Title: Beware the (log)logjam: Quantum error mitigation becomes hard at polyloglog(n) depth

Speakers: Yihui Quek

Series: Perimeter Institute Quantum Discussions

Date: March 15, 2023 - 3:30 PM

URL: https://pirsa.org/23030084

Abstract: Quantum error mitigation has been proposed as a means to combat unwanted and unavoidable errors in near-term quantum computing using no or few additional quantum resources, in contrast to fault-tolerant schemes that come along with heavy overheads. Error mitigation has been successfully applied to reduce noise in near-term applications.

In this work, however, we identify strong limitations to the degree to which quantum noise can be effectively `undone' for larger system sizes. We set up a framework that rigorously captures large classes of error mitigation schemes in use today. The core of our argument combines fundamental limits of statistical inference with a construction of families of random circuits that are highly sensitive to noise.

We show that even at poly loglog depth, a super-polynomial number of samples is needed in the worst case to estimate the expectation values of noiseless observables, the principal task of error mitigation. Notably, our construction implies that scrambling due to noise can kick in at exponentially smaller depths than previously thought. They also impact other near-term applications, constraining kernel estimation in quantum machine learning, causing an earlier emergence of noise-induced barren plateaus in variational quantum algorithms and ruling out exponential quantum speed-ups in estimating expectation values in the presence of noise or preparing the ground state of a Hamiltonian.

Zoom link: https://pitp.zoom.us/j/95736148335?pwd=akZLaHE5aStNQVZOeVFETlltNzVwdz09

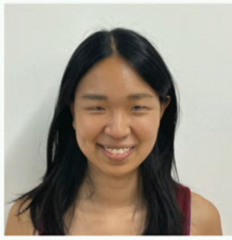
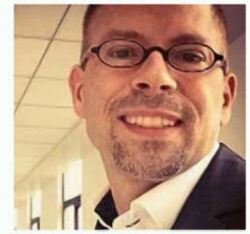
Beware the (log)logjam:
Error mitigation becomes hard at poly loglog(n) depth

Freie Universität Berlin

Yihui Quek
FU Berlin → Harvard
quekpottheories

HARVARD

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

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-

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

From https://www.wonderopolis.org/wonder/what-is-a-log-jam/

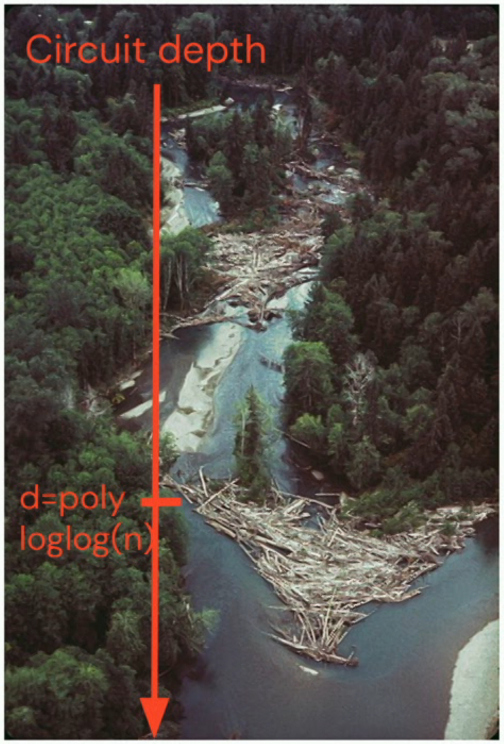
> Sometimes rivers would become so full of logs that they would pack together so tightly that nothing would move. This was known as a log jam. Log jams were common in areas with shallow water or bends in the river.

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

From https://www.wonderopolis.org/wonder/what-is-a-log-jam/

Sometimes rivers would become so full of logs that they would pack together so tightly that nothing would move. This was known as a log jam. Log jams were common in areas with shallow water or bends in the river.

Log jams were very serious situations. They blocked traffic and were very dangerous to fix. River drivers would try to move logs by hand, but they often had to resort to blasting logs with dynamite to get them moving again.



A pileup of errors making computation classically unsalvageable

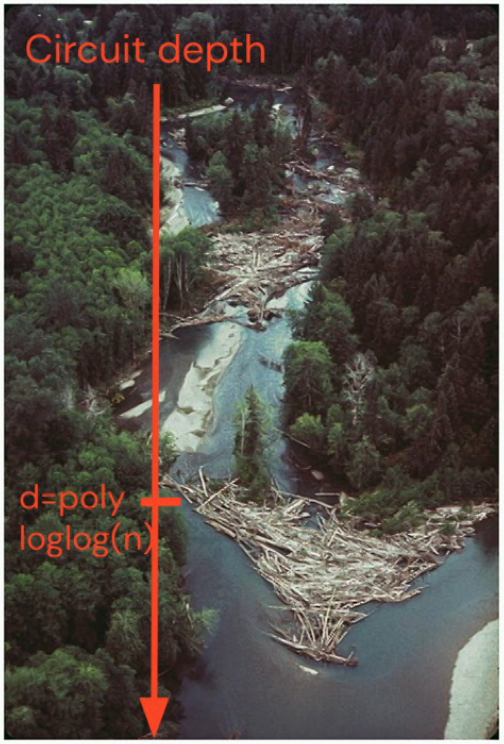
Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

From https://www.wonderopolis.org/wonder/what-is-a-log-jam/

Sometimes rivers would become so full of logs that they would pack together so tightly that nothing would move. This was known as a log jam. Log jams were common in areas with shallow water or bends in the river.

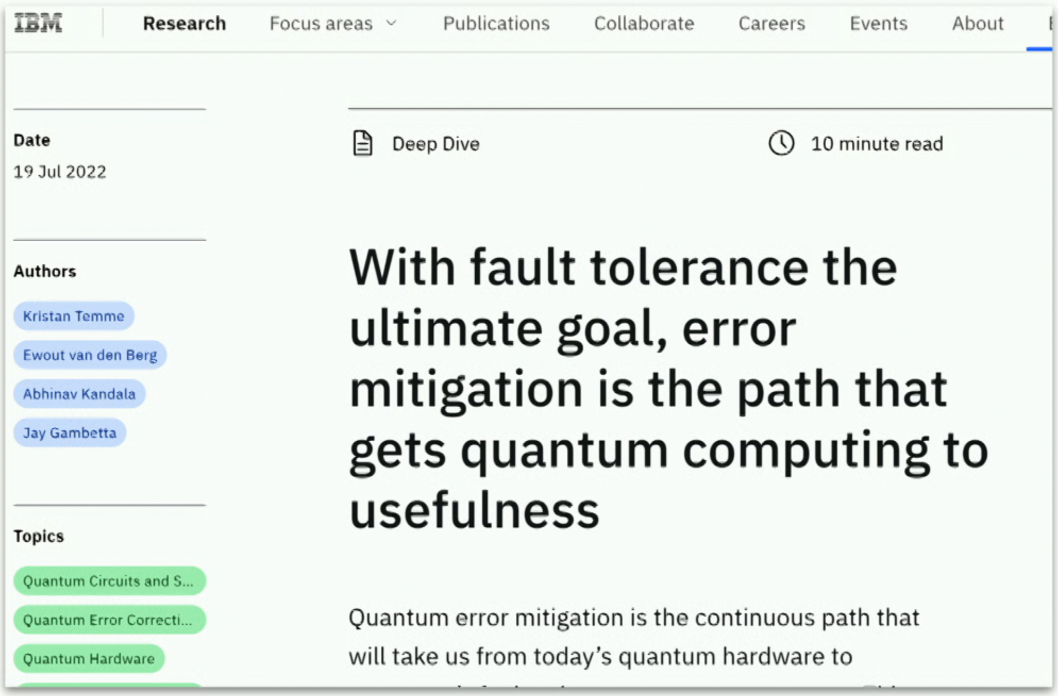
This work:

Loglogjams were very serious situations. They blocked quantum computation and were very dangerous to fix. Quantum computer scientists would try to remove errors by classical postprocessing, but they often had to resort to running the circuit **exponentially-many times** to get an accurate result.



A pileup of errors making computation classically unsalvageable

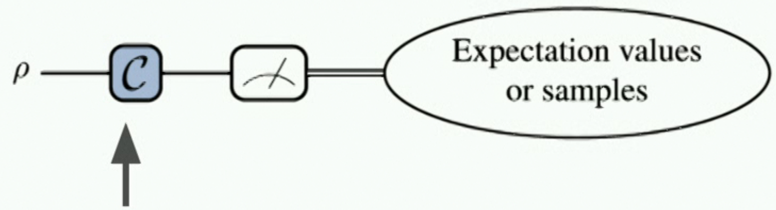
Full fault-tolerance is hard. What's achievable with limited resources?



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

What is error mitigation?

In a world with noiseless quantum computers:



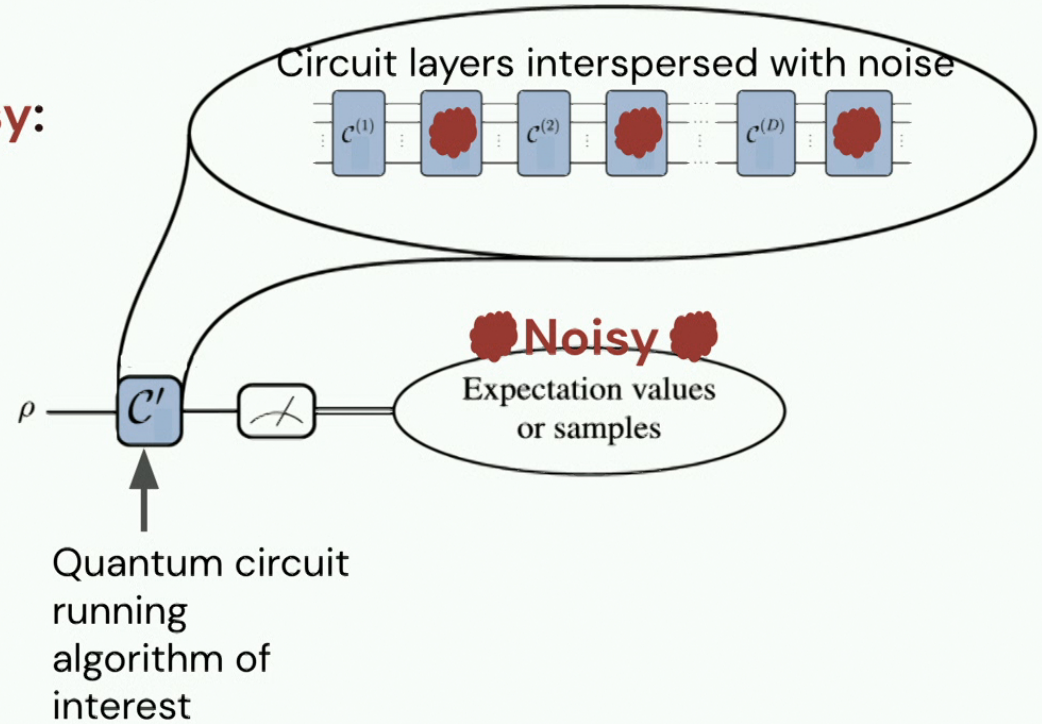
Quantum circuit running algorithm of interest

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

What is error mitigation?

In the real world, C is noisy:

- qubit decoherence
- gate errors

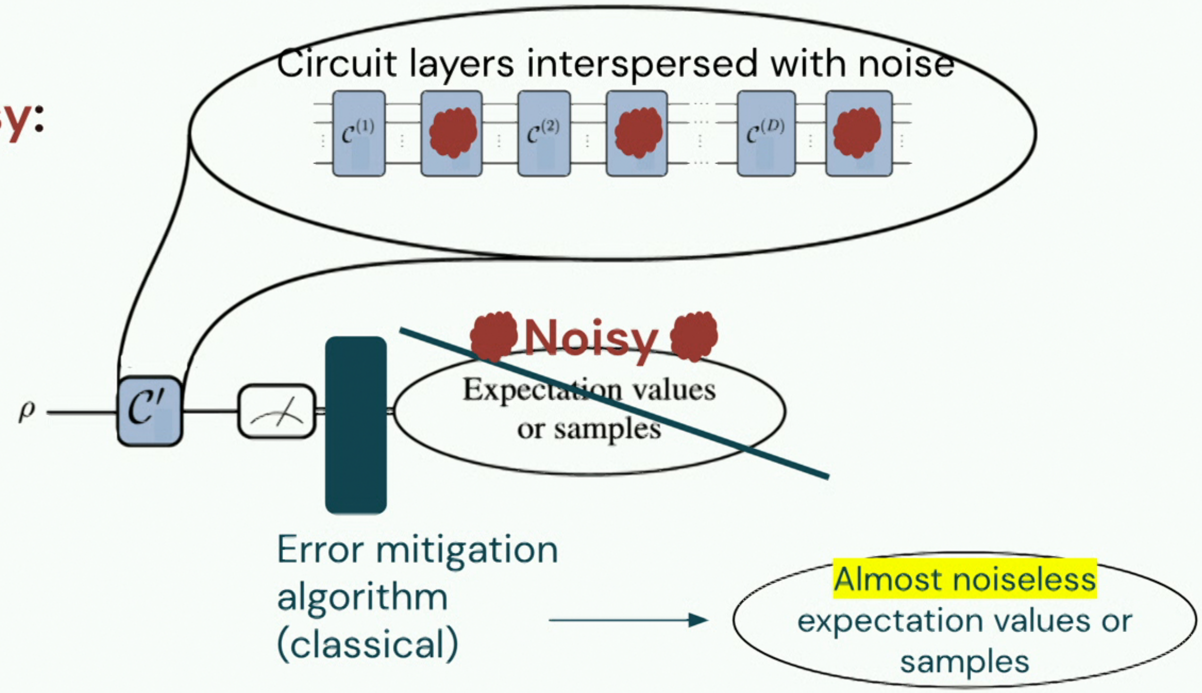


Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

What is error mitigation?

In the real world, C is noisy:

Proposal: revert the effect of noise on the computation result, with classical post-processing.



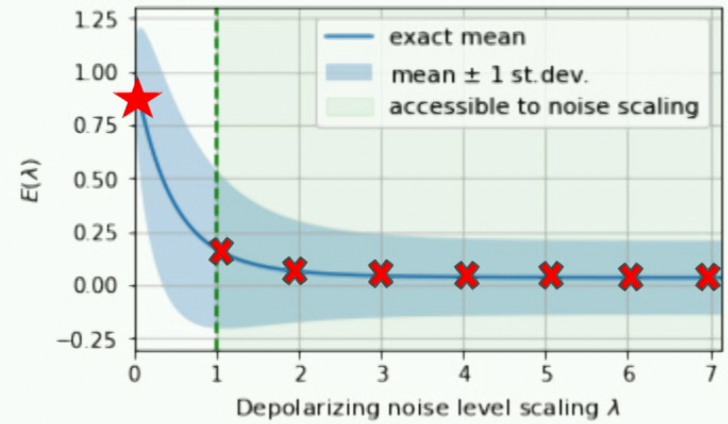
Example of error mitigation protocol

Zero-noise extrapolation:

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

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

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)

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Is removing quantum noise *classically* even possible?

learned noise model and its inversion. We can therefore conveniently express the circuit sampling overhead as a quantum runtime, J:

$$J = \gamma^{-nd} \beta d$$

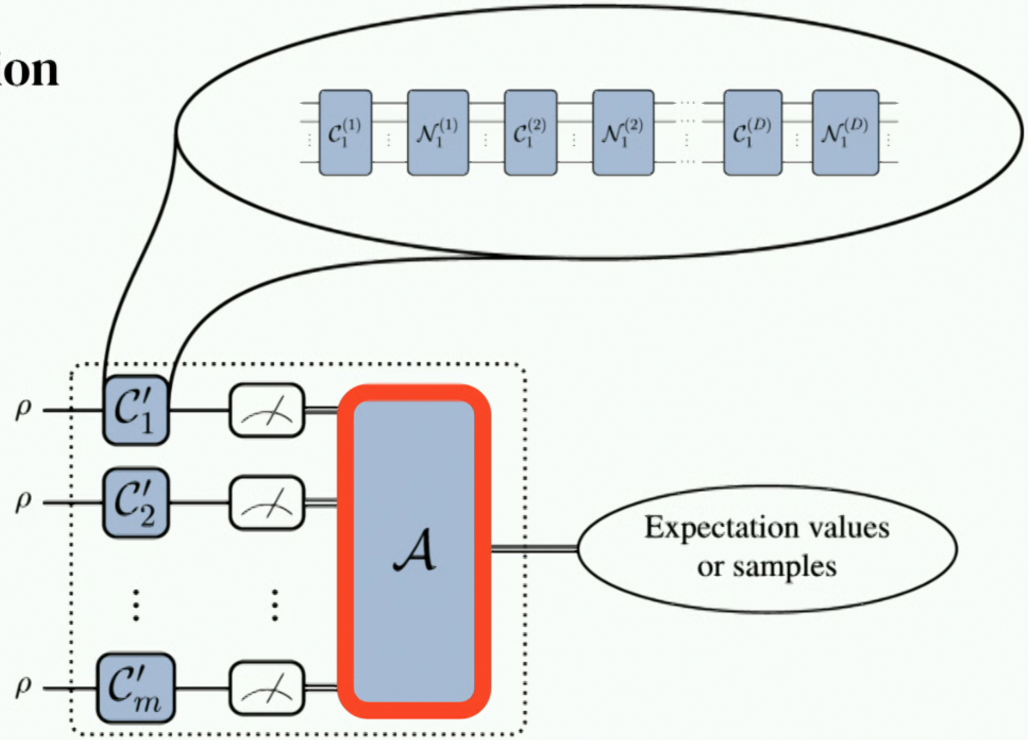
From https://research.ibm.com/blog/gammabar-for-quantum-advantage

(On *Probabilistic Error Cancellation*)

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

A closer look at our setting

$\mathcal{A}$ = error mitigation algorithm



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

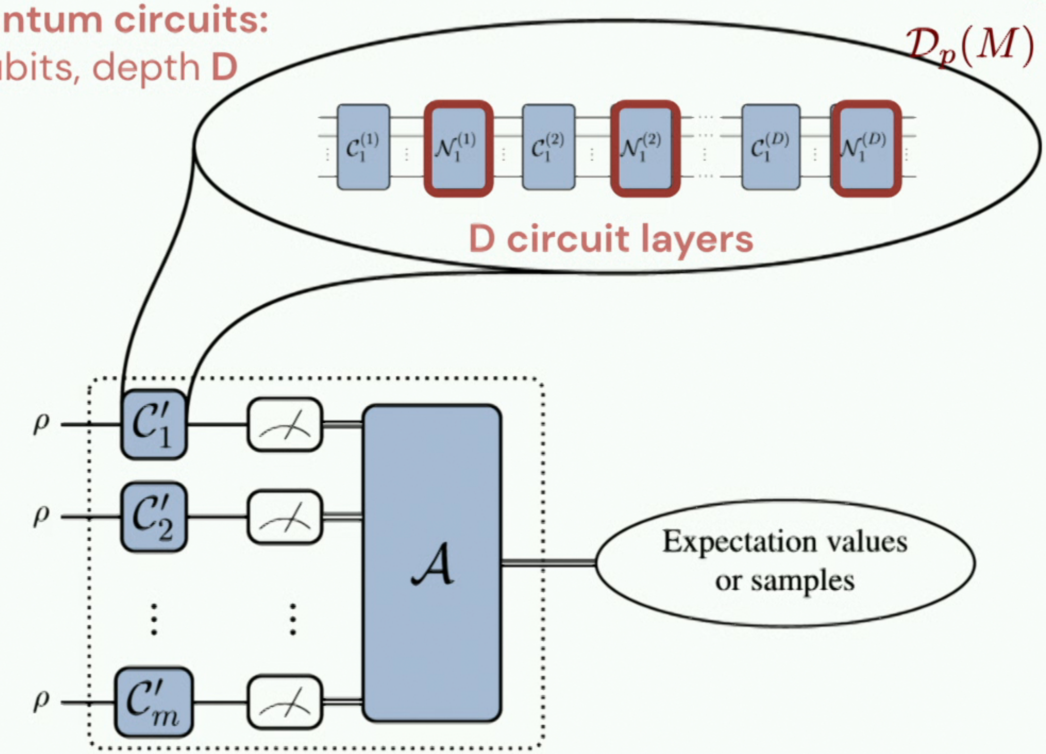
Our setting: input to $\mathcal{A}$

Noise model:
local depolarizing noise of strength **p** after each layer

$$\mathcal{N}_i^{(j)}(\cdot) = \mathcal{D}_p^{\otimes n}(\cdot)$$

$$\mathcal{D}_p(M) := pM + (1-p)\mathrm{Tr}[M]\frac{\mathbb{I}}{2},$$

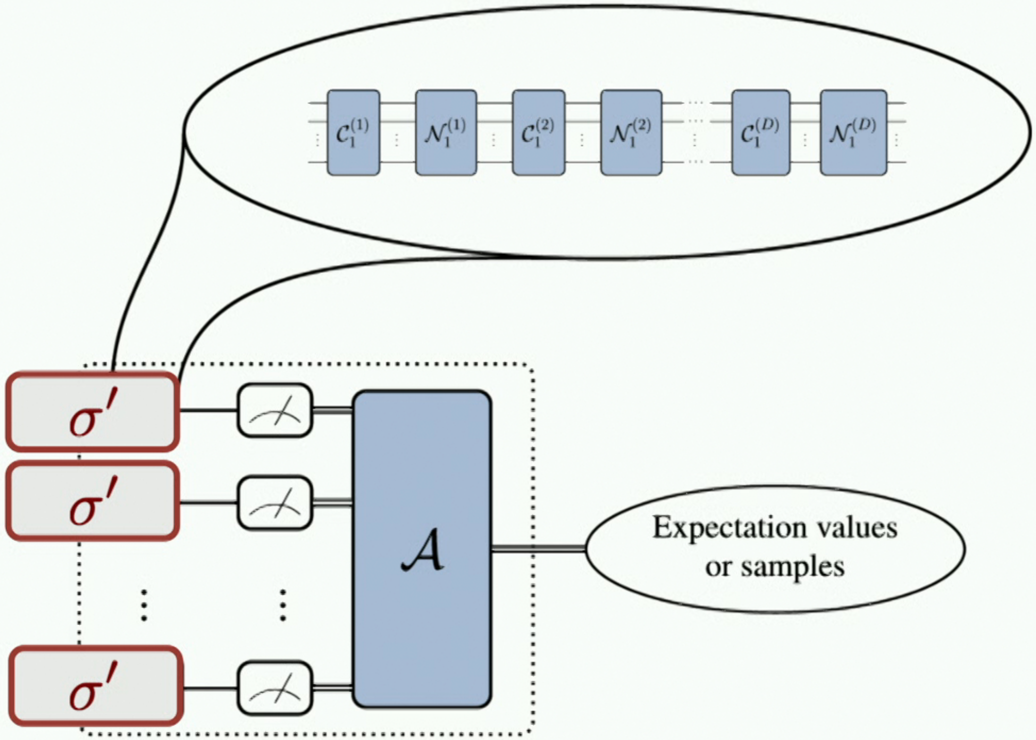
Quantum circuits:
n qubits, depth **D**



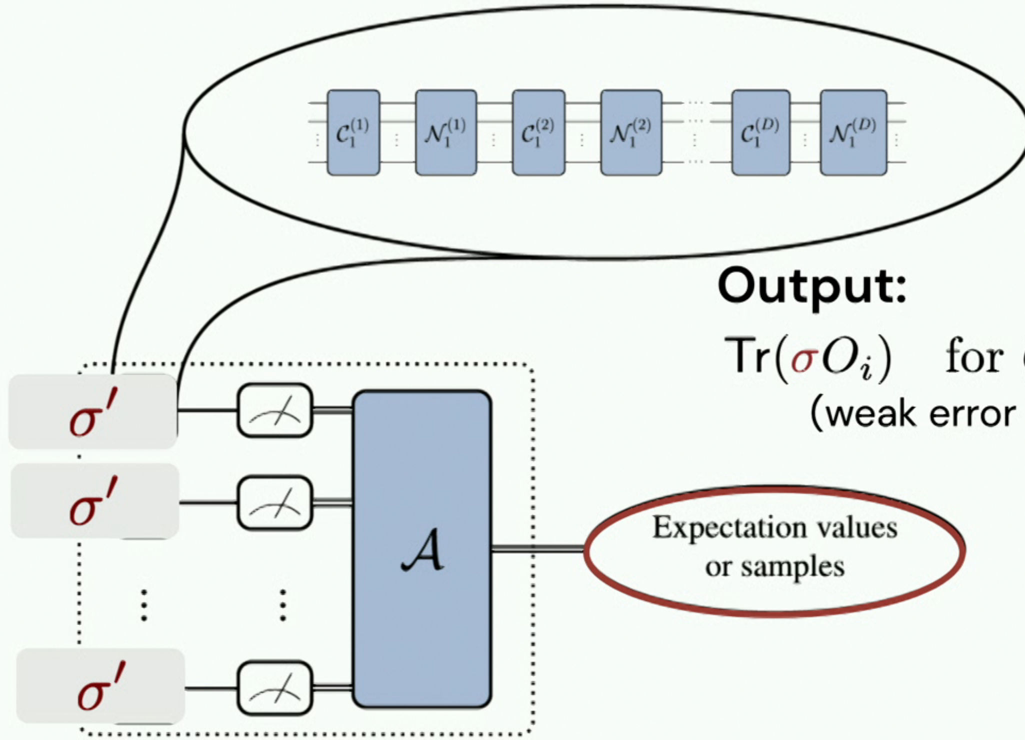
Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Our setting: input to $\mathcal{A}$

Input:
Copies of σ' output by circuits with depolarizing noise of strength p



Our setting: output of $\mathcal{A}$



Output:

Tr(σO_i) for $O_i \in \mathcal{M}$

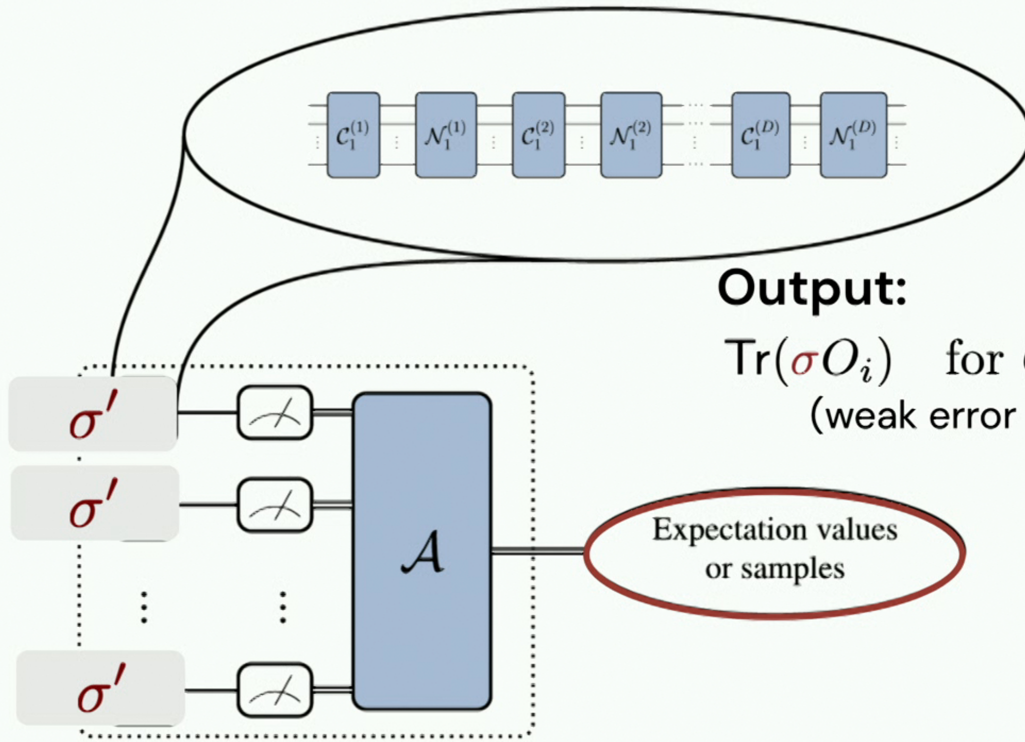
(weak error mitigation → e.g. VQE)

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Our setting: output of $\mathcal{A}$

σ = noise*less* circuit output state

σ' = noisy circuit output state



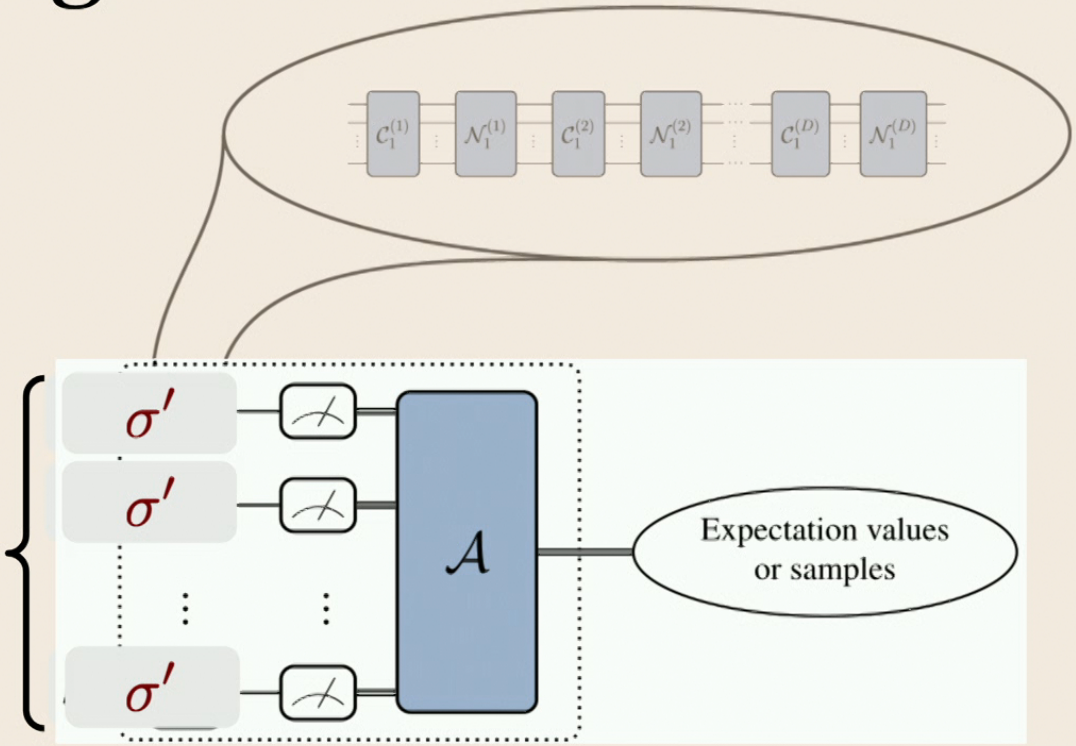
Output:

$\mathrm{Tr}(\sigma O_i)$ for $O_i \in \mathcal{M}$

(weak error mitigation → e.g. VQE)

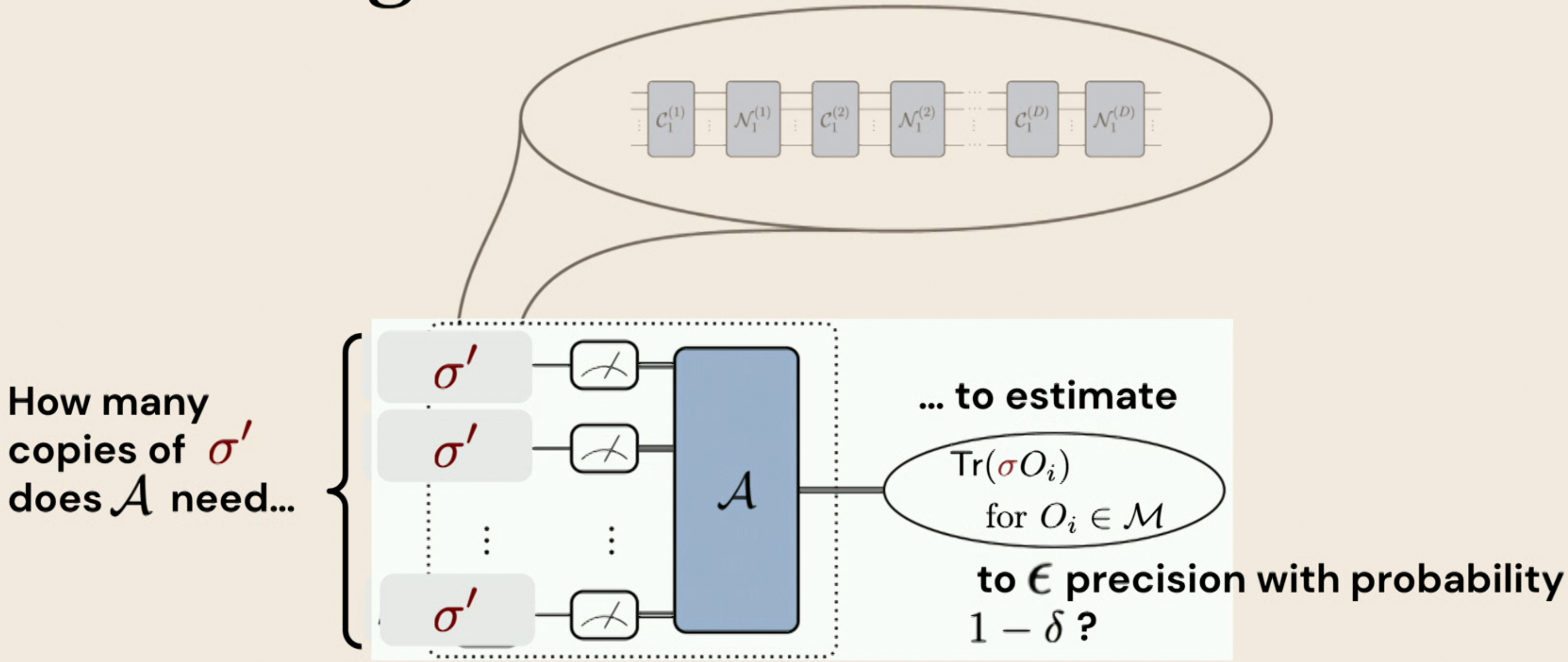
Our question: sample complexity of error mitigation?

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



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

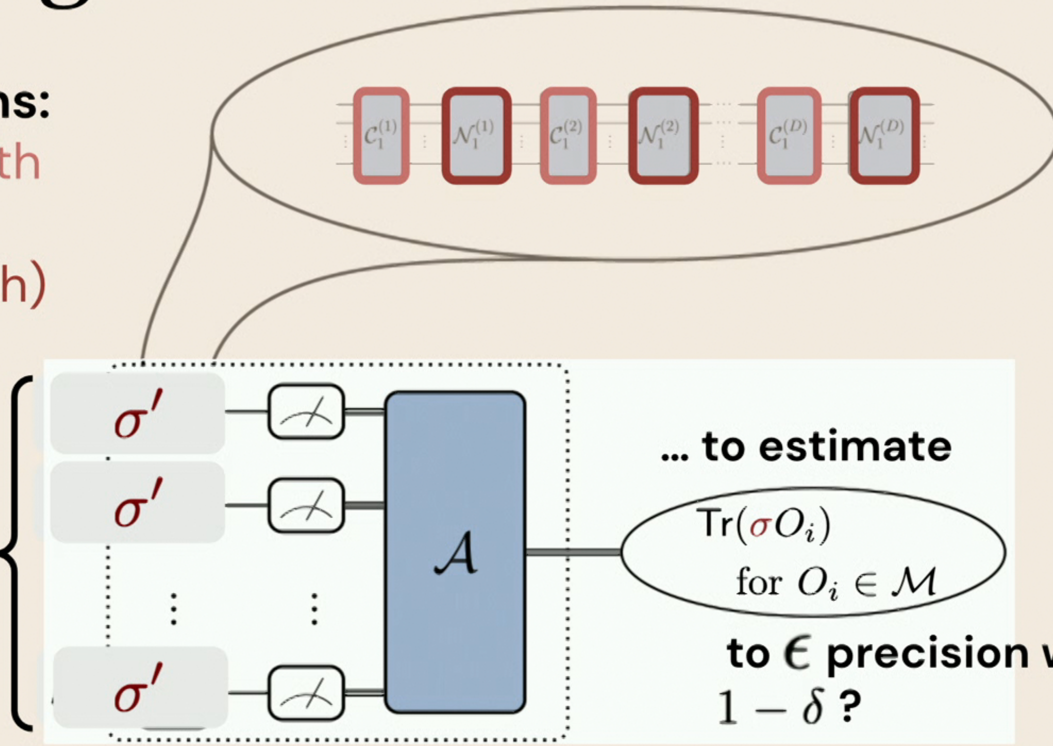
Our question: sample complexity of error mitigation?



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

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...

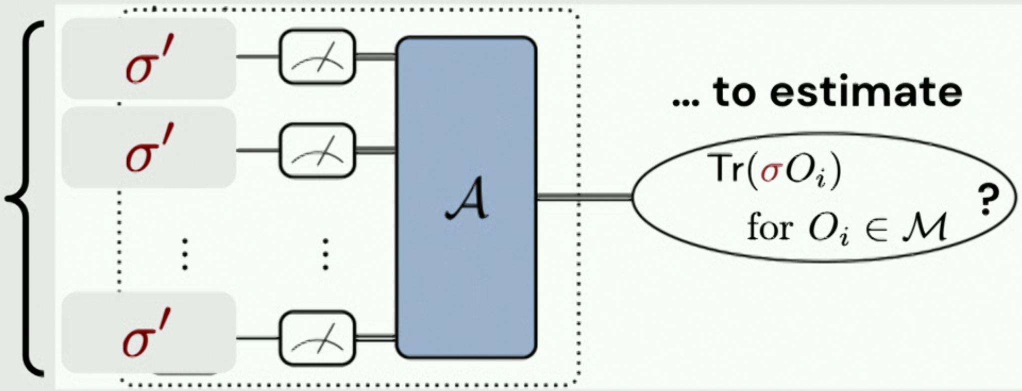
... to estimate

to ϵ precision with probability $1 - \delta$?

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Our lower bounds

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



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

Thm 1: $p^{-\Omega(n\,D)}$ **for mitigating depolarizing noise for** $D = \Omega(\log\log(n))$.

Thm 2: $c^{-\Omega(n\,D)}$ **for mitigating non-unital noise.**

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

How to interpret our results

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

- Previous belief: $\exp(\Omega(D))$ copies required.
- But NISQ circuits are depth $D = O(\log(n))$: our result is exponentially stronger.
- Loss of quantum advantage for certain QML subroutines may occur earlier than expected in the presence of noise
- No noisy circuits for ground state preparation

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Proof technique

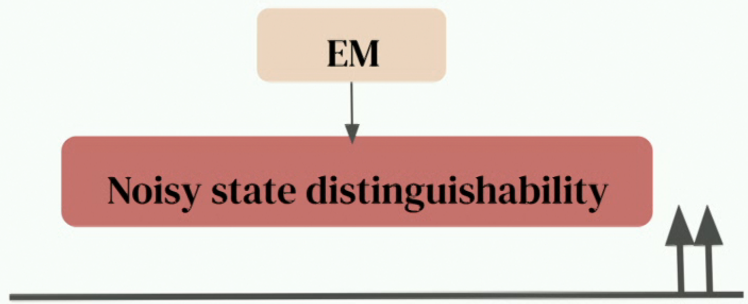
Question: How many copies of σ' are needed for EM?

Our strategy:

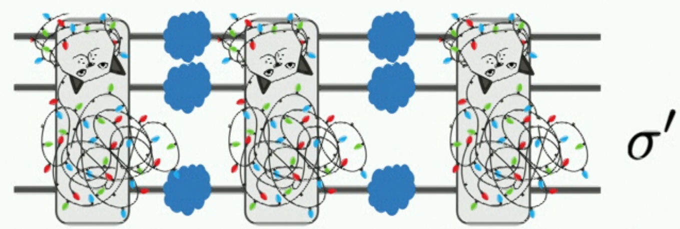
1) Define a learning problem that can be solved by error mitigation

2) Get an expression to lower bound the sample complexity

3) Construct circuits for which the bound is enormous



Critical quantity: $D(\sigma' \| \mathbb{I}/2^n)$

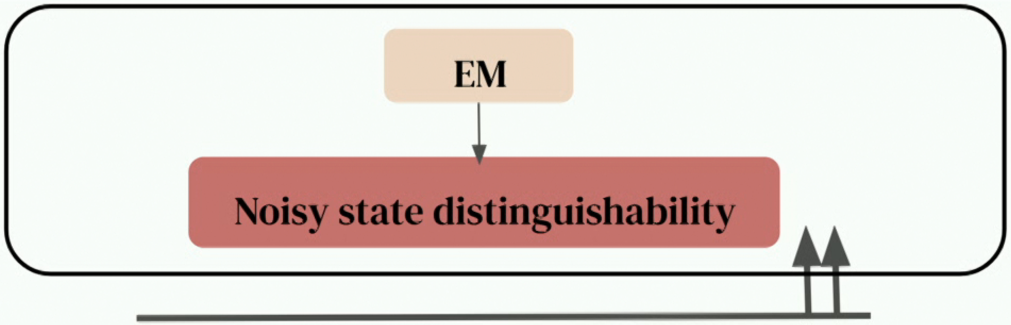


Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

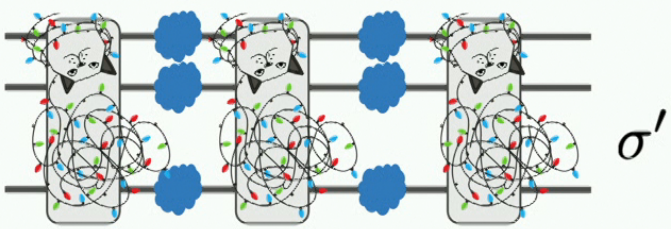
Proof technique

Question: How many copies of σ' are needed for EM?

Pt 1



Critical quantity: $D(\sigma' \| \mathbb{I}/2^n)$



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 1: Define a task solvable by EM

Problem: Noisy state distinguishability

Let S be a **known set** of states.

Given:

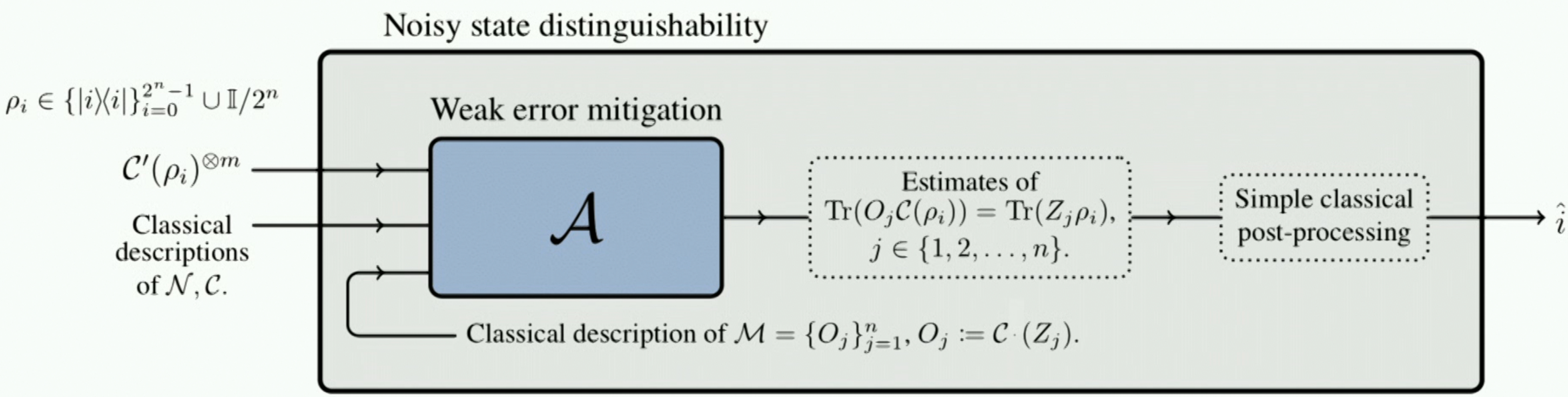
1) m copies of a **noisy** *state* $\mathcal{C}'(\sigma)$ where $\sigma \in S$

2) Knowledge of a circuit C and noise channel transforming C → C′

Output: σ

Pt 1: Define a task solvable by EM

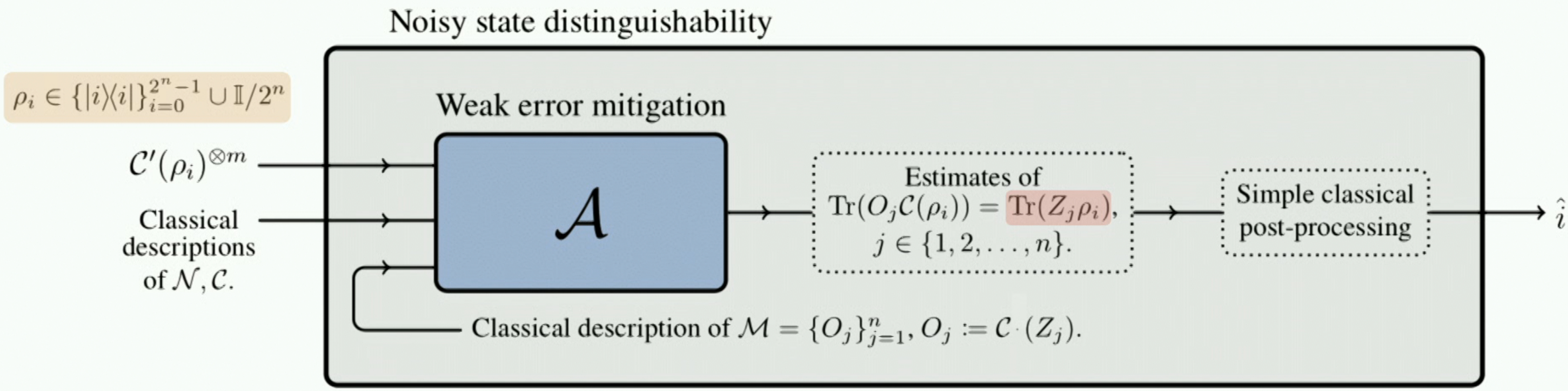
If you can error mitigate, you can distinguish (certain) noisy states



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 1: Define a task solvable by EM

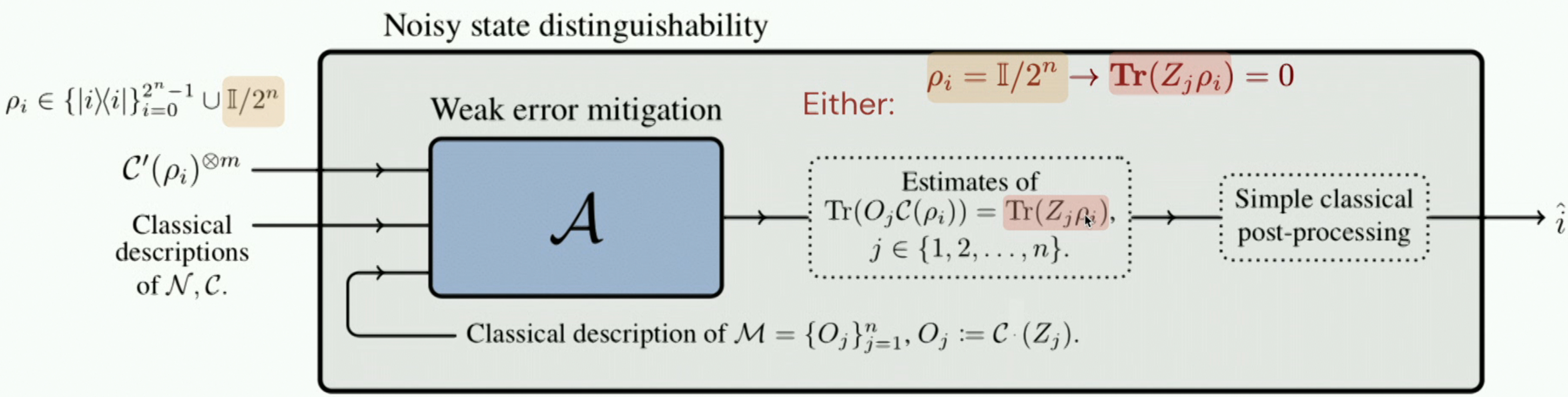
If you can error mitigate, you can distinguish (certain) noisy states



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 1: Define a task solvable by EM

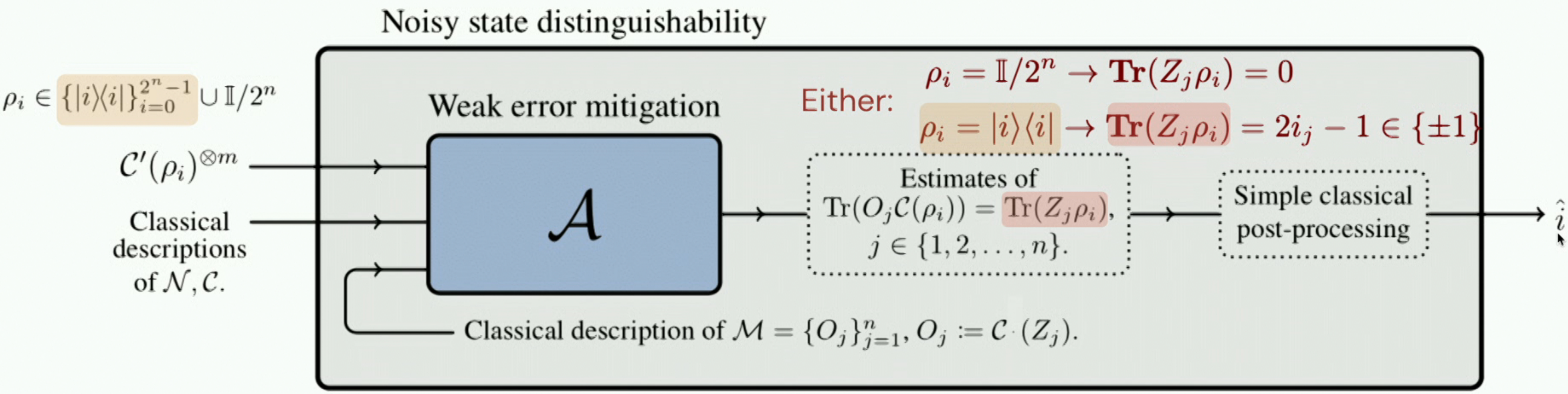
If you can error mitigate, you can distinguish (certain) noisy states



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 1: Define a task solvable by EM

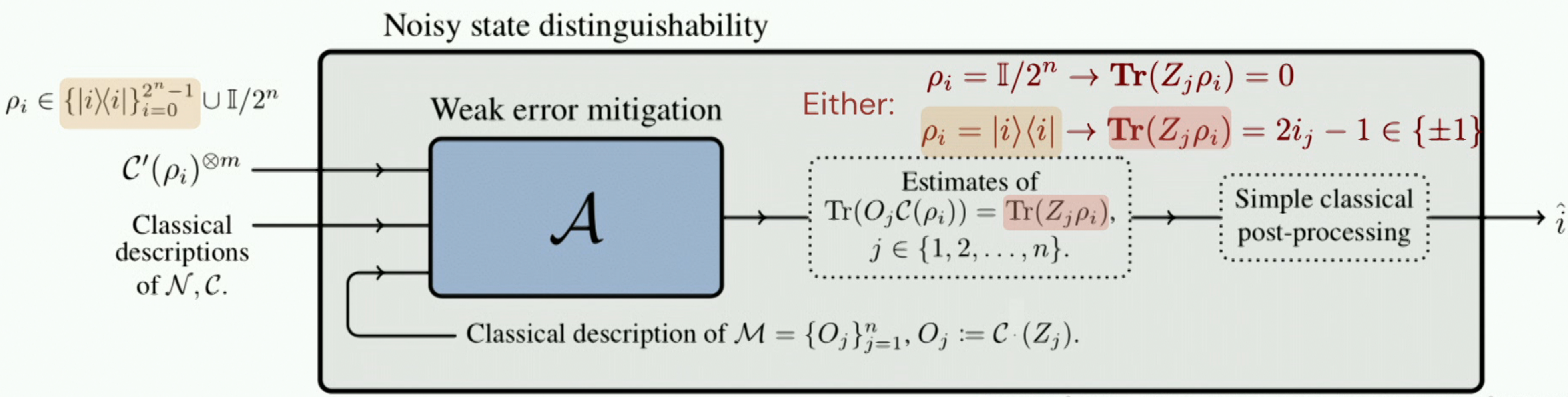
If you can error mitigate, you can distinguish (certain) noisy states



Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 1: Define a task solvable by EM

If you can error mitigate, you can distinguish (certain) noisy states



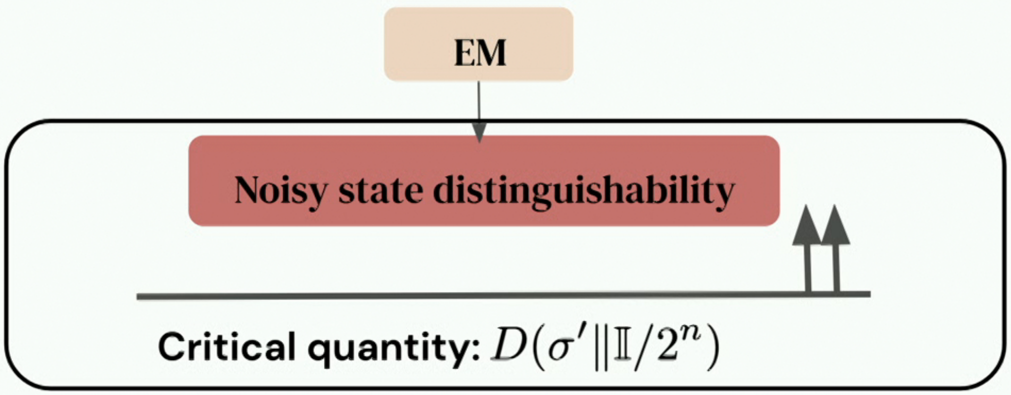
Based on EM's outputs, you can always infer the input state to the circuit!

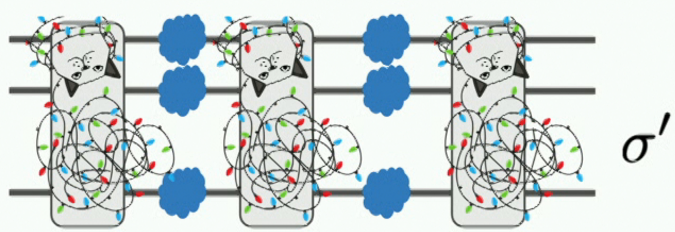
Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Proof technique

Question: How many copies of σ' are needed for EM?

Pt 2





Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 2: Lower bounding the sample complexity of distinguishing noisy states

Fano's lemma: Any single-sample test to distinguish N+1 possible probability distributions P_1, P_2...P_N must fail with probability at least $1-\alpha$, where

$$\frac{1}{\log(N)} \frac{1}{N+1} \sum_{k=0}^{N} D(P_k \| P_N) \leq \alpha$$

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 2: Lower bounding the sample complexity of distinguishing noisy states

Fano's lemma: Any single-sample test to distinguish N+1 possible probability distributions P_1, P_2...P_N must fail with probability at least $1-\alpha$, where

$$\frac{1}{\log(N)}\frac{1}{N+1}\sum_{k=0}^{N} D(P_k \| P_N) \leq \alpha$$

P_i: output of noisy state distinguisher on input $\mathcal{C}'(\rho_i) = \mathcal{C}'(|i\rangle\langle i|)$
P_N: output of noisy state distinguisher on input $\mathcal{C}'(\mathbb{I}/2^n)$

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 2: Lower bounding the sample complexity of distinguishing noisy states

Fano's lemma: Any single-sample test to distinguish N+1 possible probability distributions P_1, P_2...P_N must fail with probability at least $1-\alpha$, where

$$\frac{1}{\log(N)}\frac{1}{N+1}\sum_{k=0}^{N} D(P_k \| P_N) \leq \alpha$$

P_i: output of noisy state distinguisher on input $\mathcal{C}'(\rho_i) = \mathcal{C}'(|i\rangle\langle i|)$

P_N: output of noisy state distinguisher on input $\mathcal{C}'(\mathbb{I}/2^n)$

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 2: Lower bounding the sample complexity of distinguishing noisy states

Reminder: m = # copies of noisy state

Fano's lemma: Any single-sample test to distinguish N+1 possible probability distributions P_1, P_2 ...P_N must fail with probability at least $1-\alpha$, where

$$\frac{1}{\log(N)}\frac{1}{N+1}\sum_{k=0}^{N} D(P_k \| P_N) \leq \alpha = c^{Dn}\, m$$

P_i: output of noisy state distinguisher on input $\mathcal{C}'(\rho_i) = \mathcal{C}'(|i\rangle\langle i|)$

P_N: output of noisy state distinguisher on input $\mathcal{C}'(\mathbb{I}/2^n)$

Yields the claim:

$m = \Omega(c^{-Dn})$

Pt 2: Lower bounding the sample complexity of distinguishing noisy states

"My algorithm is only good if its output is sufficiently different from that on the maximally-mixed state" – a tale as old as time [ABIN'96, CCHL'21, DNS+'21, CBLLS'22...]

Fano's lemma: Any single-sample test to distinguish N+1 possible probability distributions P_1, P_2...P_N must fail with probability at least $1-\alpha$, where

$$\frac{1}{\log(N)}\frac{1}{N+1}\sum_{k=0}^{N} D(P_k \| P_N) \leq \alpha = c^{Dn} m$$

P_i: output of noisy state distinguisher on input $\mathcal{C}'(\rho_i) = \mathcal{C}'(|i\rangle\langle i|)$
P_N: output of noisy state distinguisher on input $\mathcal{C}'(\mathbb{I}/2^n)$

But somewhat unusually, our bounds are exponential in *both* n and D.

Proof technique

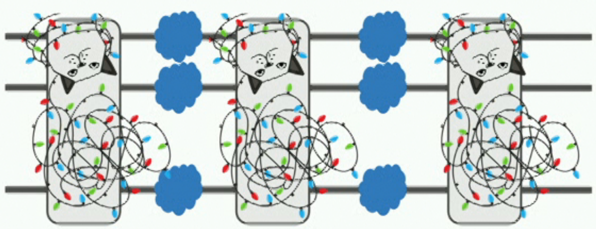
Question: How many copies of σ' are needed for EM?

EM

Noisy state distinguishability

Pt 3

Critical quantity: $D(\sigma' \| \mathbb{I}/2^n)$



σ'

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 3: constructing C' that minimizes the relative entropy

Task: construct $\mathcal{C}'$ that makes $D(\mathcal{C}'(|i\rangle\langle i|)\|\mathcal{C}'(\mathbb{I}/2^n))$ as small as possible.

[CLLW '16]: a Clifford circuit construction that achieves an exact 2-design at depth $\log^2(n)$.

2-design = ensemble of circuits that "looks" Haar random to "simple" observables

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Pt 3: constructing C' that minimizes the relative entropy

Task: construct $\mathcal{C}'$ that makes $D(\mathcal{C}'(|i\rangle\langle i|)\|\mathcal{C}'(\mathbb{I}/2^n))$ as small as possible.

Intuition 1: it suffices to make the *purity* $\mathrm{Tr}(\mathcal{C}'(|i\rangle\langle i|)^2)$ small

Intuition 2:
random quantum circuits shift `purity contribution' to higher-weight Paulis

+

depolarizing noise acts exponentially in Pauli weight

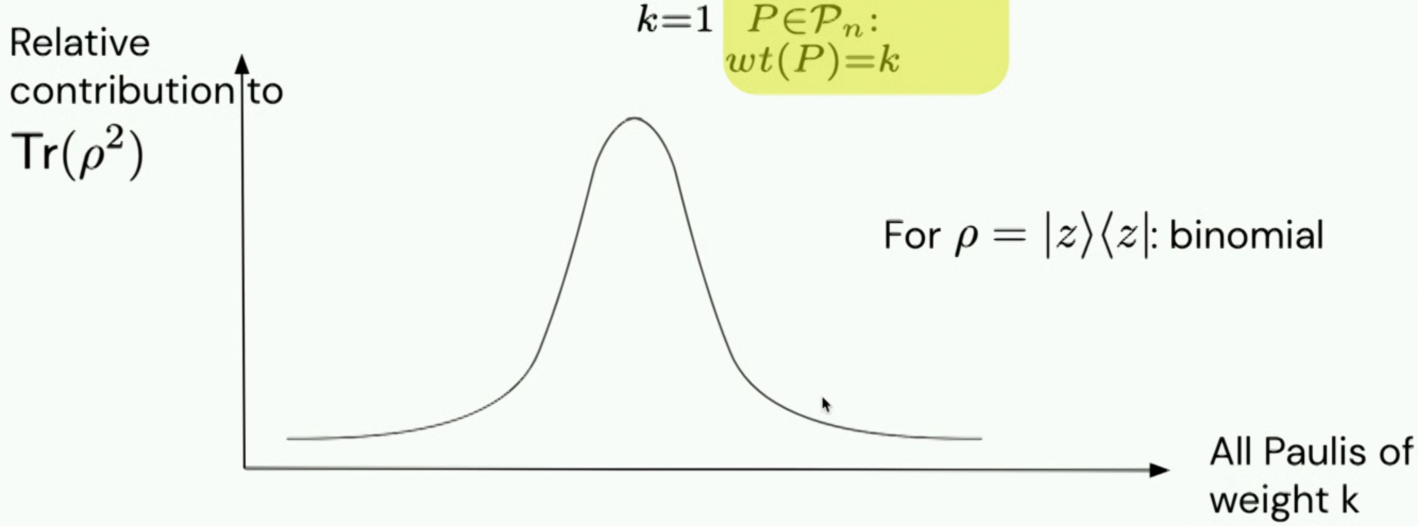
Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Why do our circuits make $\mathrm{Tr}(\mathcal{C}'(|i\rangle\langle i|)^2)$ decay rapidly?

Purity breakdown: For a given state, see how much Pauli strings of different weights contribute to the purity.

$$\mathrm{Tr}(\rho^2) = \sum_{k=1}^{n} \sum_{\substack{P \in \mathcal{P}_n: \\ wt(P)=k}} \alpha_P$$

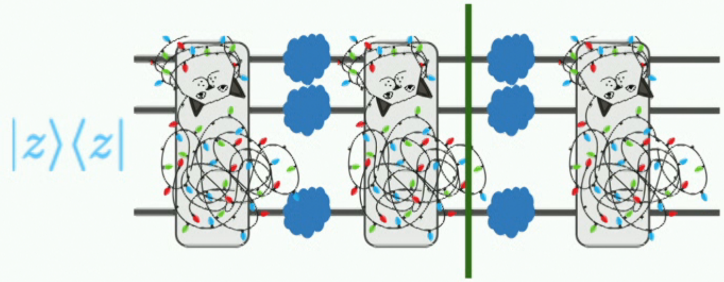
Plot this!

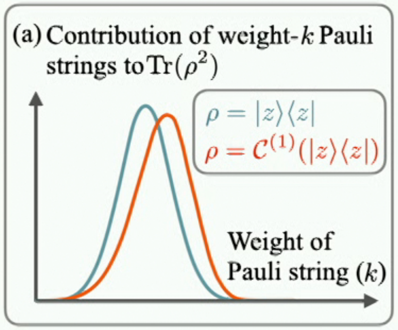


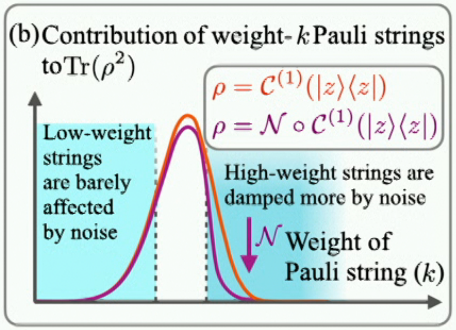
Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

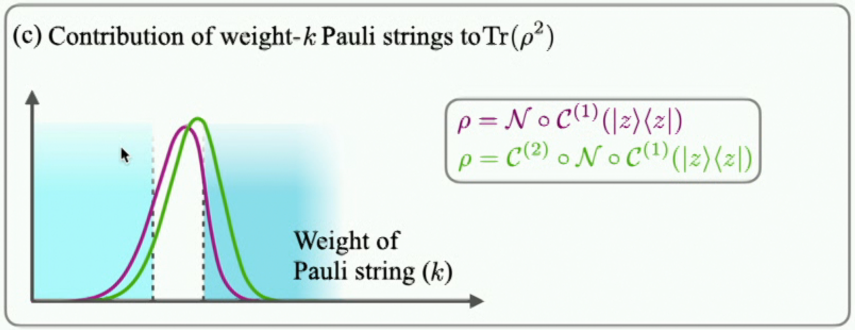
Why do our circuits make $\mathrm{Tr}(\mathcal{C}'(|i\rangle\langle i|)^2)$ decay rapidly?

Let's track $\mathrm{Tr}(\rho^2)$ of a state progressing through the circuit:









Frequently-asked questions

Why doesn't your analysis also kill error correction?

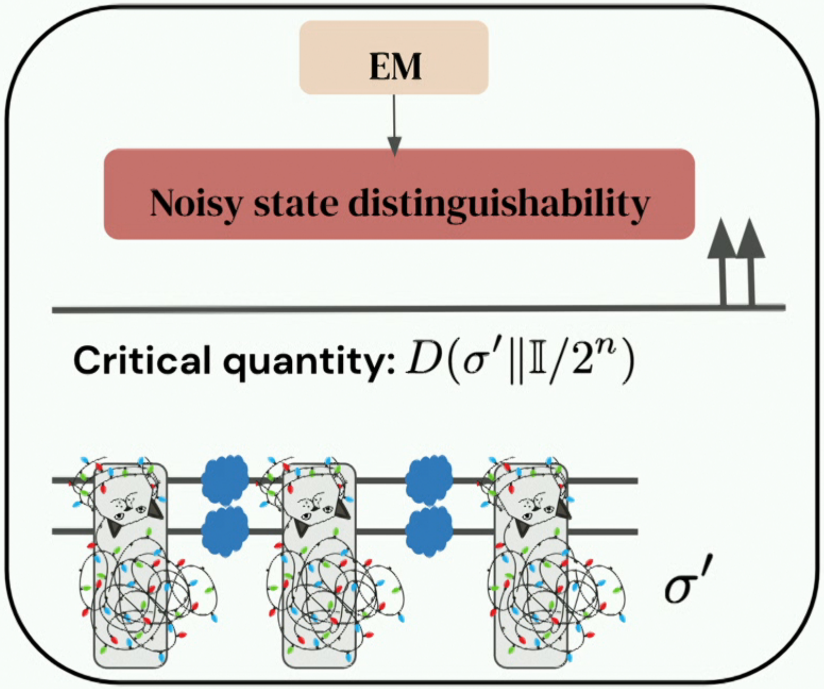
- Because we don't allow for intermediate measurements, adaptive processing and fresh ancillas in our noisy circuits.
- How 'near-term' can we push the line between error mitigation and error correction?

How do we interpret your worst-case bounds?

- You can't apply error mitigation blindly; it doesn't do well on all circuits.
- Do our conclusions hold up for the 'average' circuit?
- What is `pathological' about our circuits (they require all-to-all connectivity, for one)?

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter

Error mitigation is hopeless on circuits that scramble information rapidly.



Open questions:

1. Entanglement generation ↔ noise sensitivity?
2. Average-case (not worst-case)?
3. What's intermediate between error mitigation and error correction?

Yihui Quek | Harvard | EM becomes hard at poly loglog(n) depth (arXiv:2210.11505) | Perimeter