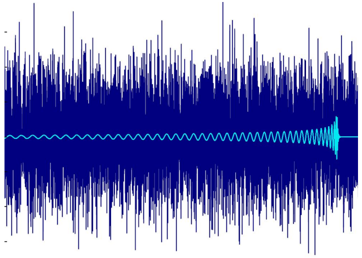
The signal and the noise

Yihui Quek
MIT
quekpottheories



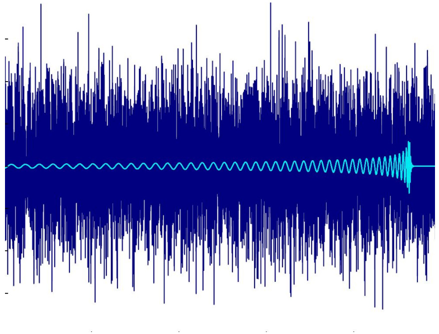
Noise is the defining characteristic of NISQ computation!

Noise is the defining characteristic of NISQ computation!

Example: Google's fidelity for their quantum advantage demonstration was just 0.002.

Noise is the defining characteristic of NISQ computation!

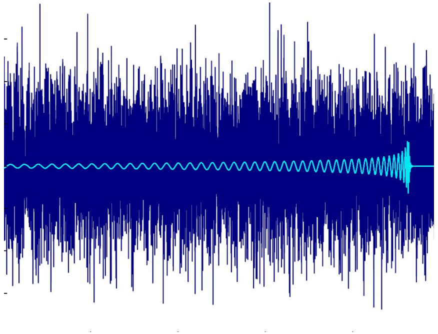
Example: Google's fidelity for their quantum advantage demonstration was just 0.002.



Can we compute with noisy devices
= can we extract any signal from the noise?

Noise is the defining characteristic of NISQ computation!

Example: Google's fidelity for their quantum advantage demonstration was just 0.002.



Can we compute with noisy devices
= can we extract any signal from the noise?

Answer: depends on the noise!

A popular noise model

Depolarizing noise: $\mathcal{D}_p(\rho) = p \cdot \rho + (1-p) \cdot \frac{\mathbb{I}}{2}$

Noisy circuit: Every gate is followed by $\mathcal{D}_p$

A popular noise model

Depolarizing noise: $\mathcal{D}_p(\rho) = p \cdot \rho + (1-p) \cdot \frac{\mathbb{I}}{2}$

Noisy circuit: Every gate is followed by $\mathcal{D}_p$

Depolarizing noise increases entropy → drives towards m.m. state

A popular noise model

Depolarizing noise: $\mathcal{D}_p(\rho) = p \cdot \rho + (1-p) \cdot \frac{\mathbb{I}}{2}$

Noisy circuit: Every gate is followed by $\mathcal{D}_p$

Depolarizing noise increases entropy → drives towards m.m. state

But what about other sources (T1, decay, readout error etc) that can decrease entropy? → drives towards non m.m. state

Depolarizing noise is usually bad news

Limitations of variational quantum algorithms: a quantum optimal transport approach

Giacomo De Palma, Milad Marvian, Cambyse Rouzé, Daniel Stilck França

The impressive progress in quantum hardware in the last years has raised the interest of the quantum computing community in harvesting the computational power of such devices. However, in the absence of error correction, these devices can only reliably implement very shallow circuits or comparatively deeper circuits at the expense of a nontrivial density of errors. In this work, we obtain extremely tight limitation bounds for standard NISQ proposals in both the noisy
with or without error-mitigation tools. The bounds limit the performance of both circuit model algorit
and also continuous-time algorithms, such as quantum annealing. In the noisy regime with local depo
prove that at depths $L = \square(p^{-1})$ it is exponentially unlikely that the outcome of a noisy quantum circ
classical algorithms for combinatorial optimization problems like Max-Cut. Although previous results

A Polynomial-Time Classical Algorithm for Noisy Random Circuit Sampling

Dorit Aharonov
Department of Computer Science and Engineering, Hebrew University
Jerusalem, Israel
dorit.aharonov@gmail.com

Xun Gao
Department of Physics, Harvard University
Cambridge, MA, USA
xungao@g.harvard.edu

Zeph Landau
Department of EECS, UC Berkeley
Berkeley, CA, USA
zeph.landau@gmail.com

Yunchao Liu
Department of EECS, UC Berkeley
Berkeley, CA, USA
yunchaoliu@berkeley.edu

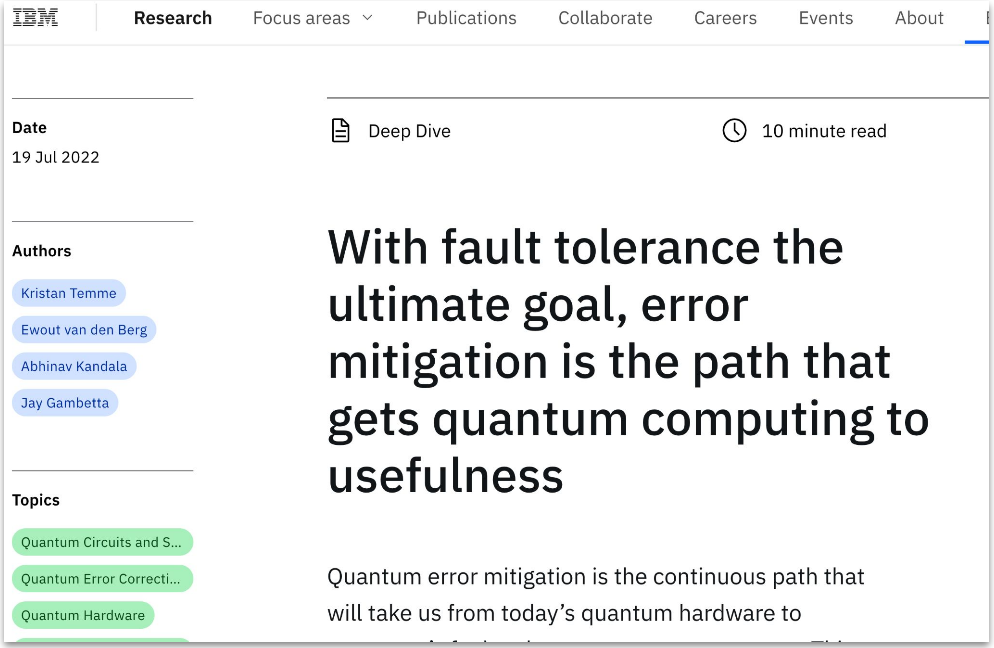
Umesh Vazirani
Department of EECS, UC Berkeley
Berkeley, CA, USA
vazirani@cs.berkeley.edu

ABSTRACT

We give a polynomial time classical algorithm for sampling from the output distribution of a noisy random quantum circuit in the regime of anti-concentration to within inverse polynomial total variation distance. The algorithm is based on a quantum analog of noise induced low degree approximations of Boolean functions, which takes the form of the truncation of a Feynman path integral in the Pauli basis.

are collected from the experimental implementation of RCS (though this number must necessarily scale exponentially in d), followed by a classical verification of these samples, using a statistical measure such as linear cross entropy (XEB), which requires classical post-processing time that is much larger and scales exponentially in n. Moreover in the experiments the depth d is sufficiently large that the output distribution of the ideal random quantum circuit (Fig. 1 (a)) is anti-concentrated[2], and indeed the output distribution tends to the Porter-Thomas distribution.

Can noise be mitigated?


IBM Research Focus areas Publications Collaborate Careers Events About

Date
19 Jul 2022

Authors
Kristan Temme
Ewout van den Berg
Abhinav Kandala
Jay Gambetta

Topics
Quantum Circuits and S...
Quantum Error Correcti...
Quantum Hardware

Deep Dive
10 minute read

With fault tolerance the ultimate goal, error mitigation is the path that gets quantum computing to usefulness

Quantum error mitigation is the continuous path that will take us from today's quantum hardware to

Short answer: no.

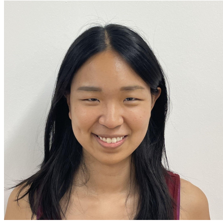
Short answer: no.

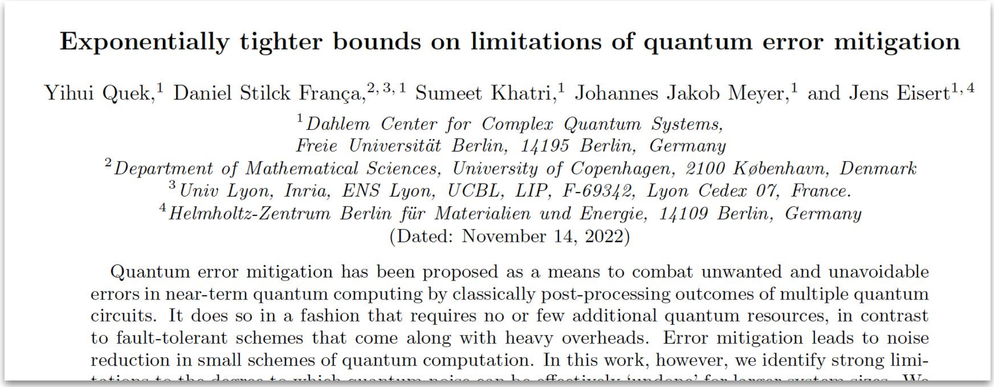
Not-so-short answer: Error mitigation is impractical (in the worst case)

Short answer: no.

Not-so-short answer: Error mitigation is impractical (in the worst case)

Vignette I: Error mitigation can fail badly





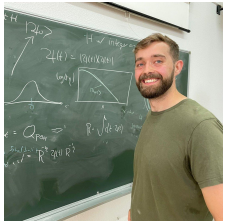
Exponentially tighter bounds on limitations of quantum error mitigation

Yihui Quek,[1] Daniel Stilck França,[2,3,1] Sumeet Khatri,[1] Johannes Jakob Meyer,[1] and Jens Eisert[1,4]

[1] *Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2] *Department of Mathematical Sciences, University of Copenhagen, 2100 København, Denmark*
[3] *Univ Lyon, Inria, ENS Lyon, UCBL, LIP, F-69342, Lyon Cedex 07, France.*
[4] *Helmholtz-Zentrum Berlin für Materialien und Energie, 14109 Berlin, Germany*
(Dated: November 14, 2022)

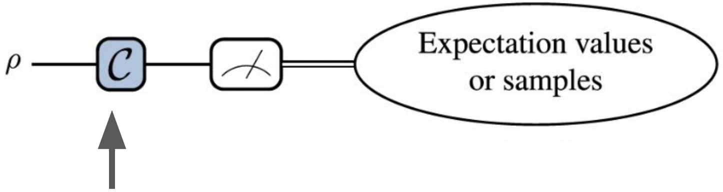
Quantum error mitigation has been proposed as a means to combat unwanted and unavoidable errors in near-term quantum computing by classically post-processing outcomes of multiple quantum circuits. It does so in a fashion that requires no or few additional quantum resources, in contrast to fault-tolerant schemes that come along with heavy overheads. Error mitigation leads to noise reduction in small schemes of quantum computation. In this work, however, we identify strong limi-

arXiv: 2210.11505



What is error mitigation?

In a world with noiseless quantum computers:

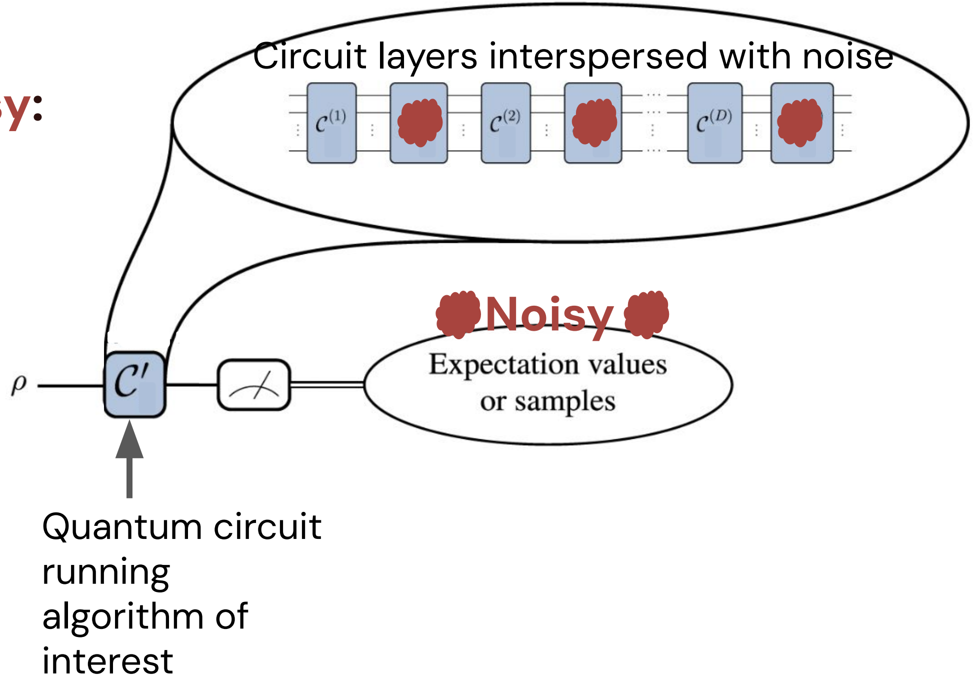


Quantum circuit running algorithm of interest

What is error mitigation?

In the real world, C is noisy:

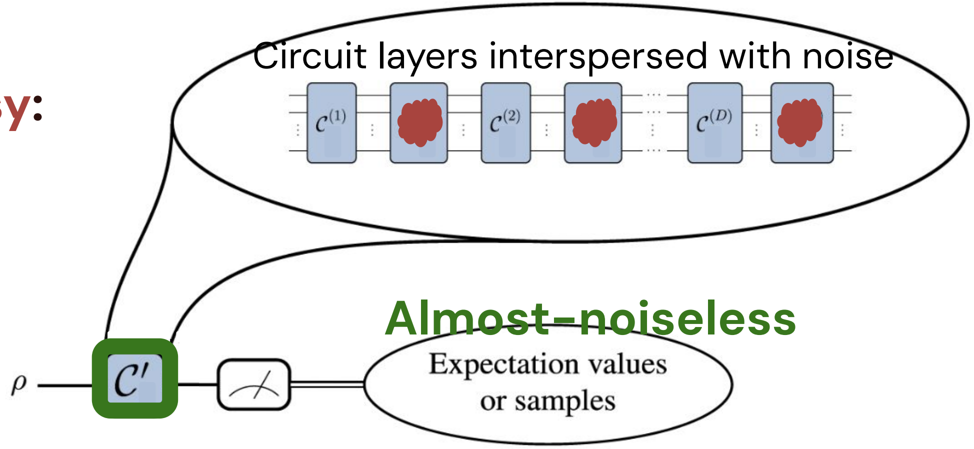
- qubit decoherence
- gate errors



What is error mitigation?

In the real world, C is noisy:

- qubit decoherence
- gate errors

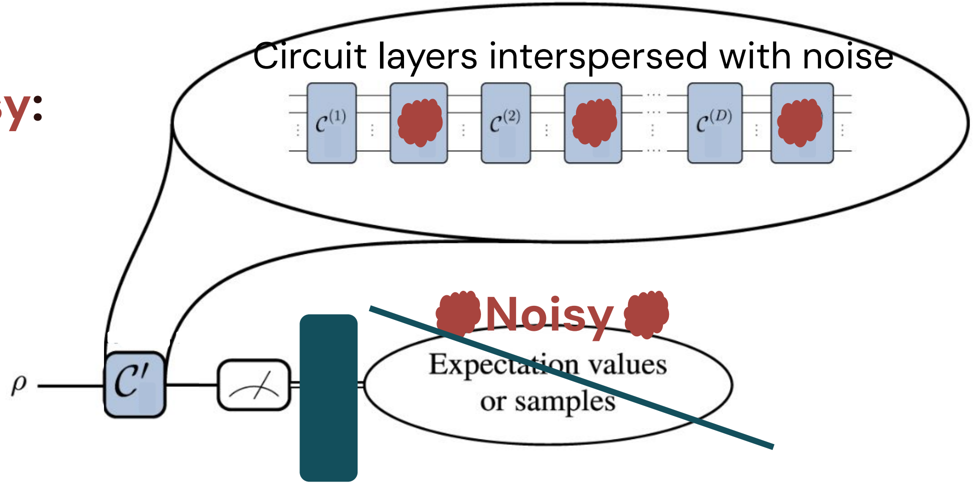


Solution 1: **fault-tolerance** (requires lots of machinery – mid-circuit measurements, auxiliary qubits, pumping out entropy)

What is error mitigation?

In the real world, C is noisy:

- qubit decoherence
- gate errors



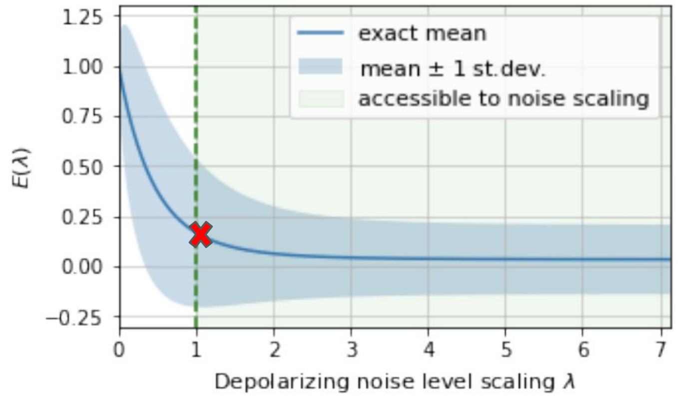
Solution 1: fault-tolerance
Solution 2: **error mitigation**
(near-term alternative)

Error mitigation algorithm (mostly classical) → Almost noiseless expectation values or samples

Example of error mitigation protocol

Zero-noise extrapolation:

1) Run the circuit of interest at amplified noise level λ (call this). C_λ



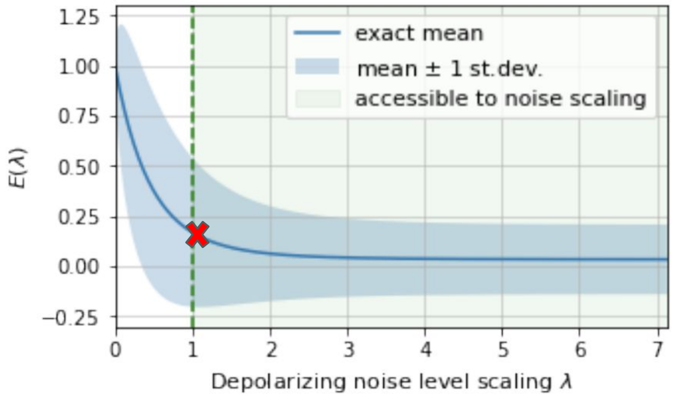
Plot taken from Giurgica-Tiron et al, 2020 IEEE International Conference on Quantum Computing and Engineering (QCE)

Example of error mitigation protocol

Zero-noise extrapolation:

1) Run the circuit of interest at amplified noise level λ (call this). C_λ
2) Measure

$$E(\lambda) = \mathrm{Tr}(C_\lambda(\rho_{\rm in})O)$$



Plot taken from Giurgica-Tiron et al, 2020 IEEE International Conference on Quantum Computing and Engineering (QCE)

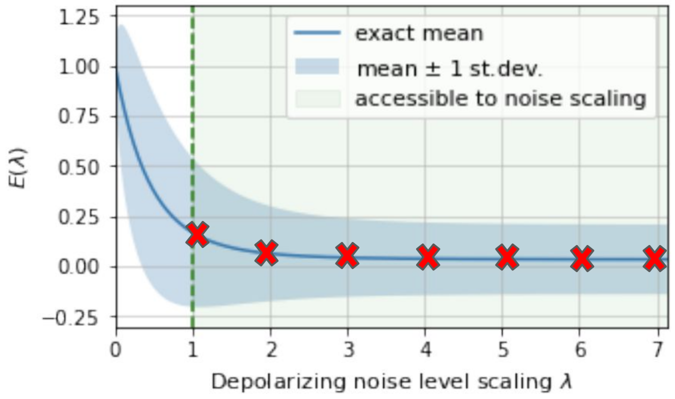
Example of error mitigation protocol

Zero-noise extrapolation:

1) Run the circuit of interest at amplified noise level λ (call this). C_λ
2) Measure

$$E(\lambda) = \mathrm{Tr}(C_\lambda(\rho_{\text{in}})O)$$

3) Repeat steps 1, 2 for different λ.



Plot taken from Giurgica-Tiron et al, 2020 IEEE International Conference on Quantum Computing and Engineering (QCE)

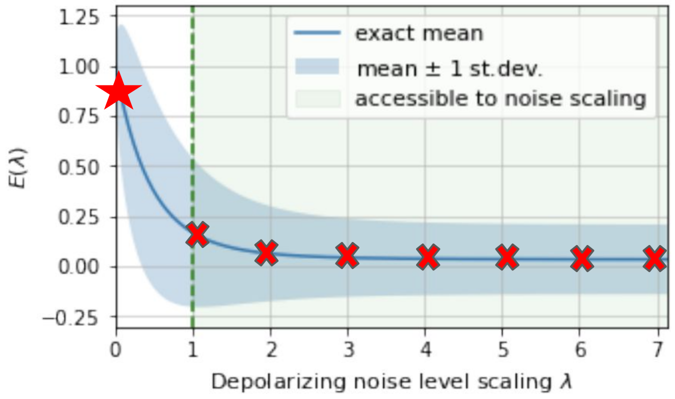
Example of error mitigation protocol

Zero-noise extrapolation:

1) Run the circuit of interest at amplified noise level λ (call this C_λ).
2) Measure

$$E(\lambda) = \mathrm{Tr}(C_\lambda(\rho_{\text{in}})O)$$

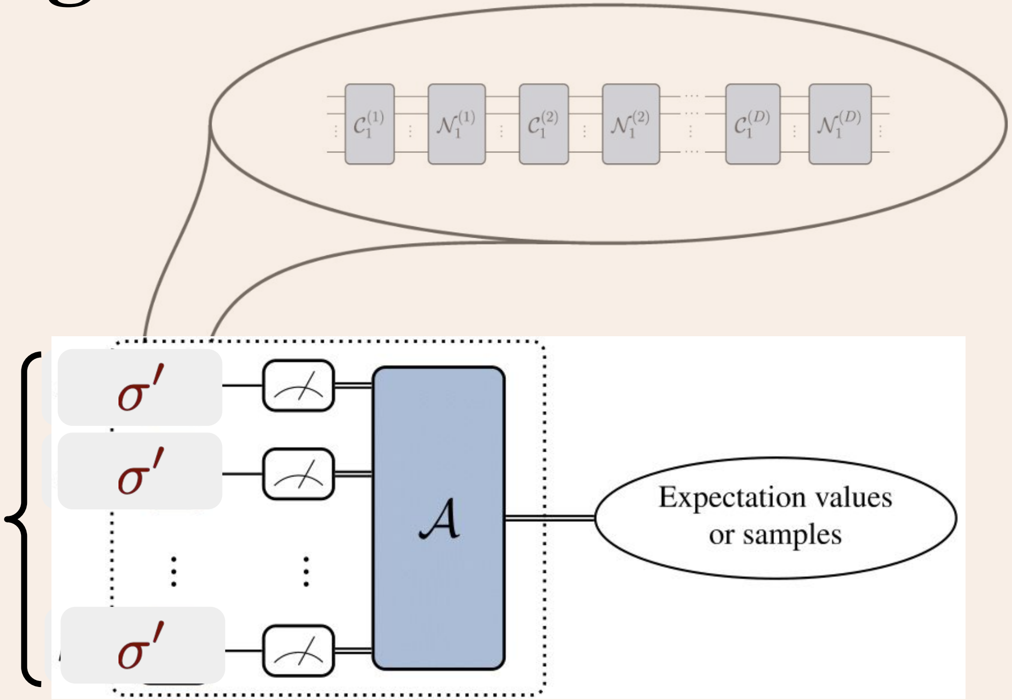
3) Repeat steps 1, 2 for different λ.
4) Output the extrapolated value $E(0)$.



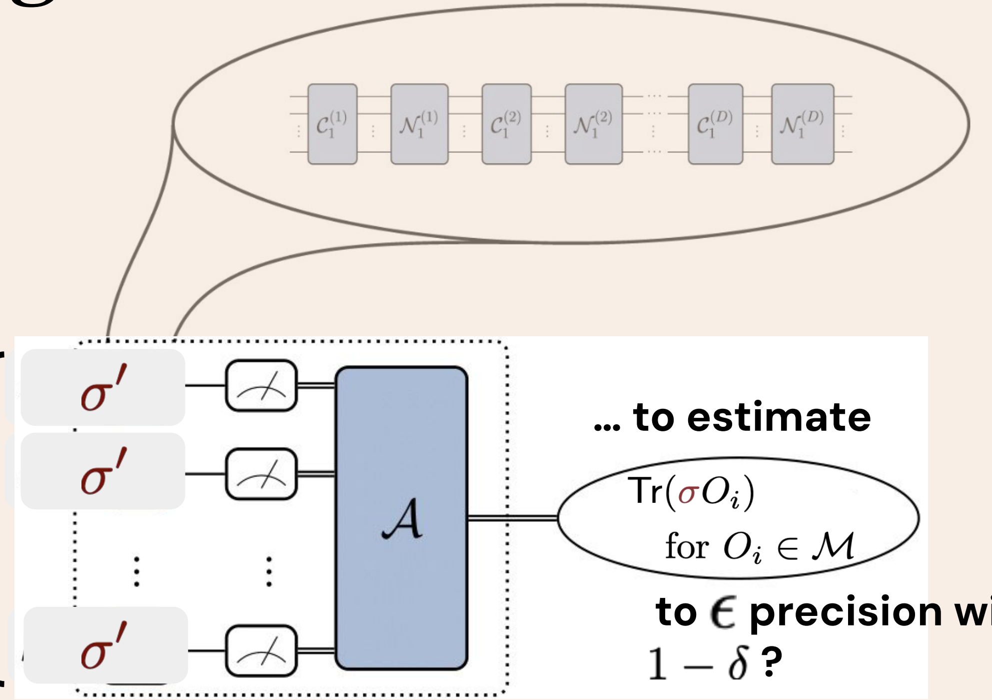
Plot taken from Giurgica-Tiron et al, 2020 IEEE International Conference on Quantum Computing and Engineering (QCE)

Our question: sample complexity of error mitigation?

How many copies of σ' does $\mathcal{A}$ need...



Our question: sample complexity of error mitigation?



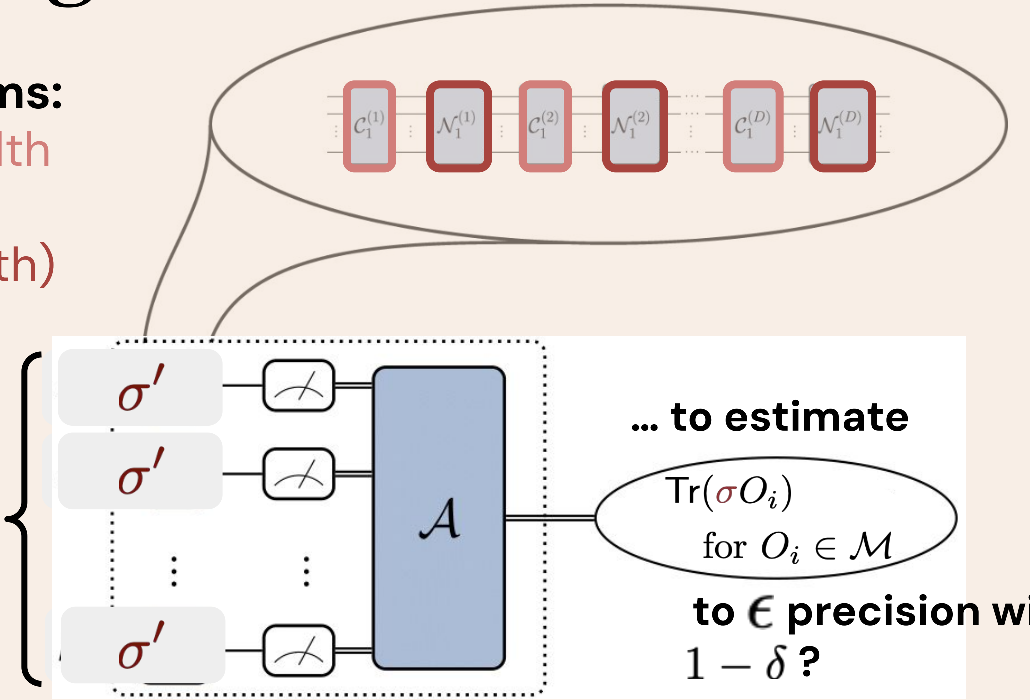
How many copies of σ' does $\mathcal{A}$ need…

… to estimate $\mathrm{Tr}(\sigma O_i)$ for $O_i \in \mathcal{M}$ to ϵ precision with probability $1-\delta$?

Our question: sample complexity of error mitigation?

Relevant params:
n, D (circuit width and depth)
p (noise strength)

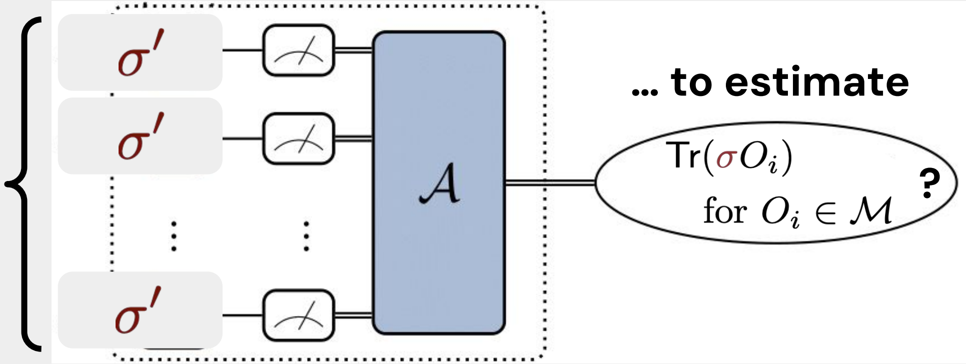
How many copies of σ' does $\mathcal{A}$ need…



… to estimate $\mathrm{Tr}(\sigma O_i)$ for $O_i \in \mathcal{M}$ **to ϵ precision with probability $1-\delta$?**

Our lower bounds

How many copies of σ' does $\mathcal{A}$ need...



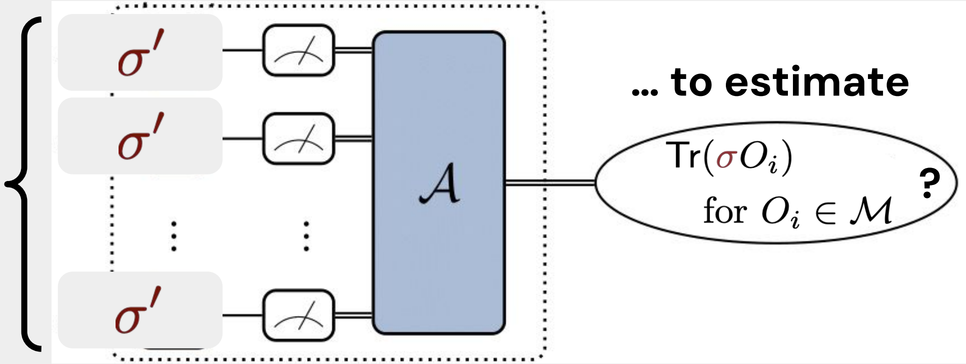
Relevant parameters:
n, D (circuit width/depth);
p (depolarizing noise strength)

For depolarizing noise:

$p^{-\Omega(n\,D)}$ copies for $D = \Omega(\log\log(n))$

Our lower bounds

How many copies of σ' does $\mathcal{A}$ need...



Relevant parameters:
n, D (circuit width/depth);
p (depolarizing noise strength)

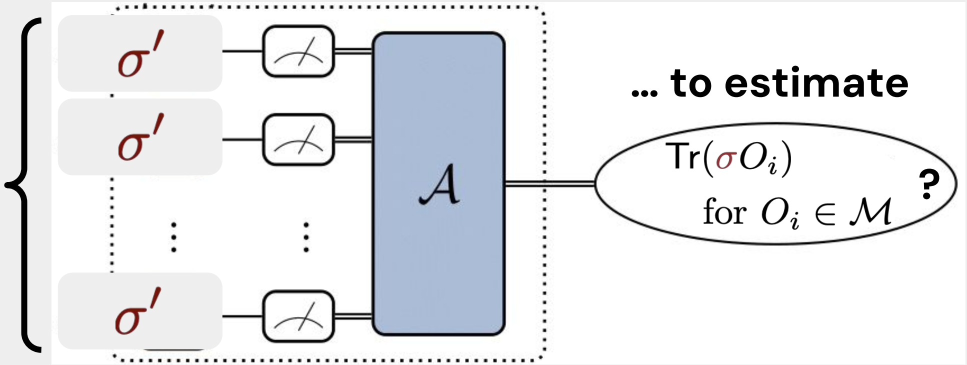
For depolarizing noise:

$p^{-\Omega(n\,D)}$ copies for $D = \Omega(\log\log(n))$

For non-unital noise (toy model): $c^{-\Omega(n\,D)}$ copies

Our lower bounds

How many copies of σ' does $\mathcal{A}$ need...



Relevant parameters:
n, D (circuit width/depth);
p (depolarizing noise strength)

Previously proven: $\Omega(\exp(D))$

Our result: $\Omega(\exp(nD)) \rightarrow$ **Exponentially stronger**

Intuition

We show: $\exp(\Omega(nD))$ runs of a noisy circuit are required for good error mitigation.

Proof intuition:

1. **Depolarizing noise + rapidly mixing circuit drives *every* input state towards the maximally-mixed state**

Intuition

We show: $\exp(\Omega(nD))$ runs of a noisy circuit are required for good error mitigation.

Proof intuition:

1. **Depolarizing noise + rapidly mixing circuit drives *every* input state towards the maximally-mixed state**
2. This makes circuit output states non-distinguishable

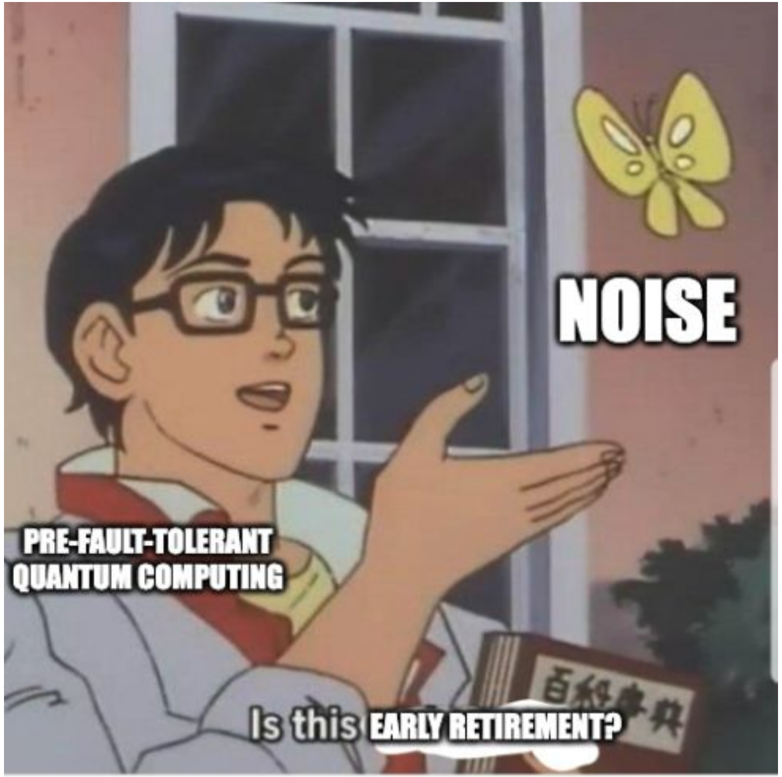
Intuition

We show: $\exp(\Omega(nD))$ runs of a noisy circuit are required for good error mitigation.

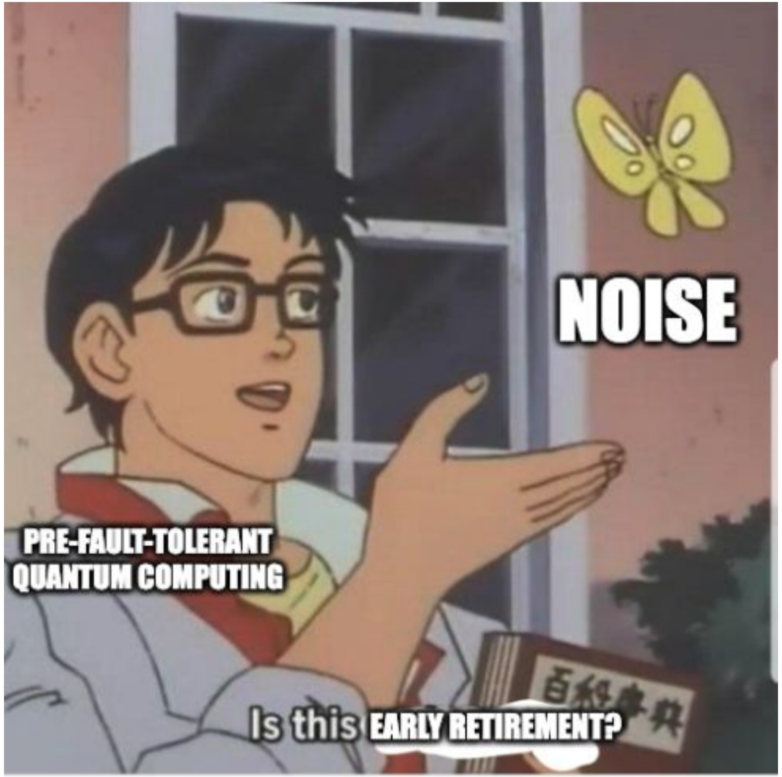
Proof intuition:

1. **Depolarizing noise + rapidly mixing circuit drives *every* input state towards the maximally-mixed state**
2. This makes circuit output states non-distinguishable
3. All information lost after circuit → cannot error mitigate

The wrong conclusion:



The wrong conclusion:



Let's now look at a different type of noise.

What about…non-unital noise?

Canonical example: **amplitude damping noise!**

$$\mathcal{N}(\rho) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \rho \begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}$$

What about...non-unital noise?

Canonical example: **amplitude damping noise!**

$|1\rangle \to |0\rangle$

$$\mathcal{N}(\rho) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \rho \begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}$$

What about…non-unital noise?

Canonical example: **amplitude damping noise!**

$|1\rangle \to |0\rangle$

$$\mathcal{N}(\rho) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \rho \begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}$$

Amplitude-damping noise ~ **reset**-to-all-0s

What about...non-unital noise?

Canonical example: **amplitude damping noise!**

$|1\rangle \rightarrow |0\rangle$

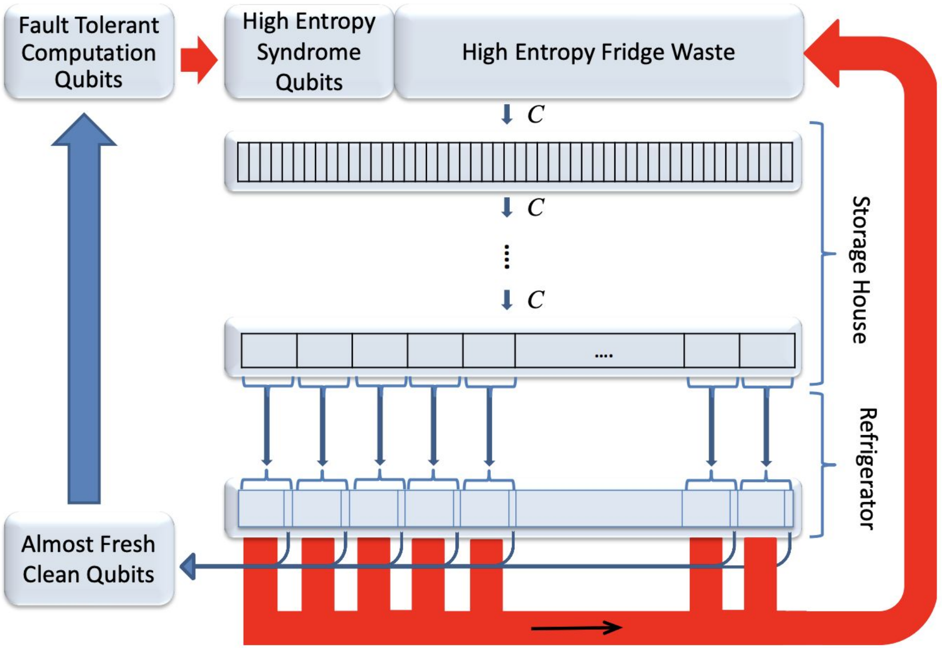
$$\mathcal{N}(\rho) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \rho \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \rho \begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}$$

Amplitude-damping noise ~ **reset**-to-all-0s

Example of amplitude damping: T1 decay, readout error etc. Dominant source of noise in superconducting qubits.

Amplitude-damping noise ~ **reset**-to-all-0s

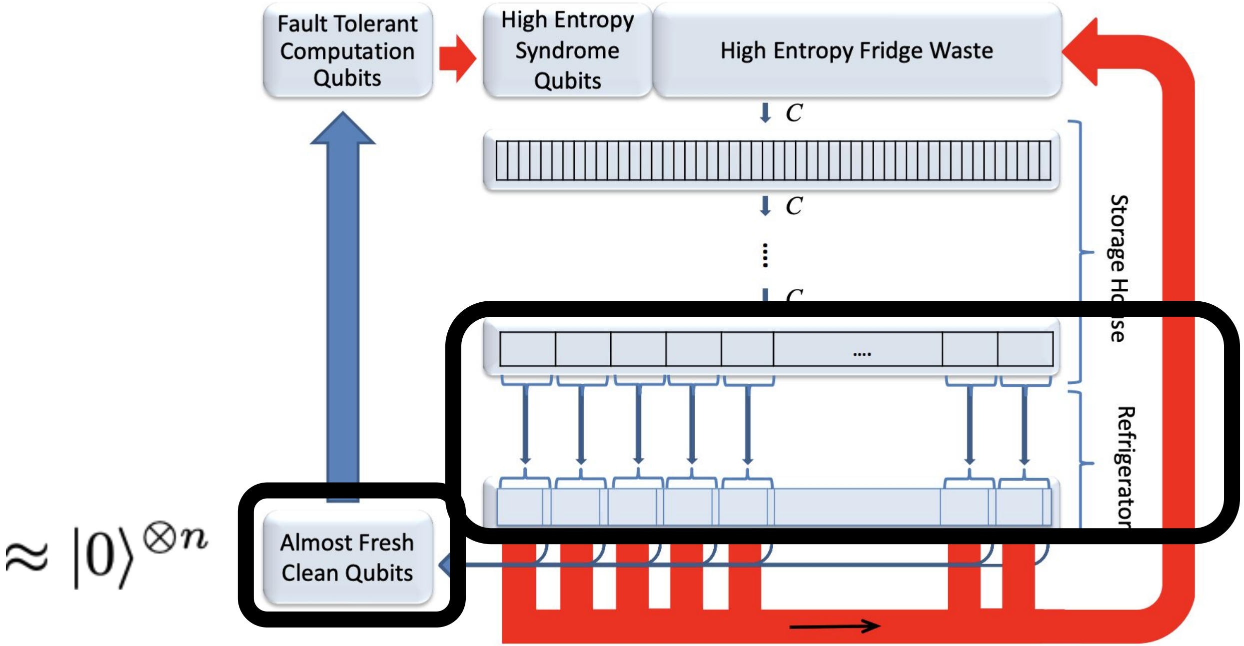
Non-unital noise: fault-tolerance "for free"?



Quantum refrigerator, Ben-Or, Gottesman, Hassidim (arXiv 1301.1995)

Amplitude-damping noise ~ **reset**-to-all-0s

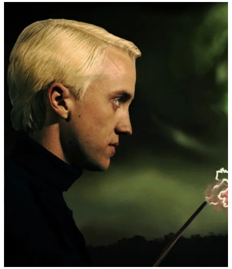
Non-unital noise: fault-tolerance "for free"?



Implemented by non-unital noise (drives towards $|0\rangle$)

Quantum refrigerator, Ben-Or, Gottesman, Hassidim (arXiv 1301.1995)

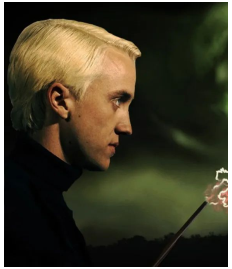
What happens when both amplitude-damping and depolarizing noise are present?



Depolarizing noise tends to “scramble” the distribution by increasing entropy.

A layer of random gates also tries to“scramble” the distribution!

What happens when both amplitude-damping and depolarizing noise are present?

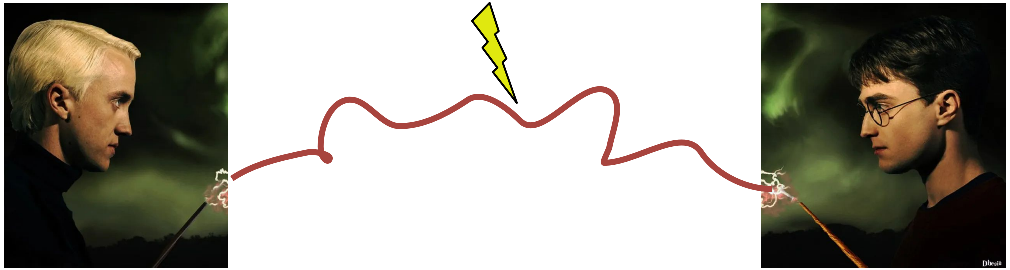


Depolarizing noise tends to “scramble” the distribution by increasing entropy.

A layer of random gates also tries to“scramble” the distribution!

Amplitude damping noise tries to “unscramble” the distribution by decreasing entropy!

What happens when both amplitude-damping and depolarizing noise are present?



Depolarizing noise tends to “scramble” the distribution by increasing entropy.

A layer of random gates also tries to“scramble” the distribution!

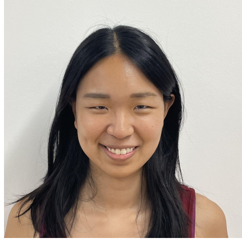
Amplitude damping noise tries to “unscramble” the distribution by decreasing entropy!

Who wins this fight?!

Vignette II: Non-unital noise induces absence of barren plateaus

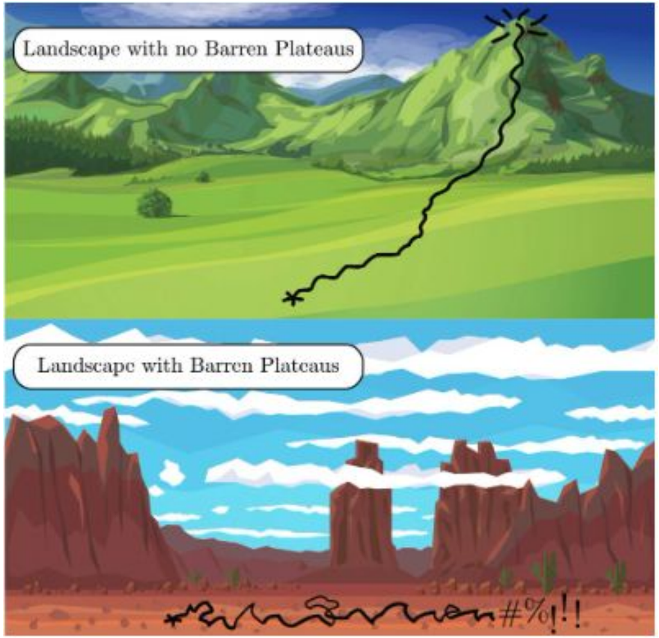







With Antonio Mele, Armando Angrisani, Soumik Ghosh, Daniel Stilck Franca, Jens Eisert

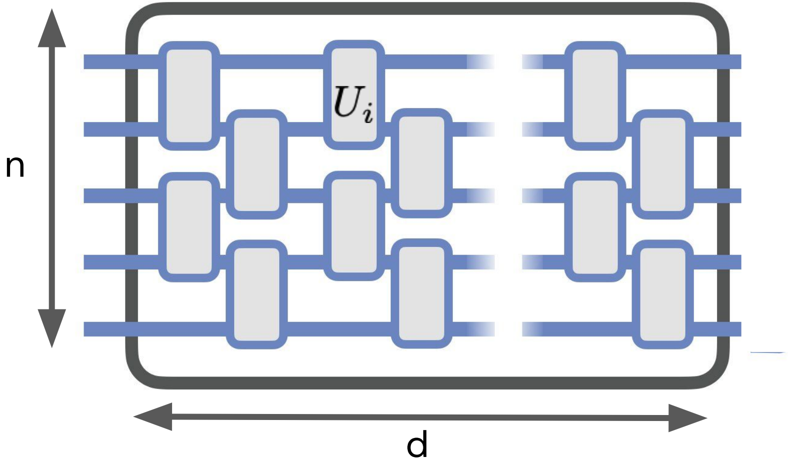
Barren plateaus are bad for variational algorithms



https://www.eurekalert.org/multimedia/739167

Gradient vanishes in all directions; can't figure out where to go!

Barren plateaus make optimization hard

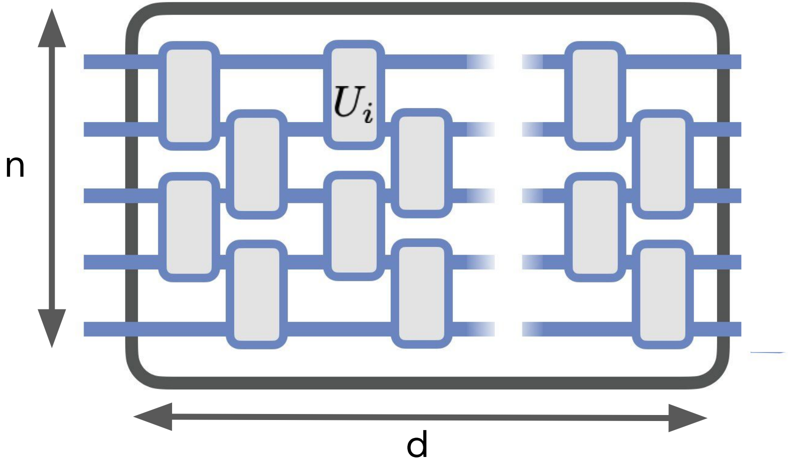


C = gradient of cost function

Math translation of ‘vanishing gradient’:

$$\mathrm{Var}_{U_1,\ldots U_m}[C] = O(\exp(-n))$$

Barren plateaus make optimization hard



C = gradient of cost function

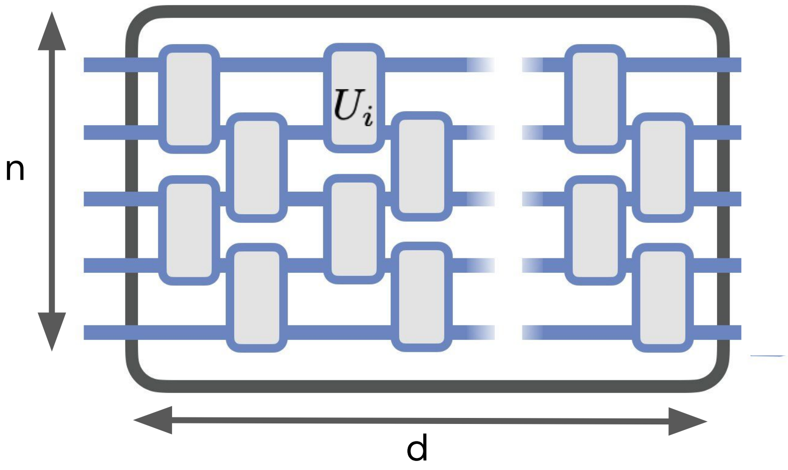
Math translation of 'vanishing gradient':

$$\mathrm{Var}_{U_1,\ldots U_m}[C] = O(\exp(-n))$$

Recall: We are optimizing over

$$U_1, \ldots U_m$$

Barren plateaus make optimization hard



C = gradient of cost function

Math translation of 'vanishing gradient':

$$\mathrm{Var}_{U_1,\ldots U_m}[C] = O(\exp(-n))$$

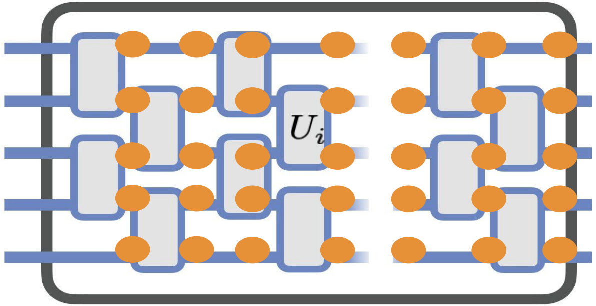
Recall: We are optimizing over

$$U_1, \ldots U_m$$

BP implies: if we initialize the optimization randomly, gradient looks the same in all directions.

BP: $\mathrm{Var}_{U_1,\ldots U_m}[C] = O(\exp(-n))$

What if the circuit is affected by noise ●?

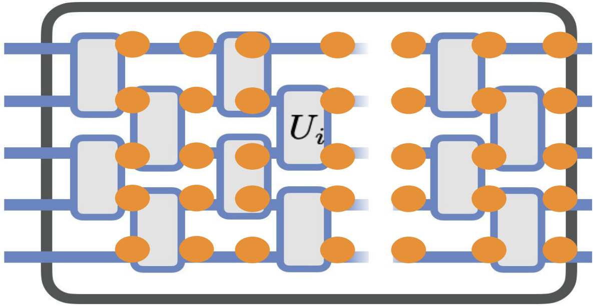


Type of noise ●	Depth at which BP happen	Cause of BP
Noiseless [1]	Linear (in n)	[BHH'13] Circuit converges to Haar-random at linear depth

[1] Barren plateaus in quantum neural network training landscapes. ***J. R. McClean et al. Nature Comm. (2018).***

BP: $\mathrm{Var}_{U_1,\ldots U_m}[C] = O(\exp(-n))$

What if the circuit is affected by noise ?

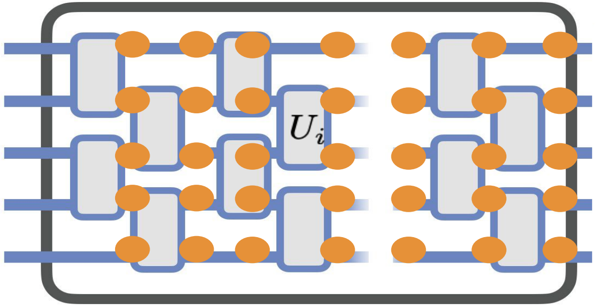


Type of noise	Depth at which BP happen	Cause of BP
Noiseless [1]	Linear (in n)	[BHH'13] Circuit converges to Haar-random at linear depth
Depolarizing noise [2]	Linear	Depolarizing noise drives output state towards identity.

[1] Barren plateaus in quantum neural network training landscapes. *J. R. McClean et al. Nature Comm. (2018).*
[2] Noise-induced barren plateaus in variational quantum algorithms. *S. Wang et al. Nature Comm. (2021).*

BP: $\mathrm{Var}_{U_1,\ldots U_m}[C] = O(\exp(-n))$

What if the circuit is affected by noise ?



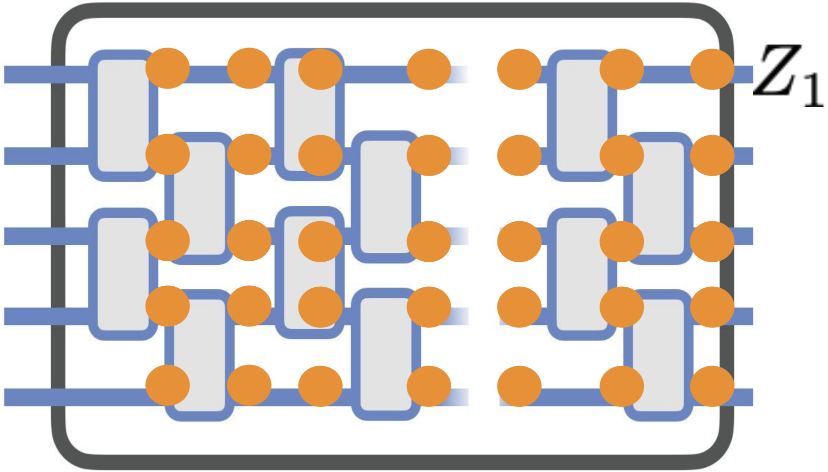
Type of noise	Depth at which BP happen	Cause of BP
Noiseless [1]	Linear (in n)	[BHH'13] Circuit converges to Haar-random at linear depth
Depolarizing noise [2]	Linear	Depolarizing noise drives output state towards identity.
Non-unital noise (Our work)	Never (Var(C) = **Ω**(1), all depths)	Non-unital noise acts like mini 'reset'.

I.e. we show noise-induced *absence of* barren plateaus! (No guarantees that the minimum is still in the right place, though…)

[1] Barren plateaus in quantum neural network training landscapes. *J. R. McClean et al. Nature Comm. (2018).*
[2] Noise-induced barren plateaus in variational quantum algorithms. *S. Wang et al. Nature Comm. (2021).*

Intuition

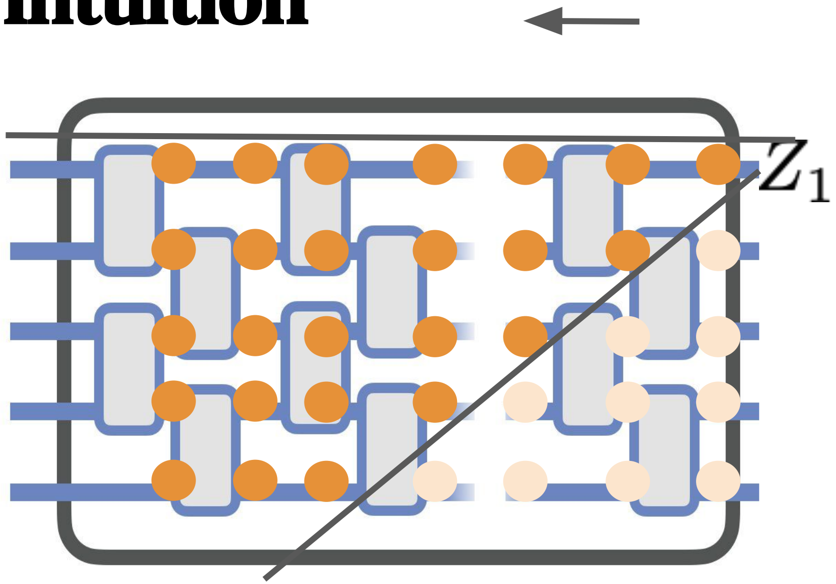
Think about $\mathrm{Var}_{\mathcal{C}}[\mathrm{Tr}(Z_1 \mathcal{C}(\rho))]$



Will show this is at least some constant.

Intuition

Claim: $\mathrm{Var}_{\mathcal{C}}[\mathrm{Tr}(Z_1\mathcal{C}(\rho))] \geq \Omega(1)$

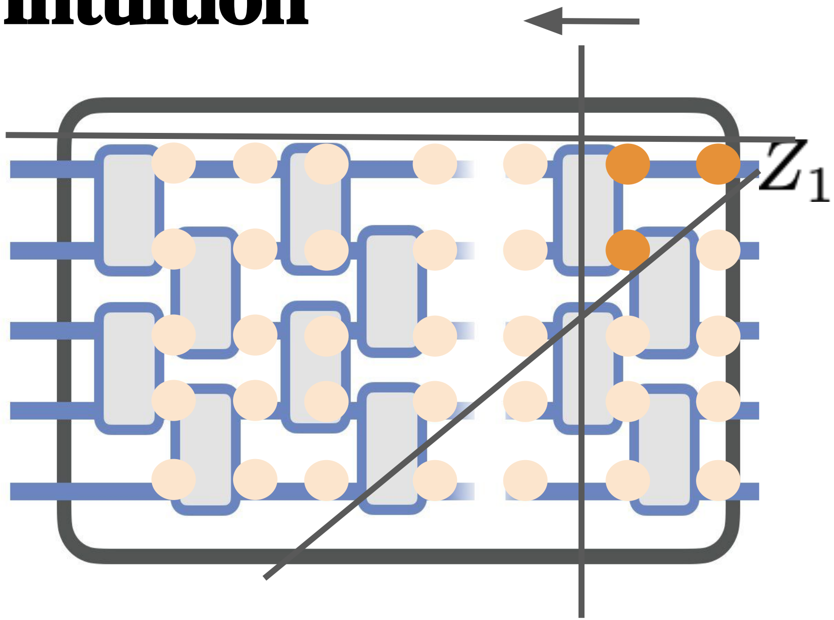


1. In Heisenberg picture:

$$\mathrm{Tr}(Z_1\mathcal{C}(\rho)) = \mathrm{Tr}(\mathcal{C}^\dagger(Z_1)\rho)$$

Intuition

Claim: $\text{Var}_{\mathcal{C}}[\text{Tr}(Z_1\mathcal{C}(\rho))] \geq \Omega(1)$



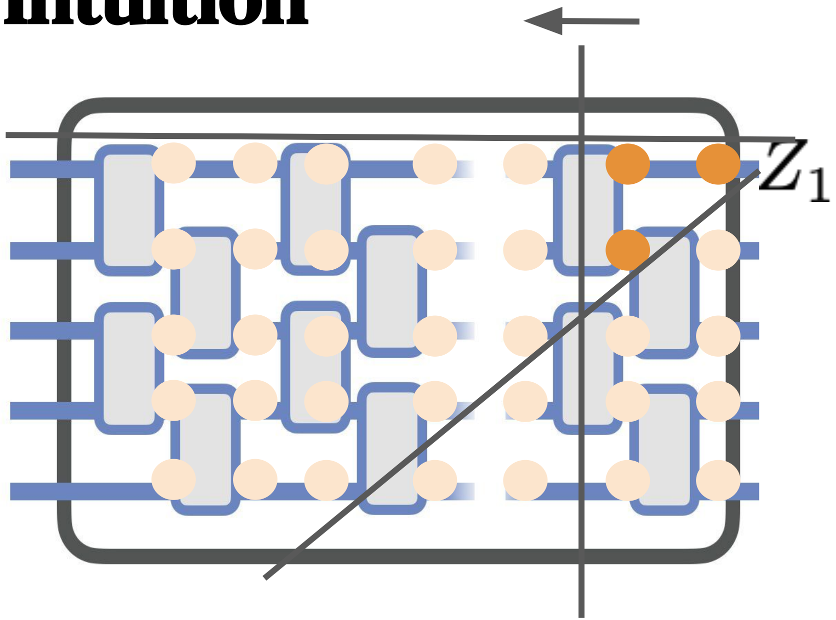
1. In Heisenberg picture:

$$\text{Tr}(Z_1\mathcal{C}(\rho)) = \text{Tr}(\mathcal{C}^\dagger(Z_1)\rho)$$

2. After constant # layers:
   - noise resets entire lightcone of Z1 to some fixed state with constant probability

Intuition

Claim: $\text{Var}_{\mathcal{C}}[\text{Tr}(Z_1\mathcal{C}(\rho))] \geq \Omega(1)$



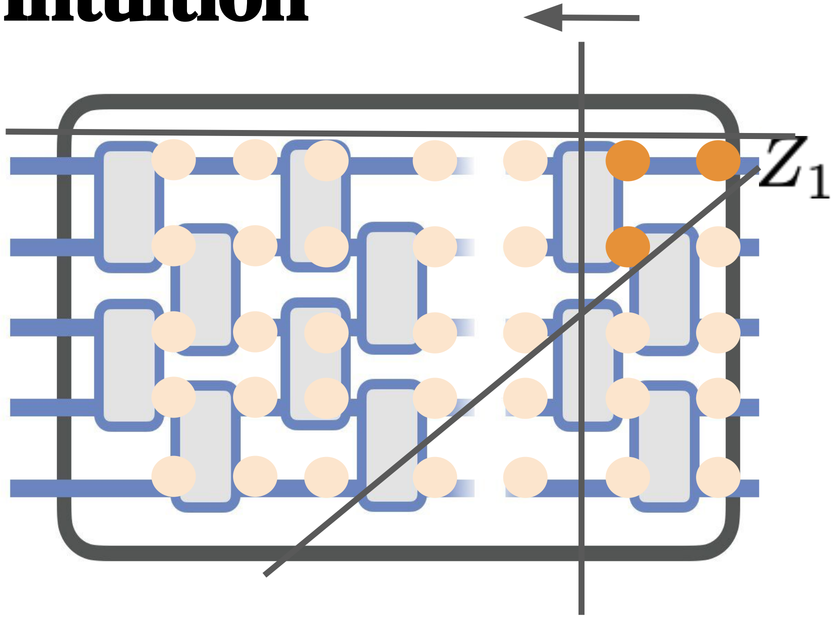
1. In Heisenberg picture:

$$\text{Tr}(Z_1\mathcal{C}(\rho)) = \text{Tr}(\mathcal{C}^\dagger(Z_1)\rho)$$

2. After constant # layers:
   - noise resets entire lightcone of Z1 to some fixed state with constant probability
   - equivalent to running a constant-depth circuit

Intuition

Claim: $\mathrm{Var}_{\mathcal{C}}[\mathrm{Tr}(Z_1\mathcal{C}(\rho))] \geq \Omega(1)$



1. In Heisenberg picture:

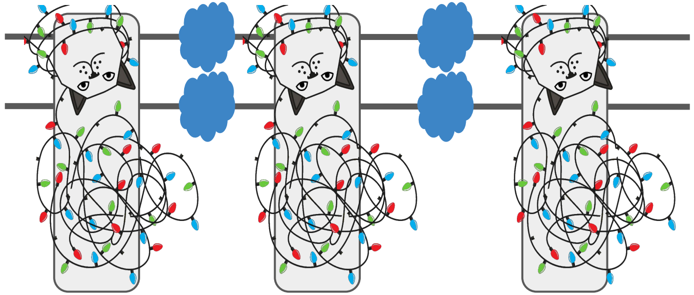
$$\mathrm{Tr}(Z_1\mathcal{C}(\rho)) = \mathrm{Tr}(\mathcal{C}^\dagger(Z_1)\rho)$$

2. After constant # layers:
   - noise resets entire lightcone of Z1 to some fixed state with constant probability
   - equivalent to running a constant-depth circuit
3. Constant depth circuits have no barren plateaus (not enough randomness)

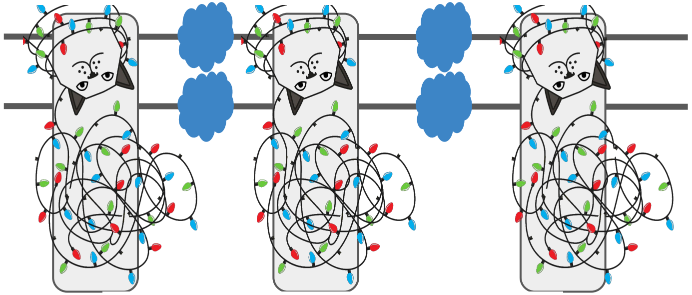
Who wins this fight: us (with some thought)

Who wins this fight: us (with some thought)

Error mitigation may be hopeless on circuits that scramble information rapidly (increase entropy fast)



Who wins this fight: us (with some thought)



Error mitigation may be hopeless on circuits that scramble information rapidly (increase entropy fast)

BUT: Non-unital noise decreases entropy. Can we take advantage of this in the intermediate term?