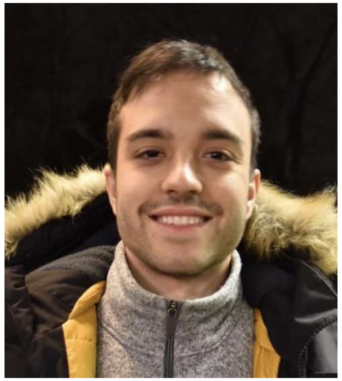
Noise-induced shallow circuits and absence of barren plateaus

Antonio Anna Mele, **Armando Angrisani**, Soumik Ghosh, Sumeet Khatri, Jens Eisert, Daniel Stilck França, Yihui Quek

arXiv:2403.13927

Joint work with:

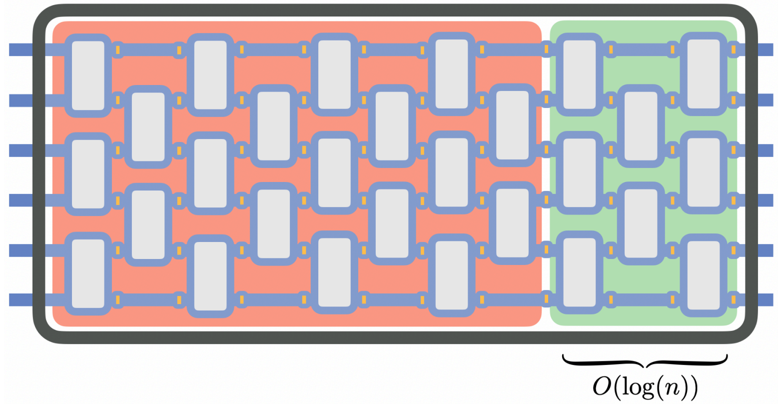
Antonio Anna Mele, Soumik Ghosh, Sumeet Khatri, Jens Eisert, Daniel Stilck França, Yihui Quek






FREIE UNIVERSITÄT BERLIN
SORBONNE UNIVERSITÉ
EPFL
THE UNIVERSITY OF CHICAGO
ENS DE LYON
Massachusetts Institute of Technology

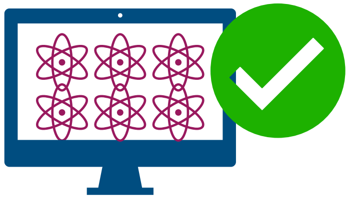
Outline



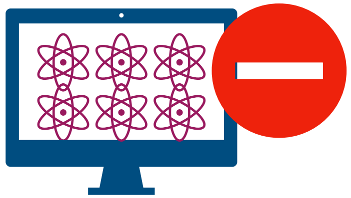
- Noise in quantum circuits
- Effective shallow circuits
- Classical simulation of Pauli expectation values of noisy random circuits
- Barren plateaus

- Understanding **noise impact** in current quantum devices is important for:

Capabilities



Limitations



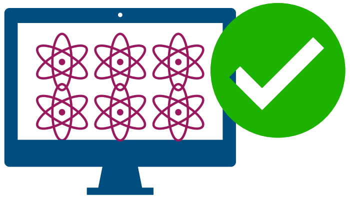
- Understanding **noise impact** in current quantum devices is important for:

- Many previous works modeled noise as solely **depolarizing.**

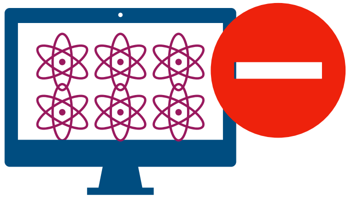
$$\mathcal{N}(\sigma) = (1-p)\sigma + p\frac{I}{2} \quad \text{with } p \in [0,1]$$

- Understanding **noise impact** in current quantum devices is important for:

Capabilities



Limitations

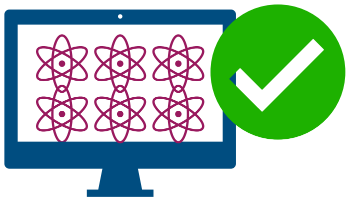


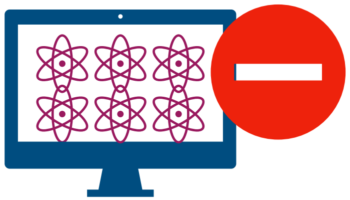
- Many previous works modeled noise as solely **depolarizing.**

$$\mathcal{N}(\sigma) = (1-p)\sigma + p\frac{I}{2} \quad \text{with } p \in [0,1]$$

- **Depolarizing noise**, e.g., induces:

- Understanding **noise impact** in current quantum devices is important for:





- Many previous works modeled noise as solely **depolarizing**.

$$\mathcal{N}(\sigma) = (1-p)\sigma + p\frac{I}{2} \quad \text{with } p \in [0,1]$$

- **Depolarizing noise**, e.g., induces:
    - Barren plateaus in variational quantum algorithms [1],

[1] Noise-induced barren plateaus in variational quantum algorithms. Wang *et al,* Nature Comm. *(2021).*

- Understanding **noise impact** in current quantum devices is important for:

- Many previous works modeled noise as solely **depolarizing**.

$$\mathcal{N}(\sigma) = (1-p)\sigma + p\frac{I}{2} \quad \text{with } p \in [0,1]$$

- **Depolarizing noise**, e.g., induces:
  - Barren plateaus in variational quantum algorithms [1],
  - Efficient classical simulation of **'supremacy' sampling experiments** [2].

[1] Noise-induced barren plateaus in variational quantum algorithms. Wang *et al,* Nature Comm. *(2021).*
[2] A polynomial-time classical algorithm for noisy random circuit sampling, Aharonov *et al*., STOC (2023)

- Understanding **noise impact** in current quantum devices is important for:

- Many previous works modeled noise as solely **depolarizing**.

$$\mathcal{N}(\sigma) = (1-p)\sigma + p\frac{I}{2} \quad \text{with } p \in [0,1]$$

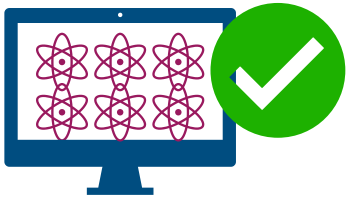
- **Depolarizing noise**, e.g., induces:
  - Barren plateaus in variational quantum algorithms [1],
  - Efficient classical simulation of **'supremacy' sampling experiments** [2].

- But, **tiny departure** from this model could lead to **different conclusions** [3,4].

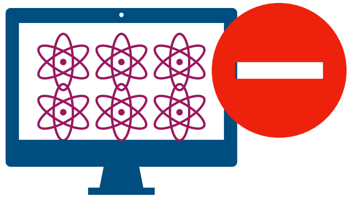
[1] Noise-induced barren plateaus in variational quantum algorithms. Wang *et al,* Nature Comm. *(2021).*
[2] A polynomial-time classical algorithm for noisy random circuit sampling, Aharonov *et al*., STOC (2023)

- Understanding **noise impact** in current quantum devices is important for:

Capabilities



Limitations



- Many previous works modeled noise as solely **depolarizing**.

$$\mathcal{N}(\sigma) = (1-p)\sigma + p\frac{I}{2} \quad \text{with } p \in [0,1]$$

- **Depolarizing noise**, e.g., induces:
  - Barren plateaus in variational quantum algorithms [1],
  - Efficient classical simulation of **'supremacy' sampling experiments** [2].

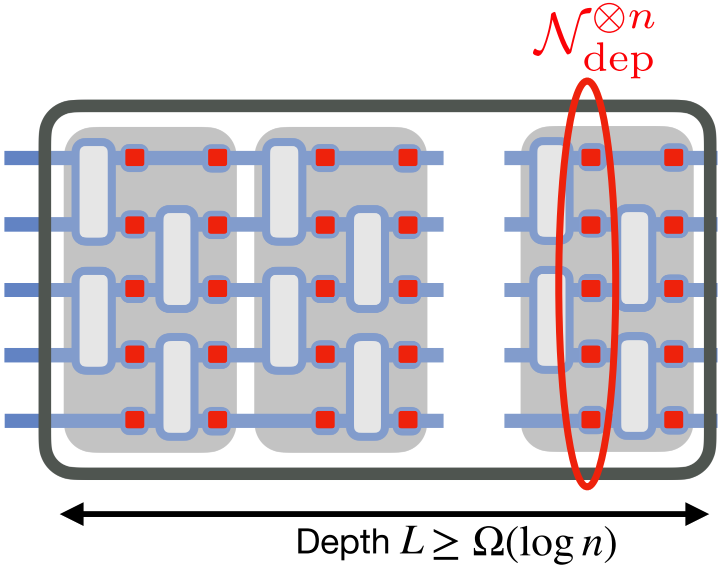
- But, **tiny departure** from this model could lead to **different conclusions** [3,4].

[1] Noise-induced barren plateaus in variational quantum algorithms. Wang *et al,* Nature Comm. *(2021).*
[2] A polynomial-time classical algorithm for noisy random circuit sampling, Aharonov *et al*., STOC (2023)
[3] Quantum refrigerator, Ben-Or *et al*, QIP (2013).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

$$\Phi_L =$$

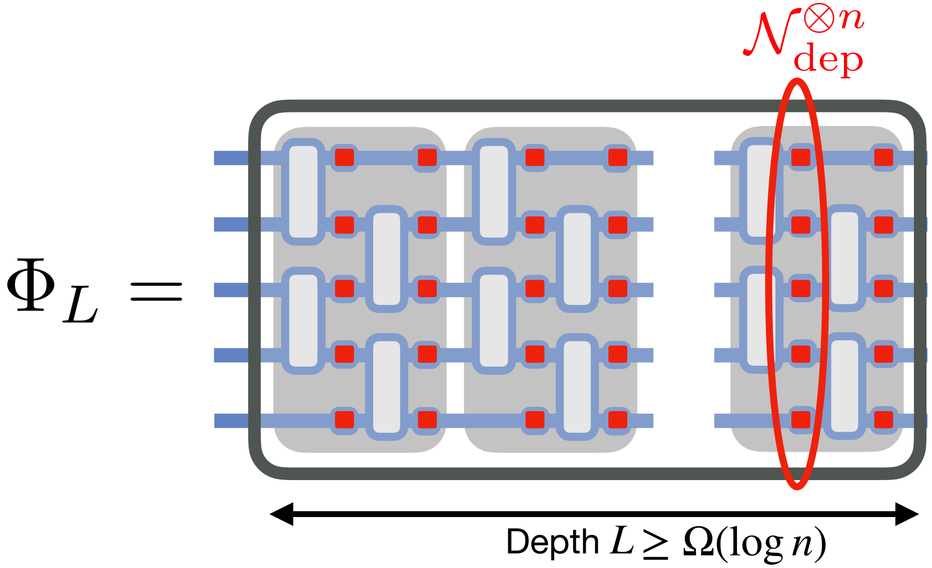


$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)), \quad \text{[5]}$$

Quantum computation possible only for $\log(n)$ depth! [6]

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:



$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)),$$ [5]

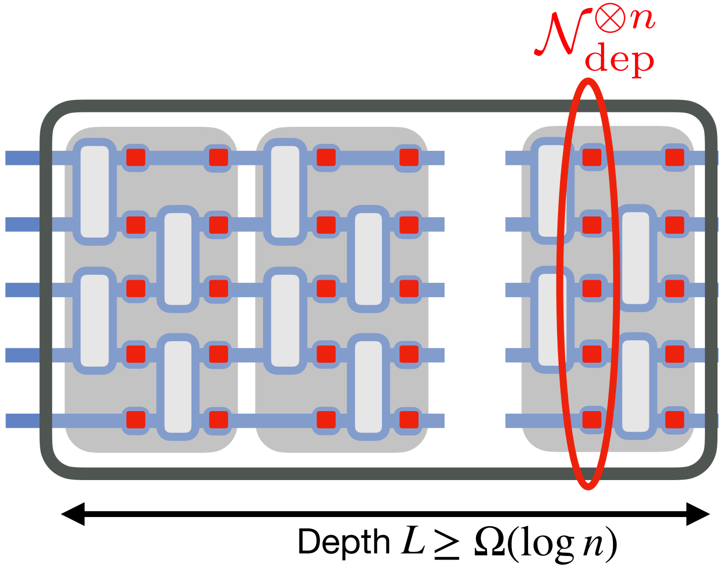
Quantum computation possible only for $\log(n)$ depth! [6]

However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

$\Phi_L =$



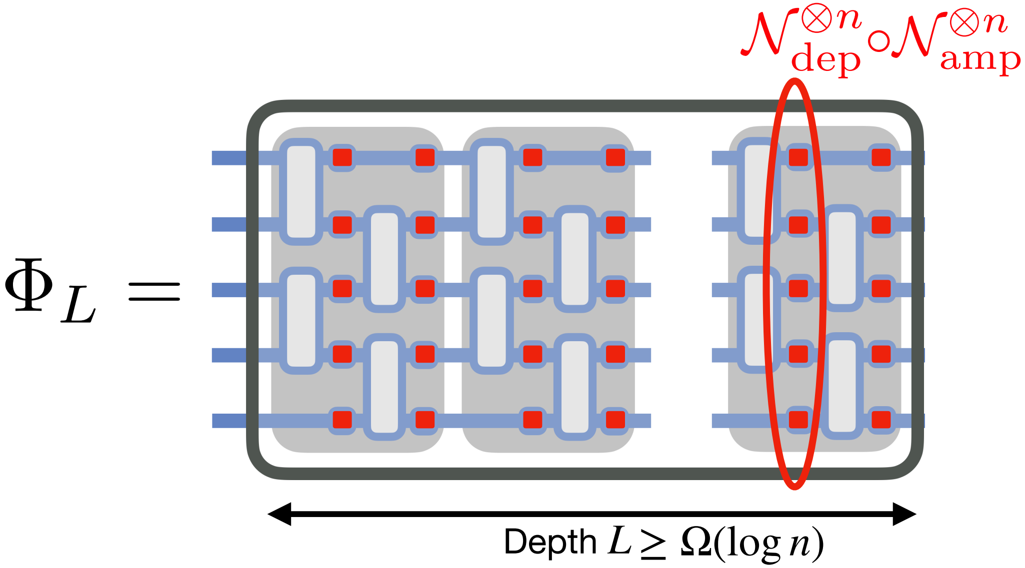
$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)), \text{ [5]}$$

Quantum computation possible only for $\log(n)$ depth! [6]

However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:



$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)),$$ [5]

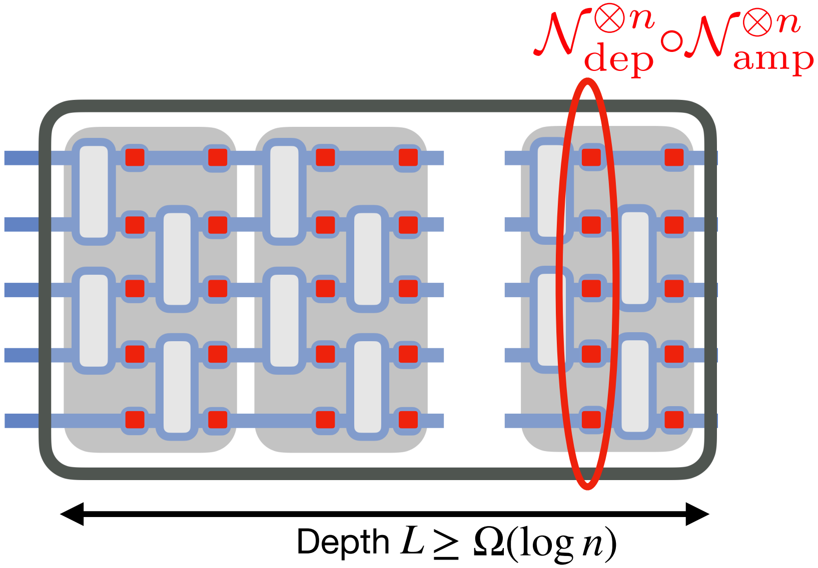
Quantum computation possible only for $\log(n)$ depth! [6]

However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

$$\Phi_L =$$



$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)),$$ [5]

Quantum computation possible only for $\log(n)$ depth! [6]

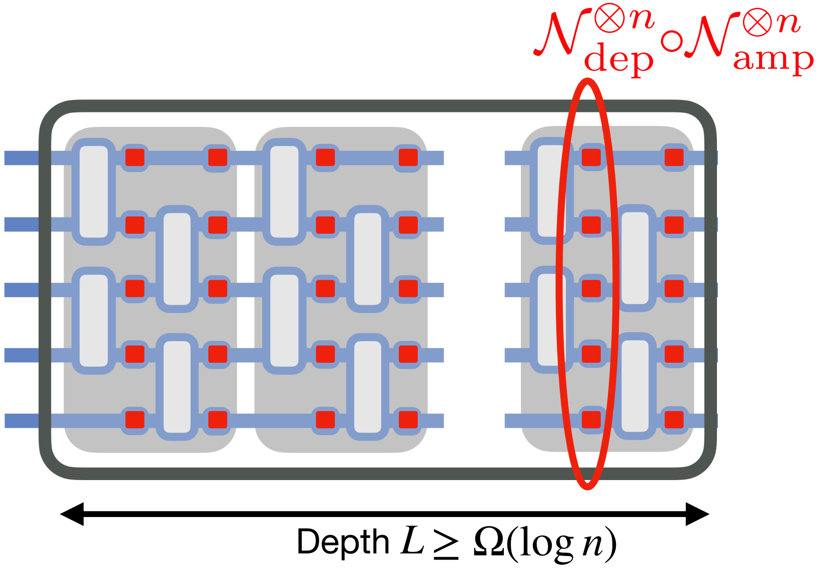
However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

- Real **hardware noise** is **not solely depolarizing** (which is 'unital', i.e., $\mathcal{N}(I) = I$)

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

$\Phi_L =$



$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)),$$ [5]

Quantum computation possible only for $\log(n)$ depth! [6]

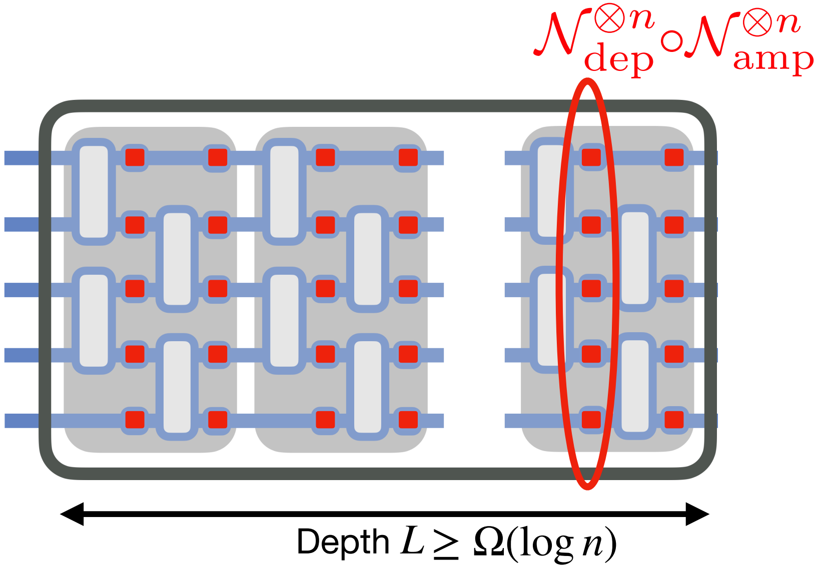
However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

- Real **hardware noise** is **not solely depolarizing** (which is ‘unital’, i.e., $\mathcal{N}(I) = I$)
- Other noise components present (T1-error, amplitude damping, …) are **non-unital**.

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

$\Phi_L =$



$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)),$$ [5]

Quantum computation possible only for $\log(n)$ depth! [6]

However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

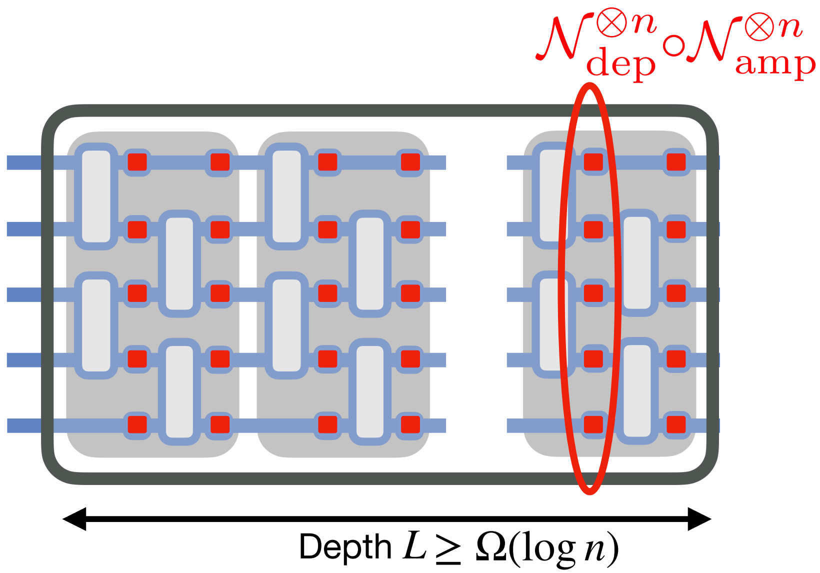
- Real **hardware noise** is **not solely depolarizing** (which is ‘unital’, i.e., $\mathcal{N}(I) = I$)
- Other noise components present (T1-error, amplitude damping, …) are **non-unital**.

[3] Quantum refrigerator, Ben-Or *et al*, QIP (2013).

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

- The **depolarizing** assumption, if true, **enormously constrains noisy computation**:

$\Phi_L =$



$$\left\| \Phi_L(\rho_0) - \frac{\mathbb{I}}{2^n} \right\|_1 \leq \exp(-\Theta(L)),$$ [5]

Quantum computation possible only for $\log(n)$ depth! [6]

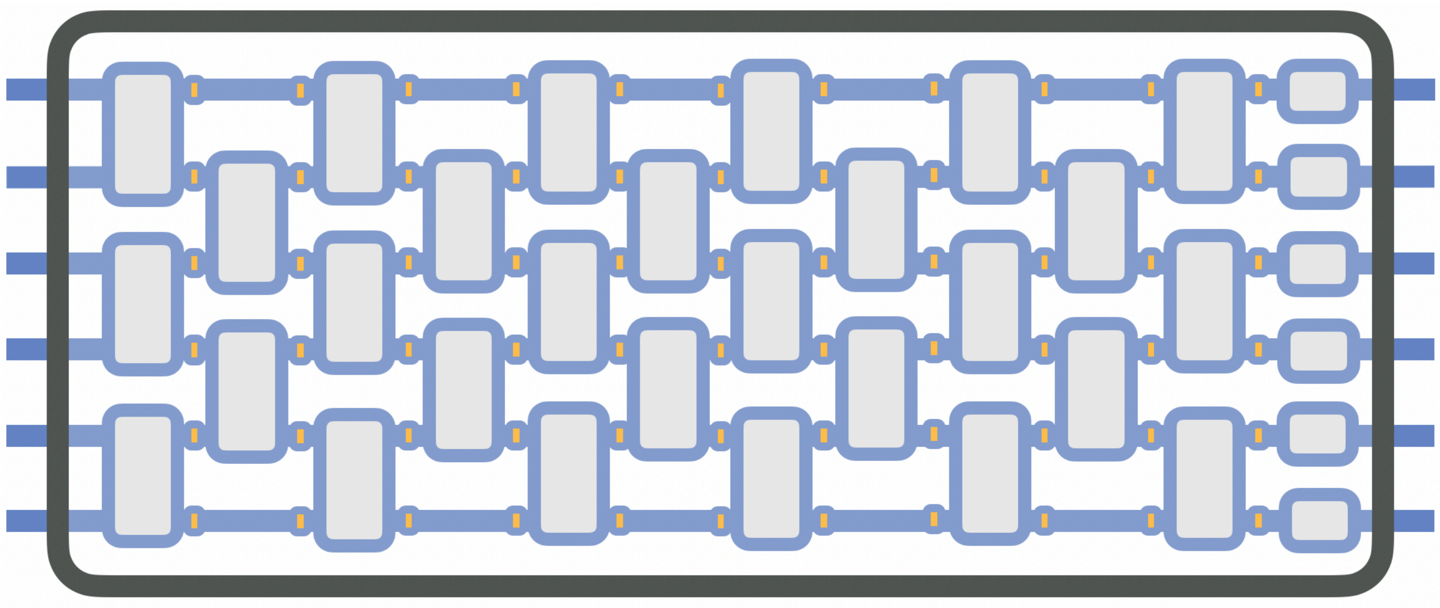
However, for <u>noises different from depolarizing</u>, not all circuits become trivial after log-depth!

- Real **hardware noise** is **not solely depolarizing** (which is 'unital', i.e., $\mathcal{N}(I) = I$)
- Other noise components present (T1-error, amplitude damping, …) are **non-unital**.

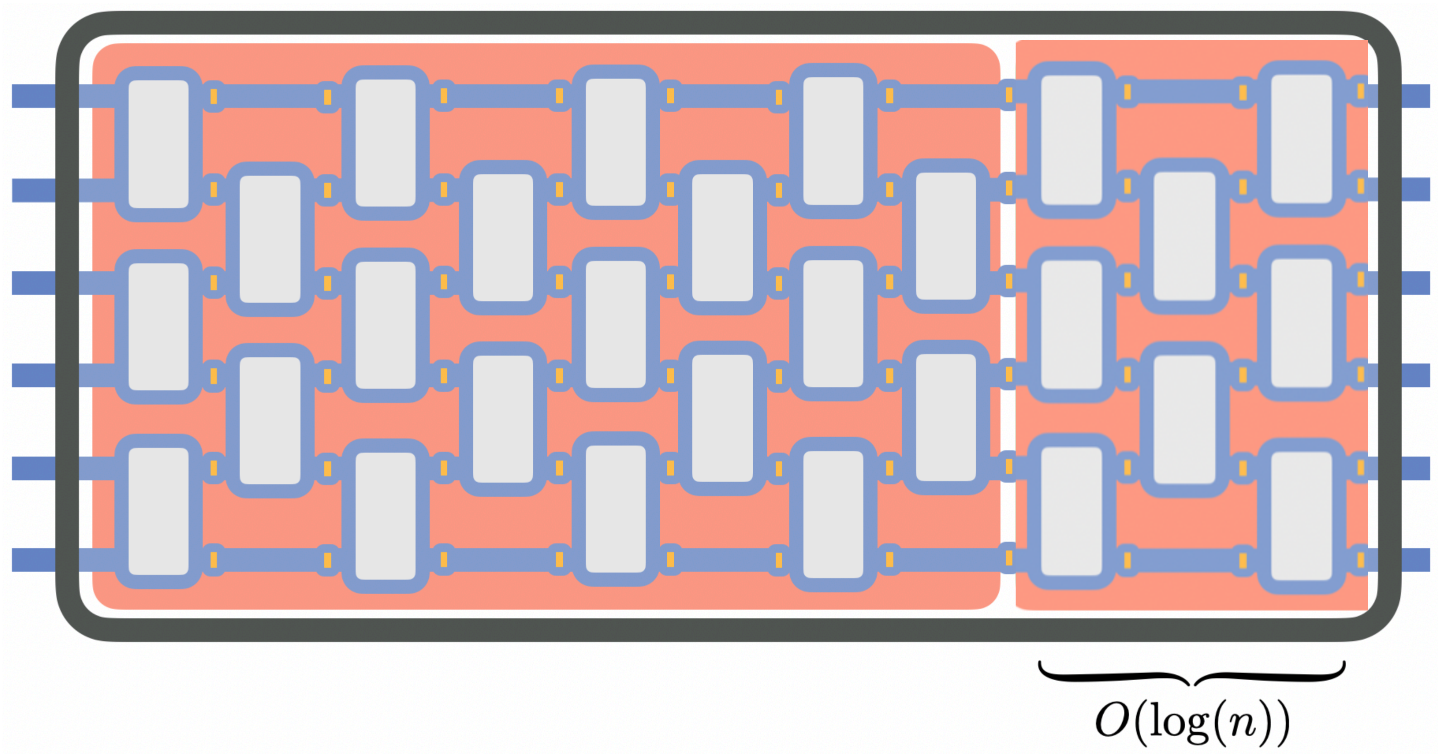
[3] Quantum refrigerator, Ben-Or *et al*, QIP (2013). → **'Non-unital' noisy circuits can be made fault-tolerant, without fresh auxiliary qubits!**

[5] Relative entropy convergence for depolarizing channels, Muller-Hermes *et al,* J. Math. Phys. (2016).
[6] Limitations of noisy reversible computation, Aharonov et al, ArXiv (1996).

What happens for generic noisy circuits?

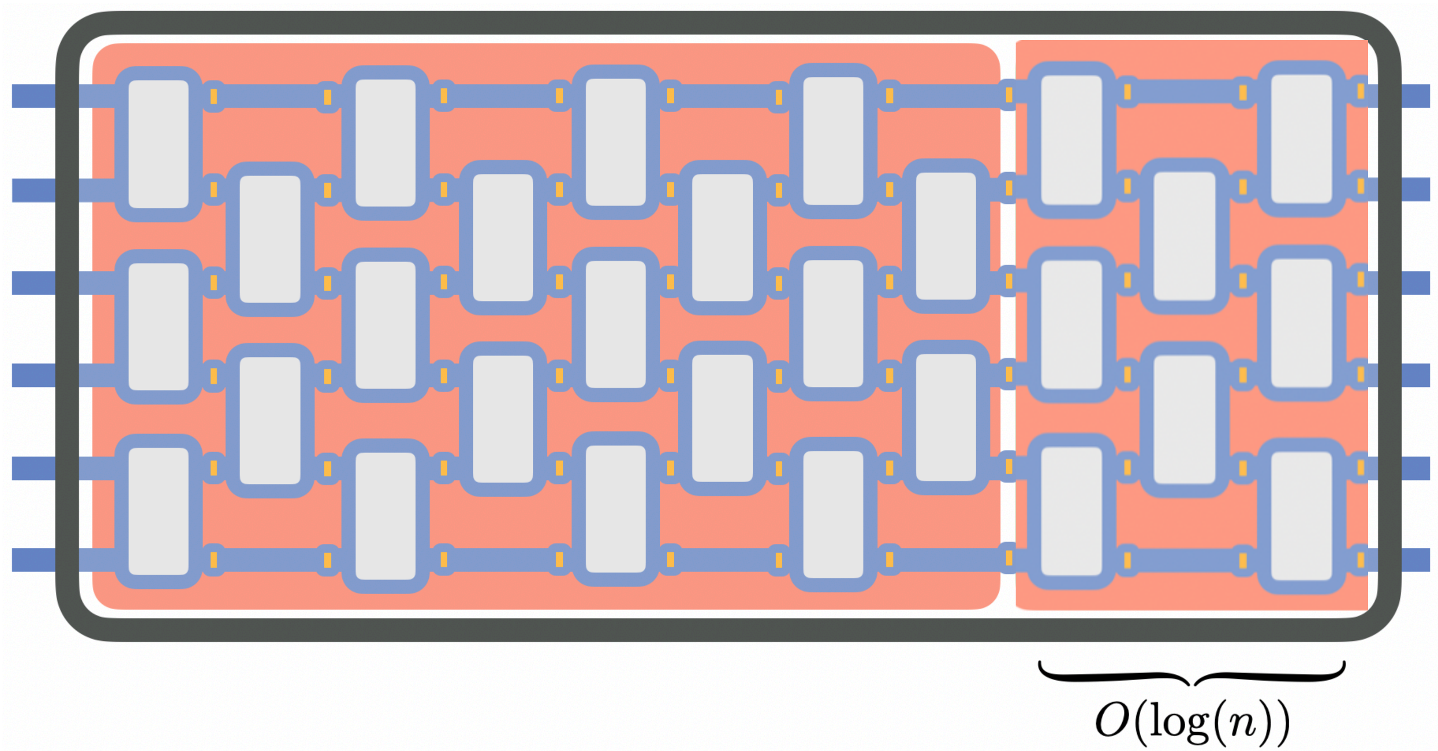


If the noise is **depolarizing**, no layers matter.

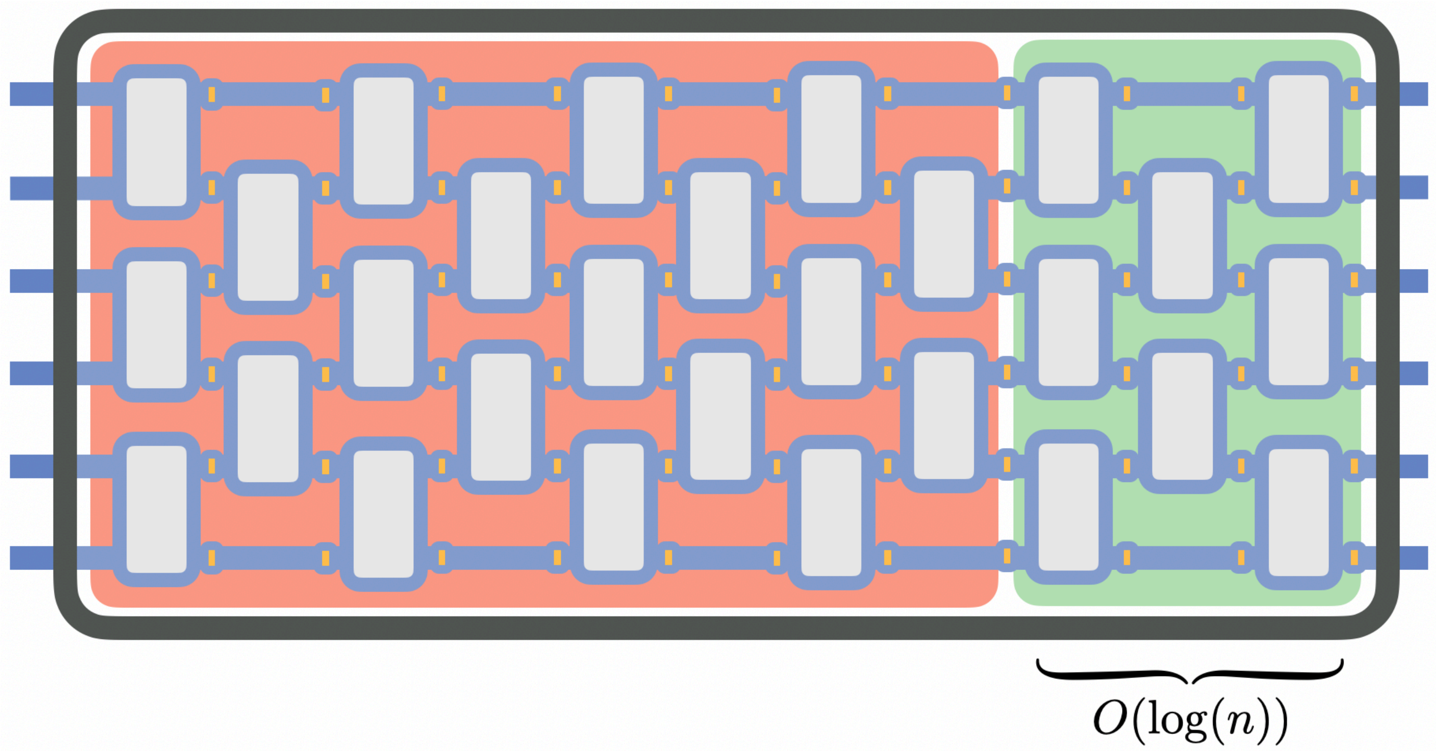


If the noise is **depolarizing**, no layers matter.

(Even for ALL circuits)

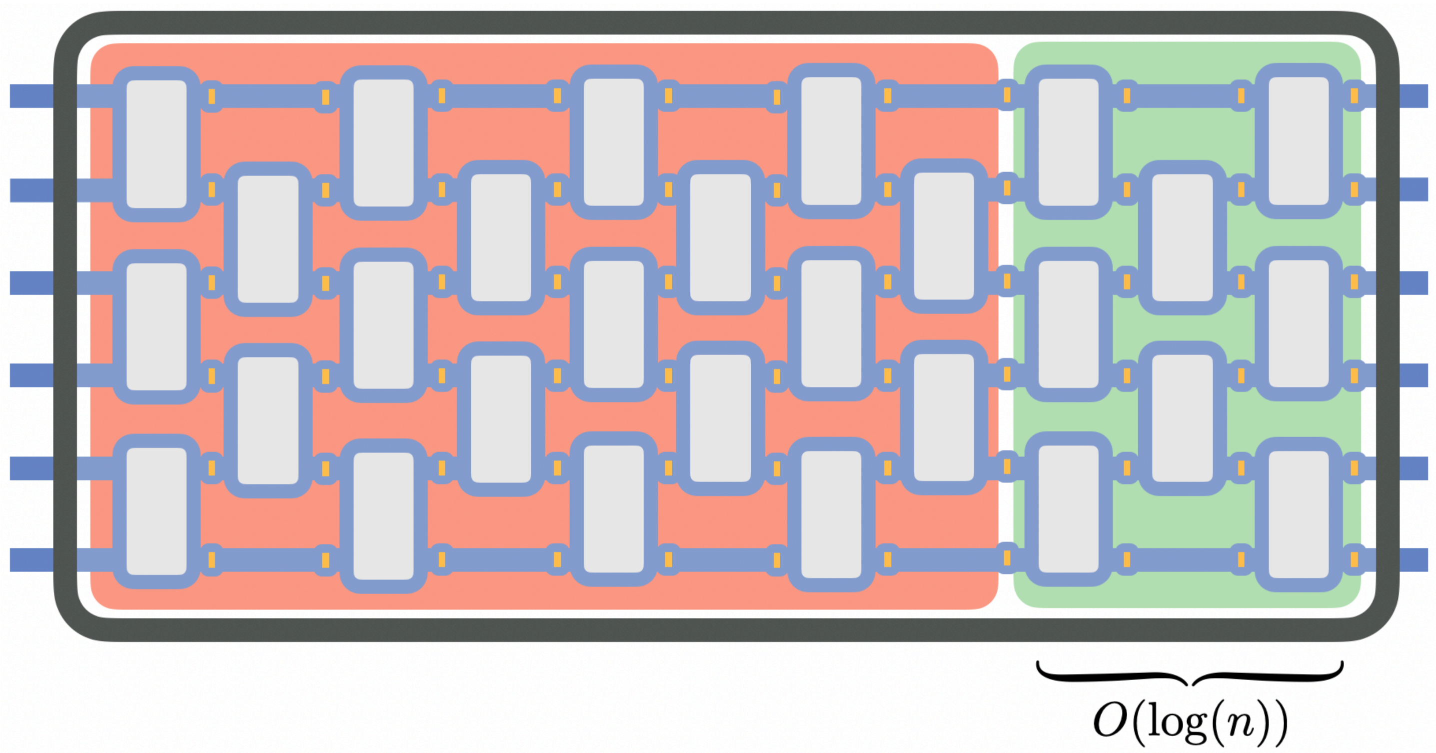


If the noise is **non-unital**, the last $O(\log(n))$ do matter!



If the noise is **non-unital**, the last $O(\log(n))$ do matter!

(for most circuits)



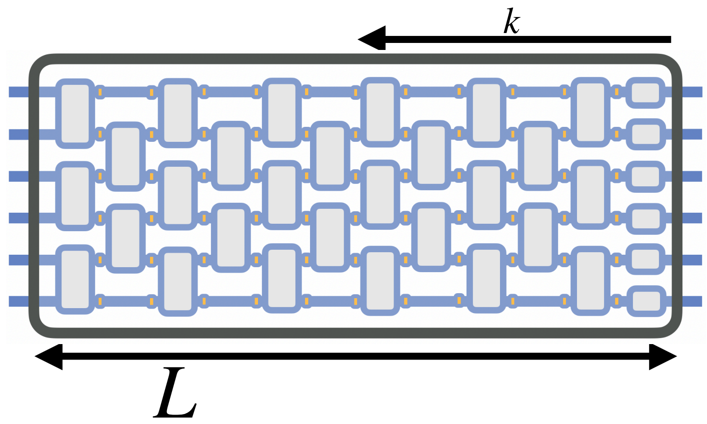
Effective shallow circuits

Proposition.

For any initial state ρ_0 (possibly complex) and observable O,

$$\mathbb{E}_{\Phi}\left[\left|\mathrm{Tr}\left(O\Phi(\rho_0)\right) - \mathrm{Tr}\left(O\Phi_{[L-k,L]}(\sigma_0)\right)\right|\right] \leq \|O\|_\infty \exp(-\Omega(k)).$$

$\Phi =$



Effective shallow circuits

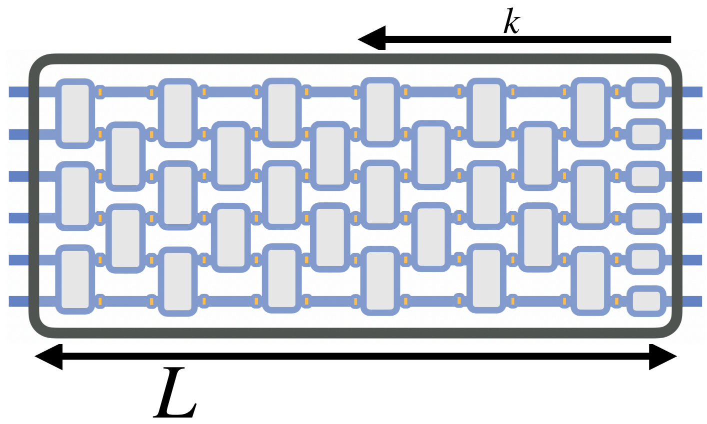
Proposition.

For any initial state ρ_0 (possibly complex) and observable O,

$$\mathbb{E}_\Phi \left[\left| \mathrm{Tr}\left(O\Phi(\rho_0)\right) - \mathrm{Tr}\left(O\Phi_{[L-k,L]}(\sigma_0)\right) \right| \right] \leq \|O\|_\infty \exp(-\Omega(k)).$$

Target expectation value

$\Phi =$



Effective shallow circuits

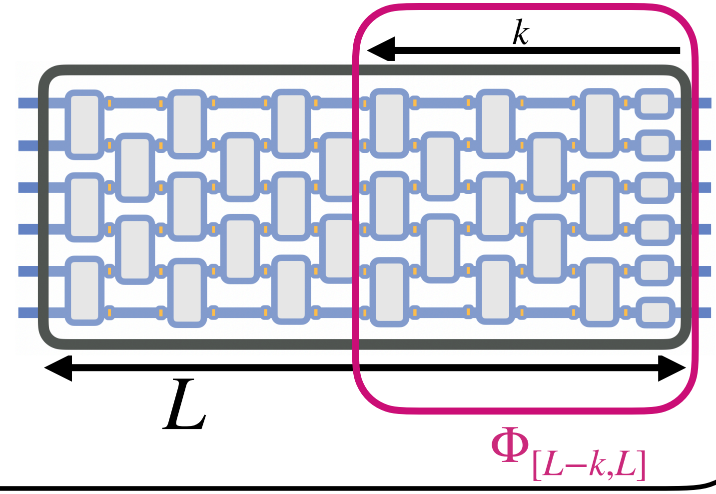
Proposition.

For any initial state ρ_0 (possibly complex) and observable O,

$$\mathbb{E}_{\Phi}\left[\left|\mathrm{Tr}\left(O\Phi(\rho_0)\right) - \mathrm{Tr}\left(O\Phi_{[L-k,L]}(\sigma_0)\right)\right|\right] \leq \|O\|_\infty \exp(-\Omega(k)).$$

Target expectation value

Last k layers

$\Phi =$ 

Effective shallow circuits

Proposition.

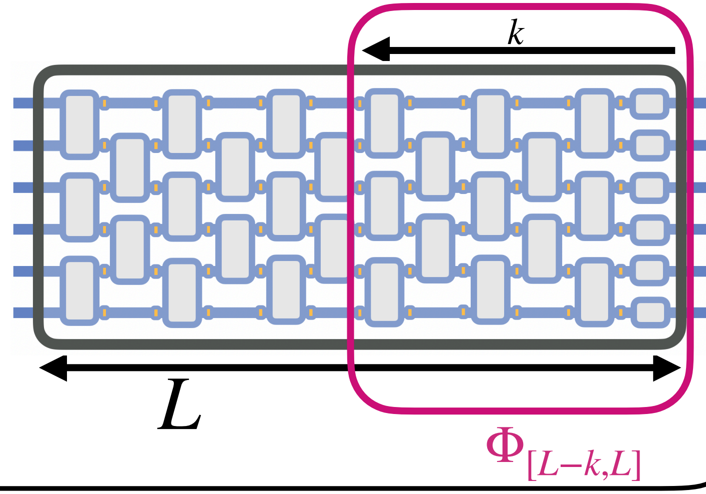
For any initial state ρ_0 (possibly complex) and observable O,

$$\mathbb{E}_\Phi \left[\left| \mathrm{Tr}\left(O\Phi(\rho_0)\right) - \mathrm{Tr}\left(O\Phi_{[L-k,L]}(\sigma_0)\right) \right| \right] \leq \|O\|_\infty \exp(-\Omega(k)).$$

Target expectation value

Last k layers

Any preferred initial state

$\Phi =$ 

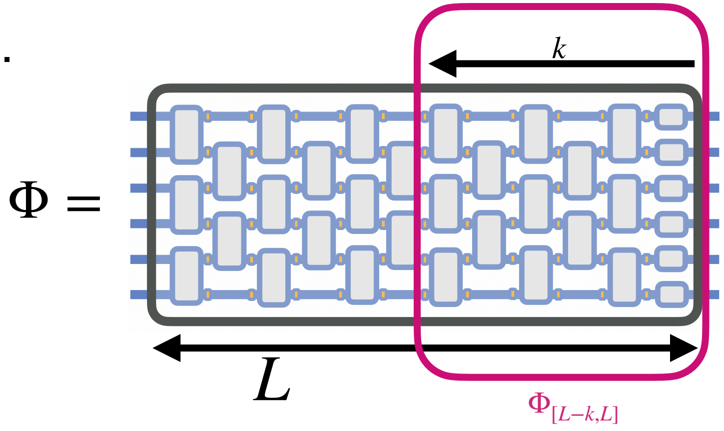
Classical simulation of noisy random circuits

Task: Estimate $\mathrm{Tr}(P\Phi(\rho_0))$, with high probability over the choice of Φ.

Classical simulation of noisy random circuits

Task: Estimate $\mathrm{Tr}(P\Phi(\rho_0))$, with high probability over the choice of Φ.

Solution: It suffices to output $\mathrm{Tr}\big(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\big)$.



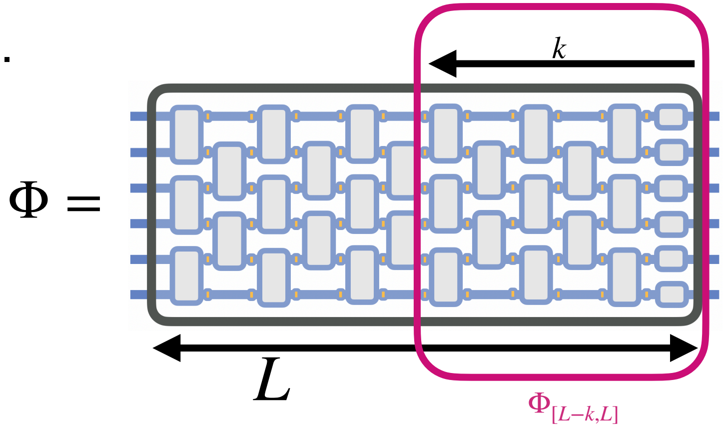
Classical simulation of noisy random circuits

Task: Estimate $\mathrm{Tr}(P\Phi(\rho_0))$, with high probability over the choice of Φ.

Solution: It suffices to output $\mathrm{Tr}\left(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\right)$.

Previous proposition.

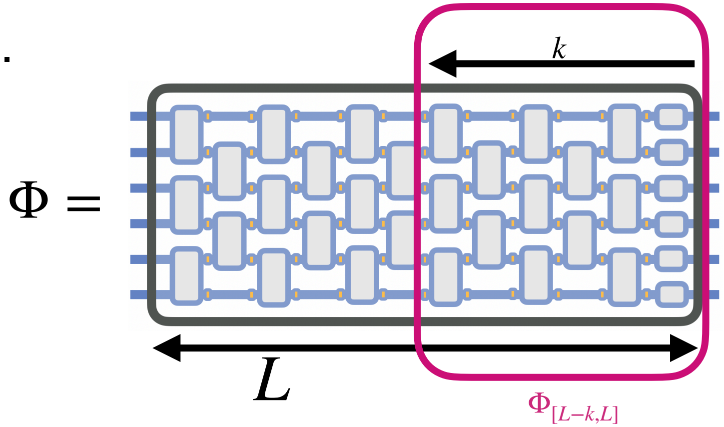
$$\mathbb{E}_{\Phi}\left[\left|\mathrm{Tr}\left(P\Phi(\rho_0)\right) - \mathrm{Tr}\left(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\right)\right|\right] \leq \exp(-\Omega(k+|P|)).$$



Classical simulation of noisy random circuits

Task: Estimate $\mathrm{Tr}(P\Phi(\rho_0))$, with high probability over the choice of Φ.

Solution: It suffices to output $\mathrm{Tr}\big(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\big)$.



Previous proposition.

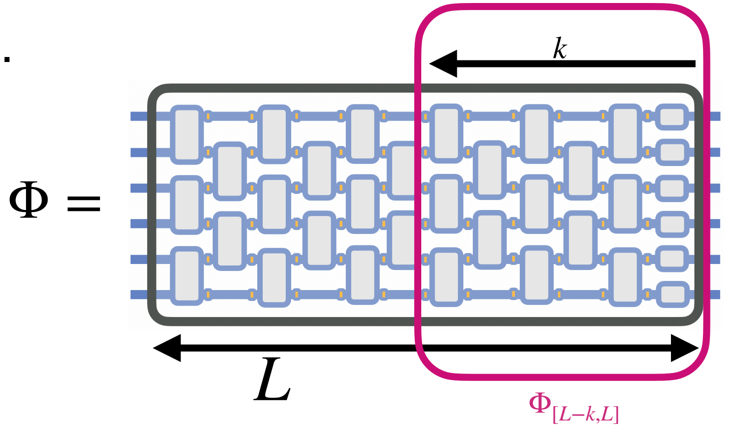
$$\mathbb{E}_{\Phi}\left[\left|\mathrm{Tr}\left(P\Phi(\rho_0)\right) - \mathrm{Tr}\left(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\right)\right|\right] \leq \exp(-\Omega(k+|P|)).$$

Choosing $k = O(\log(\varepsilon^{-1}))$ suffices to have: $\leq \varepsilon$

Classical simulation of noisy random circuits

Task: Estimate $\mathrm{Tr}(P\Phi(\rho_0))$, with high probability over the choice of Φ.

Solution: It suffices to output $\mathrm{Tr}\big(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\big)$.



Previous proposition.

$$\mathbb{E}_\Phi\left[\left|\mathrm{Tr}\left(P\Phi(\rho_0)\right) - \mathrm{Tr}\left(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\right)\right|\right] \leq \exp(-\Omega(k + |P|)).$$

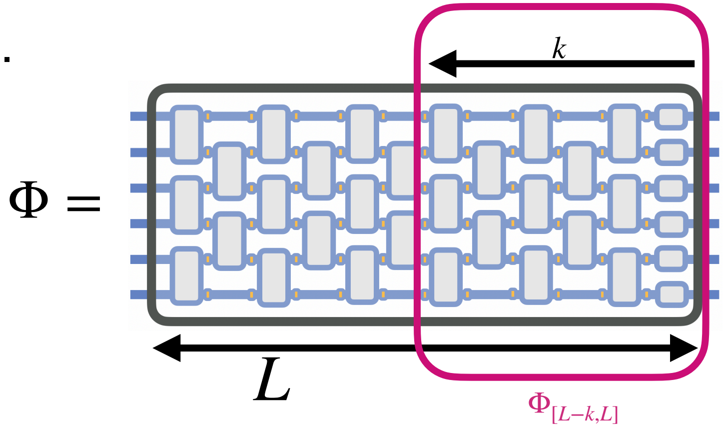
Choosing $k = O(\log(\varepsilon^{-1}))$ suffices to have: $\leq \varepsilon$

- We have $\mathrm{Tr}\big(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\big) = \mathrm{Tr}\big({\Phi_{[L-k,L]}}^*(P)|0^n\rangle\langle 0^n|\big)$

Classical simulation of noisy random circuits

Task: Estimate $\mathrm{Tr}(P\Phi(\rho_0))$, with high probability over the choice of Φ.

Solution: It suffices to output $\mathrm{Tr}\big(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\big)$.



Previous proposition.

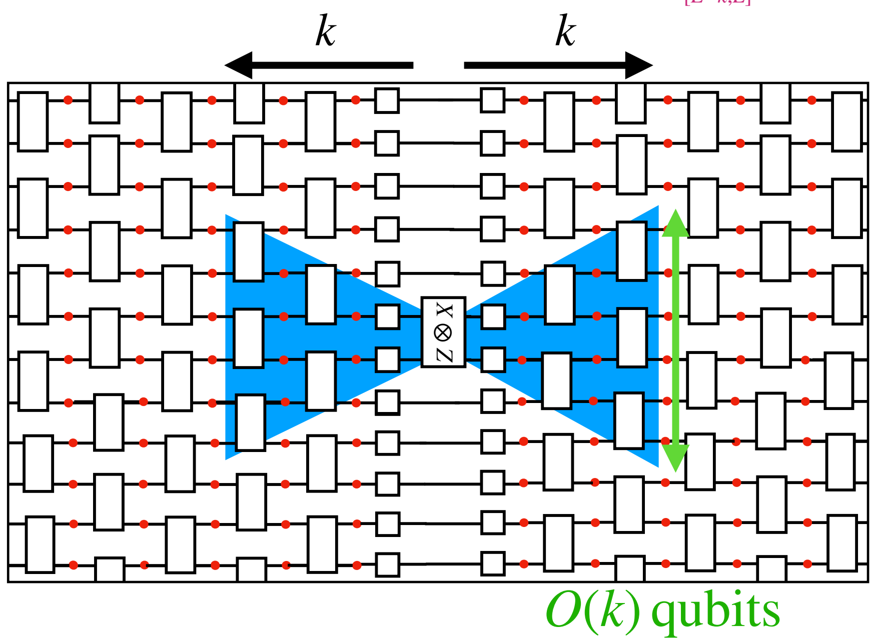
$$\mathbb{E}_\Phi\left[\left|\mathrm{Tr}\left(P\Phi(\rho_0)\right) - \mathrm{Tr}\left(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\right)\right|\right] \leq \exp(-\Omega(k+|P|)).$$

Choosing $k = O(\log(\varepsilon^{-1}))$ suffices to have: $\leq \varepsilon$

- We have $\mathrm{Tr}\big(P\Phi_{[L-k,L]}(|0^n\rangle\langle 0^n|)\big) = \mathrm{Tr}\big(\Phi_{[L-k,L]}{}^*(P)|0^n\rangle\langle 0^n|\big)$

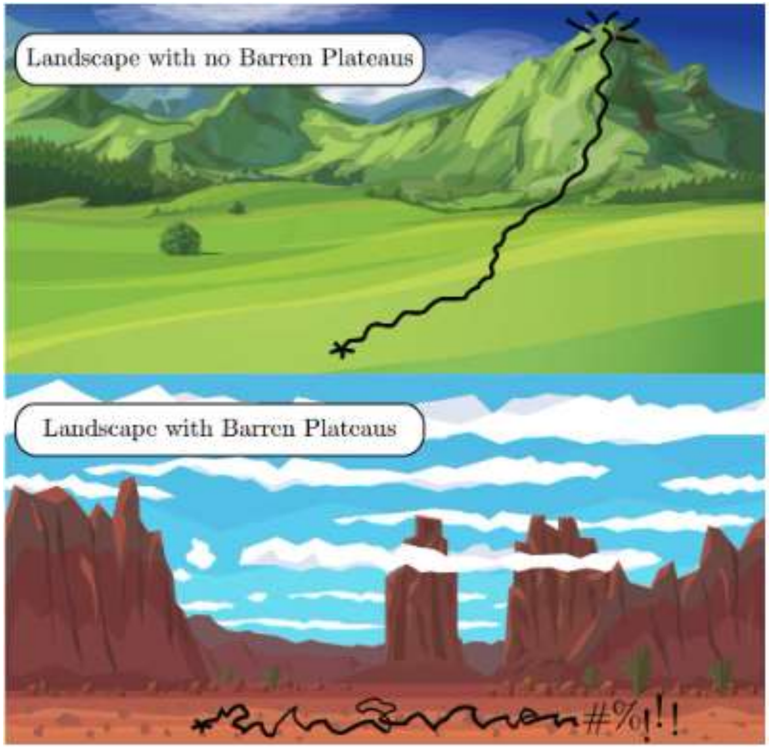
  If P is local, then this is also local

- Thus, this can be computed efficiently for light-cone arguments.



Computational time (for 1D): $\exp(O(k)) = \mathrm{poly}(\varepsilon^{-1})$

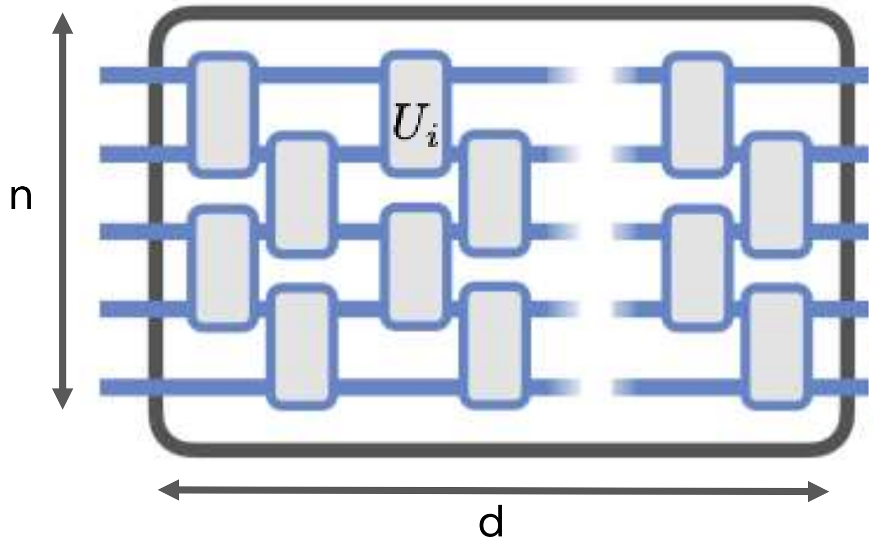
Barren plateaus are bad for variational algorithms



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

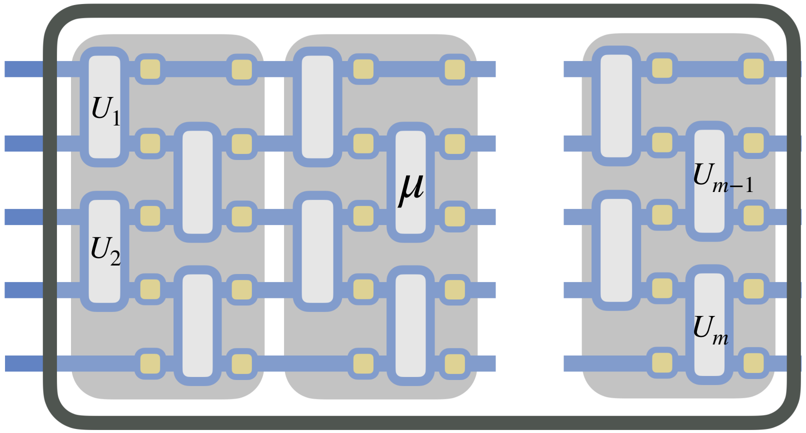
Math translation of ‘vanishing gradient’:

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

Recall: We are optimizing over

$$U_1, \ldots U_m$$

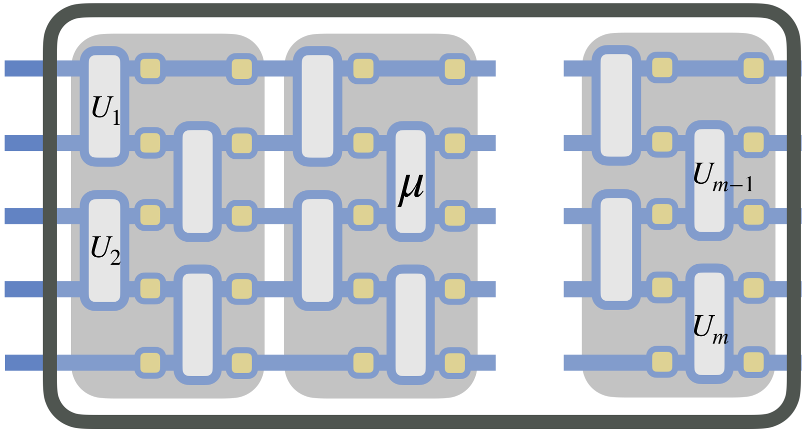
$\Phi =$



Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

$\Phi =$



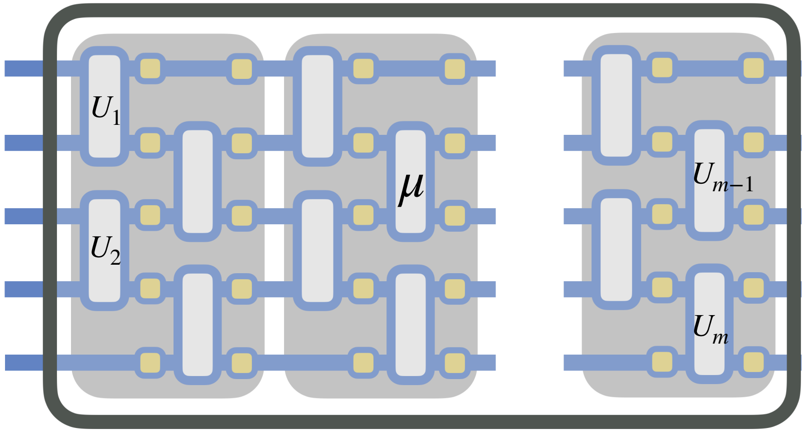
Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

Noiseless [1]**:**

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

$\Phi =$



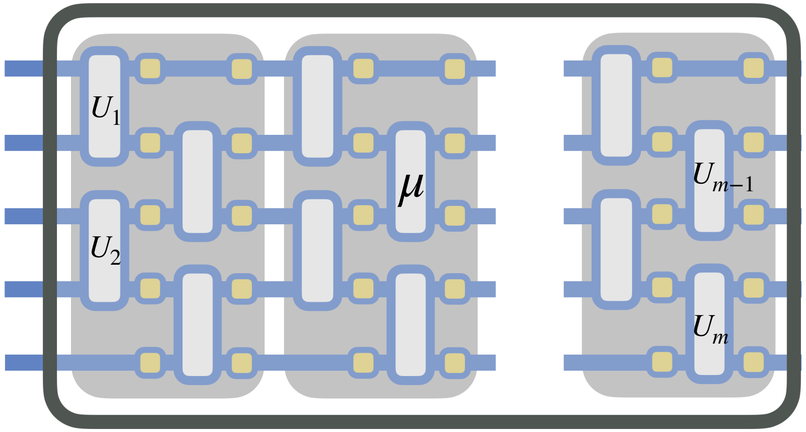
Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

✺ **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

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

$$\Phi =$$



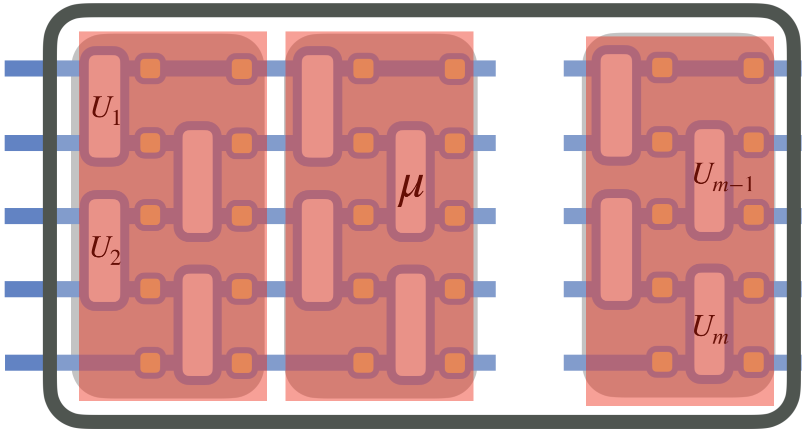
Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

Noiseless [1]: Barren plateaus for depth $\geq \Omega(n)$. $\mathrm{Var}[C] \leq \exp(-\Theta(n))$,

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

$\Phi =$



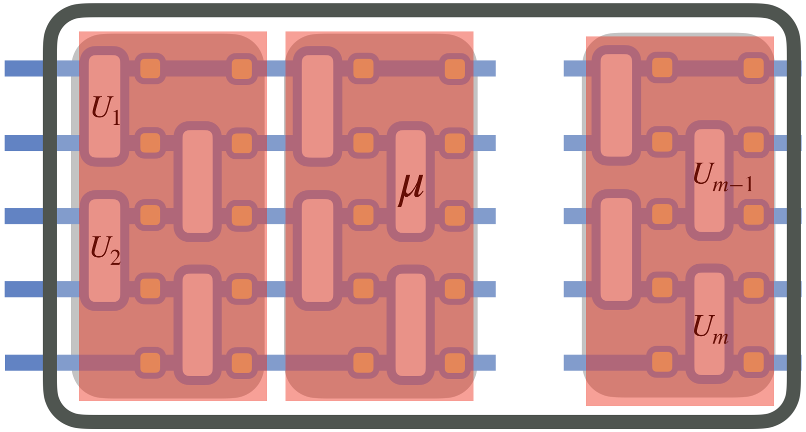
Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

Noiseless [1]: Barren plateaus for depth $\geq \Omega(n)$.

$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$
$$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))\,.$$

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

$\Phi =$ 

Cost function:

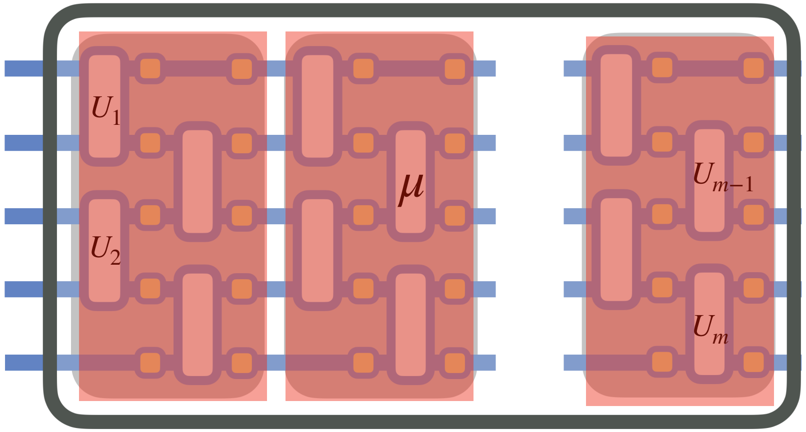
$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

Noiseless [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$

(All the gates are not-trainable) $$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n)).$$

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

$\Phi =$ 

Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

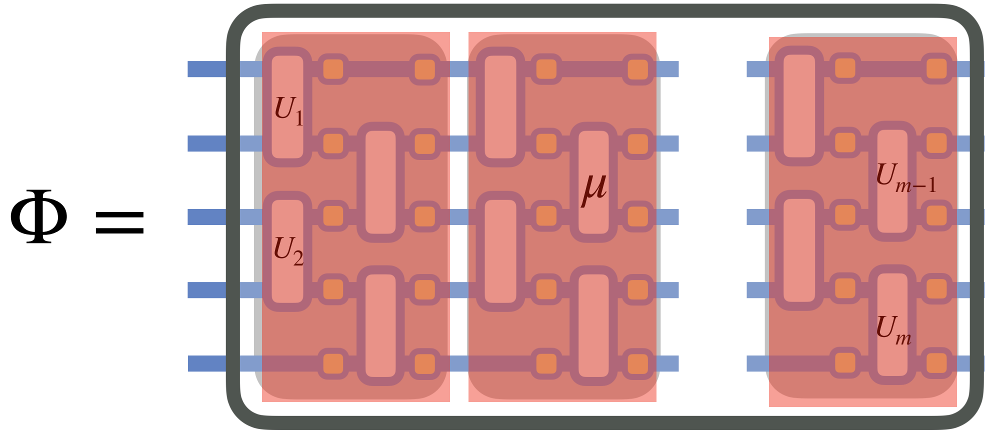
✺ **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

(All the gates are not-trainable)

$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$
$$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n)).$$

✺ **Depolarizing noise** [2]**:**

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



Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

- **Noiseless** [1]: Barren plateaus for depth $\geq \Omega(n)$.

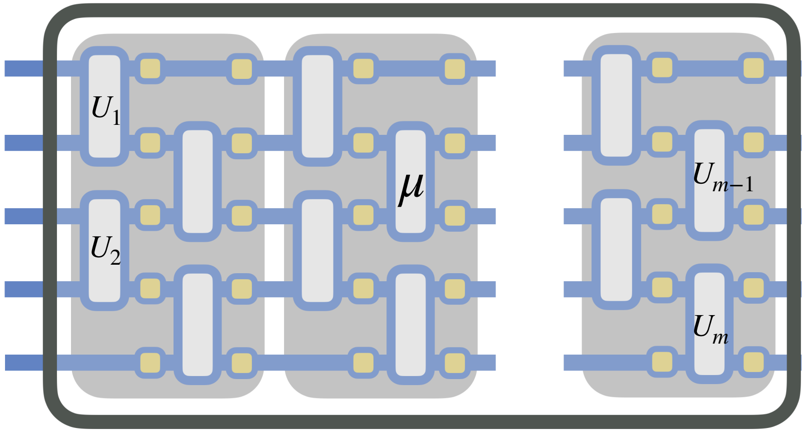
$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$

(All the gates are not-trainable)

$$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))\,.$$

- **Depolarizing noise** [2]: Same of before.

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

$\Phi =$ 

Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

- **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

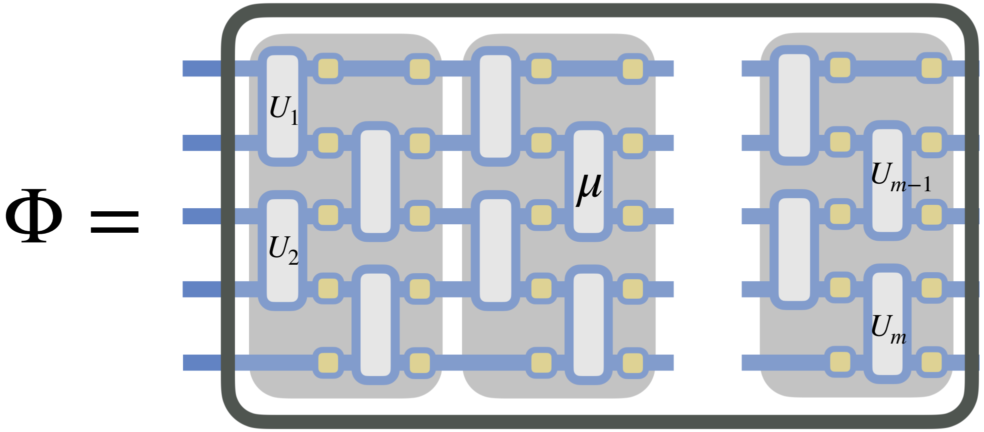
$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$

(All the gates are not-trainable) $$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))\,.$$

- **Depolarizing noise** [2]**:** Same of before.

- **Non-unital noise [This work]:**

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



Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

- **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$. $\mathrm{Var}[C] \leq \exp(-\Theta(n))$,

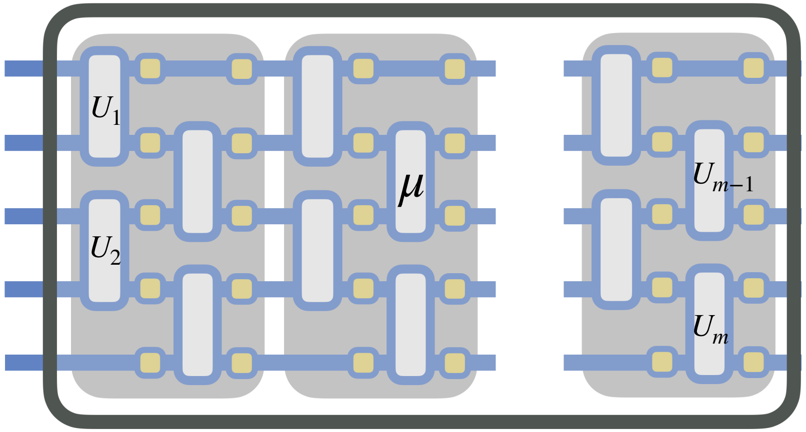
  (All the gates are not-trainable) $\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))$.

- **Depolarizing noise** [2]**:** Same of before.

- **Non-unital noise [This work]:** No Barren plateaus at any depth for local cost functions.

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

$\Phi =$



Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

- **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$

(All the gates are not-trainable) $\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))$.

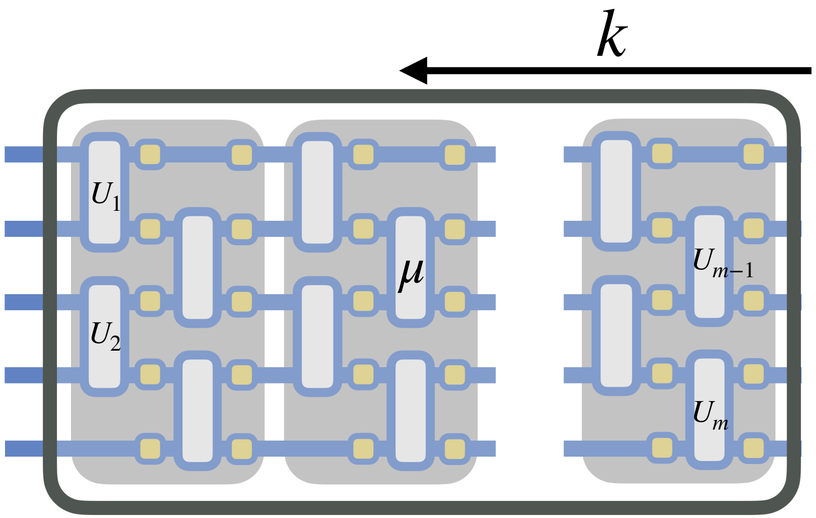
- **Depolarizing noise** [2]**:** Same of before.

- **Non-unital noise [This work]:** No Barren plateaus at any depth for local cost functions.

$$\mathrm{Var}[C] = \Omega(1),$$

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

$$\Phi =$$



Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

- **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

(All the gates are not-trainable)

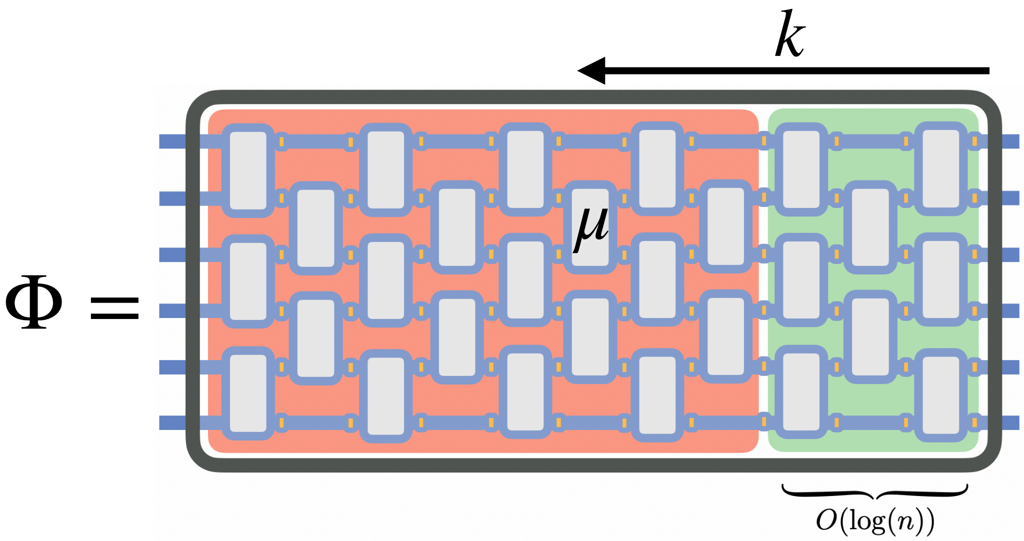
$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$
$$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n)).$$

- **Depolarizing noise** [2]**:** Same of before.

- **Non-unital noise [This work]:** No Barren plateaus at any depth for local cost functions.

$$\mathrm{Var}[C] = \Omega(1),$$
$$\mathrm{Var}[\partial_\mu C] = \exp(-\Theta(k)).$$

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



Cost function:
$C := \mathrm{Tr}(P\Phi(\rho_0))$

- **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$

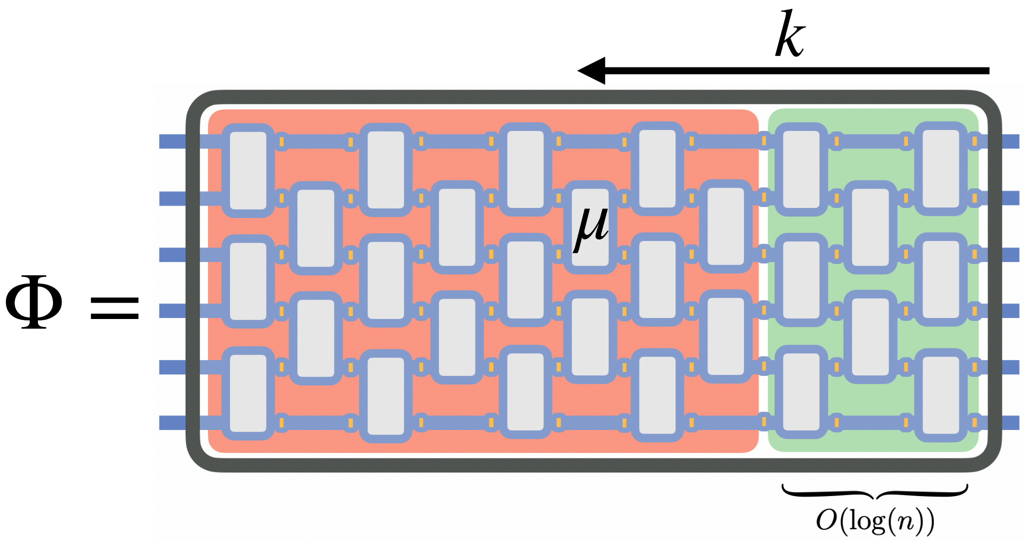
(All the gates are not-trainable) $$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))\,.$$

- **Depolarizing noise** [2]**:** Same of before.

- **Non-unital noise [This work]:** No Barren plateaus at any depth for local cost functions.

$$\mathrm{Var}[C] = \Omega(1),$$
$$\mathrm{Var}[\partial_\mu C] = \exp(-\Theta(k))\,.$$

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



Cost function:

$C := \mathrm{Tr}(P\Phi(\rho_0))$

✺ **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$

(All the gates are not-trainable) $\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n))\,.$

✺ **Depolarizing noise** [2]**:** Same of before.

✺ **Non-unital noise [This work]:** No Barren plateaus at any depth for local cost functions.

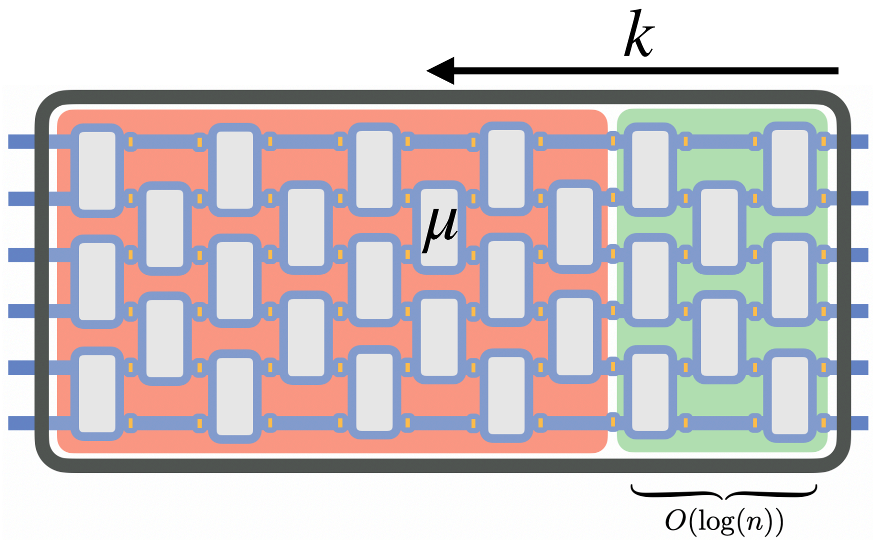
$\mathrm{Var}[C] = \Omega(1),$

$\mathrm{Var}[\partial_\mu C] = \exp(-\Theta(k))\,.$

Only the last $\Theta(\log(n))$ layers are trainable!

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

$\Phi =$



Cost function:

$$C := \mathrm{Tr}(P\Phi(\rho_0))$$

- **Noiseless** [1]**:** Barren plateaus for depth $\geq \Omega(n)$.

$$\mathrm{Var}[C] \leq \exp(-\Theta(n)),$$

(All the gates are not-trainable) $$\mathrm{Var}[\partial_\mu C] \leq \exp(-\Theta(n)).$$

- **Depolarizing noise** [2]**:** Same of before.

- **Non-unital noise [This work]:** No Barren plateaus at any depth for local cost functions.

$$\mathrm{Var}[C] = \Omega(1),$$

$$\mathrm{Var}[\partial_\mu C] = \exp(-\Theta(k)).$$

<u>Only</u> the last $\Theta(\log(n))$ layers are trainable!

[1] Barren plateaus in quantum neural network training landscapes. McClean *et al.* Nature Comm. (2018).
[2] Noise-induced barren plateaus in variational quantum algorithms. Wang *et al,* Nature Comm. (2021).
<u>**Compare with**</u>: [3] Beyond unital noise in variational quantum algorithms: noise-induced barren plateaus and fixed points. Singkanipa *et al.*, ArXiv. (2024).

Proof ingredients

Proof ingredients

- Work in the Heisenberg picture.

Proof ingredients

- Work in the Heisenberg picture. $\mathrm{Tr}(P\Phi(\rho_0)) = \mathrm{Tr}(\Phi^*(P)\rho_0)$

Proof ingredients

- Work in the Heisenberg picture. $\mathrm{Tr}(P\Phi(\rho_0)) = \mathrm{Tr}(\Phi^*(P)\rho_0)$

- 2-design properties (of the 2-qubits gates).

Proof ingredients

- Work in the Heisenberg picture. $\mathrm{Tr}(P\Phi(\rho_0)) = \mathrm{Tr}(\Phi^*(P)\rho_0)$

- 2-design properties (of the 2-qubits gates).

- ‘Normal form’ of noisy channels.

Proof ingredients

- Work in the Heisenberg picture. $\mathrm{Tr}(P\Phi(\rho_0)) = \mathrm{Tr}(\Phi^*(P)\rho_0)$

- 2-design properties (of the 2-qubits gates).

- ‘Normal form’ of noisy channels. Any 1-qubit channel is unitarily equivalent to a channel parametrized by:

Proof ingredients

- Work in the Heisenberg picture. $\mathrm{Tr}(P\Phi(\rho_0)) = \mathrm{Tr}(\Phi^*(P)\rho_0)$

- 2-design properties (of the 2-qubits gates).

- 'Normal form' of noisy channels. Any 1-qubit channel is unitarily equivalent to a channel parametrized by:

$$(D_X, D_Y, D_Z), (t_X, t_Y, t_Z) \in \mathbb{R}^3$$

$$\mathcal{N}(I) = I + t_X X + t_Y Y + t_Z Z$$

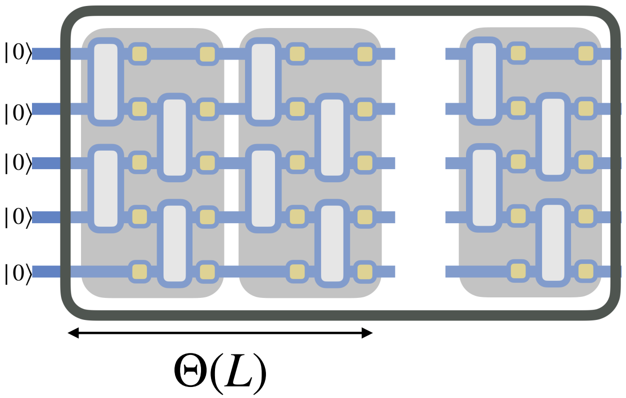
$$\mathcal{N}(X) = D_X X$$

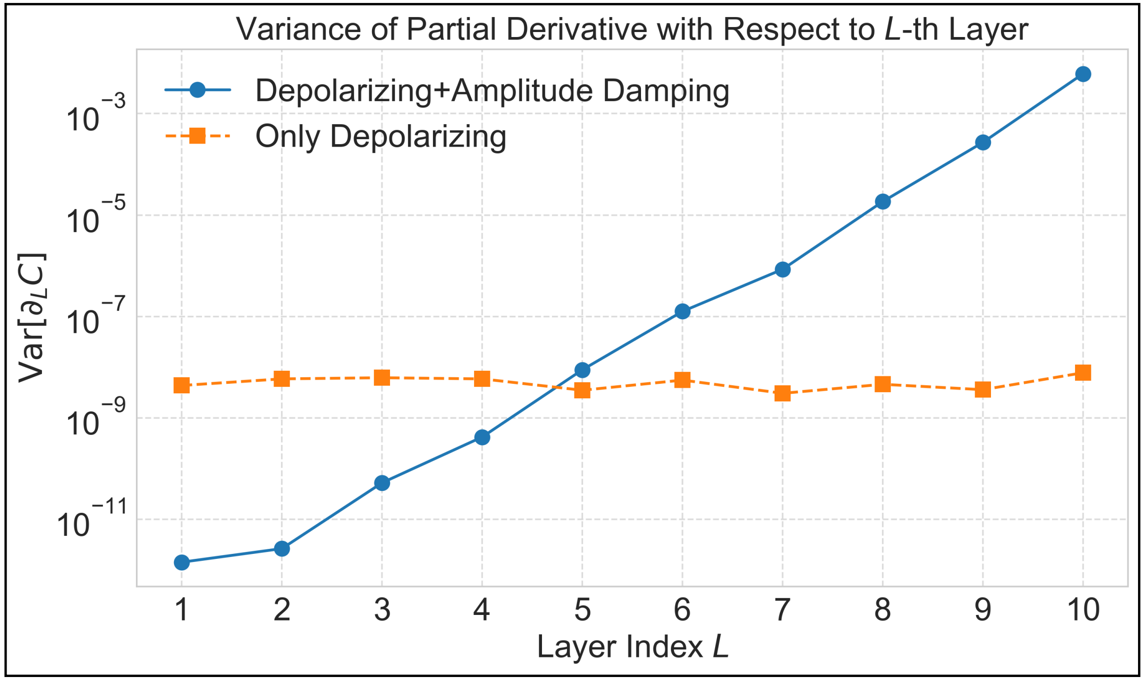
$$\mathcal{N}(Y) = D_Y Y$$

$$\mathcal{N}(Z) = D_Z Z$$

Numerical experiments confirm our results for more structured circuits.

Numerical experiments confirm our results for more structured circuits.





Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it ‘truncates’ most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it ‘truncates’ most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it 'truncates' most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it ‘truncates’ most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).
- Assuming overly simplistic noise models could be misleading.

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it 'truncates' most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).
- Assuming overly simplistic noise models could be misleading.

Open questions

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it 'truncates' most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).
- Assuming overly simplistic noise models could be misleading.

Open questions

- Simulating Pauli expectation values of noisy random quantum circuits with $\mathrm{poly}(n^{-1})$ accuracy in 2D and all-to-all connectivity in polynomial time.

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it 'truncates' most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).
- Assuming overly simplistic noise models could be misleading.

Open questions

- Simulating Pauli expectation values of noisy random quantum circuits with $\mathrm{poly}(n^{-1})$ accuracy in 2D and all-to-all connectivity in polynomial time.
- Complexity of classical simulation of random quantum circuit sampling with non-unital noise [1].

[1] Effect of non-unital noise on random circuit sampling, Fefferman *et al*, QIP (2023).

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it 'truncates' most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).
- Assuming overly simplistic noise models could be misleading.

Open questions

- Simulating Pauli expectation values of noisy random quantum circuits with $\mathrm{poly}(n^{-1})$ accuracy in 2D and all-to-all connectivity in polynomial time.
- Complexity of classical simulation of random quantum circuit sampling with non-unital noise [1].

(The depolarizing case was addressed in [2])

[1] Effect of non-unital noise on random circuit sampling, Fefferman *et al*, QIP (2023).
[2] A polynomial-time classical algorithm for noisy random circuit sampling, Aharonov *et al.*, STOC (2023)

Conclusions

- Any amount of **non-unital noise induces lack of Barren plateaus**, but it 'truncates' most circuits to **effectively log-depth**, allowing **efficient classical simulation** in estimating Pauli expectation values.
- **Unless** we carefully **engineer the circuits to take advantage of the noise**, it is **unlikely** that **non-unital** noisy circuits are **preferable over depolarizing** noisy circuits (which are restricted to log depth).
- Assuming overly simplistic noise models could be misleading.

Open questions

- Simulating Pauli expectation values of noisy random quantum circuits with $\mathrm{poly}(n^{-1})$ accuracy in 2D and all-to-all connectivity in polynomial time.
- Complexity of classical simulation of random quantum circuit sampling with non-unital noise [1].

(The depolarizing case was addressed in [2])

Thanks a lot for your attention!

[1] Effect of non-unital noise on random circuit sampling, Fefferman *et al*, QIP (2023).
[2] A polynomial-time classical algorithm for noisy random circuit sampling, Aharonov *et al.*, STOC (2023)