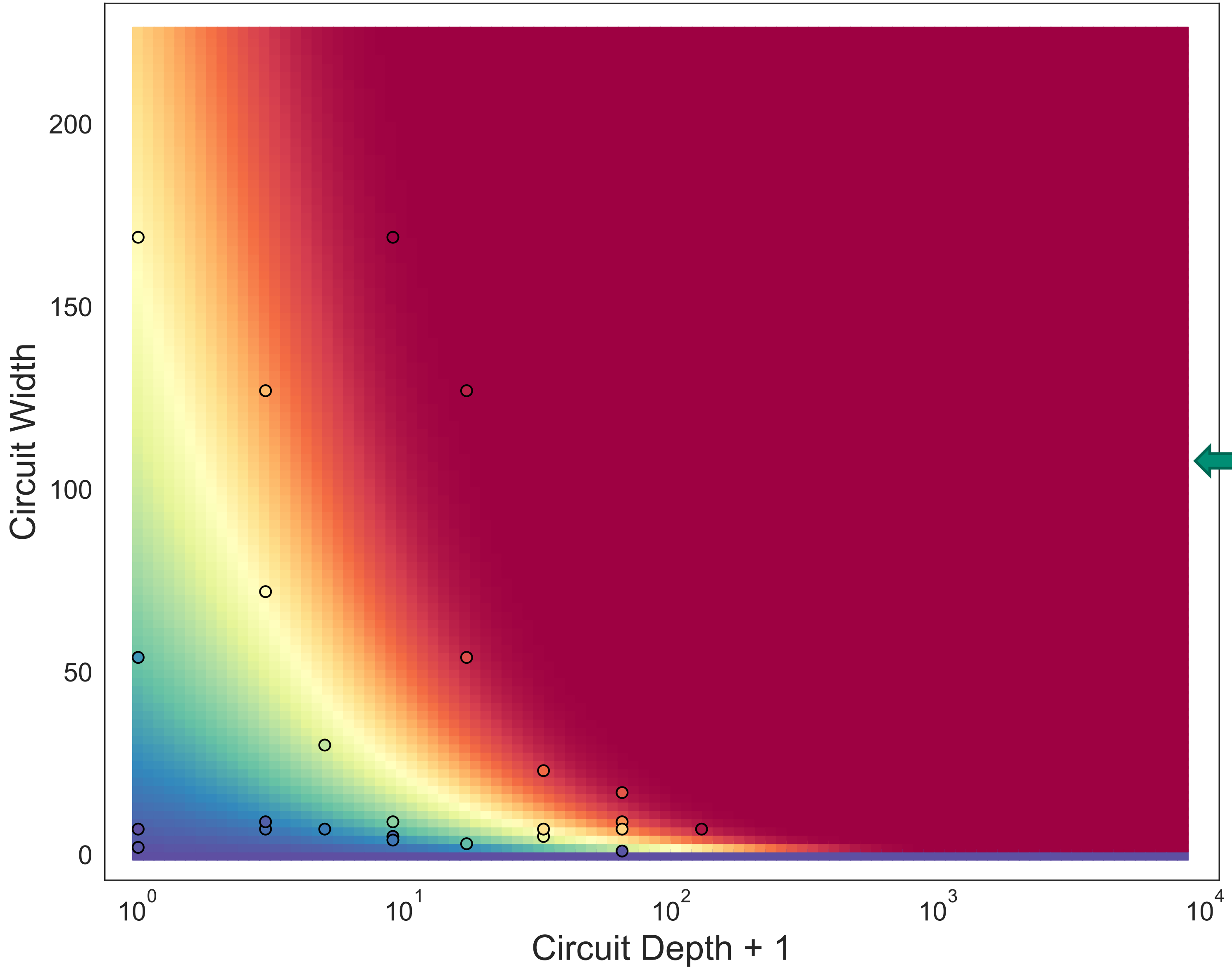
Benchmarking 225 qubits with only 25 circuits (simulation)



- How can we learn the full performance map from this small set of data?
- We use a few-parameter predictive model, and fit it to the data.
- What's a good model?
    - Each data point is a sample from an unknown and (w, d) dependent distribution over [0, 1].
    - We model this distribution by a *beta distribution* with a (w, d) dependent mean and variance.
    - We pick a simple few-parameter function of w and d for the mean and variance.

Technical note: effective polarizations are not bounded below by zero, so we rescale the beta distribution.

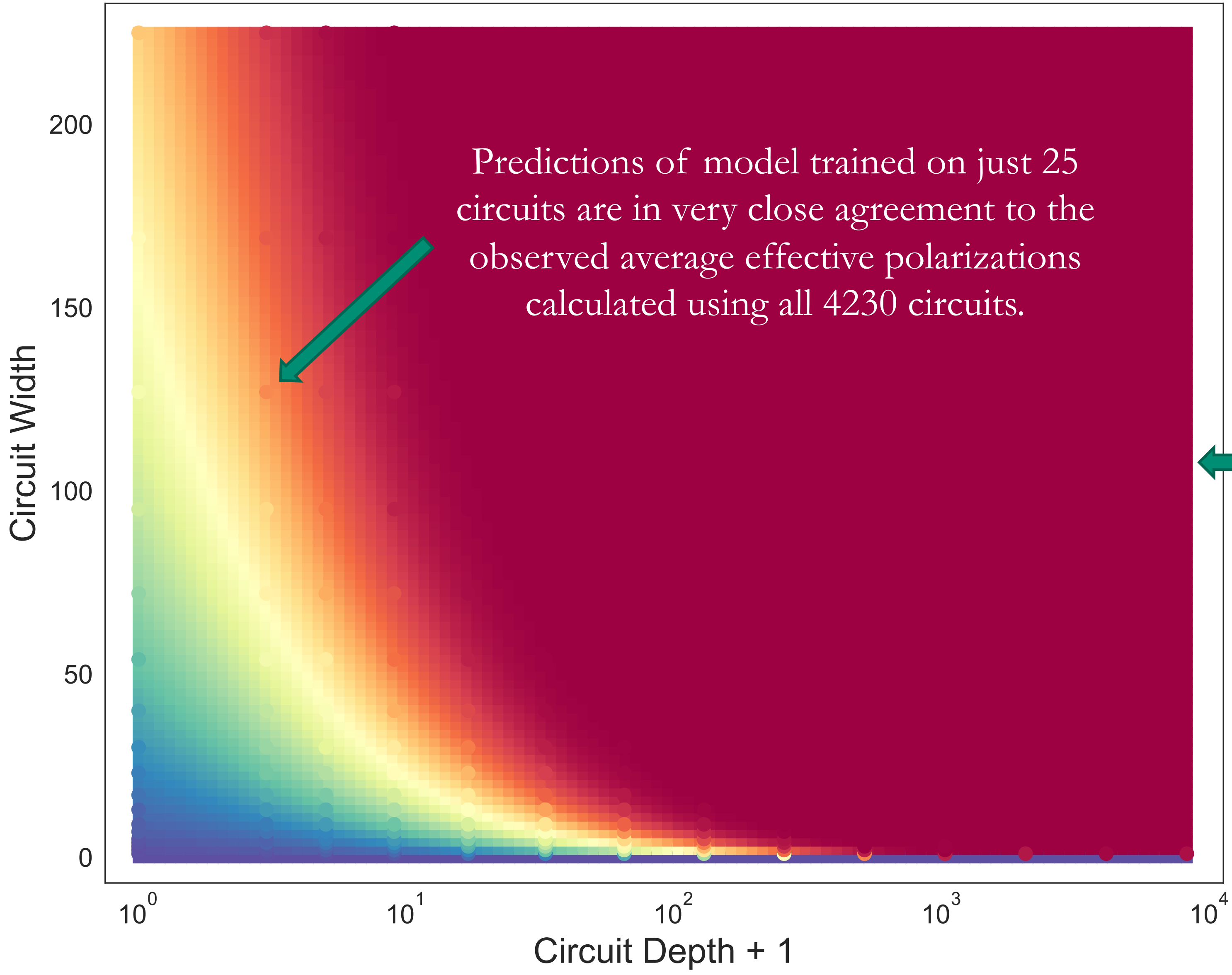
Benchmarking 225 qubits with only 25 circuits (simulation)



- How can we learn the full performance map from this small set of data?
- We use a few-parameter predictive model, and fit it to the data (using maximum likelihood estimation).

Heatmap is the predicted average effective polarization as a *continuous* function of (width, depth).

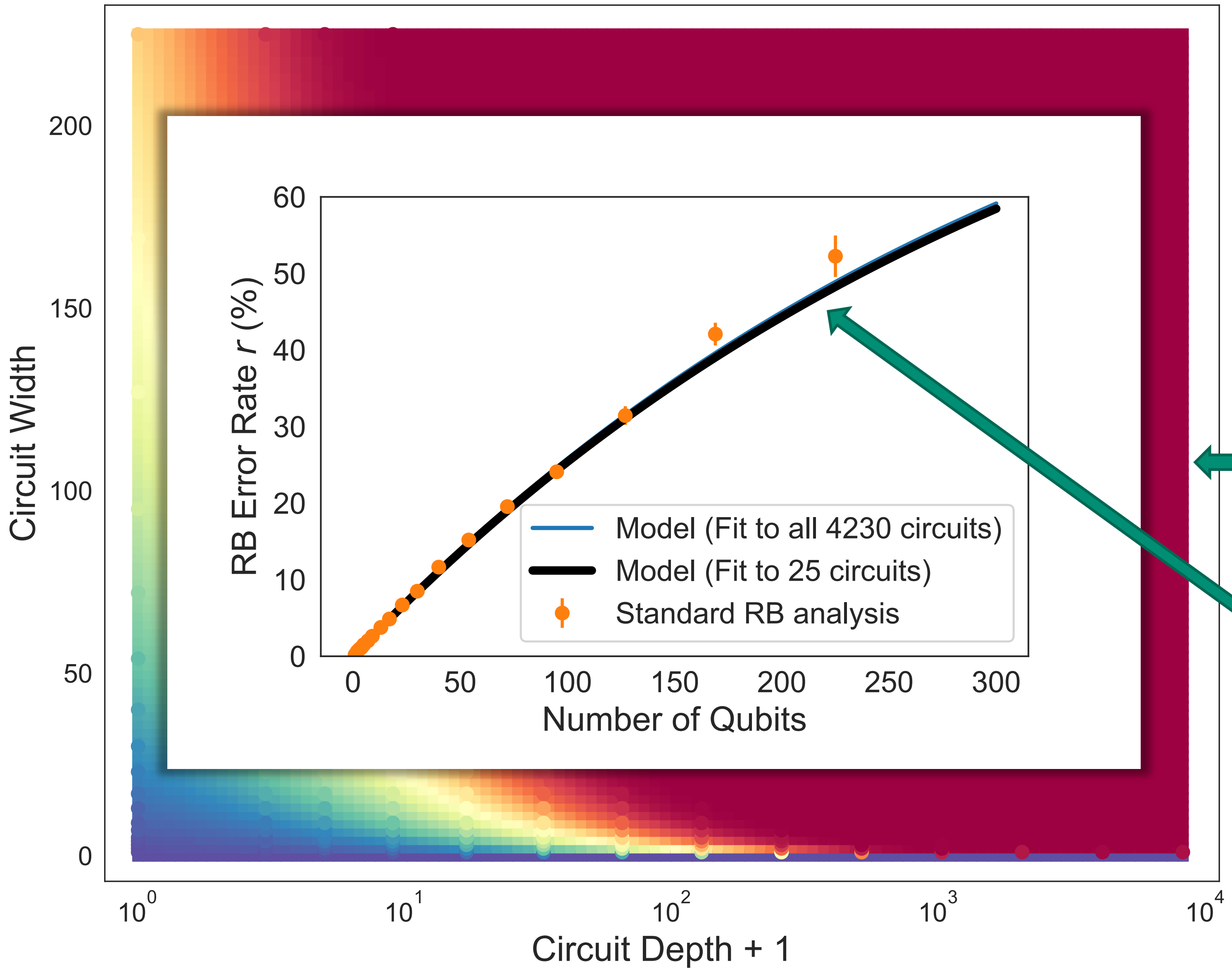
Benchmarking 225 qubits with only 25 circuits (simulation)



- How can we learn the full performance map from this small set of data?
- We use a few-parameter predictive model, and fit it to the data (using maximum likelihood estimation).

Heatmap is the predicted average effective polarization as a *continuous* function of (width, depth).

Benchmarking 225 qubits with only 25 circuits (simulation)

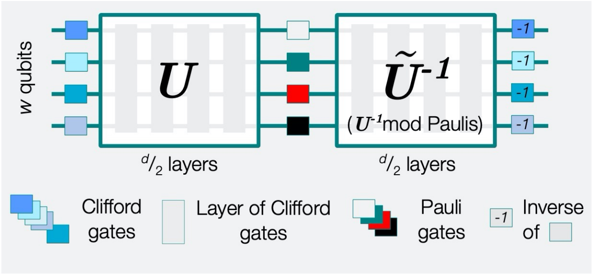


- How can we learn the full performance map from this small set of data?
- We use a few-parameter predictive model, and fit it to the data (using maximum likelihood estimation).

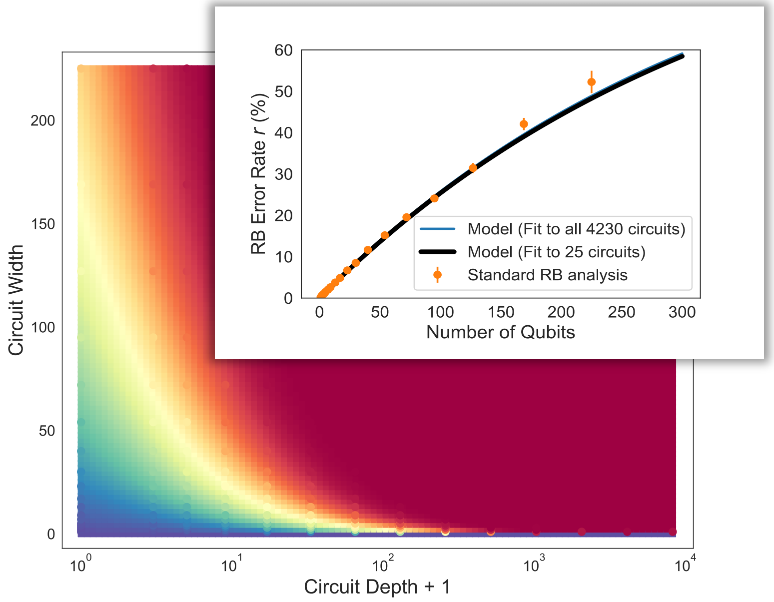
Heatmap is the predicted average effective polarization as a *continuous* function of (width, depth).

We've learnt a good approximation to the per-layer error rate error as function of the number of qubits – using data from only 25 circuit!

Summary



- We need scalable and efficient holistic benchmarks for quantum computers.
  - Many popular benchmarks require exponentially expensive classical computation
- Circuit mirroring can convert an arbitrary circuit into an efficiently verifiable circuit. It enables:
  - Scalable benchmarks built from any circuits, including algorithm circuits (see Stefan Seritan's talk, N38.00008).
  - Scalable randomized benchmarking of Clifford gates (this talk) and universal gate sets (see Jordan Hines' talk, N38.00011).
  - Scalable algorithm verification (see Mohan Sarovar's talk, N38.00010).
- We can benchmark 100+ qubits using only a handful of randomized mirror circuits.
- Techniques for interpolating data from general benchmarking circuits would be a really powerful tool for super-efficient benchmarking.



Where can I read more?

Circuit mirroring: T. Proctor *et al.* Nature Physics 18, 75-79 (2022).

Randomized benchmarking using mirror circuits: T. Proctor *et al.* arXiv:2112.09853 (2021).

Efficient extrapolation of benchmarking data: look out for an arXiv posting soon-ish.

The Team

Kenneth Rudinger, Stefan Seritan, Daniel Hothem, Jordan Hines, Thomas Catanach, Robin Blume-Kohout and Kevin Young

Quantum Performance Laboratory

Thanks!

Many thanks to IBM Quantum Experience for access to their quantum computing platform.

Get Your Capabilities Checked Now!

If you'd like to run mirror circuit benchmarks to test a processor's capabilities:

- Get in contact with me (tjproct@sandia.gov) or anyone at Sandia's QPL.
- Code for running experiments like these is in `pyGSTi` (*www.pygsti.info*).

This work was funded in part by Sandia's Laboratory Directed Research and Development program. This material was funded in part by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research Quantum Testbed Program. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of the U.S. Government. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.