

Randomized benchmarking of many-qubit devices

Timothy Proctor¹, Kenneth Rudinger¹, Robin Blume-Kohout¹,
Arnaud Carignan-Dugas², Erik Nielsen¹, and Kevin Young¹

¹Sandia National Laboratories.

²University of Waterloo.

SQuInT 2019



Sandia National Laboratories



**U.S. DEPARTMENT OF
ENERGY**



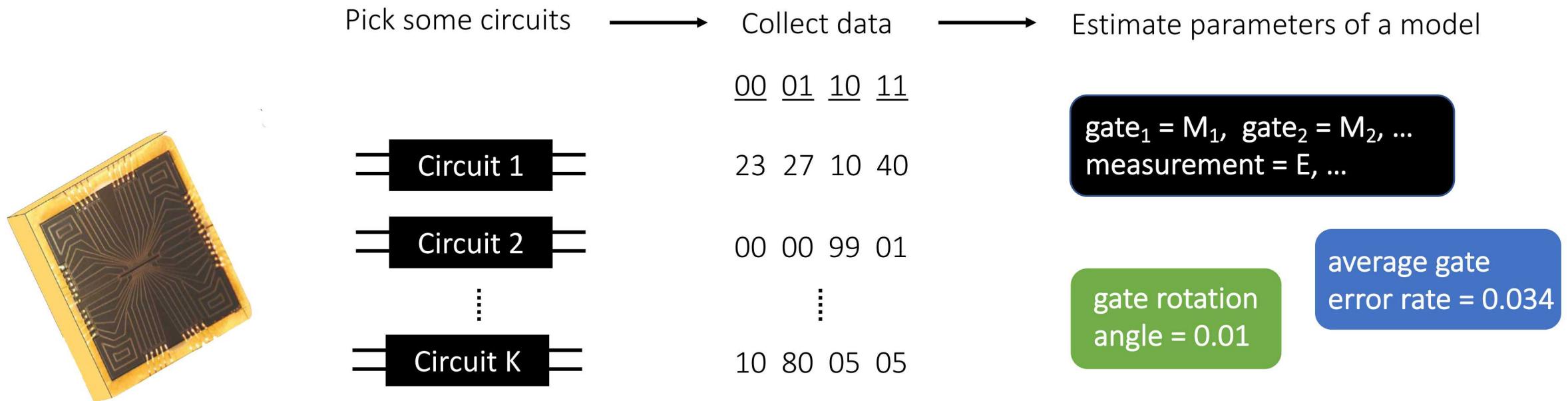
IARPA
BE THE FUTURE

Office of the Director of National Intelligence

Part I – Background

Device characterization and benchmarking

- Characterization & benchmarking = Techniques for understanding device performance.



- Examples: randomized benchmarking¹⁻³, tomographic methods, robust phase estimation, ...

¹Magesan *et al.*, PRL 106, 180504 (2011), ²Emerson *et al.*, Science 317, 1893 (2007), ³Knill *et al.*, PRA 77, 012307 (2008).

Randomized benchmarking (RB)

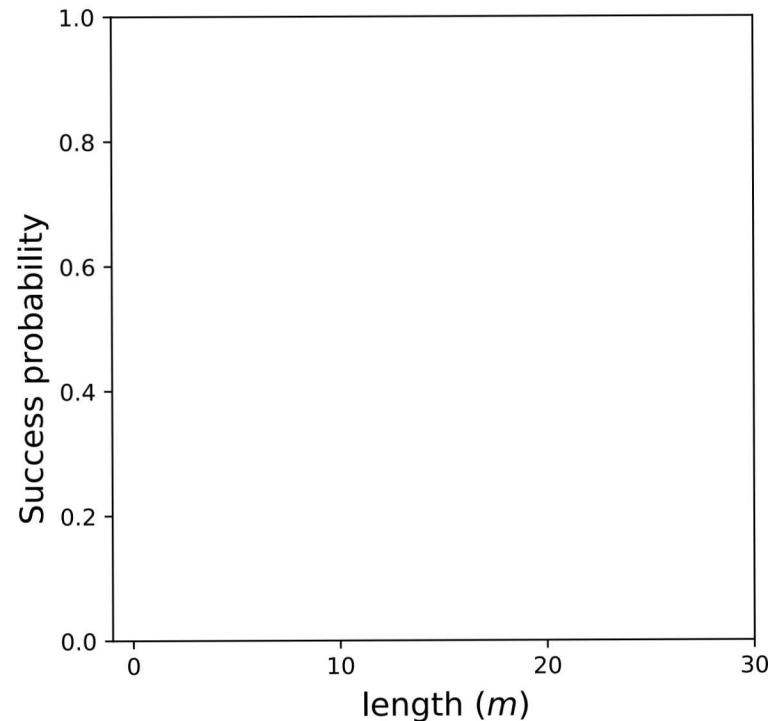
- Benchmarks gates using random circuits, returning an average error rate for the gates.
- Arguably the first method for characterizing a quantum device that is not prone to catastrophic systematic errors (unlike, e.g., “ordinary” process tomography).
- Easy and fast to use.
- Has become a *de facto* standard technique.
- The current standard RB method¹ benchmarks the Clifford group, and is often called “Clifford RB”.

(Clifford group = gates that map Pauli gates to Pauli gates under conjugation.)

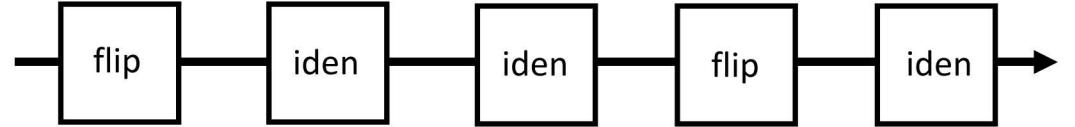
¹Magesan *et al.*, PRL 106, 180504 (2011), ²Emerson *et al.*, Science 317, 1893 (2007), ³Knill *et al.*, PRA 77, 012307 (2008).

Clifford RB on n qubits

- Pick a set of non-negative integers m , and for each m :

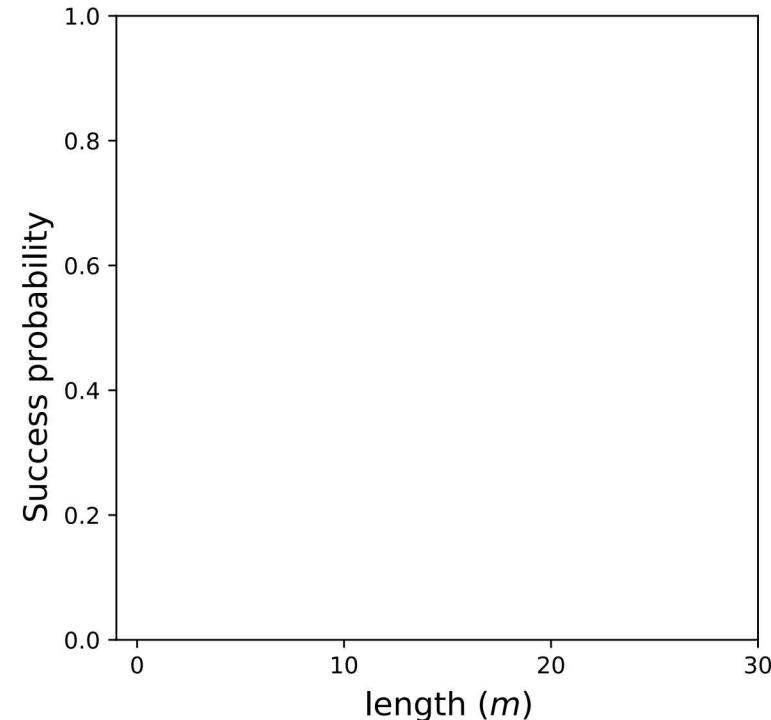


Classical analogy gates = { bit flip, identity }



Clifford RB on n qubits

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.

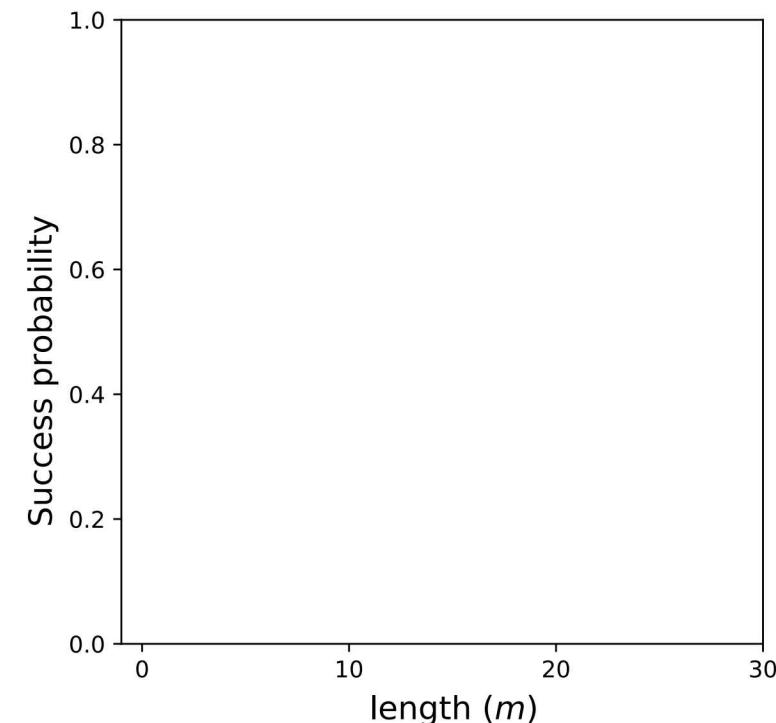


Clifford RB on n qubits

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.



- For each circuit, repeat 1+ times:
 - Prepare all n qubits in $|0\rangle$
 - Run the random circuit on the qubits.
 - Record “success” if all n qubits return 0.

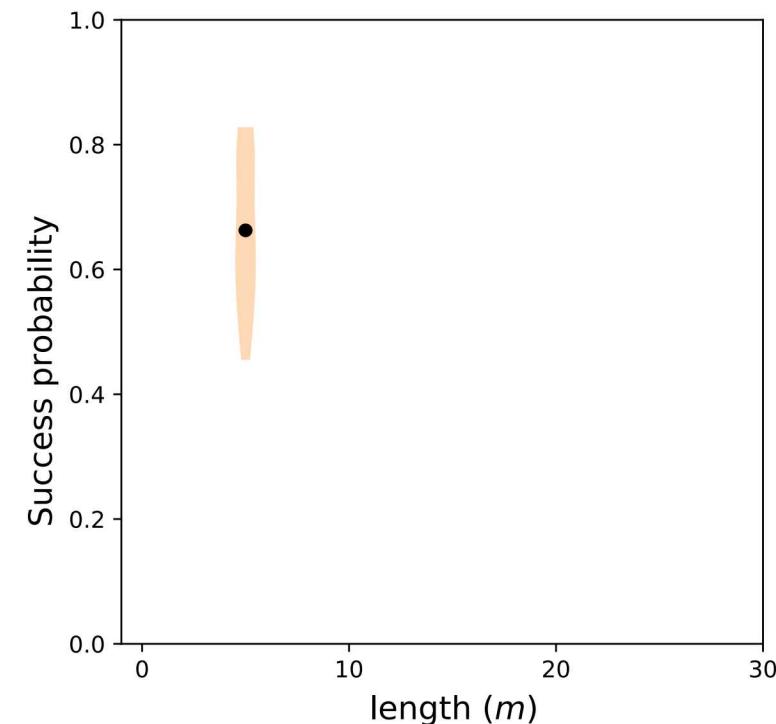


Clifford RB on n qubits

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.



- For each circuit, repeat 1+ times:
 - Prepare all n qubits in $|0\rangle$
 - Run the random circuit on the qubits.
 - Record “success” if all n qubits return 0.
- Calculate the average success probability P_m over the K circuits at this length m .

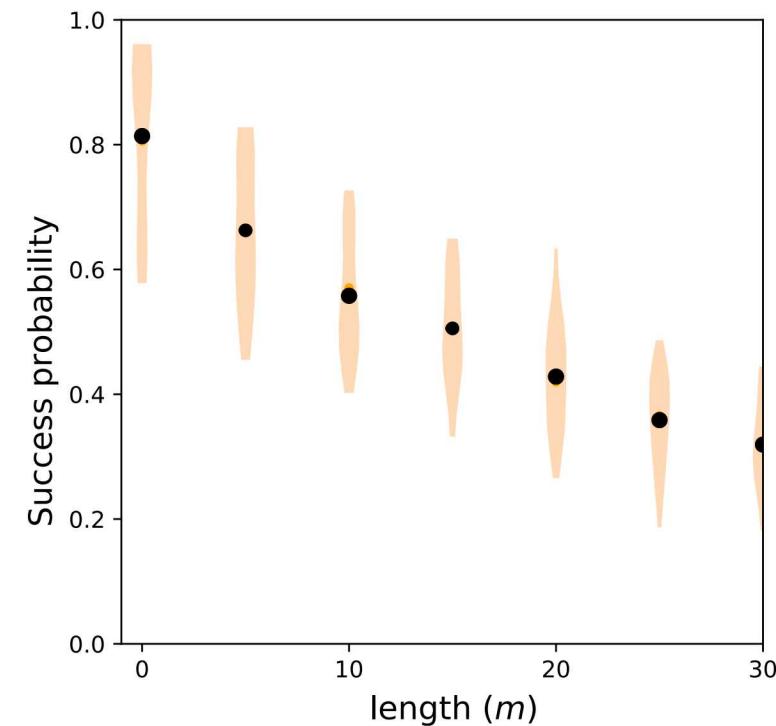


Clifford RB on n qubits

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.



- For each circuit, repeat 1+ times:
 - Prepare all n qubits in $|0\rangle$
 - Run the random circuit on the qubits.
 - Record “success” if all n qubits return 0.
- Calculate the average success probability P_m over the K circuits at this length m .

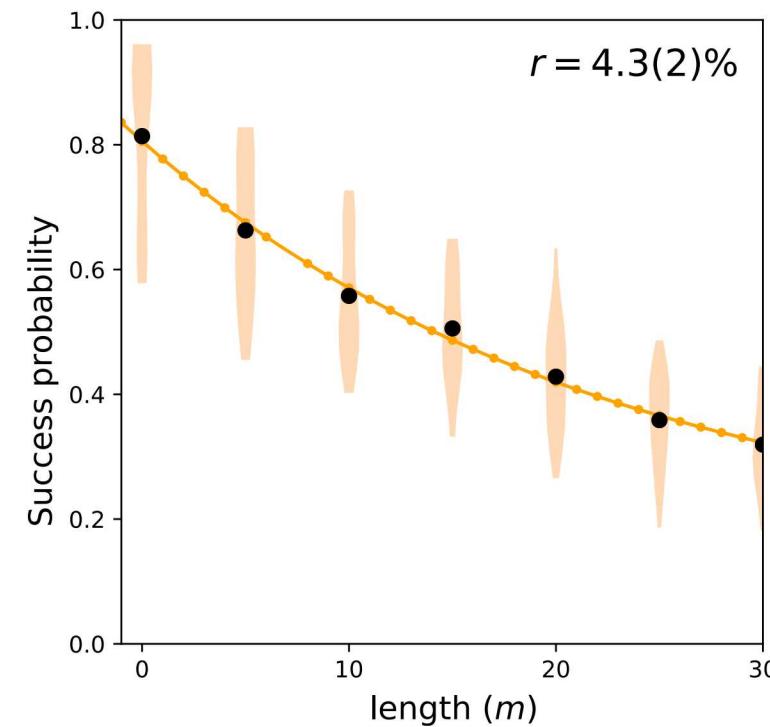


Clifford RB on n qubits

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.



- For each circuit, repeat 1+ times:
 - Prepare all n qubits in $|0\rangle$
 - Run the random circuit on the qubits.
 - Record “success” if all n qubits return 0.
- Calculate the average success probability P_m over the K circuits at this length m .
- Fit P_m to $P_m = A + Bp^m$, obtaining an error rate via $r \propto 1 - p$.



Clifford RB cannot holistically benchmark near-term devices

- Clifford RB scales well with the number of qubits n in *some* ways:
 - The number of experiments needed is independent of n .¹
 - The data analysis complexity is independent of n .²

¹Magesan *et al.*, PRA 85, 042311 (2012), ²Helsen *et al.*, arXiv:1701.04299 (2017), ³McKay *et al.*, arXiv:1712.06550 (2017).

Clifford RB cannot holistically benchmark near-term devices

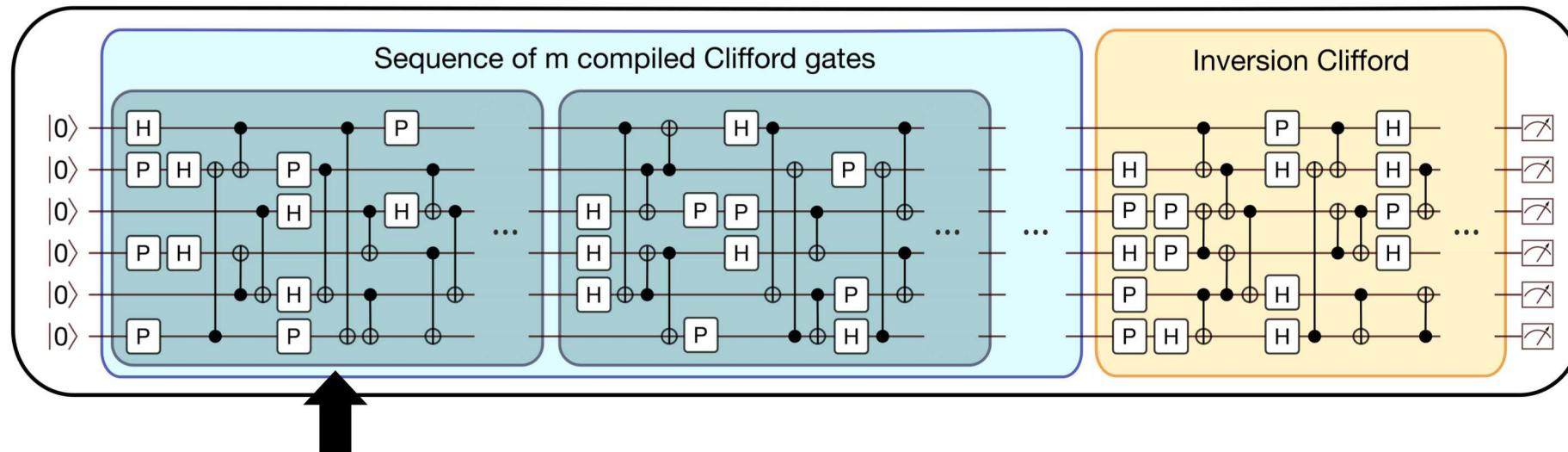
- Clifford RB scales well with the number of qubits n in *some* ways:
 - The number of experiments needed is independent of n .¹
 - The data analysis complexity is independent of n .²
- Yet, there is only one reported case of 3-qubit Clifford RB³, and none of 4+ qubit Clifford RB. Because...

Scaling Clifford RB up to many qubits is difficult because of
compilation.

¹Magesan *et al.*, PRA 85, 042311 (2012), ²Helsen *et al.*, arXiv:1701.04299 (2017), ³McKay *et al.*, arXiv:1712.06550 (2017).

Clifford RB cannot holistically benchmark near-term devices

- The full set of n -qubit Clifford gates are not computational primitives.
- Each Clifford gate has to be compiled into many 1- and 2-qubit gates ($O[n^2/\log(n)]$ of them¹).

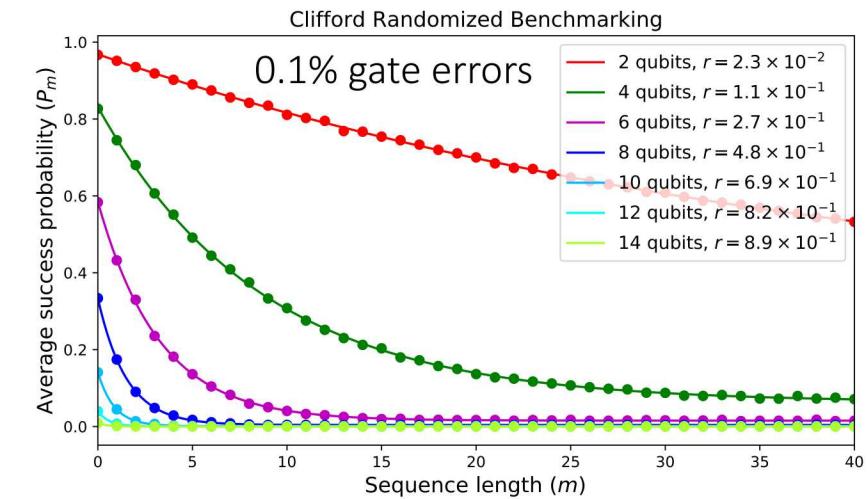


- A typical n -qubit Clifford is a circuit with a size that grows quickly with n .

¹Patel *et al.*, QIC 8, 282 (2008), Aaronson and Gottesman, PRA 70, 052328 (2014).

Clifford RB cannot holistically benchmark near-term devices

- The full set of n -qubit Clifford gates are not computable.
- Each Clifford gate has to be compiled into many 1-qubit gates (and then them¹).



Problem: The fidelity of an n -qubit Clifford quickly decreases with increasing n , severely limiting the number of qubits Clifford RB can handle.

- A typical n -qubit Clifford is a circuit with a size that grows quickly with n .

¹Patel *et al.*, QIC 8, 282 (2008), Aaronson and Gottesman, PRA 70, 052328 (2014).

Part II – Streamlining randomized benchmarking

Streamlining RB

We've developed a streamlined version of RB, called *direct RB*.¹

- Direct RB is feasible on more qubits than Clifford RB.
- Direct RB retains the core robustness and simplicity of Clifford RB.
- Direct RB directly measures an average error rate of the native gates.
- Direct RB is flexible, as it can be used to extract a variety of error rates.

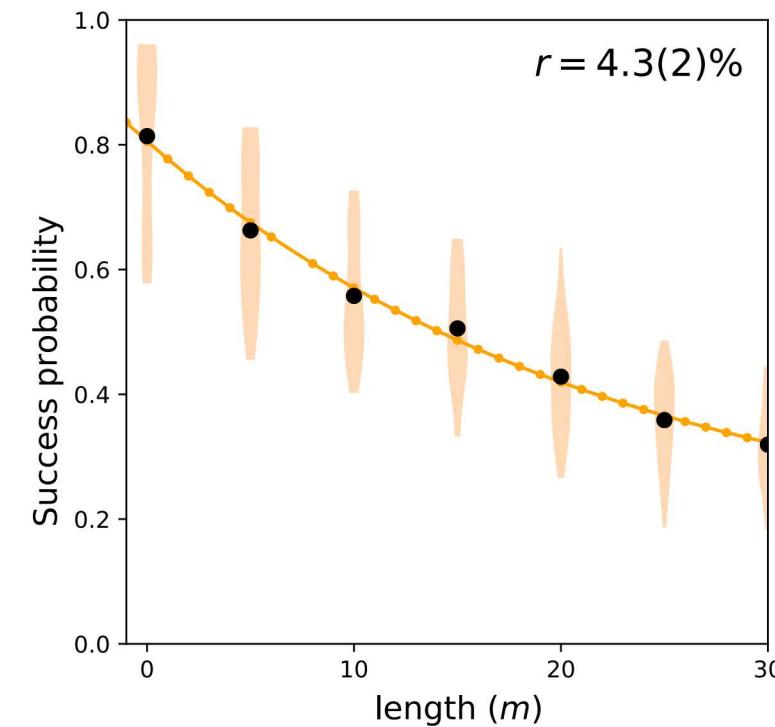
¹Proctor et al., arXiv:1807.07975 (2018).

Clifford randomized benchmarking

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.

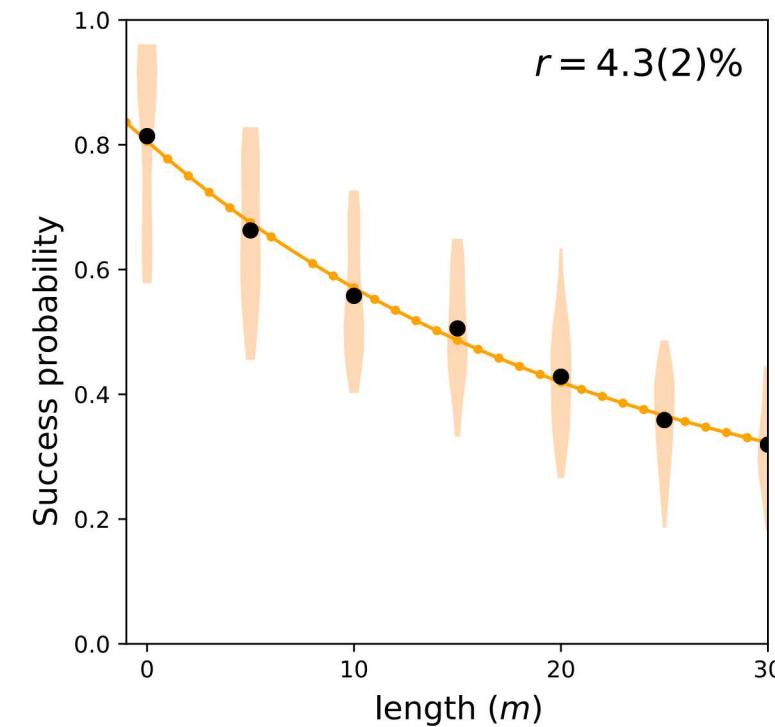


- For each circuit, repeat 1+ times:
 - Prepare all n qubits in $|0\rangle$
 - Run the random circuit on the qubits.
 - Record “success” if all n qubits return 0.
- Calculate the average success probability P_m over the K circuits at this length m .
- Fit P_m to $P_m = A + Bp^m$, obtaining an error rate via $r \propto 1 - p$.



Direct randomized benchmarking

- Pick a set of non-negative integers m , and for each m :
 - Randomly sample K length- $(m+2)$ identity Clifford circuits.
- For each circuit, repeat 1+ times:
 - Prepare all n qubits in $|0\rangle$
 - Run the random circuit on the qubits.
 - Record “success” if all n qubits return 0.
- Calculate the average success probability P_m over the K circuits at this length m .
- Fit P_m to $P_m = A + Bp^m$, obtaining an error rate via $r \propto 1 - p$.



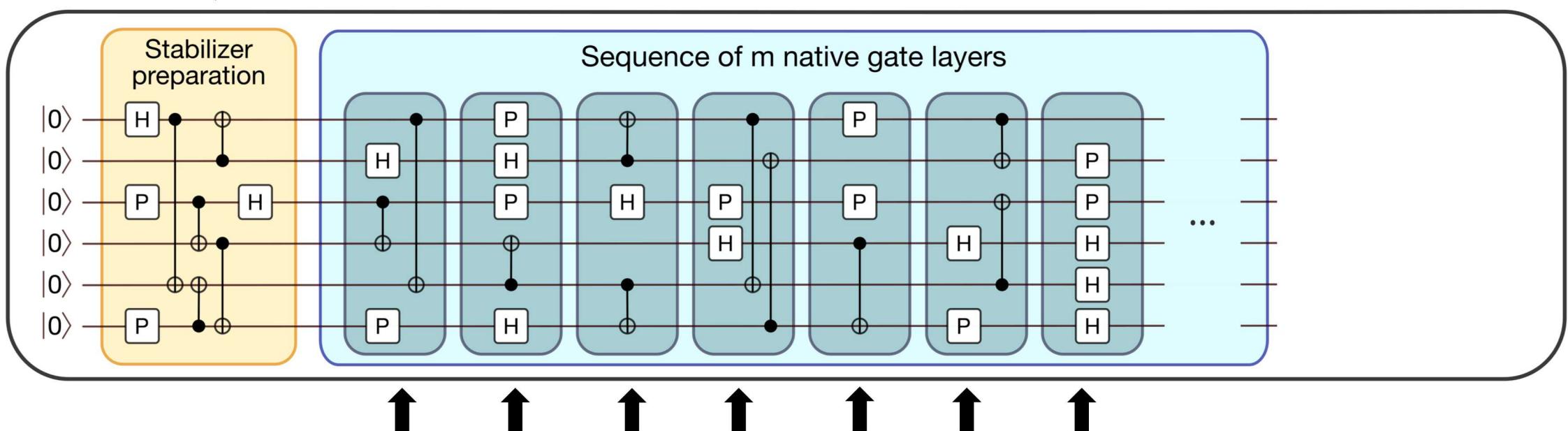
Direct randomized benchmarking

Prepare a uniformly random
 n -qubit stabilizer state. *Purpose: port the
robustness of Clifford RB to the new method.*



Direct randomized benchmarking

Prepare a uniformly random
 n -qubit stabilizer state. *Purpose: port the
robustness of Clifford RB to the new method.*

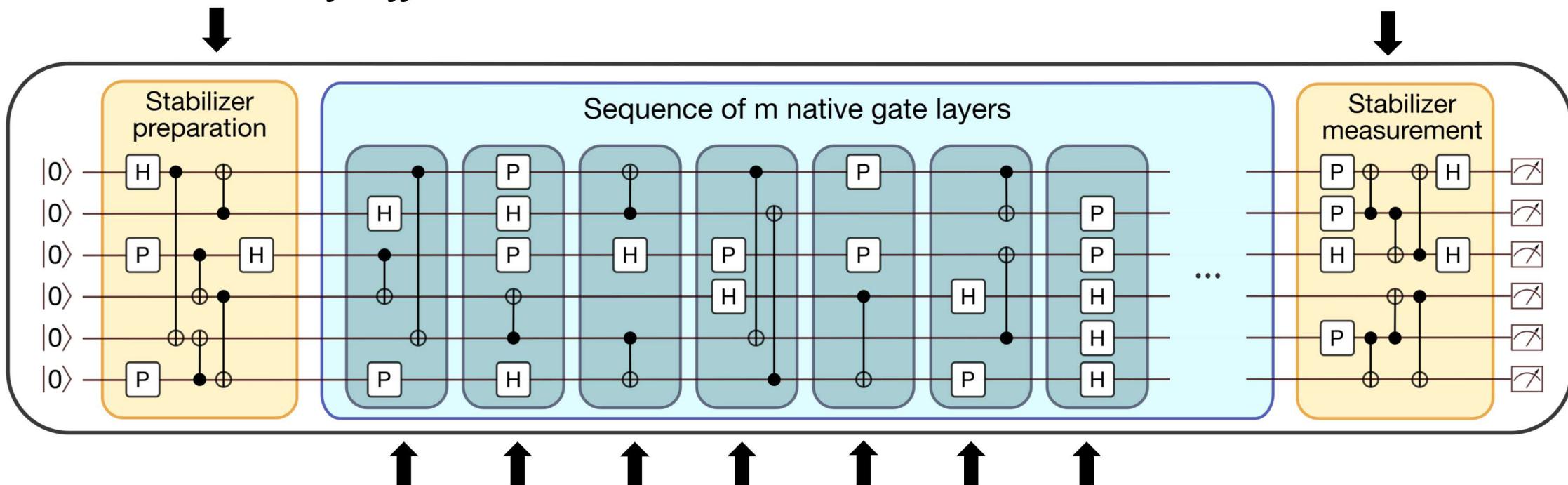


Depth-1 circuit layers sampled according to a user-specified and flexible distribution. *Purpose: streamlined layers mean (1) better scalability and (2) estimation of a per-native-gate error rate.*

Direct randomized benchmarking

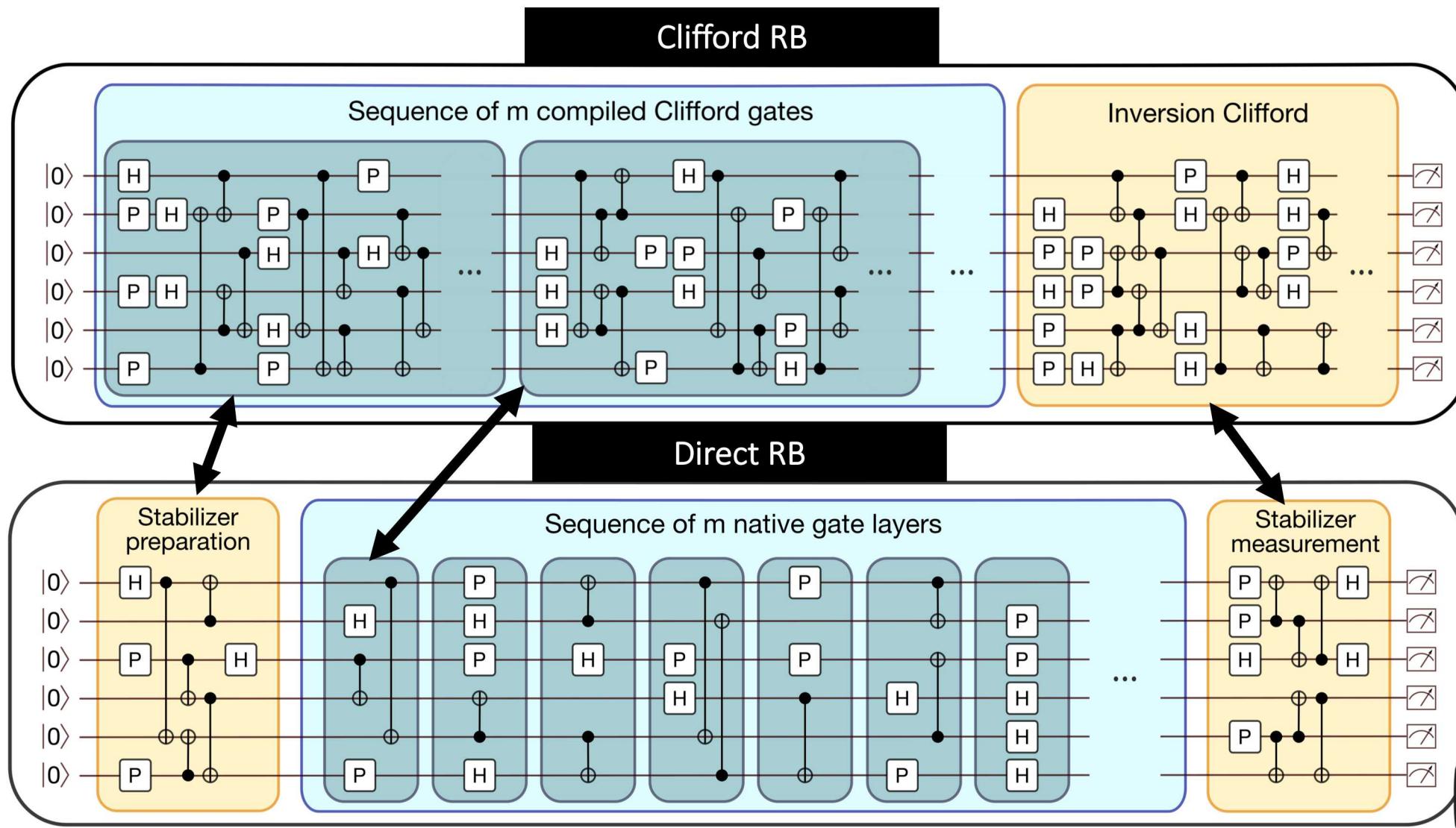
Prepare a uniformly random n -qubit stabilizer state. *Purpose: port the robustness of Clifford RB to the new method.*

Return the ideal output to a computational basis state. *Purposes: as in all RB, make errors detectable.*



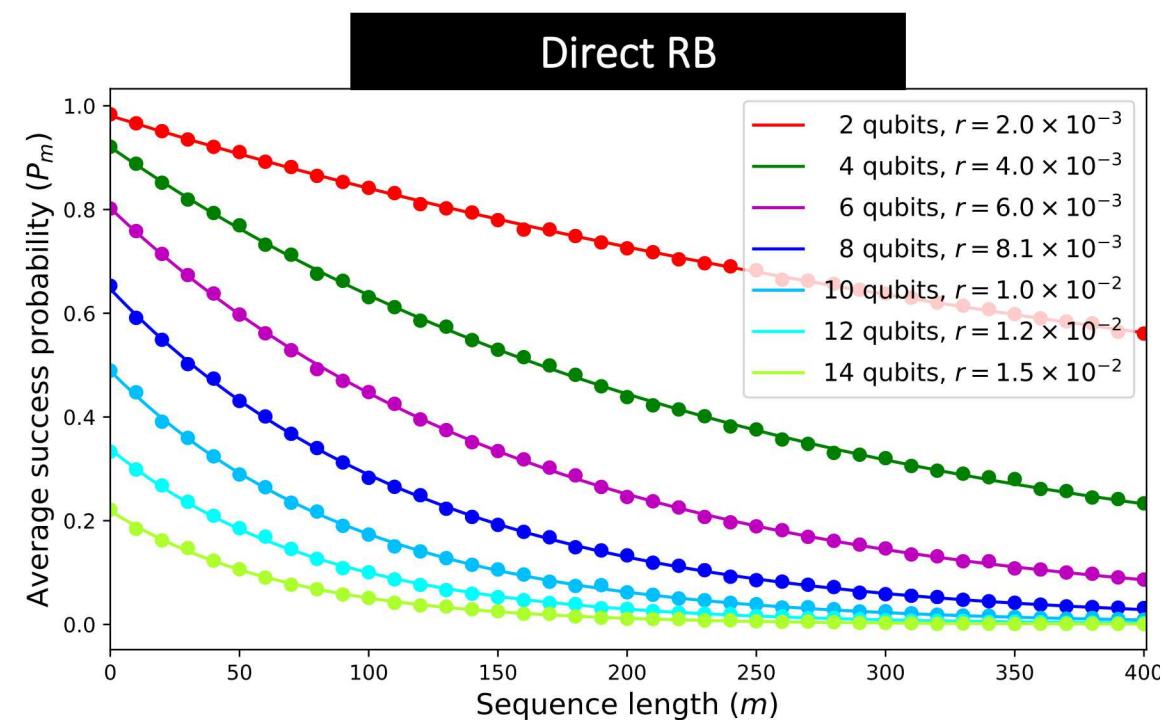
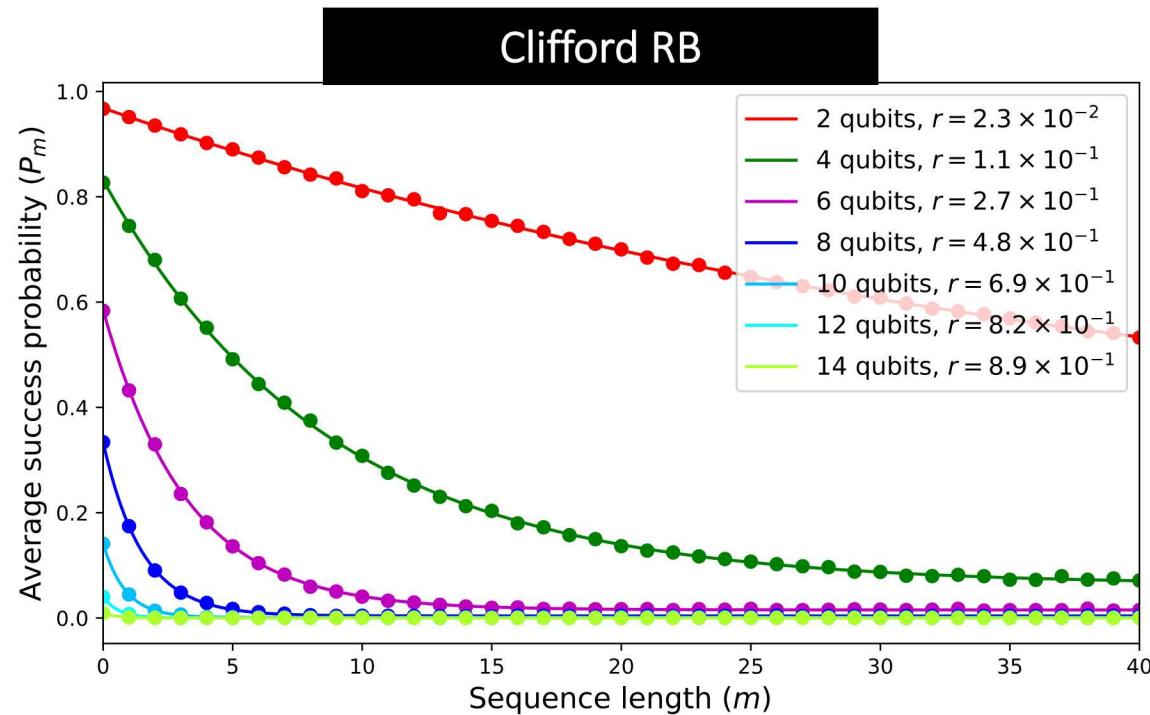
Depth-1 circuit layers sampled according to a user-specified and flexible distribution. *Purpose: streamlined layers mean (1) better scalability and (2) estimation of a per-native-gate error rate.*

Direct RB versus Clifford RB



Direct RB versus Clifford RB

An illustrative simulation, with 0.1% i.i.d depolarizing errors on each qubit

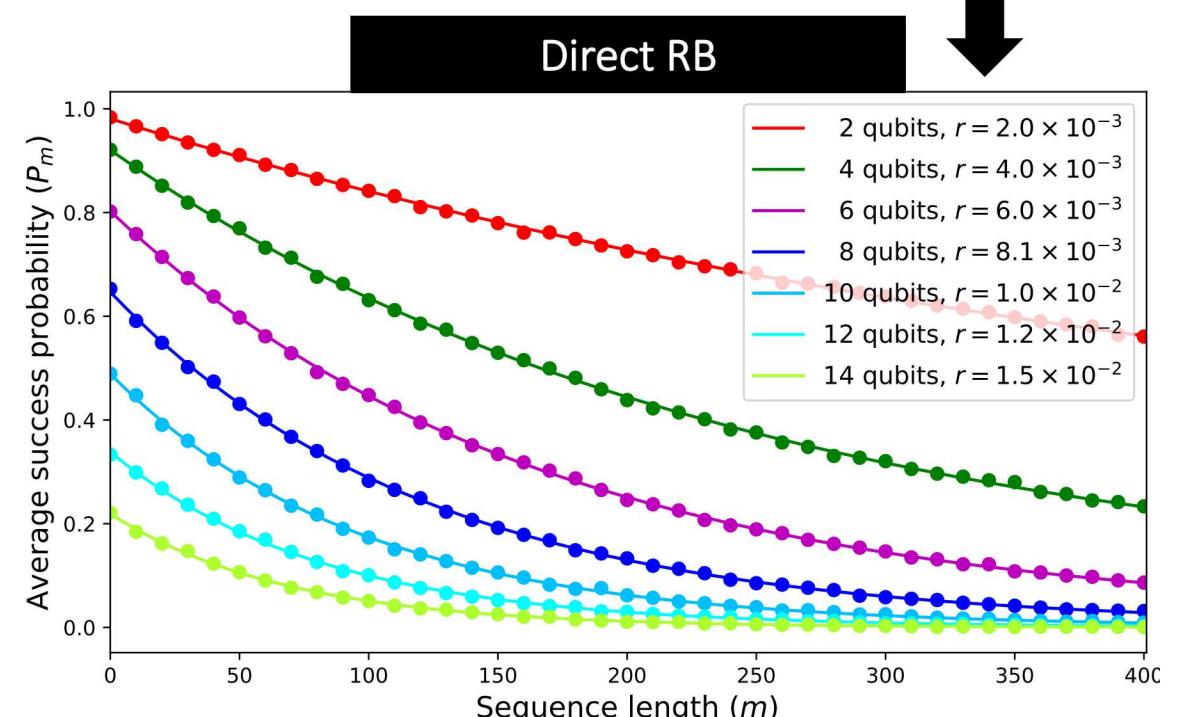
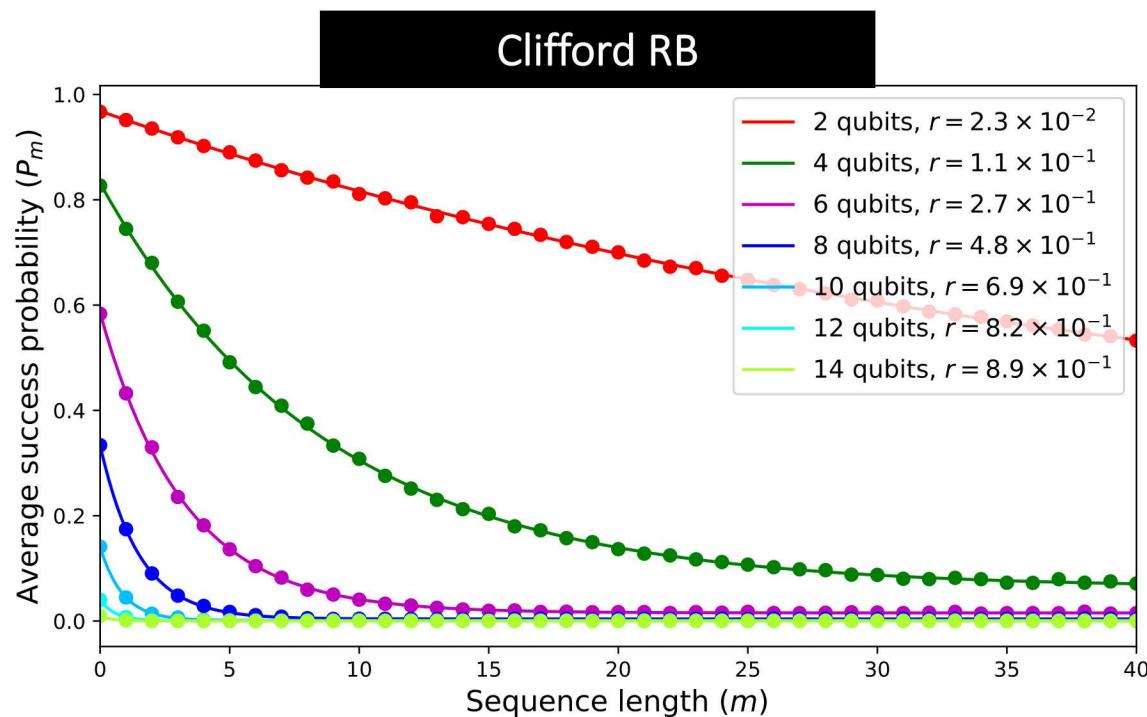


Simulation demonstration better scaling with n for direct RB!

Direct RB versus Clifford RB

Bonus: Direct RB error rates are directly linked to the gate error rate! $r \approx n \times 0.001$.

An illustrative simulation, with 0.1% i.i.d depolarizing errors on each qubit



Simulation demonstration better scaling with n for direct RB!

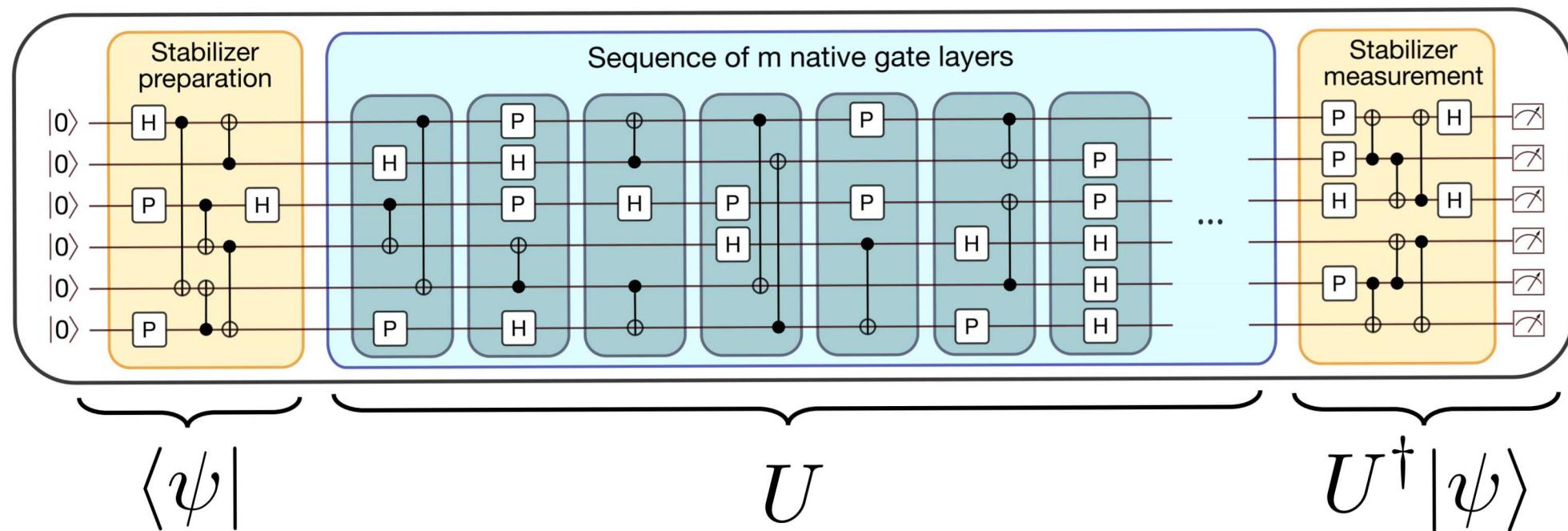
Direct RB is robust

- Clifford RB is known to be reliable, from theory¹ and experience.
- To justify replacing Clifford RB with direct RB, it is important to know that direct RB is equally reliable.
- We've developed a theory that proves that direct RB has equivalent robustness to Clifford RB (similar to the best Clifford RB theories¹).
- Here we'll only outline the fundamental ideas, sticking to *Pauli stochastic* errors.

¹Proctor *et al.*, PRL 119, 130502 (2017), Wallman, Quantum 2, 47 (2018), Merkel *et al.*, arXiv:1804.05951 (2018).

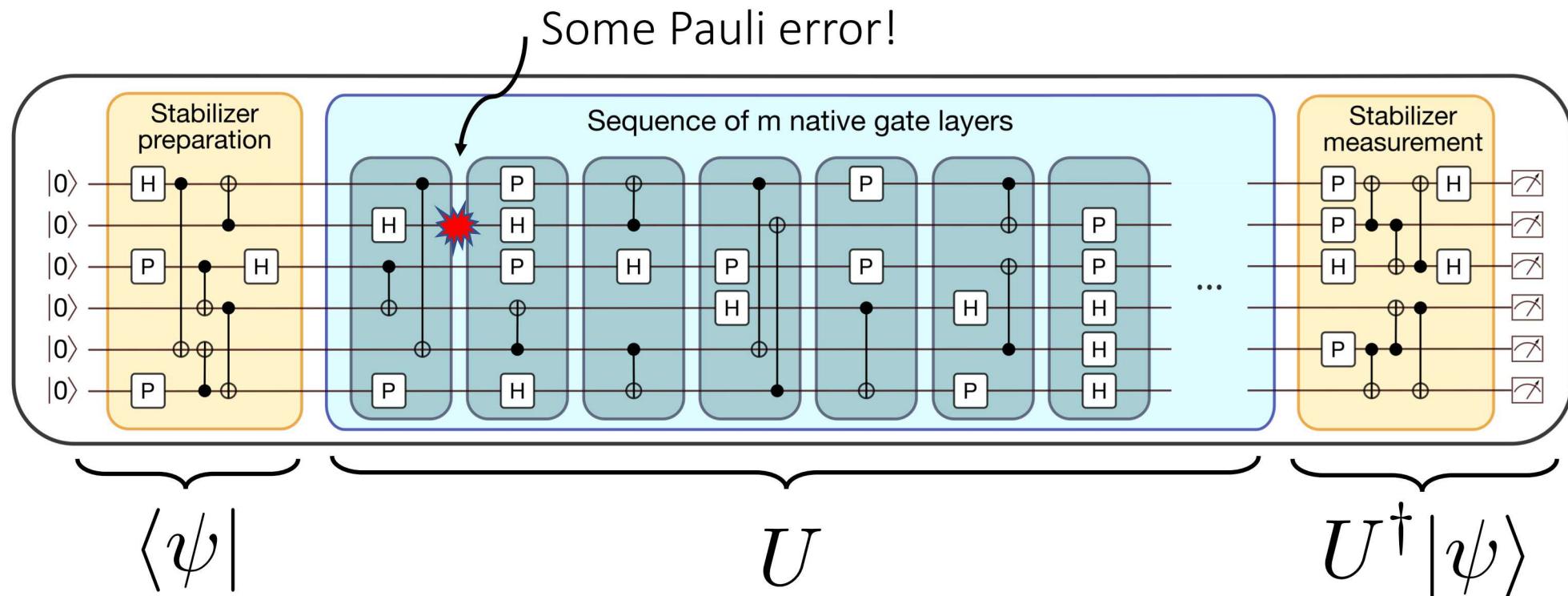
Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).



Direct RB is robust

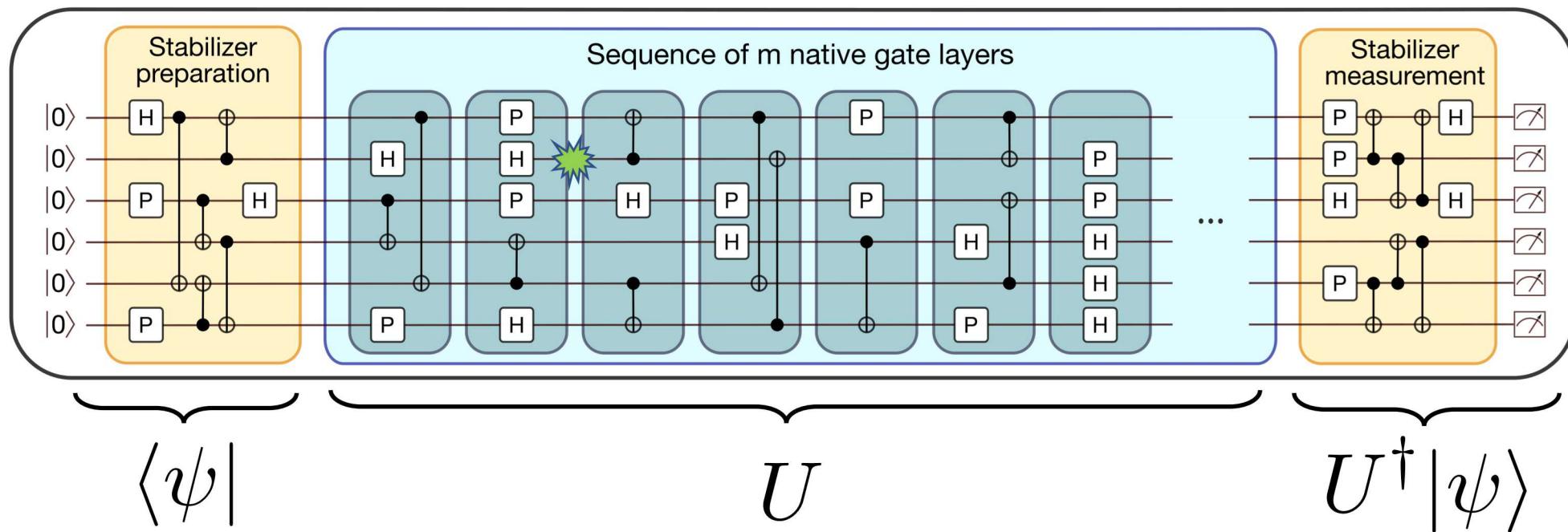
- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

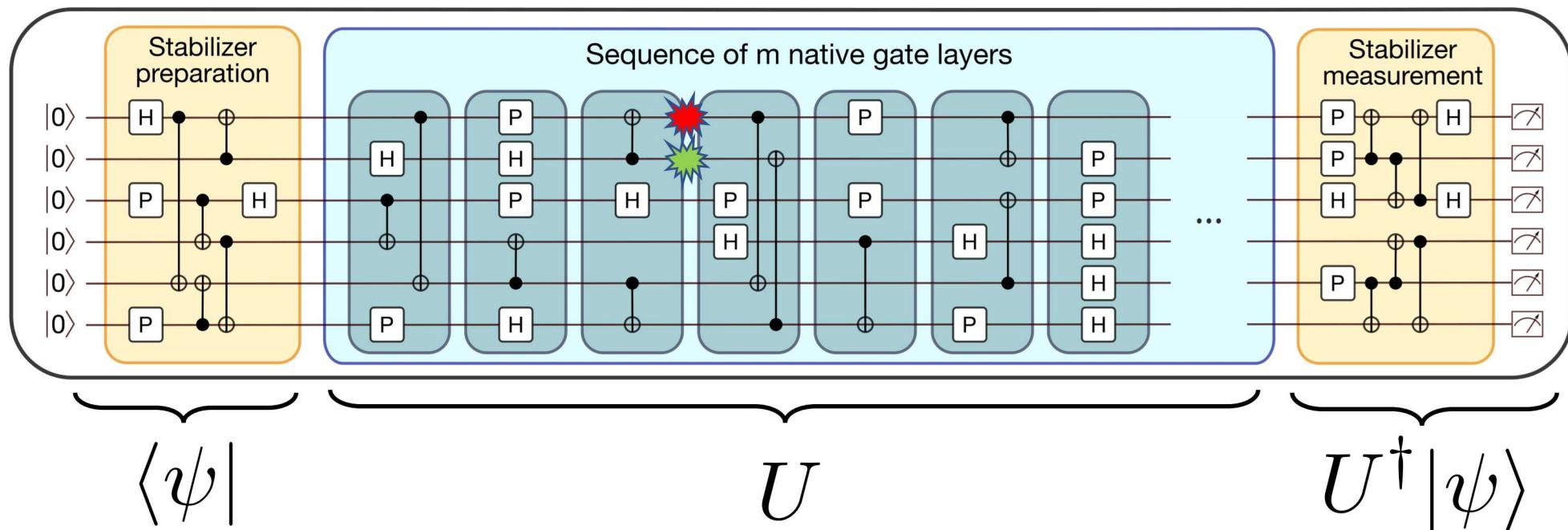
This error almost certainly spreads as it propagates through the random circuit layers.



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

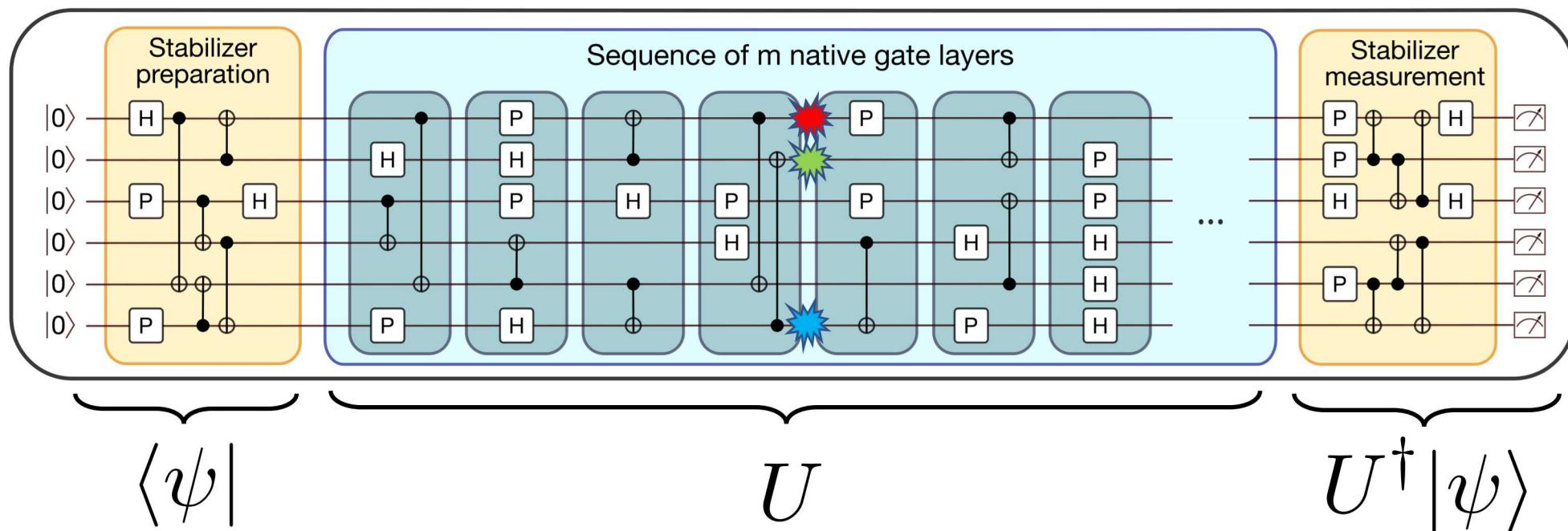
This error almost certainly spreads as it propagates through the random circuit layers.



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

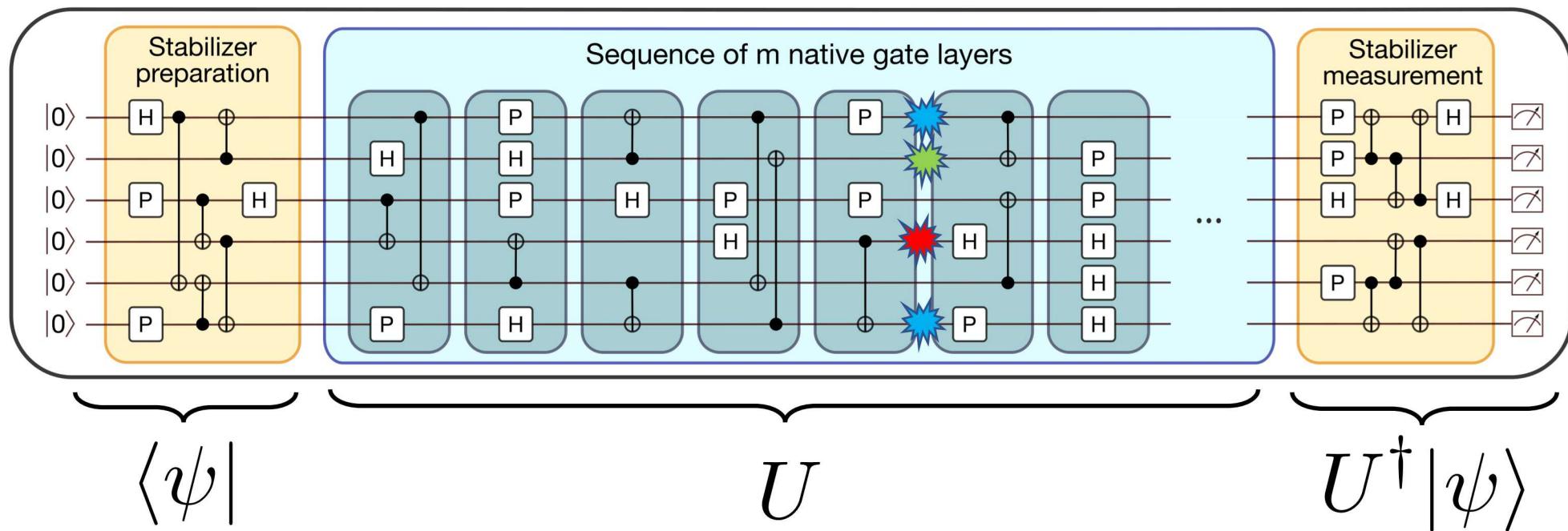
This error almost certainly spreads as it propagates through the random circuit layers.



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

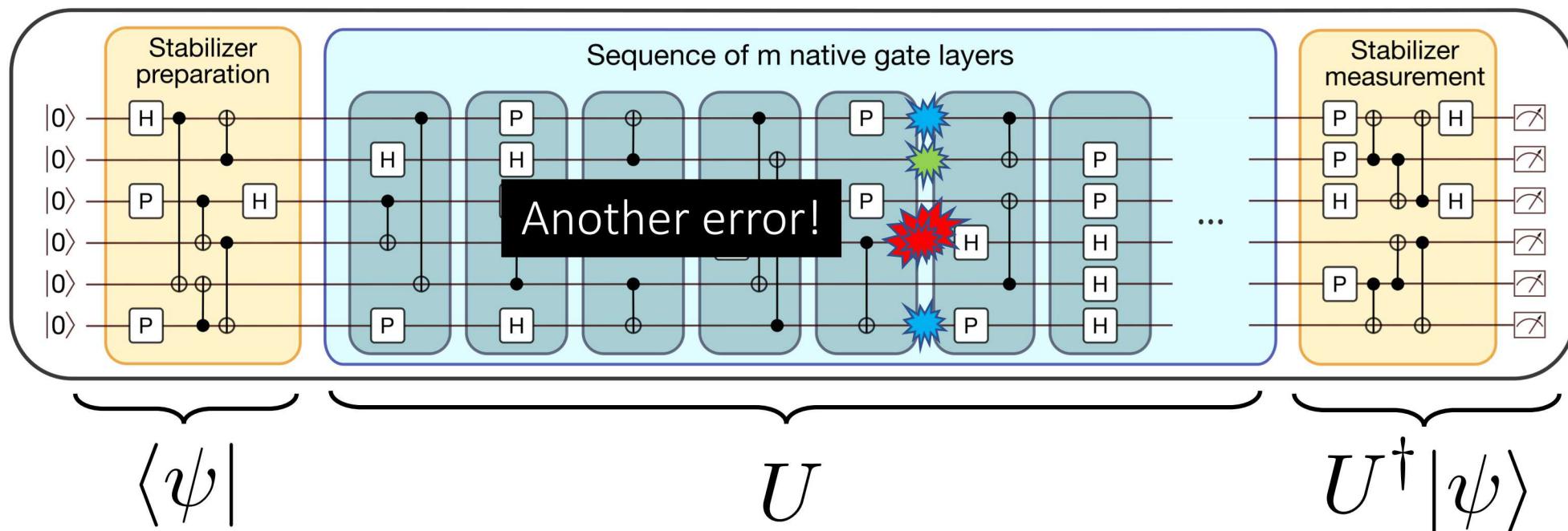
This error almost certainly spreads as it propagates through the random circuit layers.



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

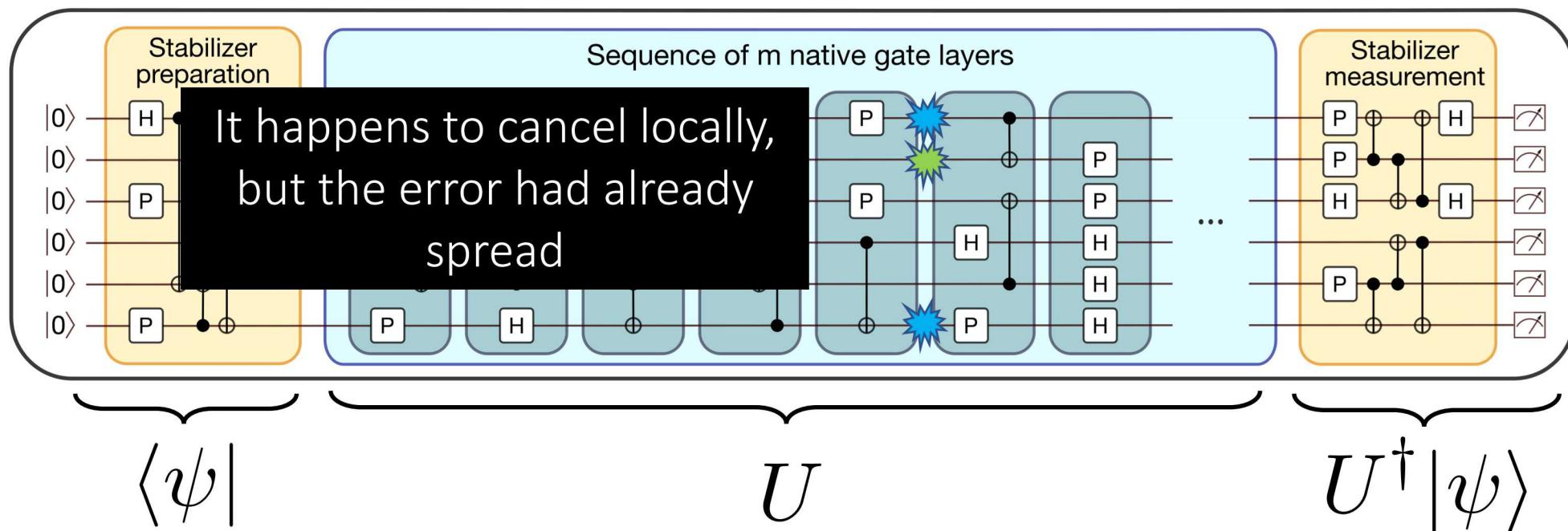
And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

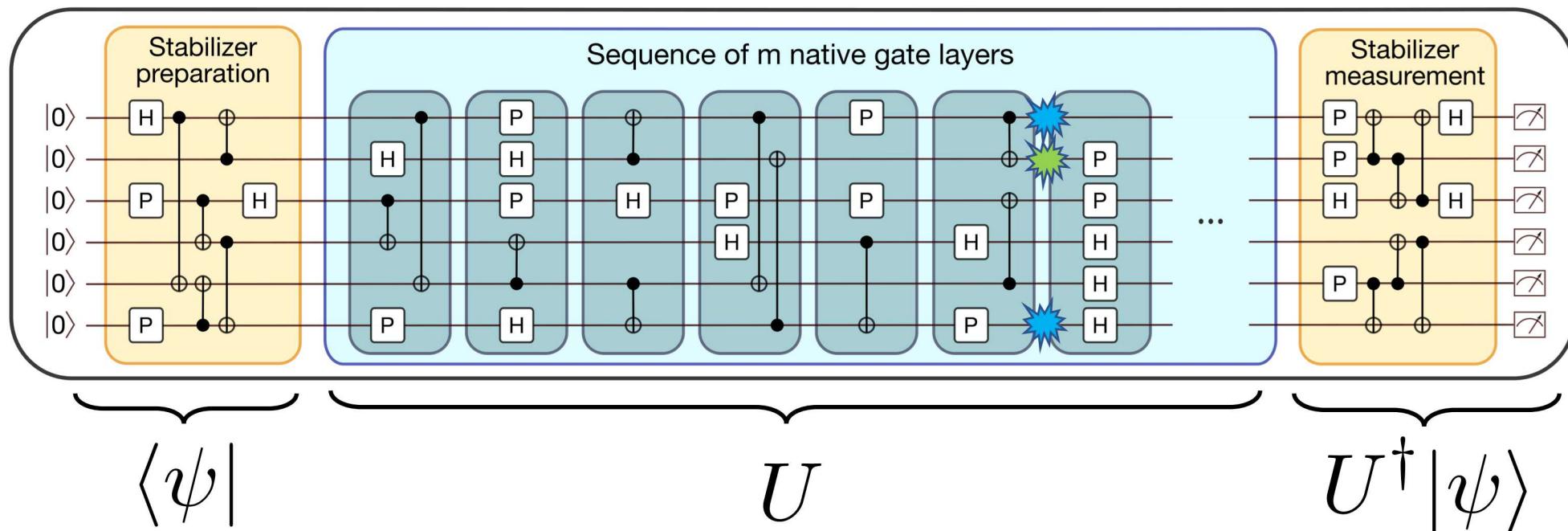
And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

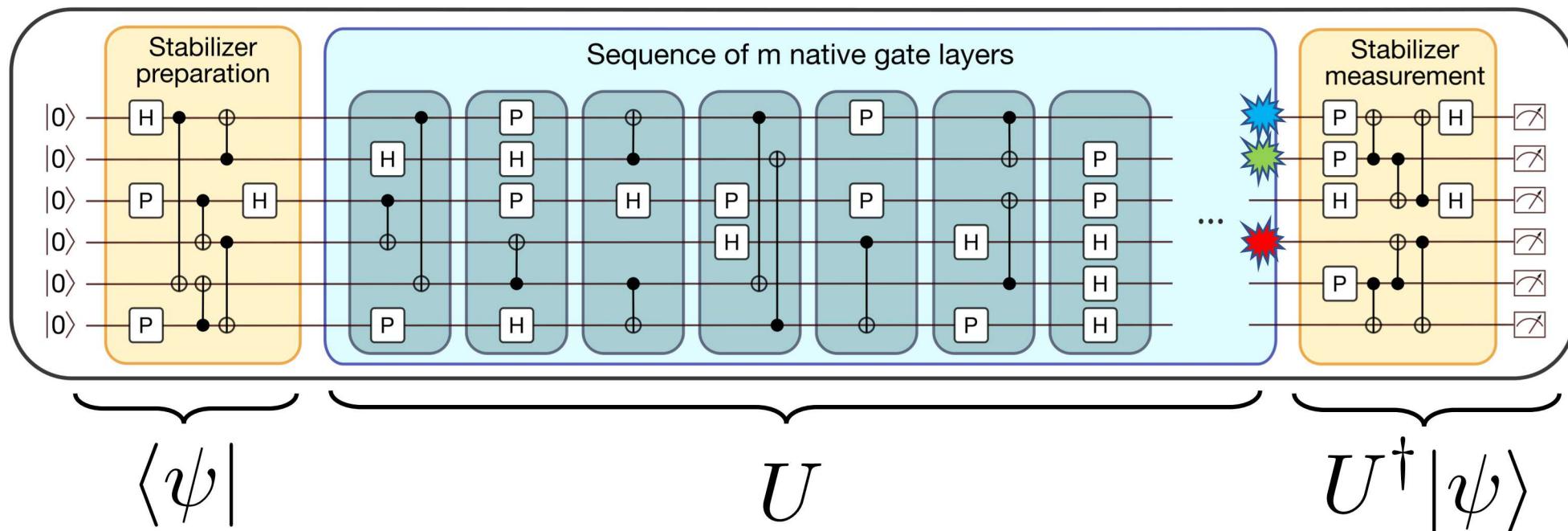
And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

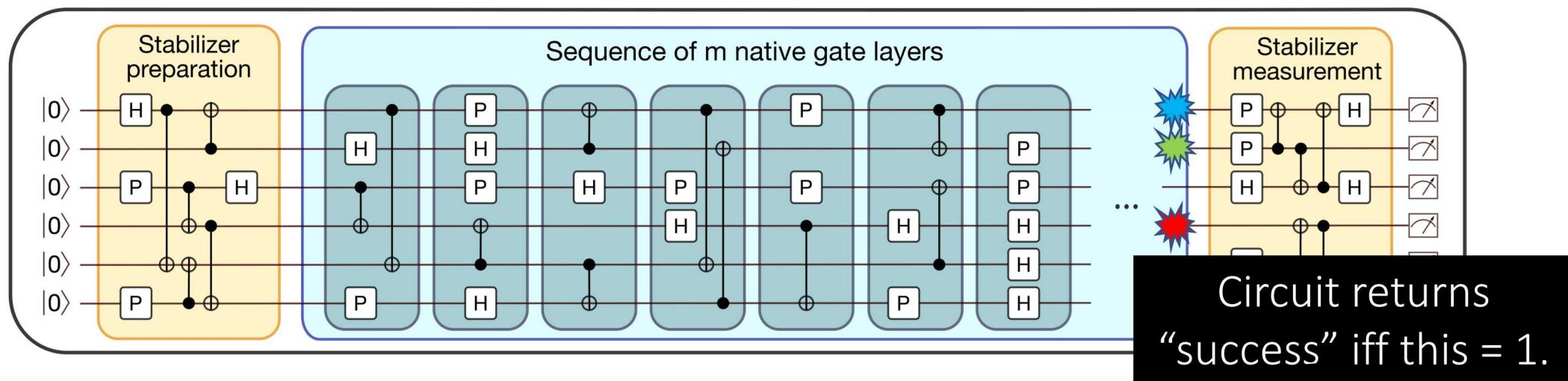
And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

And if another error happens it will almost certainly not entirely cancel the previous error

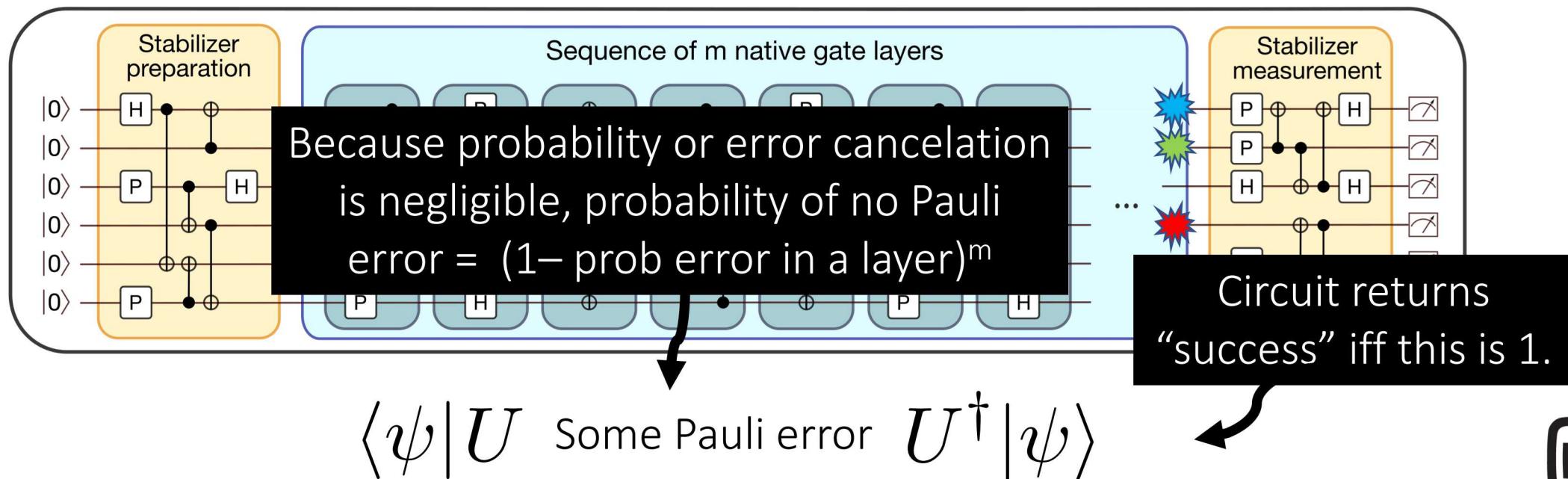


$$\langle \psi | U \text{ Some Pauli error } U^\dagger | \psi \rangle$$

Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

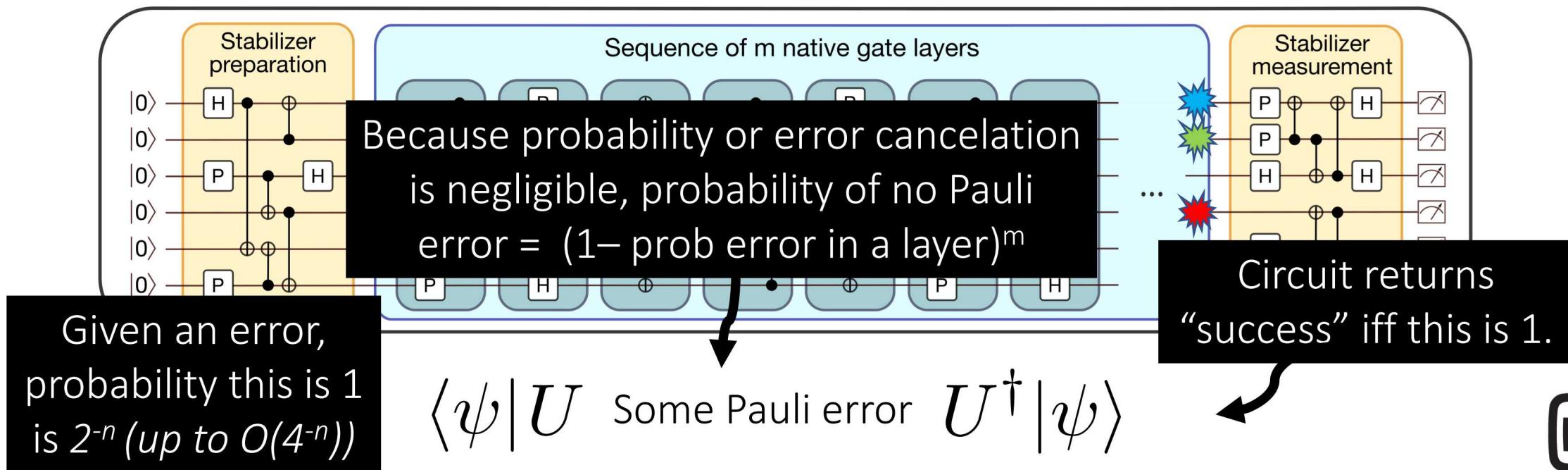
And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

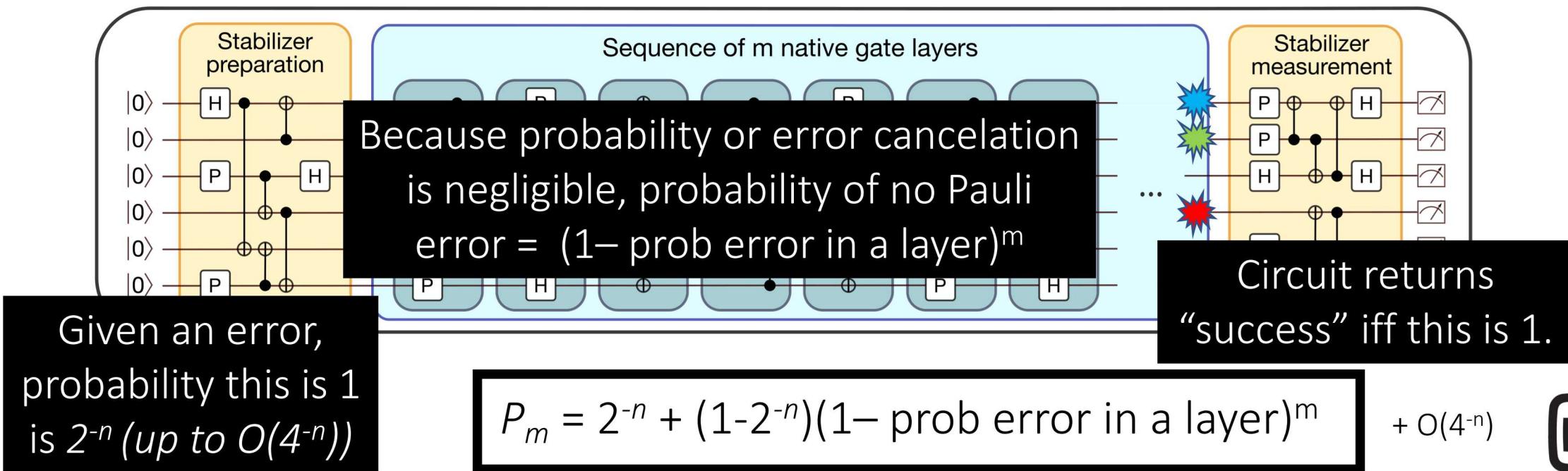
And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

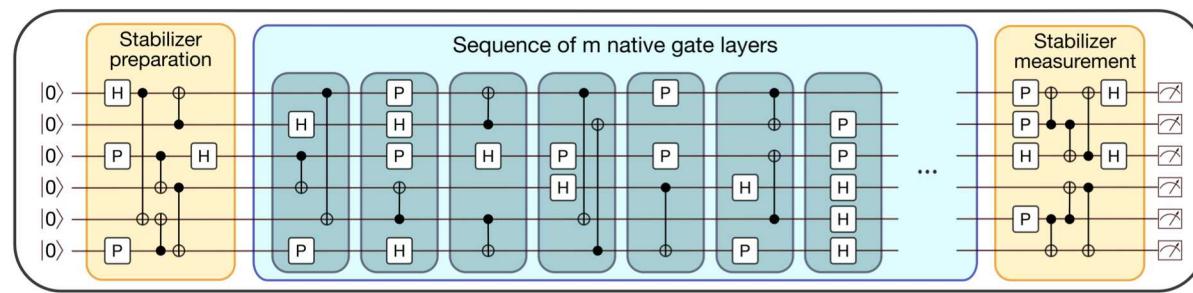
- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- Let’s look at what happens when there is an error in a direct RB circuit.

And if another error happens it will almost certainly not entirely cancel the previous error



Direct RB is robust

- The key for reliable RB is for the probability of circuit “success” to depend only on the rate of errors, and not their exact form (e.g., X versus Z errors).
- We’ve shown that direct RB measures the weighted-average error rate of the circuit layers, whenever our reasoning holds.
- The argument relies on quick error “scrambling” in the direct RB circuits, and almost all random circuits quickly scramble errors.¹



¹Sekino and Susskind, J. High Energy Phys. 10, 065 (2008), and many many other papers.

2 – 5 qubit direct RB on IBM Q Experience

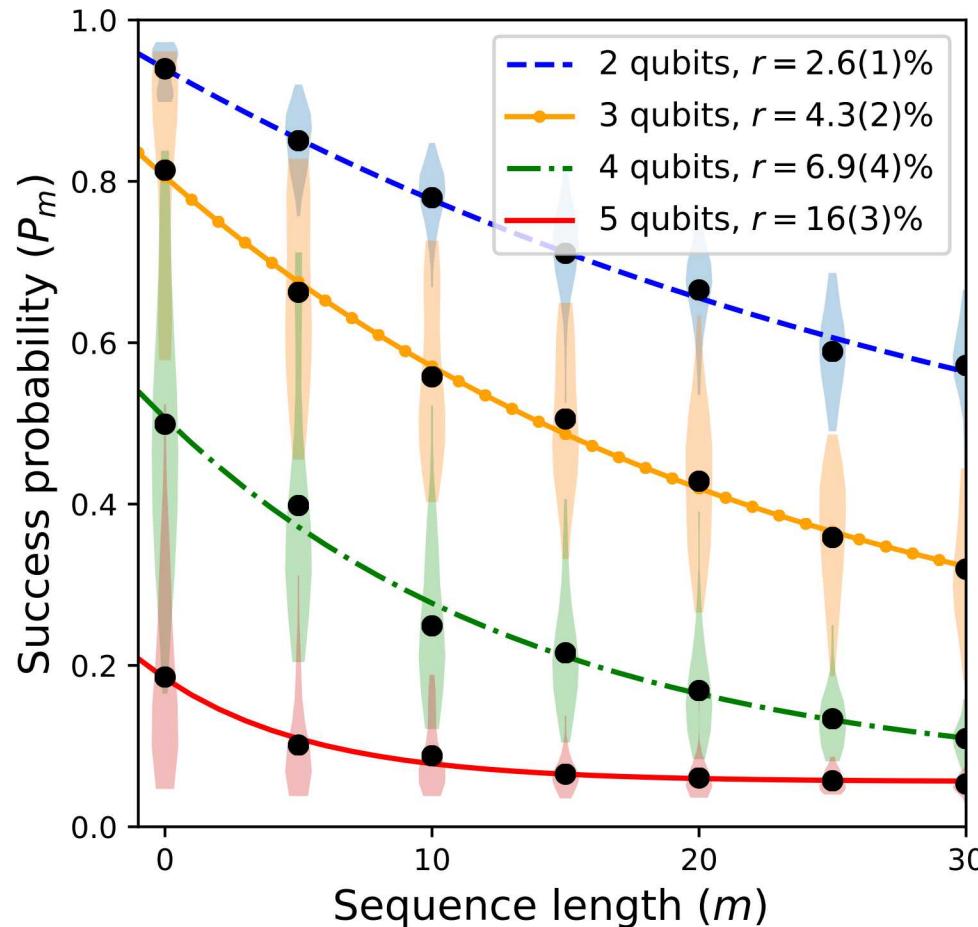
The aim is to:

- (1) Demonstrate that direct RB works.
- (2) Illustrate how direct RB it can be used to provide interesting insight into a device.

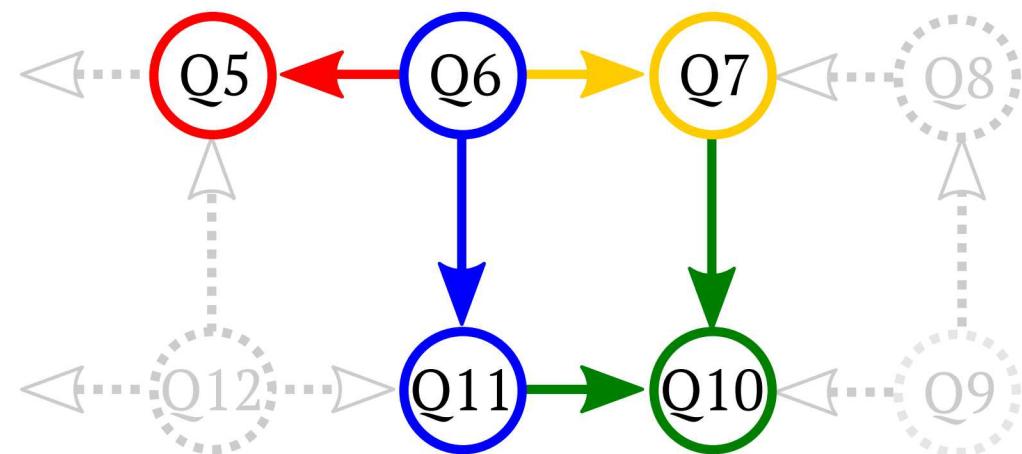
This is not intended as a comprehensive benchmarking of IBMQX.

2 – 5 qubit direct RB on IBM Q Experience

Successful 2 – 5 qubit direct RB!



Benchmarked portion of ibmqx5

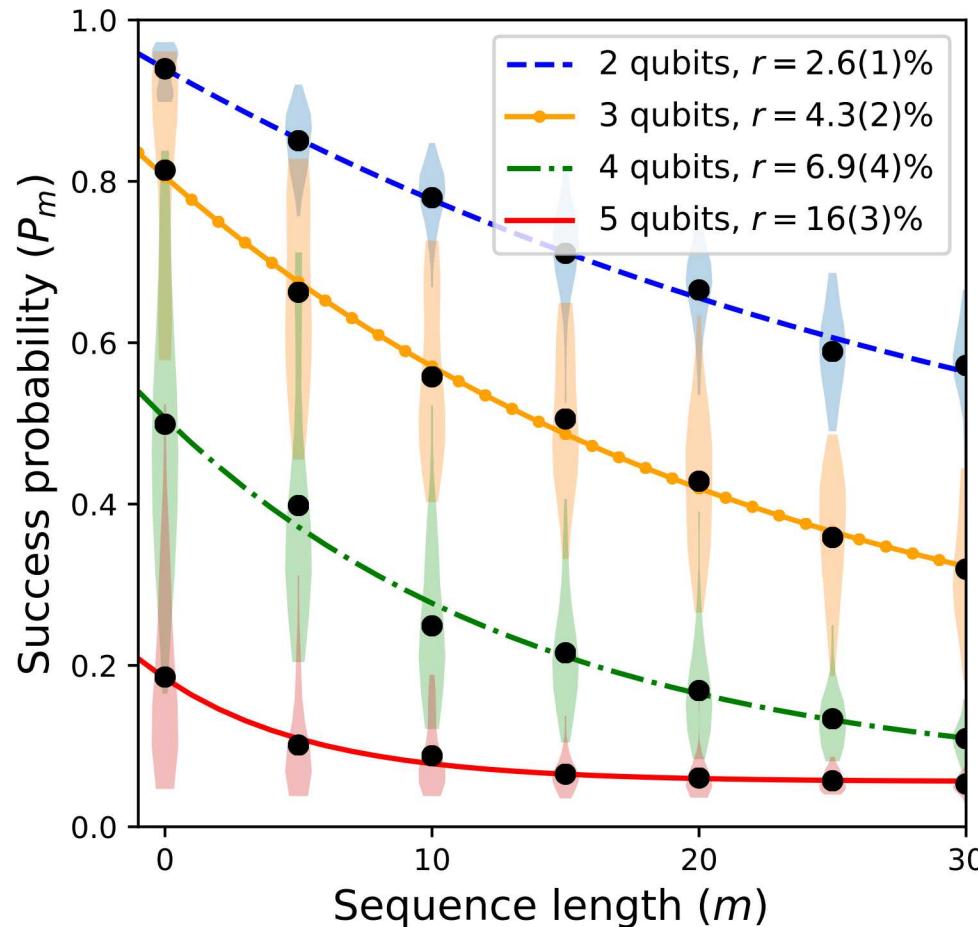


Circuit layer sampler:

- With prob. p apply 1 CNOT + local gates.
- With prob. $1 - p$ apply only local gates.
- We chose $p = 0.75$.

2 – 5 qubit direct RB on IBM Q Experience

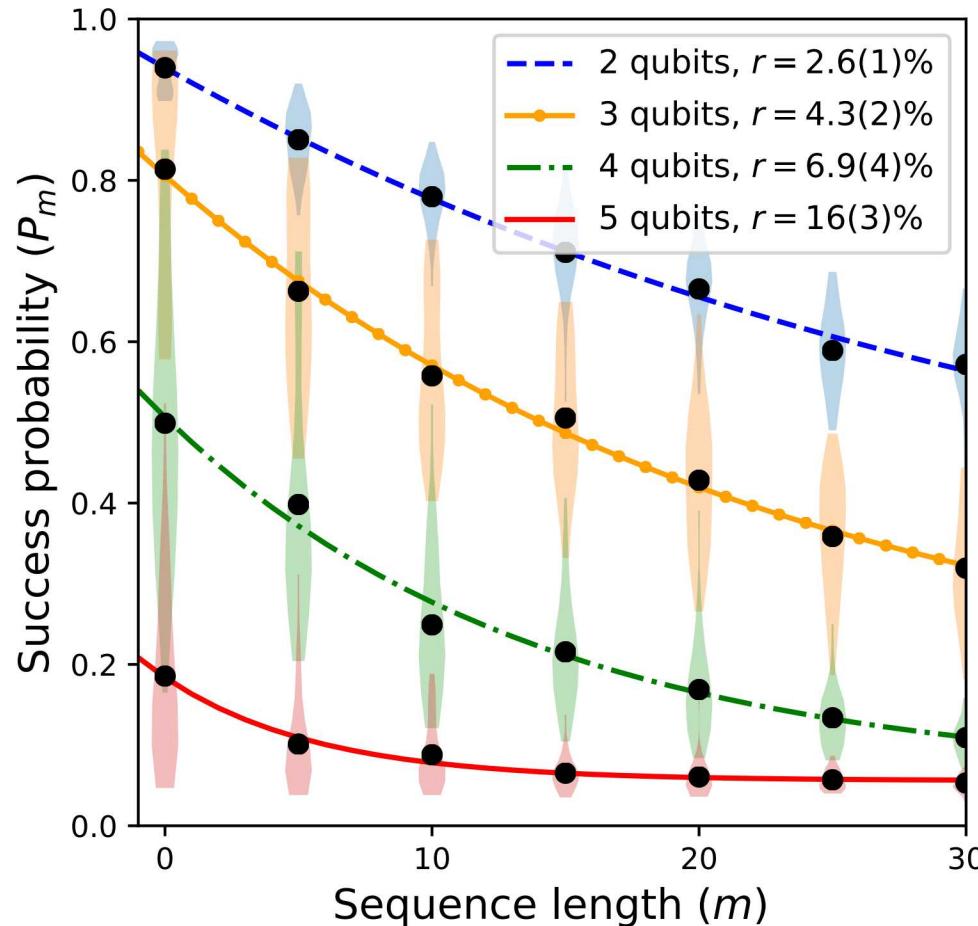
Successful 2 – 5 qubit direct RB!



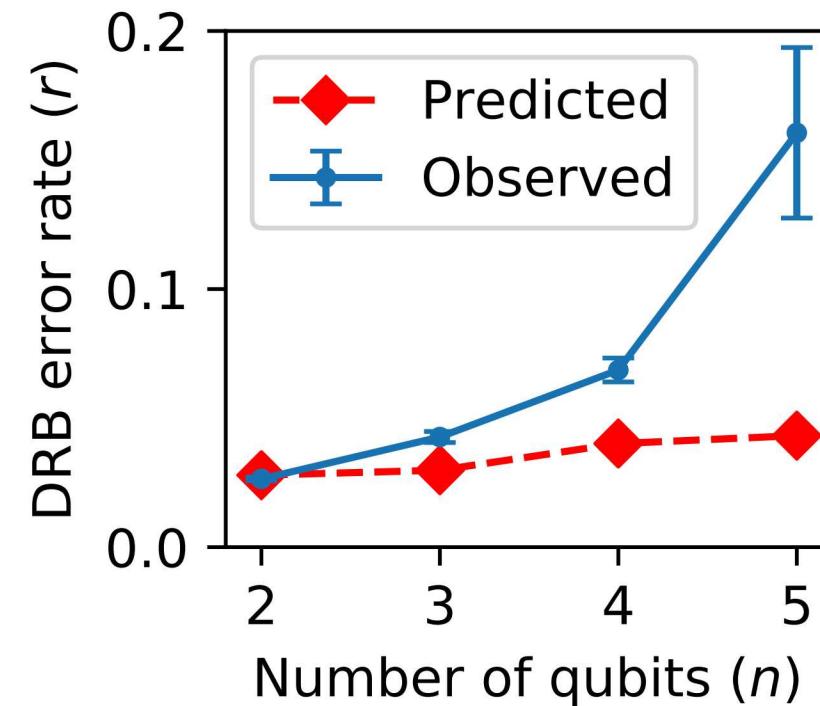
- The average number of CNOTs per layer is independent of n .
- Therefore r should grow slowly with n if:
 - CNOT errors dominate and
 - n -qubit RB is predictive of $m > n$ qubit circuits *in this device*.
- This is not what we observe, which is suggestive of crosstalk. Let's quantify this...

2 – 5 qubit direct RB on IBM Q Experience

Successful 2 – 5 qubit direct RB!



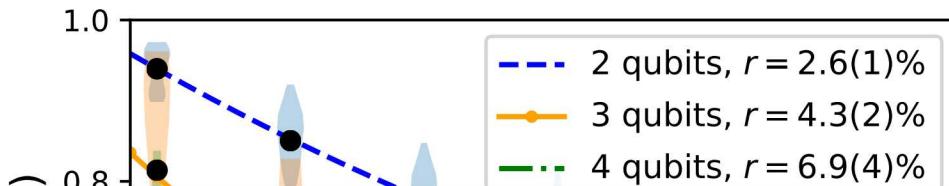
Comparing predictions from Clifford RB calibration data and direct RB results



Discrepancy shows that n -qubit RB reveals errors not predicted by 1- and 2-qubit RB (i.e., crosstalk).

2 – 5 qubit direct RB on IBM Q Experience

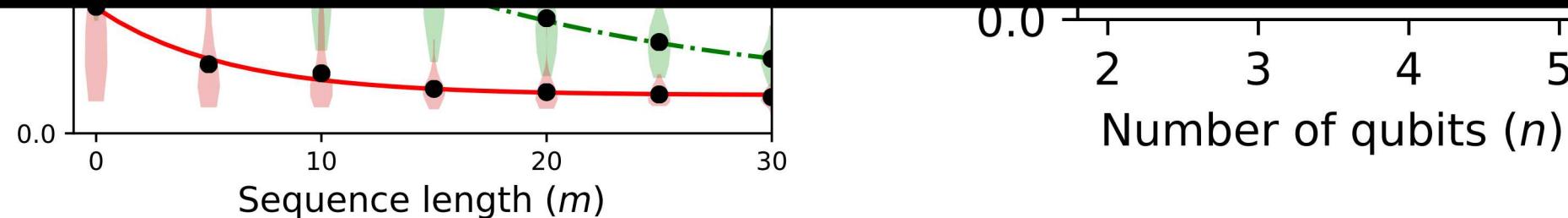
Successful 2 – 5 qubit direct RB!



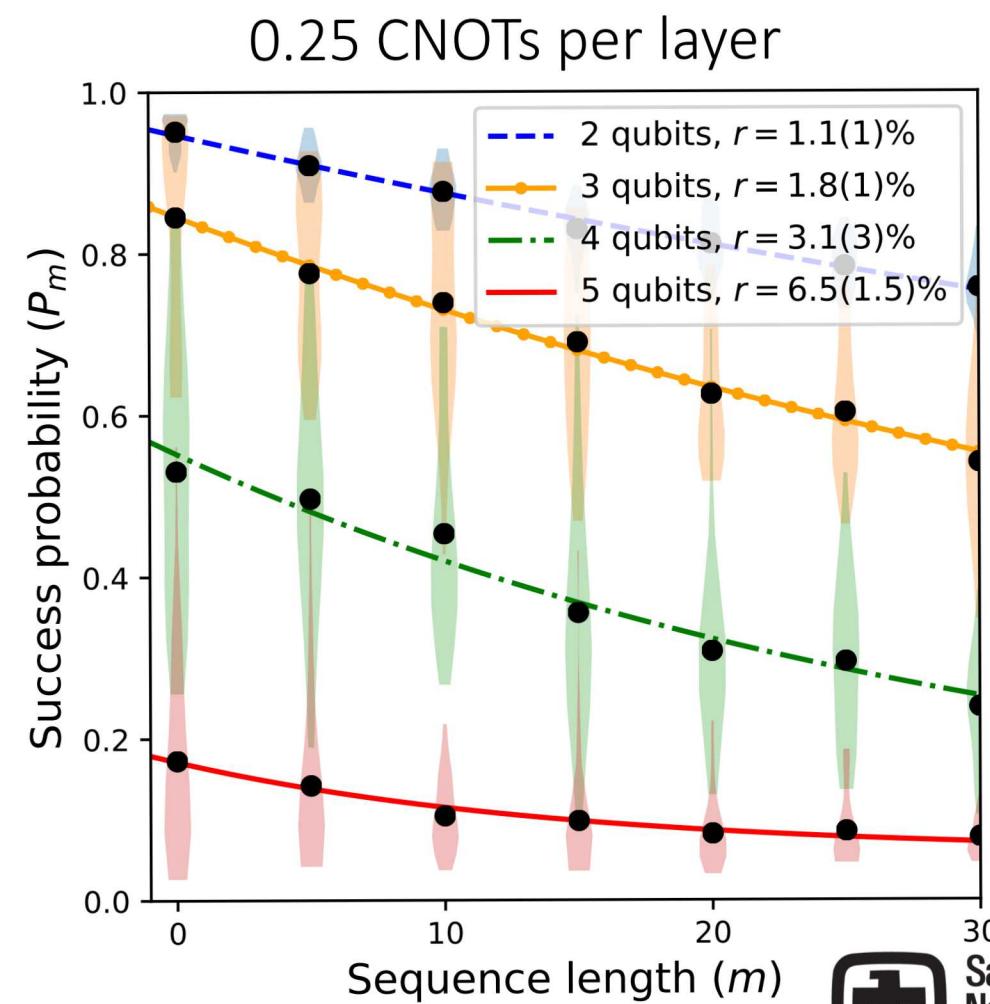
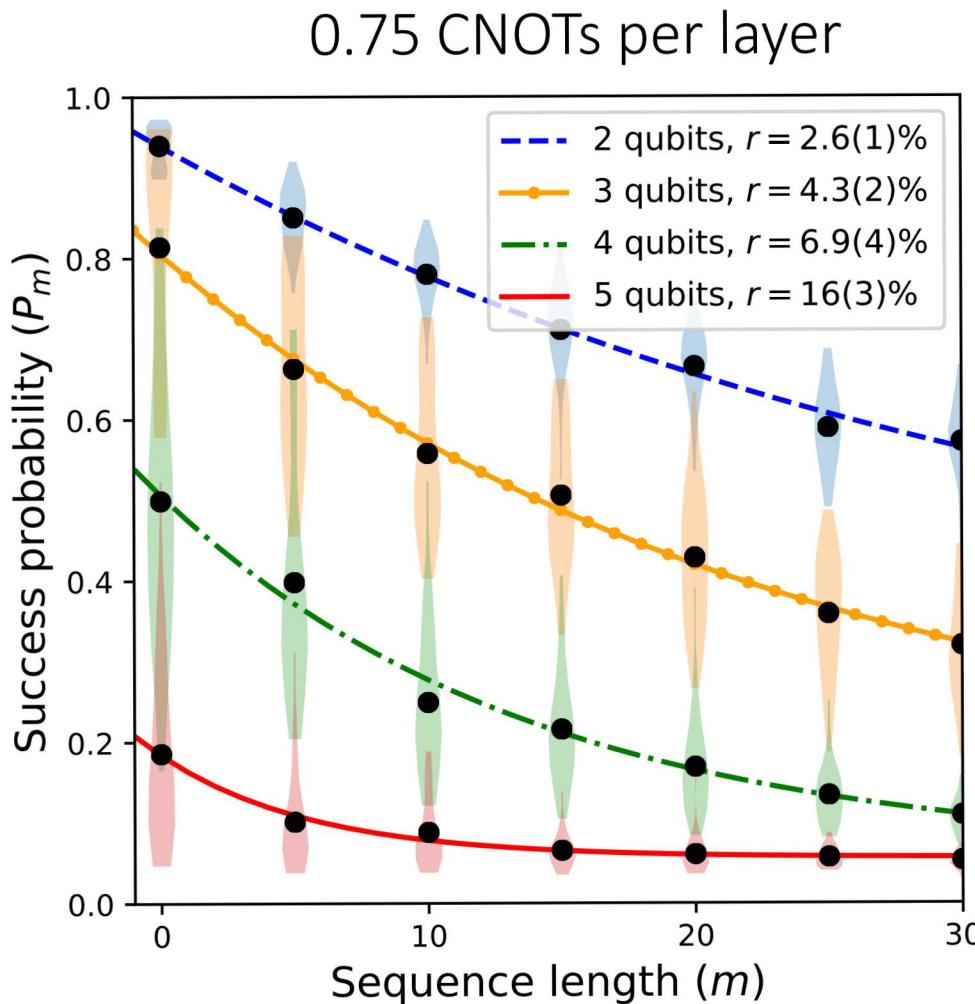
Comparing predictions from Clifford RB calibration data and direct RB results



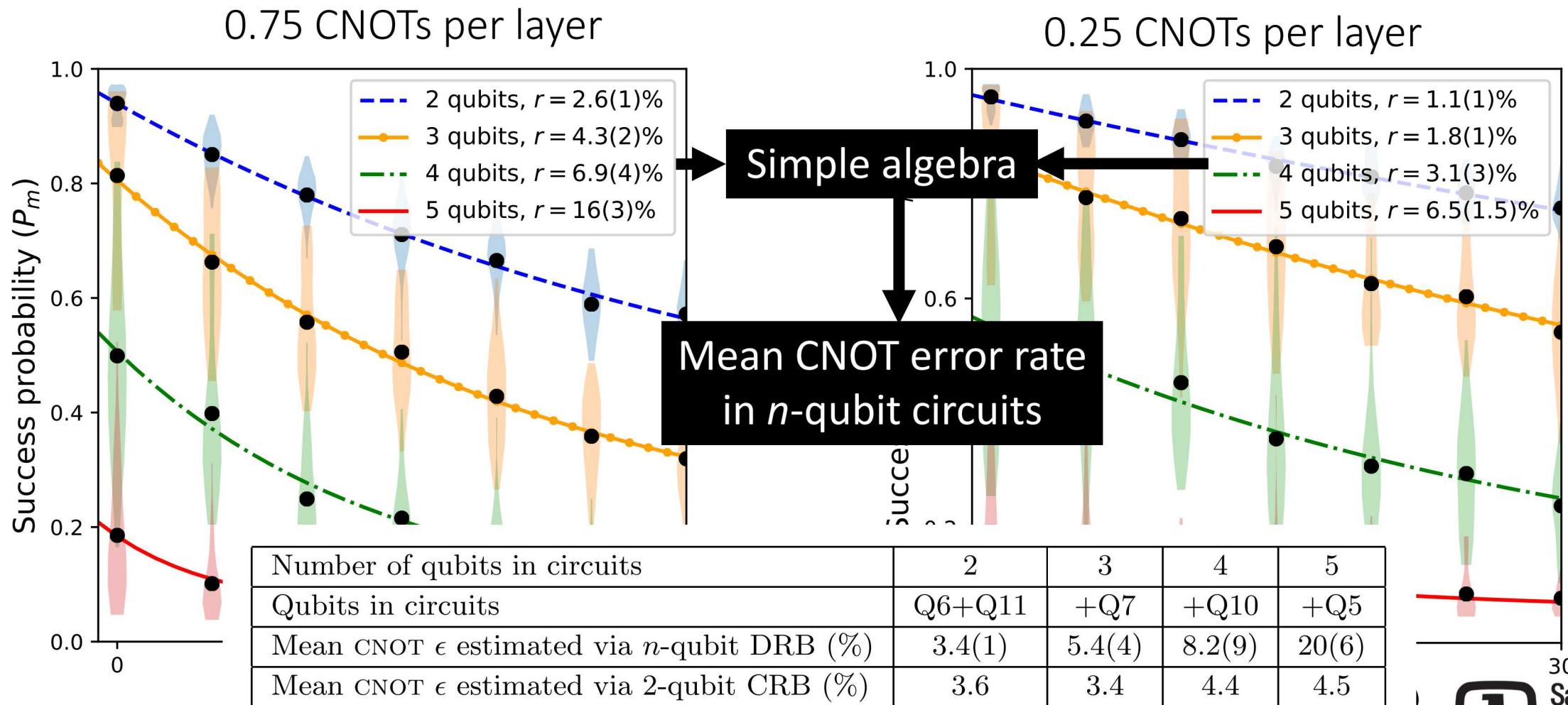
We can disentangle the contributions of 1-qubit gates and CNOTs to the total error by also implementing direct RB with different circuit layer sampling



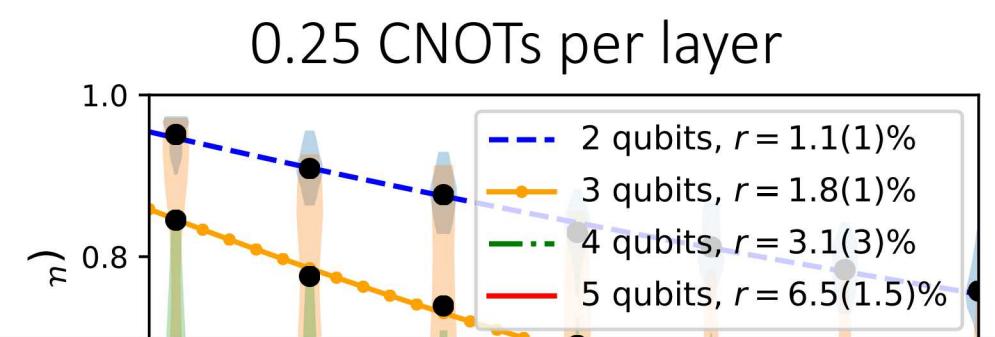
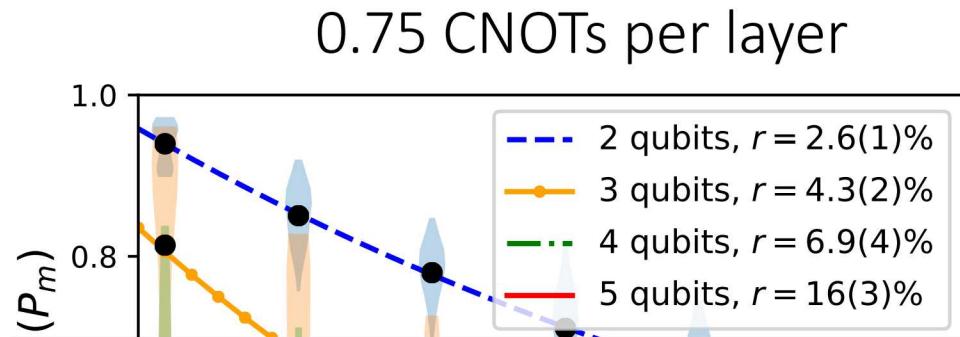
2 – 5 qubit direct RB on IBM Q Experience



2 – 5 qubit direct RB on IBM Q Experience

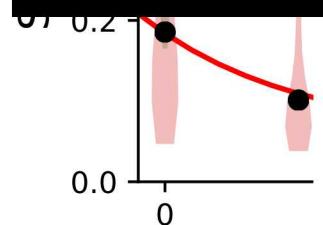


2 – 5 qubit direct RB on IBM Q Experience

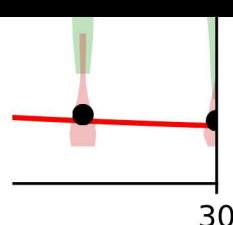


This illustrates that holistic qubit characterization is important for testing the applicability of 1- and 2-qubit characterization results.

(see also McKay *et al.*, arXiv:1712.06550).

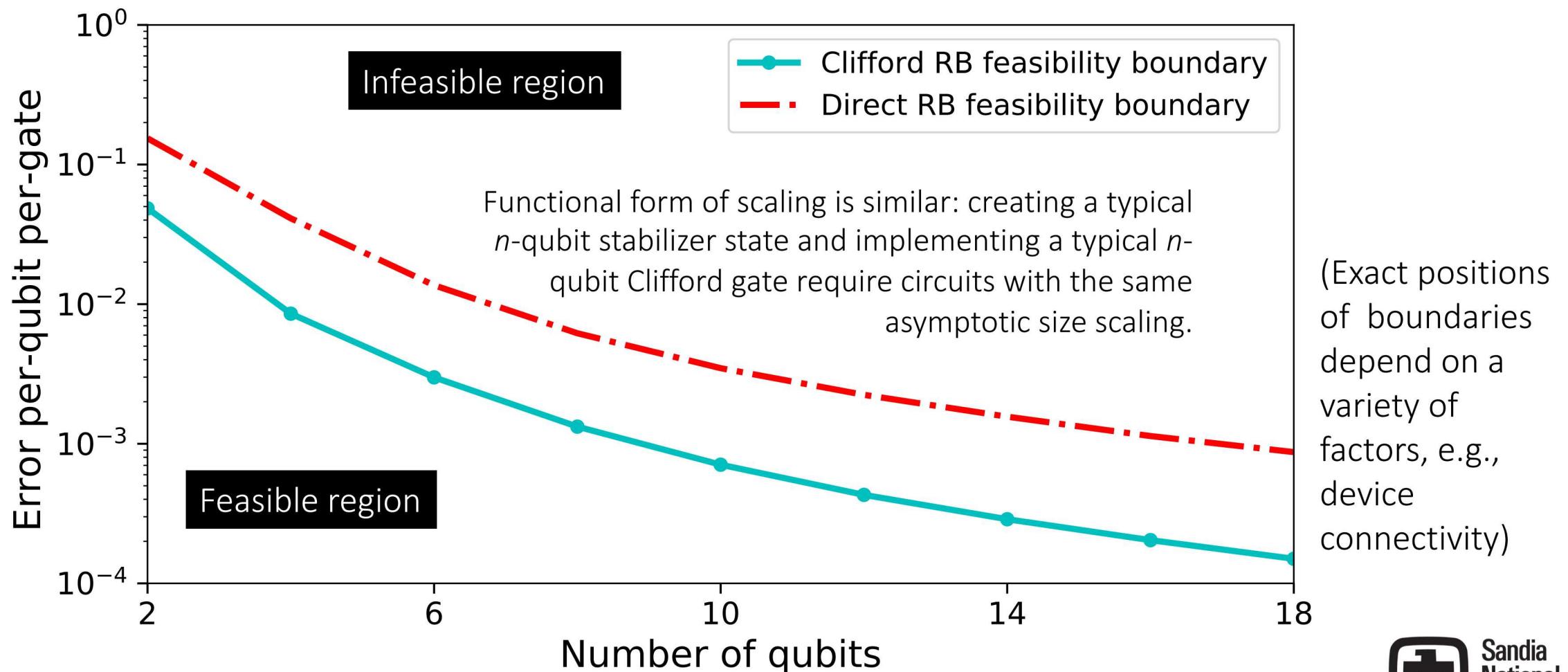


Number of qubits in circuits	2	3	4	5
Qubits in circuits	Q6+Q11	+Q7	+Q10	+Q5
Mean CNOT ϵ estimated via n -qubit DRB (%)	3.4(1)	5.4(4)	8.2(9)	20(6)
Mean CNOT ϵ estimated via 2-qubit CRB (%)	3.6	3.4	4.4	4.5

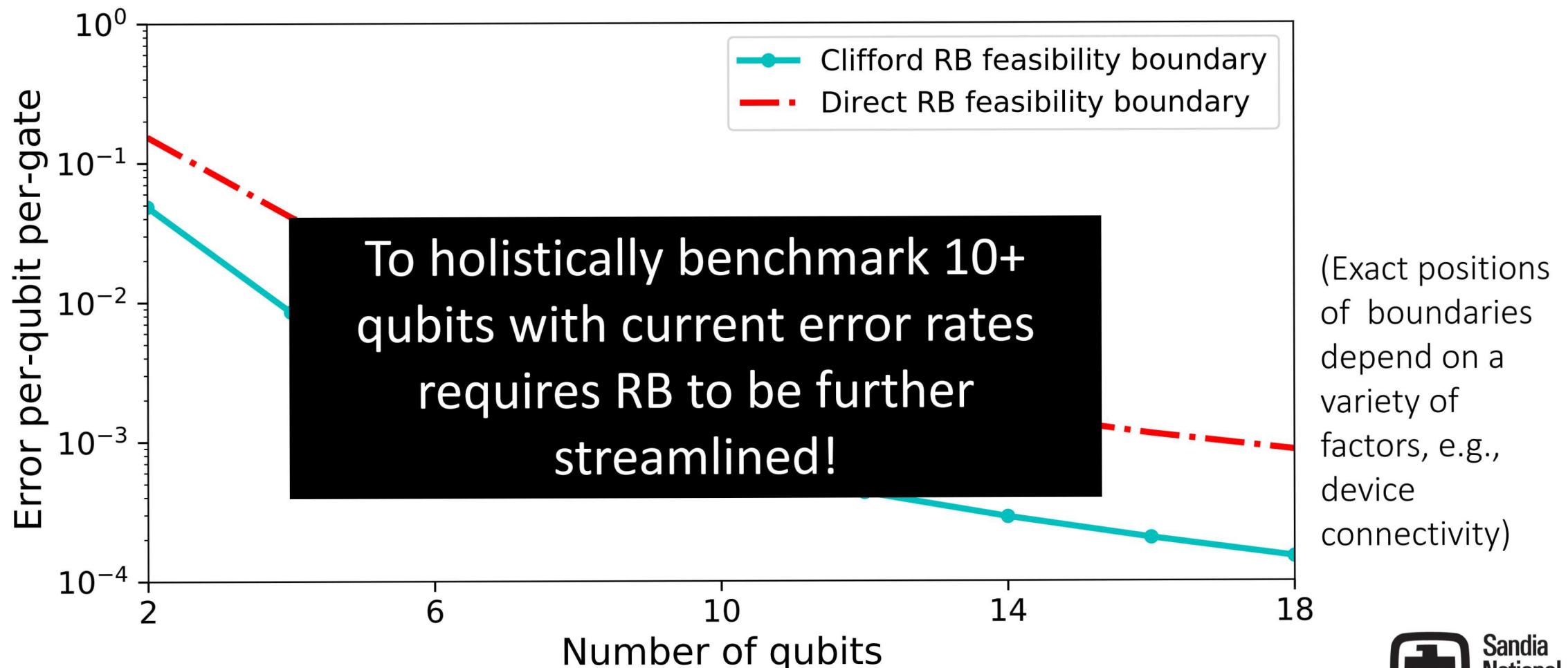


Part III – Optimally scalable RB

Gate fidelity requirements for holistic RB



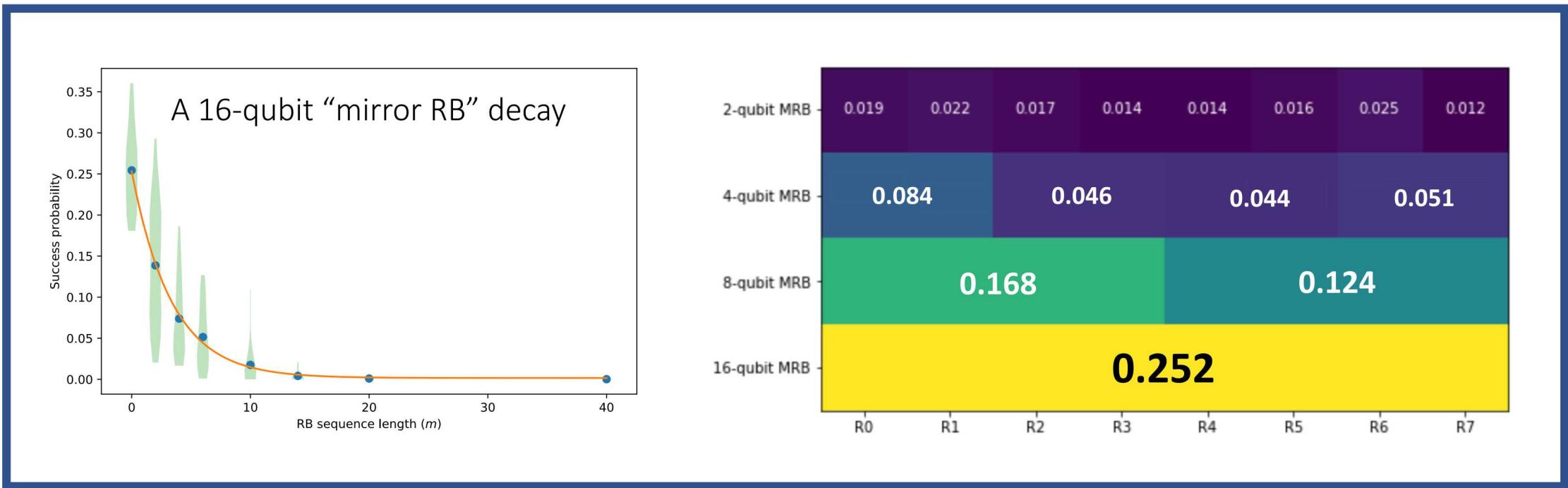
Gate fidelity requirements for holistic RB



Adapting direct RB for optimal scalability

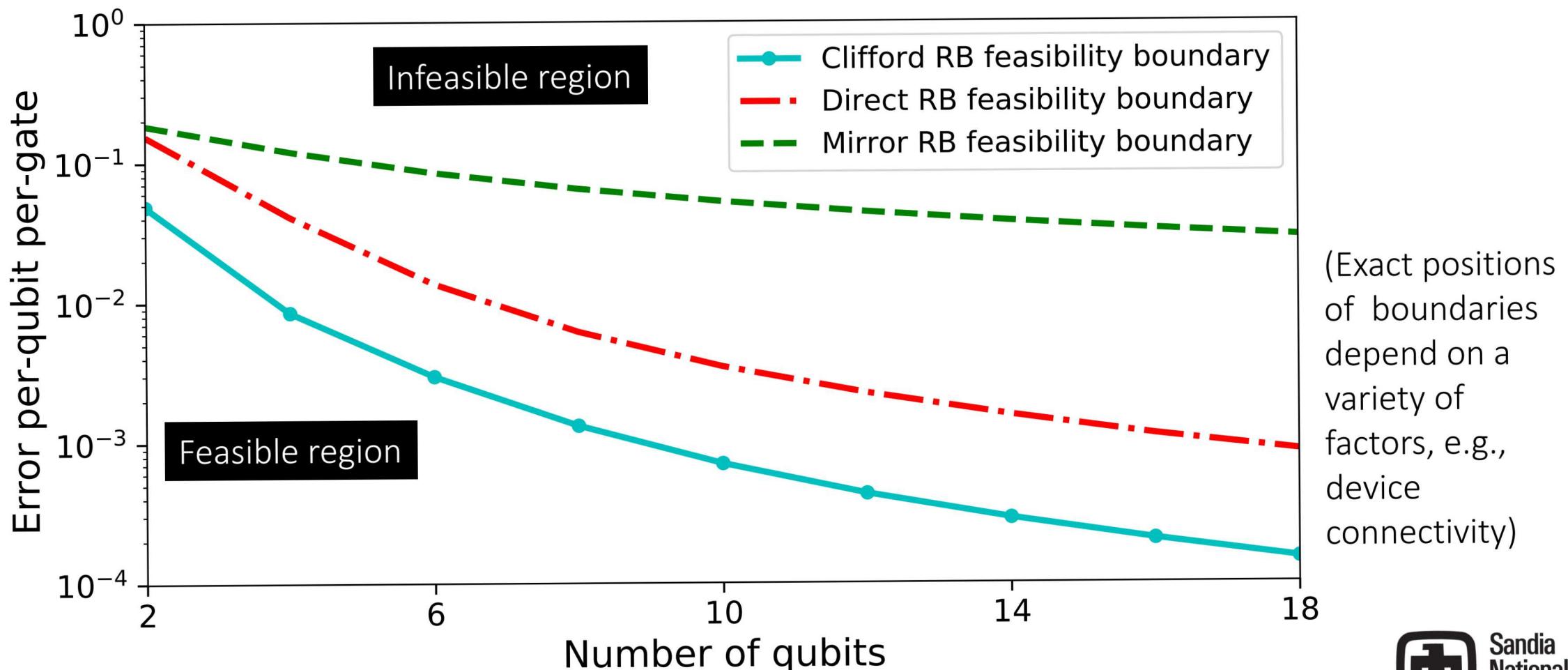
- We're adapting direct RB so that it is optimally scalable (the new method's working title is "mirror RB").
- This method unavoidably trades off some of the robustness of direct RB for increased scalability. But we expect the method to be sufficiently robust to be useful.
- We've used the method to holistically benchmark an entire 16-qubit device (IBMQX5), with promising results.

Adapting direct RB for optimal scalability

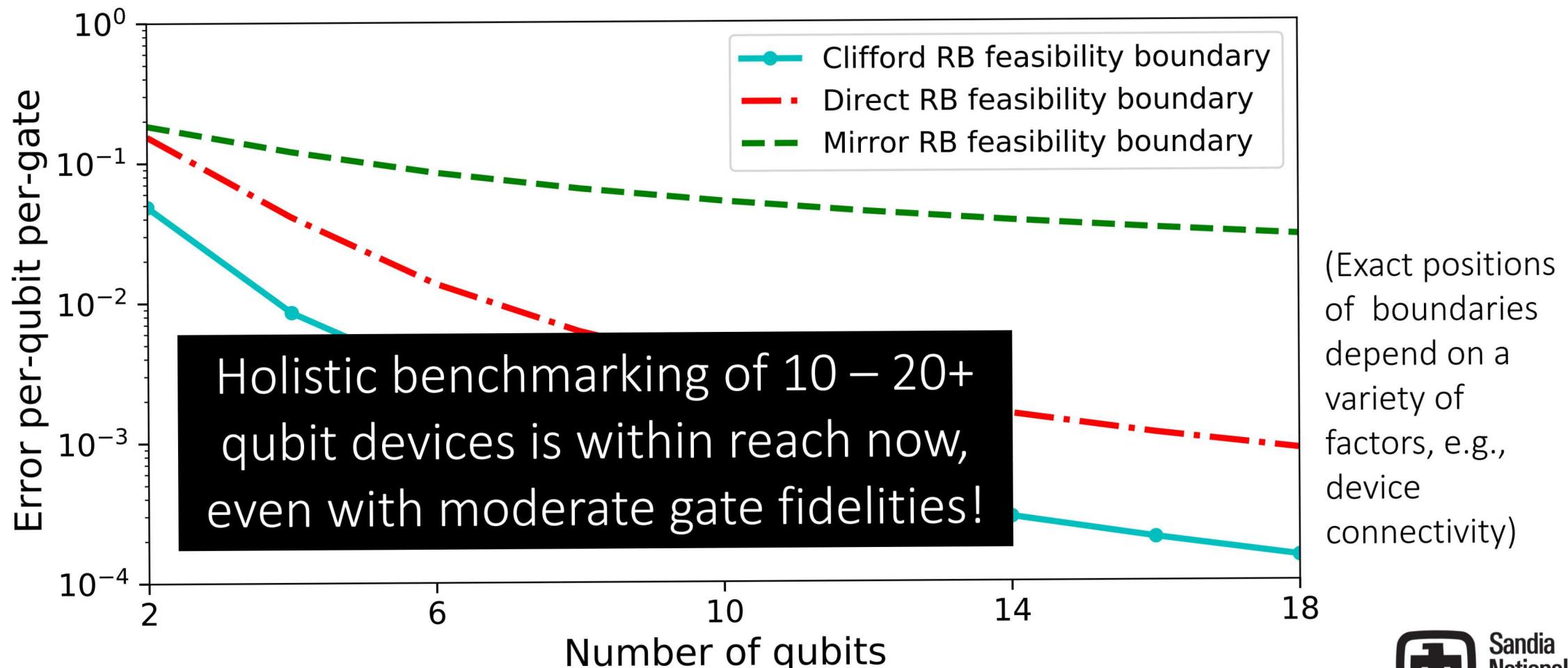


- We've used the method to holistically benchmark an entire 16-qubit device (IBMQX5), with promising results.

Gate fidelity requirements for holistic RB



Gate fidelity requirements for holistic RB



Want to run direct (or mirror) RB?

- Install “pyGSTi” using pip (see www.pygsti.info).
- Input your device’s specification (number of qubits, the gates, connectivity).

```
nQubits = 17
gate_names = ['Gi', 'Gxpi2', 'Gxmpi2', 'Gypi2', 'Gympi2', 'Gcphase']
availability = {'Gcphase':[(i,i+1) for i in range(nQubits-1)]}
pspec = pygsti.obj.ProcessorSpec(nQubits, gate_names, availability)
```

- Pick direct or “mirror” RB, and sample the circuits.

```
lengths = [0,10,20,30,40]
circuitsPerLength = 10
expDict = rb.sample.direct_rb_experiment(pspec, lengths, circuitsPerLength)
```

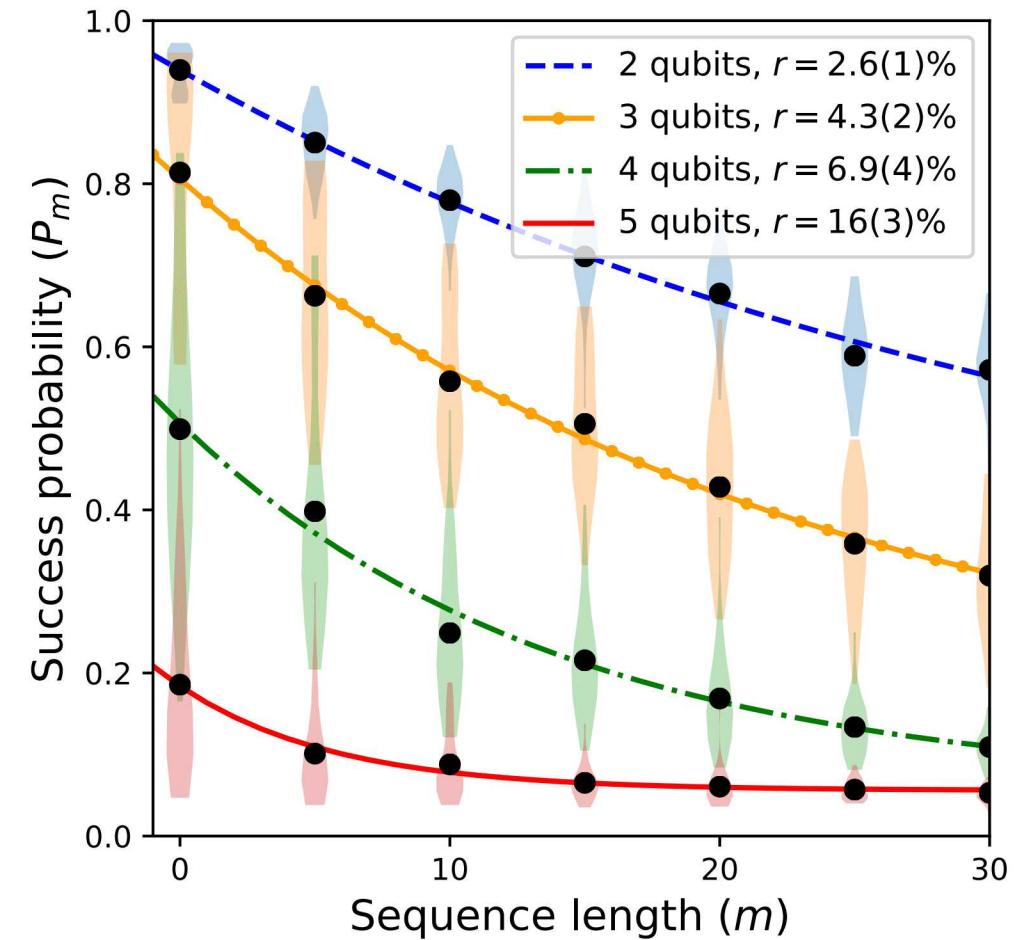
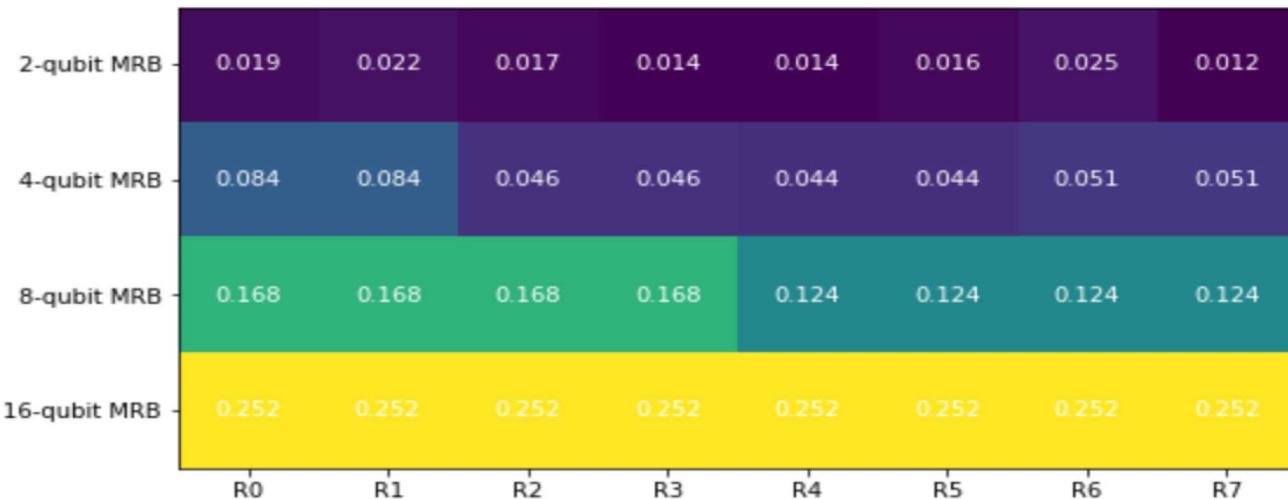
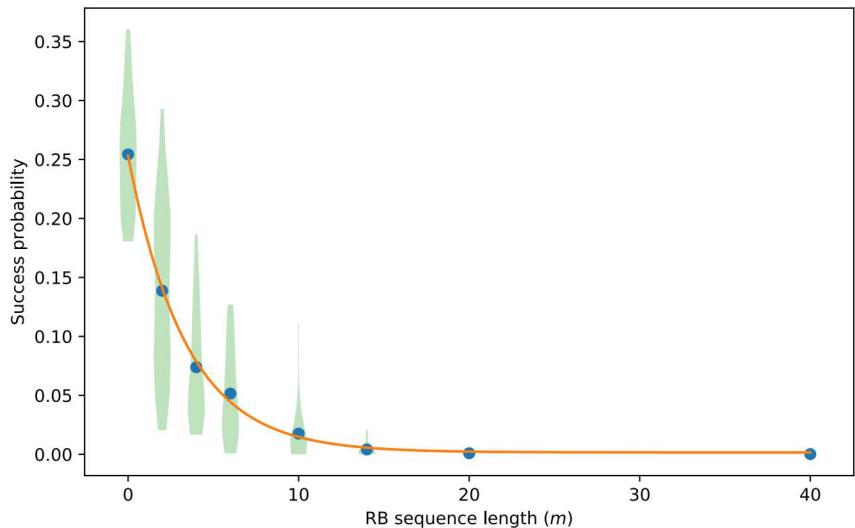
- Obtain and import data, and run the analysis.

```
rbresults = rb.analysis.std_practice_analysis(rbdata)
```

(some of this code will likely be out of date soon. Check the pyGSTi tutorials for the latest version!)

Want to run direct (or mirror) RB?

And this is
the sort of
output
you'll get...



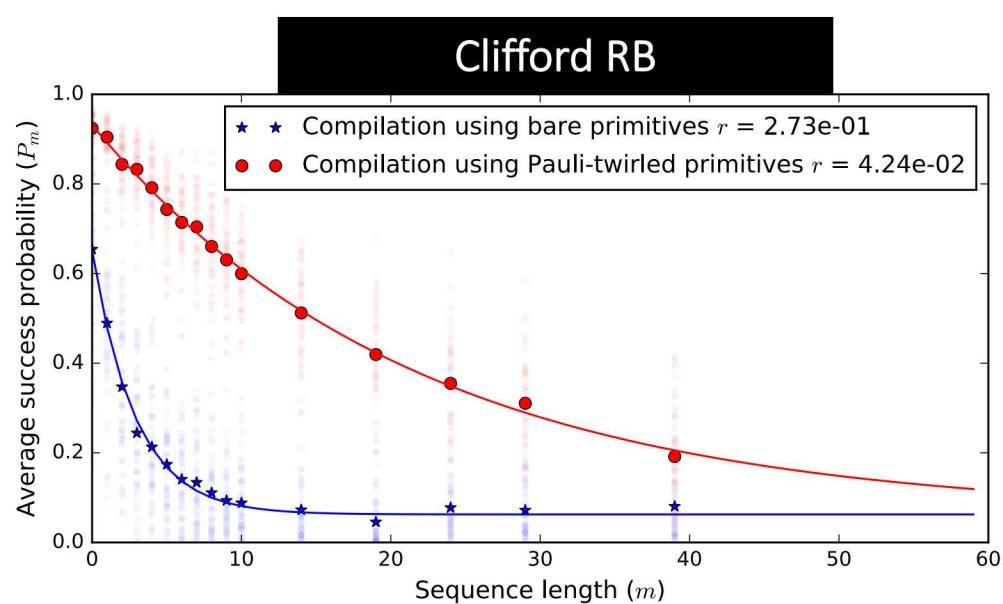
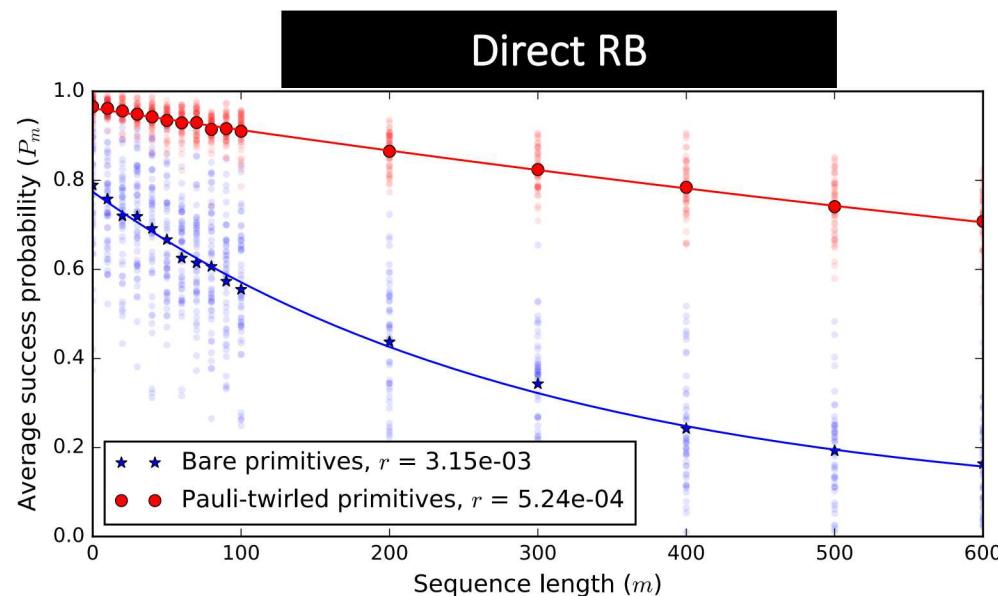
Summary

- RB is a standard-practice tool for assessing qubit performance, but the current *de facto* standard version (“Clifford RB”) is not easily applied to more than 3^{ish} qubits.
- We’ve developed streamlined RB techniques – direct and mirror RB – that facilitate benchmarking of 1 - 10s of qubits, without sacrificing the core simplicity of RB.
- Benchmarking of 10+ qubits is now feasible with current or near-term error rates!
- If you’re interested in using or learning more about direct RB or beta-stage mirror RB:
 - Come talk to me, or email me at tjproct@sandia.gov
 - Have a play around with the Jupyter notebooks on direct RB at www.pygsti.info
 - Have a read of arXiv:1807.07975 (2018).

Bonus material

Direct RB with coherent errors

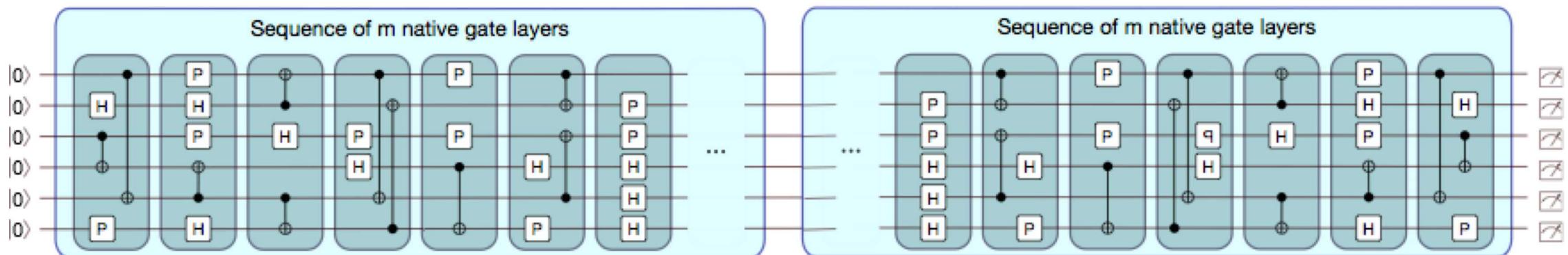
- Direct RB can have complex behavior with strongly coherent errors, but this is directly analogous to the behavior of Clifford RB under coherent errors.



- If you want to avoid this sort of behavior (this is a subjective choice) then add Pauli-randomization in between every direct RB layer.

Mirror randomized benchmarking

Bare mirror RB – a random circuit followed by it's the circuit in reverse with each gates replaced by it's inverse.¹



But this is insensitive to some types or error (e.g., if every gate was accidentally an identity gate)... So let's adjust the circuits!

¹An idea along these lines appears in Emerson *et al.*, JOB 7, S347 (2005). In that paper the gates are specified as Haar-random.

Mirror randomized benchmarking

(Pauli-twirled) mirror RB:

