

Scalable Classical Shadow Tomography with Shallow Circuits and Quantum Dynamics

Yi-Zhuang You (尤亦庄)

University of California San Diego

[1] H.-Y. Hu, Y.-Z. You. arXiv:2102.10132

[2] H.-Y. Hu, S. Choi, Y.-Z. You. arXiv:2107.04817

[3] A. A. Akhtar, H.-Y. Hu, Y.-Z. You. arXiv:2209.02093



Ahmed Akhtar
(UCSD)



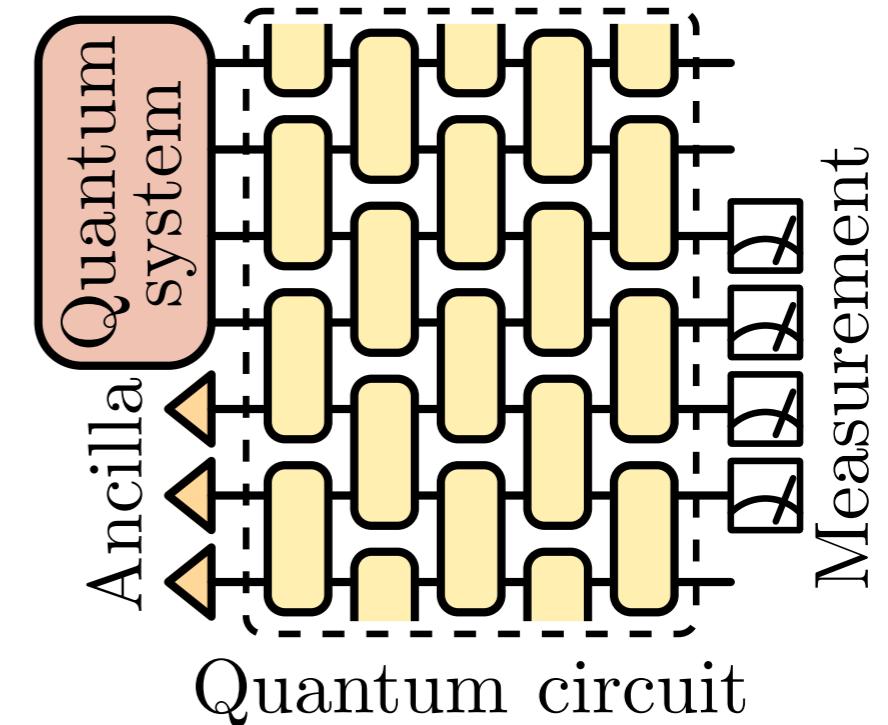
Hong-Ye Hu
(UCSD → Harvard)



Soonwon Choi
(MIT)

Quantum Dynamics meets Shadow Tomography

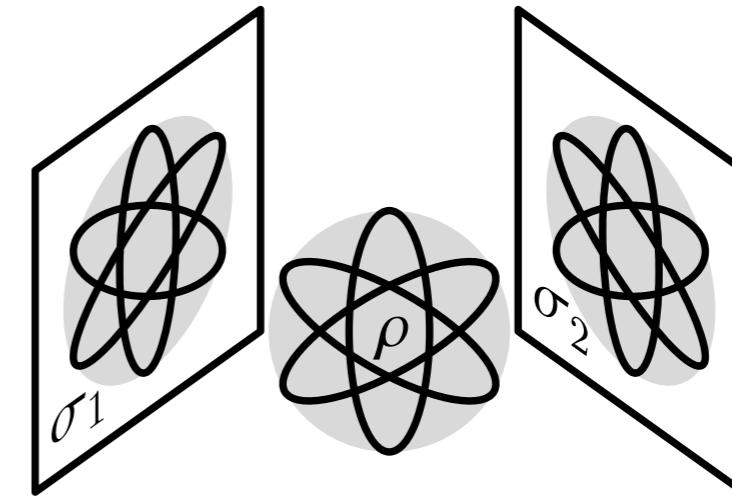
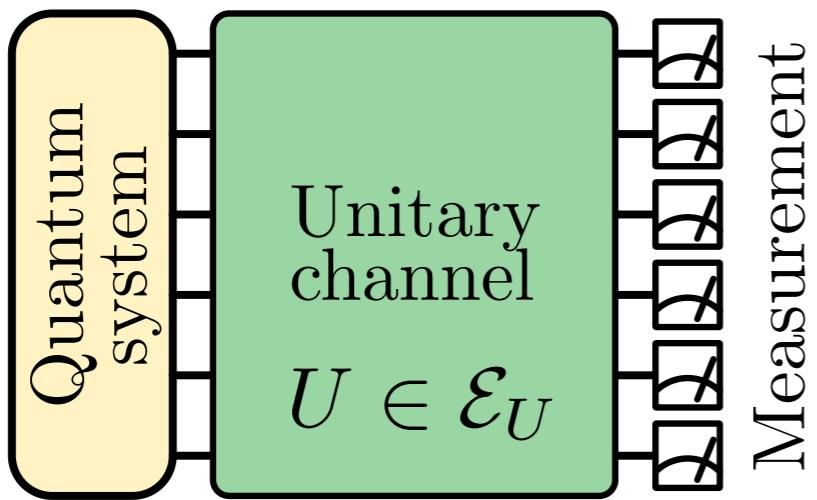
- The goal of this program is “to explore connections between NISQ devices and quantum dynamics.”
- Classical shadow tomography is one such example.
 - What do we study in quantum dynamics?
A quantum many-body system (+ ancilla) under unitary evolution and measurements
 - Is the randomized measurement tomographically complete?
 - How to reconstruct the quantum information from measurements?
 - What is the sample complexity for different tasks?
 - How to optimize tomography efficiency on NISQ devices?



Review: Elben, Flammia, Huang, Kueng, Preskill, Vermersch, Zoller. arXiv:2203.11374 (2022)

Randomized Measurement Protocol

- Randomized measurement
= **random unitary** + **computational basis** measurement



- Step 1 (Random scrambling): $\rho \rightarrow \rho' = U\rho U^\dagger$
- Step 2 (Z-basis measurement): $\rho' \rightarrow |b\rangle\langle b|$
- Step 3 (Collect classical snapshot): $\hat{\sigma}_{U,b} = U^\dagger |b\rangle\langle b| U$
- Probability to obtain a classical snapshot (Stored classically)

$$P(\hat{\sigma}_{U,b}|\rho) = P(b|\rho, U)P(U) = \text{Tr}(\hat{\sigma}_{U,b}\rho)P(U)$$

Ensembles of Classical Snapshots

- Classical snapshot states $\hat{\sigma}$ form two types of ensembles:
 - **Posterior** snapshot ensemble:

$$\mathcal{E}_{\sigma|\rho} = \{\hat{\sigma}_{U,b} \mid P(\hat{\sigma}_{U,b}|\rho) = \text{Tr}(\hat{\sigma}_{U,b}\rho)P(U)\}$$

- **Prior** snapshot ensemble:

$$\mathcal{E}_\sigma = \{\hat{\sigma}_{U,b} \mid P(\hat{\sigma}_{U,b}) = 2^{-N}P(U)\}$$

- As if $\rho = \mathbb{1}/2^N$ is maximally mixed (without any knowledge about the quantum system)
- The prior ensemble \mathcal{E}_σ only depends on the random unitary ensemble $\mathcal{E}_U = \{U|P(U)\}$ (encodes statistical features of the randomized measurement scheme)

Measurement Channel and Reconstruction Map

- **Measurement channel:** a quantum channel mapping from observed **quantum** state ρ to average **classical** snapshot σ

$$\sigma \equiv \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \hat{\sigma} = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_\sigma} \hat{\sigma} \text{Tr}(\hat{\sigma}\rho)2^N = \mathcal{M}[\rho]$$

- Assuming the measurement basis is topographically complete \rightarrow the measurement channel \mathcal{M} is invertible (the inverse map \mathcal{M}^{-1} exists)
- **Reconstruction map:** the inverse map from average **classical** snapshot σ to observed **quantum** state ρ

$$\rho = \mathcal{M}^{-1}[\sigma] = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \mathcal{M}^{-1}[\hat{\sigma}]$$

- \mathcal{M}^{-1} is the **key**! It specifies how to post-process the classical data σ to make predictions.

Not a physical channel

Measurement Channel and Reconstruction Map

- **Reconstruction map:** the inverse map from average **classical** snapshot σ to observed **quantum** state ρ

$$\rho = \mathcal{M}^{-1}[\sigma] = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \mathcal{M}^{-1}[\hat{\sigma}]$$

- \mathcal{M}^{-1} is the key! It specifies how to post-process the classical data σ to make predictions.
- Examples: predict physical properties about ρ

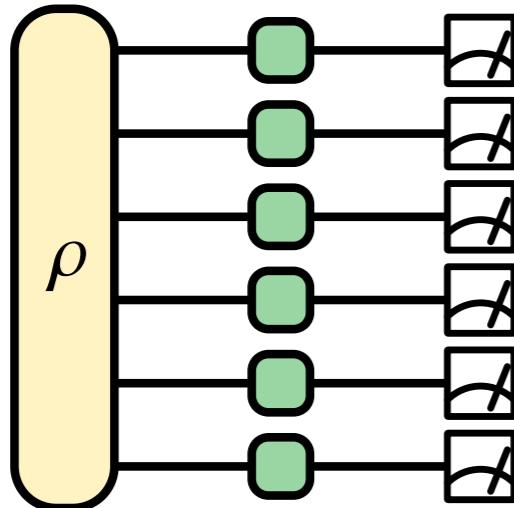
$$\langle O \rangle_\rho = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \text{Tr}(O \mathcal{M}^{-1}[\hat{\sigma}]) \text{ (Linear property)}$$

$$e^{-S_\rho^{(2)}(A)} = \mathbb{E}_{\hat{\sigma}, \hat{\sigma}' \in \mathcal{E}_{\sigma|\rho}} \text{Tr}_A \left(\text{Tr}_{\bar{A}} \mathcal{M}^{-1}[\hat{\sigma}] \text{Tr}_{\bar{A}} \mathcal{M}^{-1}[\hat{\sigma}'] \right) \text{ (Nonlinear property)}$$

- However, known results about \mathcal{M} and \mathcal{M}^{-1} were limited to random Pauli and Clifford measurements previously.

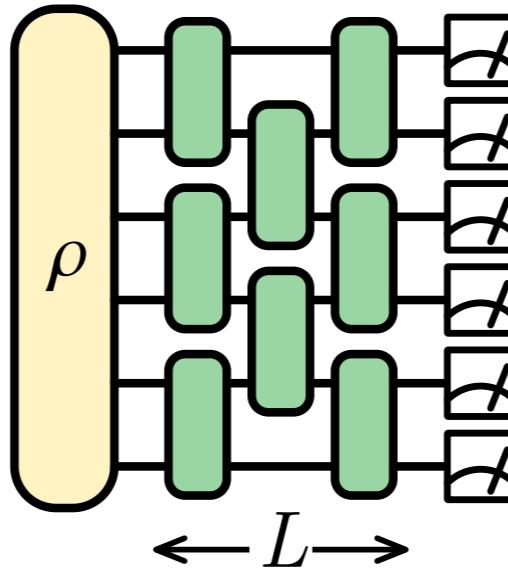
Local v.s. Global Scrambling

- Pauli Measurement

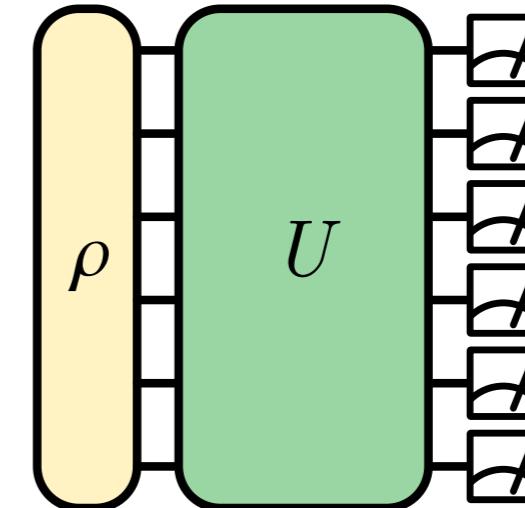


$$\mathcal{M}^{-1}[\sigma] = \bigotimes_i (3\sigma_i - 1)$$

???



- Clifford Measurement



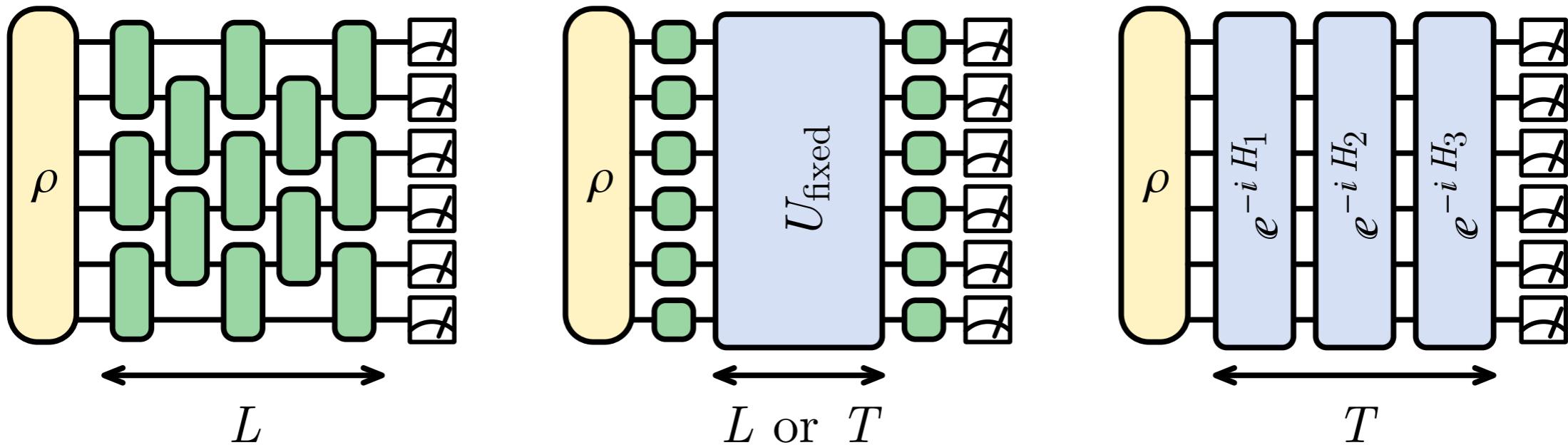
$$\mathcal{M}^{-1}[\sigma] = (2^N + 1)\sigma - 1$$

- Efficient for **local** observables (e.g. local Hamiltonian, two-point correlation)
- **Low** circuit complexity

- Efficient for **low-rank** observables (e.g. fidelity estimation)
- **High** circuit complexity

Main Result

- We made progress in formulating the **reconstruction map** \mathcal{M}^{-1} for a large class of **randomized measurements** based on **locally scrambled** quantum dynamics/circuits



$$\rho = \mathcal{M}^{-1}[\sigma] = \sum_A r_A \sigma_A$$

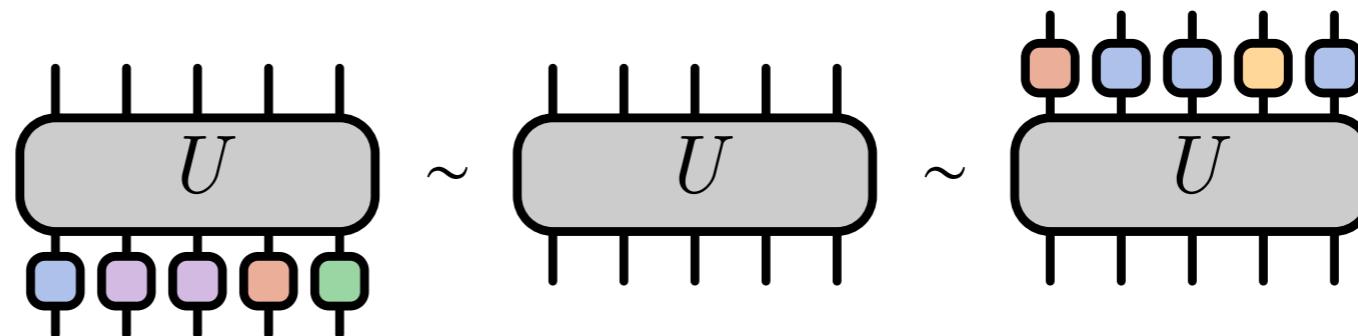
H.-Y. Hu, S. Choi, Y.-Z. You. arXiv:2107.04817 (2021)

- The key is to understand the **entanglement dynamics** under different scrambling unitaries.

Locally Scrambled Quantum Dynamics

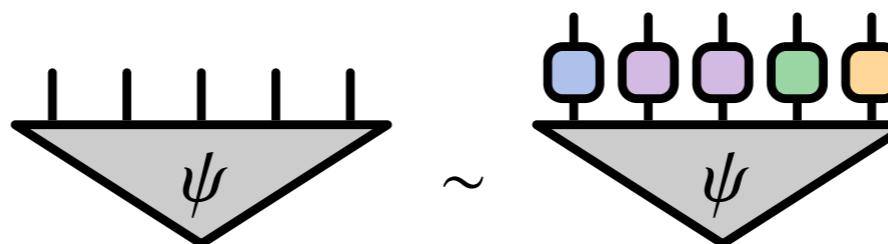
- A random unitary ensemble \mathcal{E}_U is **locally scrambled** iff

$$\forall V = \prod_i V_i : P(VU) = P(U) = P(UV)$$



- Refined from Haar (globally scrambled) random unitary
- Can be constructed by assembling locally scrambled unitary gates (e.g. random Clifford circuits of any structure)
- A random state ensemble \mathcal{E}_σ is **locally scrambled** iff

$$\forall V = \prod_i V_i : P(\sigma) = P(V^\dagger \sigma V)$$



For pure state
 $\sigma = |\psi\rangle\langle\psi|$

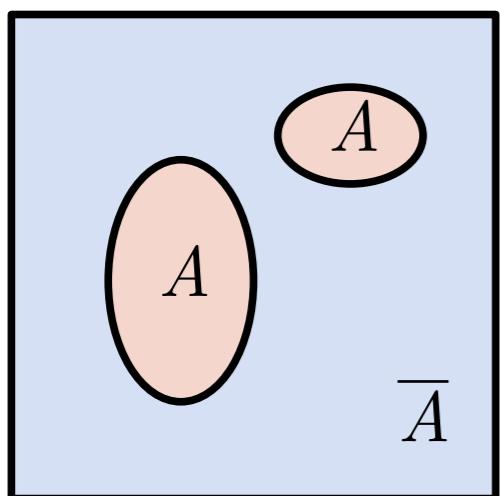
Locally Scrambled Random States

- Classical snapshots collected from locally scrambled randomized measurements are also locally scrambled!
- Only **local-basis invariant** information matters
→ All such information are **entanglement features (EF)**

You, Gu (2018); Kuo, Akhtar, Arovas, You (2019)

$$W_{\mathcal{E}_\sigma, g}^{(n)} = \text{Tr } \sigma^{\otimes n} g \quad (g \in S_n^{\times N})$$

- At the 2-replica level, EF is just the **purity** in all possible subsystems



$$\begin{aligned} W_{\mathcal{E}_\sigma, A}^{(2)} &= \mathbb{E}_{\sigma \in \mathcal{E}_\sigma} \text{Tr } \sigma^{\otimes 2} \chi_A \leftarrow \text{Swap operator} \\ &= \mathbb{E}_{\sigma \in \mathcal{E}_\sigma} \text{Tr}_A (\text{Tr}_{\bar{A}} \sigma)^2 \\ &= \mathbb{E}_{\sigma \in \mathcal{E}_\sigma} e^{-S_\sigma^{(2)}(A)} \leftarrow \text{Entanglement entropy} \end{aligned}$$

Entanglement Feature Formalism

- N-qubit system: totally 2^N possible choices of subsystems, how to encode their entanglement features efficiently?
- Idea: use a fictitious quantum state - the **EF state**

$$|W_{\mathcal{E}_\sigma}\rangle = \sum_A W_{\mathcal{E}_\sigma, A} |A\rangle = \sum_A e^{-S_\sigma(A)} |A\rangle \quad (\text{R\'enyi index will be omitted})$$

- Basis states are labeled by entanglement regions \sim Ising configurations

$$|A\rangle = |s_1 s_2 s_3 \cdots s_N\rangle \quad s_i = \begin{cases} 0 & i \in \bar{A} \\ 1 & i \in A \end{cases}$$

- Every locally scrambled **random state ensemble** \mathcal{E}_σ is characterized by a corresponding **EF state** $|W_{\mathcal{E}_\sigma}\rangle$
- Are there any fictitious operators that can act on the EF state? - Yes, the **EF operator**

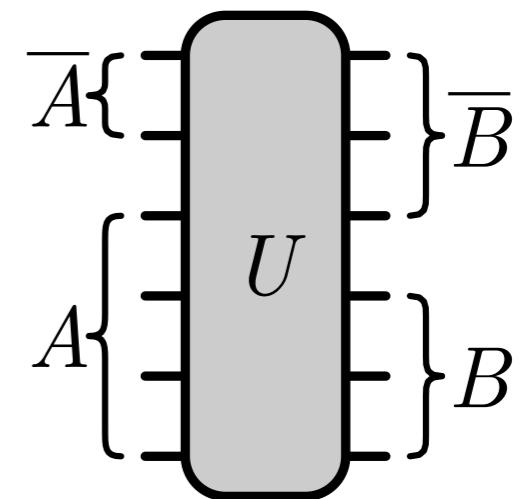
Entanglement Feature Formalism

- Entanglement features of random unitary ensemble

$$W_{\mathcal{E}_U, A, B} = \mathbb{E}_{U \in \mathcal{E}_U} \text{Tr } U^{\otimes 2} \mathcal{X}_A (U^\dagger)^{\otimes 2} \mathcal{X}_B$$

which define the **EF operator**

$$\begin{aligned}\hat{W}_{\mathcal{E}_U} &= \sum_{A, B} |A\rangle W_{\mathcal{E}_U, A, B} \langle B| \\ &= \sum_{A, B} |A\rangle e^{-S_U(A \cup B)} \langle B|\end{aligned}$$



- A and B label the entanglement regions on the past and the future sides of the unitary.
- Every locally scrambled **random unitary** ensemble \mathcal{E}_U is characterized by a corresponding **EF operator** $\hat{W}_{\mathcal{E}_U}$

Entanglement Dynamics

- **Unitary evolution** of a locally scrambled random state by a locally scrambled random unitary

$$\sigma \rightarrow U^\dagger \sigma U \quad (\sigma \in \mathcal{E}_\sigma, U \in \mathcal{E}_U)$$

- The resulting state is still locally scrambled
- Its entanglement feature evolves as

$$|W_{\mathcal{E}_\sigma}\rangle \rightarrow \hat{W}_{\mathcal{E}_U} \hat{W}_1^{-1} |W_{\mathcal{E}_\sigma}\rangle \qquad \text{Kuo, Akhtar, Arovas, You (2019)}$$

where \hat{W}_1 denotes the EF operator for the identity channel

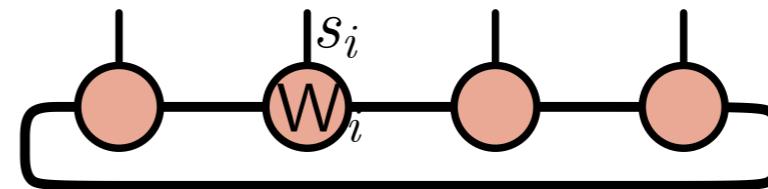
- Generalizable to locally scrambled quantum channels ...
- **Entanglement dynamics** = time evolution of EF state.
- **Entanglement transition** = quantum phase transition of EF state.

Fan, Vijay, Vishwanath, You (2020); Akhtar, You (2020)

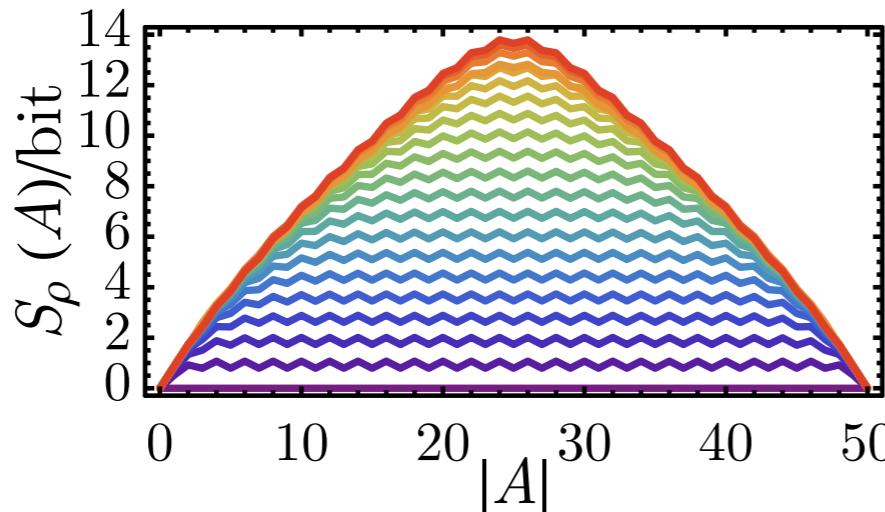
Advantage of Entanglement Feature

- EF state $|W_{\mathcal{E}_\sigma}\rangle$: A systematic organization of all single/multi-region entanglements.
Solve it once, get them all!
- Admits efficient matrix product state (MPS) representation Akhtar, You (2020)

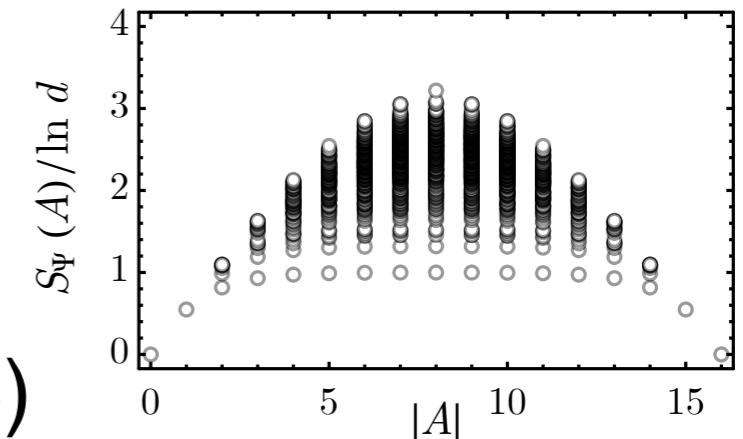
$$W_{\mathcal{E}_\sigma}[s] \propto \text{Tr} \left(\prod_i W_i^{s_i} \right)$$



since $|W_{\mathcal{E}_\sigma}\rangle$ is *almost* always an **area-law** state, even if the underlying physical state σ is highly-entangled.



Entanglement growth in random unitary circuits by TEBD method



Other Related Works

- Relation to other approaches
 - The entanglement feature formalism is in line with the statistical mechanics / conformal field theory approach towards entanglement dynamics

Znidaric (2008); Hayden, Nezami, Qi, Thomas, Walter, Yang (2016); Nahum, Ruhman, Vijay, Haah (2017); Chan, Luca, Chalker (2018); Vasseur, Potter, You, Ludwig (2019); Zhou, Nahum (2019); Lensky, Qi (2019); Bao, Choi, Altman (2019); Jian et.al. (2019); Li, Chen, Ludwig, Fisher (2020) ...
 - It reproduces results on random unitary circuits (RUC) and Brownian dynamics.

Nahum, Vijay, Haah (2017); Zhou, Xiao (2019); Xu, Swingle (2019)
 - It is consistent with entanglement line tension and capillary wave theories.

Jonay, Huse, Nahum (2018); Mezei (2018); Skinner, Ruhman, Nahum (2019); Li, Fisher (2020)

Locally Scrambled Shadow Tomography (LSST)

- How are these useful in classical shadow tomography?
- In the randomized measurement, if the prior classical snapshots \mathcal{E}_σ are locally scrambled, the **measurement channel** is given by the **entanglement feature** of \mathcal{E}_σ

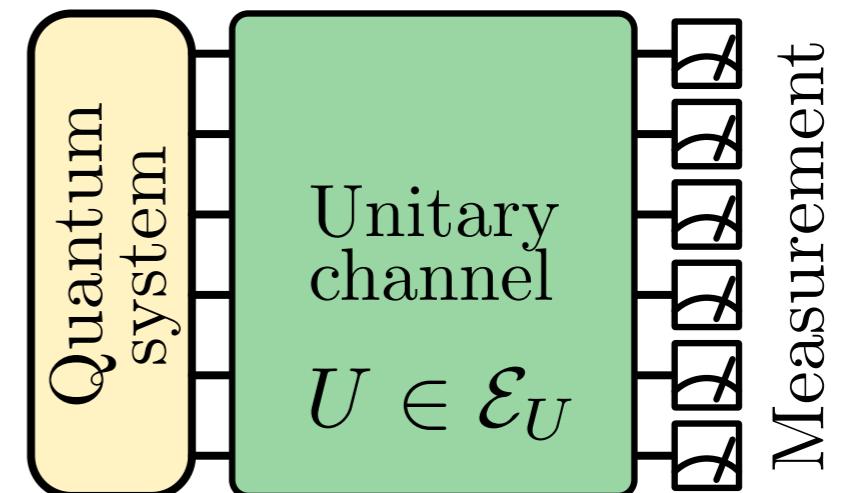
$$\mathcal{M}[\rho] = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_\sigma} \hat{\sigma} \text{Tr}(\hat{\sigma} \rho) 2^N$$

$$= \mathbb{E}_{V \in \text{U}(2)^N} \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_\sigma} (V^\dagger \hat{\sigma} V) \text{Tr}(V^\dagger \hat{\sigma} V \rho) 2^N$$

$$= \sum_A 2^{2N-|A|} \langle A | \hat{W}_1^{-1} | W_{\mathcal{E}_\sigma} \rangle \rho_A$$

$$= \sum_A m_A \rho_A$$

Coefficients only depend
on entanglement feature



$$\rho_A = (\text{Tr}_{\bar{A}} \rho) \otimes (\mathbf{1}_{\bar{A}} / 2^{|\bar{A}|})$$

Reduced density matrix
(embedded in the full Hilbert space)

Locally Scrambled Shadow Tomography (LSST)

- Inverting the measurement channel gives the **reconstruction map**

$$\sigma = \mathcal{M}[\rho] = \sum_A m_A \rho_A \quad \Leftrightarrow \quad \rho = \mathcal{M}^{-1}[\sigma] = \sum_A r_A \sigma_A$$

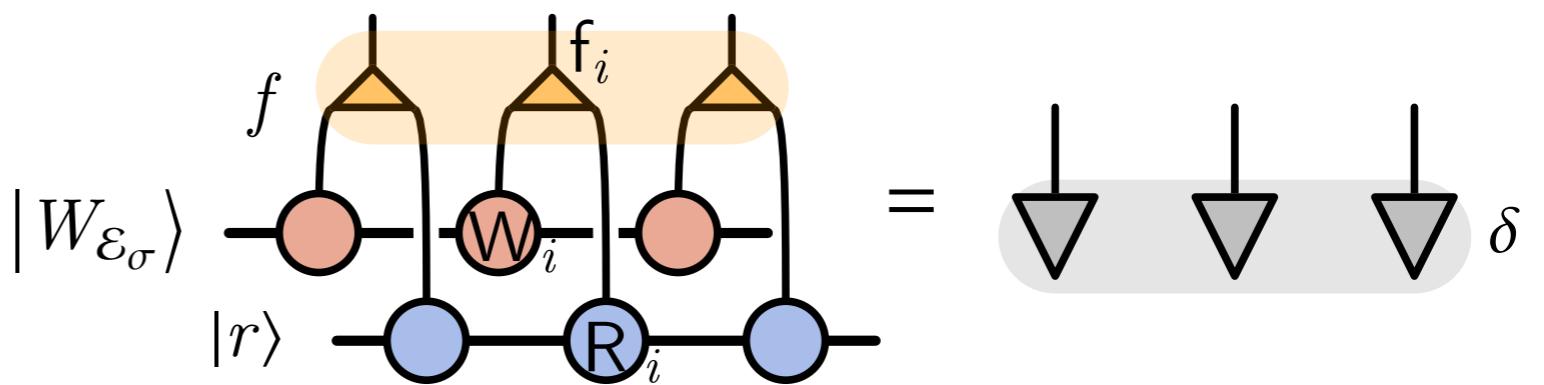
- Reconstruction coefficients given by entanglement features $W_{\mathcal{E}_\sigma}$ of the prior snapshot ensemble

$$r_A = (-1)^{|A|} \sum_{A \subseteq C} \frac{3^{|C|}}{\sum_{B \subseteq C} (-2)^{|B|} W_{\mathcal{E}_\sigma, B}}$$

Bu, Koh, Garcia, Jaffe (2022)

- These coefficients r_A can be encoded as an MPS too

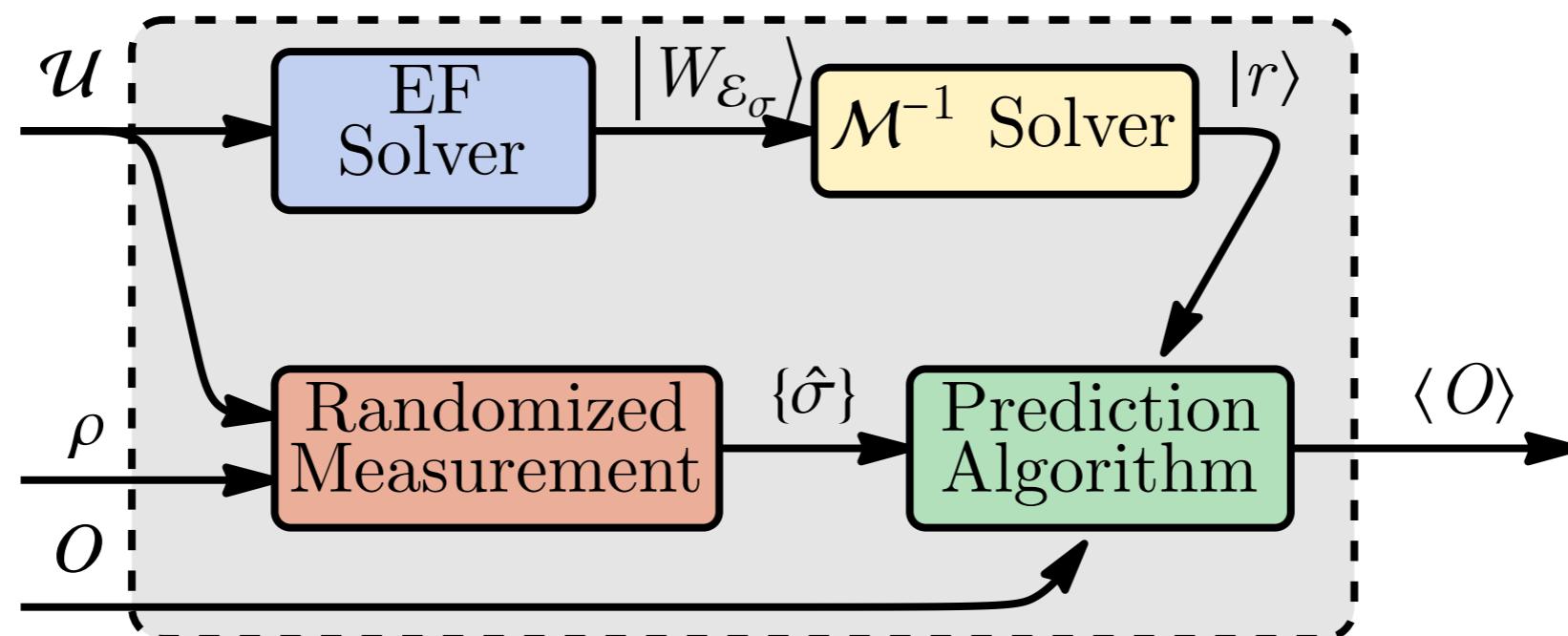
$$|r\rangle = \sum_A r_A |A\rangle$$



A.A.Akhtar, H.-Y. Hu, Y.-Z.You. arXiv:2209.02093 (2022)

Locally Scrambled Shadow Tomography (LSST)

- Given a randomized measurement scheme, specified by the **prior** snapshot ensemble \mathcal{E}_σ
 - Compute $|W_{\mathcal{E}_\sigma}\rangle$ by **entanglement dynamics**
 - Solve **reconstruction coefficients** $|r\rangle$ from $|W_{\mathcal{E}_\sigma}\rangle$
- Collect **posterior** snapshot ensemble $\mathcal{E}_{\sigma|\rho}$ from measurement
- Predict: $\langle O \rangle_\rho = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \text{Tr}(\mathcal{M}^{-1}[\hat{\sigma}]O) = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \sum_A r_A \text{Tr}(\hat{\sigma}_A O)$



Locally Scrambled Shadow Norm

- In practice, the expectation value is estimated on a **finite number** M of samples

$$\langle O \rangle = \frac{1}{M} \sum_{m=1}^M \text{Tr}(\mathcal{M}^{-1}[\hat{\sigma}^{(m)}]O)$$

- The **variance** of the finite-sample mean

$$\text{var } \langle O \rangle \lesssim \|O\|_{\text{shd}}^2/M \quad \text{Huang, Kueng, Preskill (2020)}$$

is controlled by the operator **shadow norm** $\|O\|_{\text{shd}}^2$ (assuming O is traceless)

- Shadow norm = single-shot variance (maximized over all possible states ρ)

$$\|O\|_{\text{shd}}^2 = \max_{\rho} \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} (\text{Tr } \mathcal{M}^{-1}[\hat{\sigma}]O)^2$$

Locally Scrambled Shadow Norm

- An alternative definition: **locally scrambled shadow norm**

$$\begin{aligned}\|O\|_{\mathcal{E}_\sigma}^2 &= \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_\sigma} (\mathrm{Tr} \mathcal{M}^{-1}[\hat{\sigma}] O)^2 \\ &= \sum_A 2^{|A|-2N} r_A(W_O)_A\end{aligned}$$

- Depends only on (i) the measurement scheme r_A and (ii) the operator entanglement

$$(W_O)_A = \mathrm{Tr}_A (\mathrm{Tr}_{\bar{A}} O)^2.$$

- Advantage: easy to compute (no maximization over states)
- Given the sample number M , determine the typical variance of the estimation $\mathrm{var} \langle O \rangle \simeq \|O\|_{\mathcal{E}_\sigma}^2 / M$
- Given the desired prediction accuracy $\mathrm{var} \langle O \rangle \leq \epsilon^2$, roughly bound the number of experiments $M \gtrsim \|O\|_{\mathcal{E}_\sigma}^2 / \epsilon^2$

Scalable Tomography on Clifford Circuits

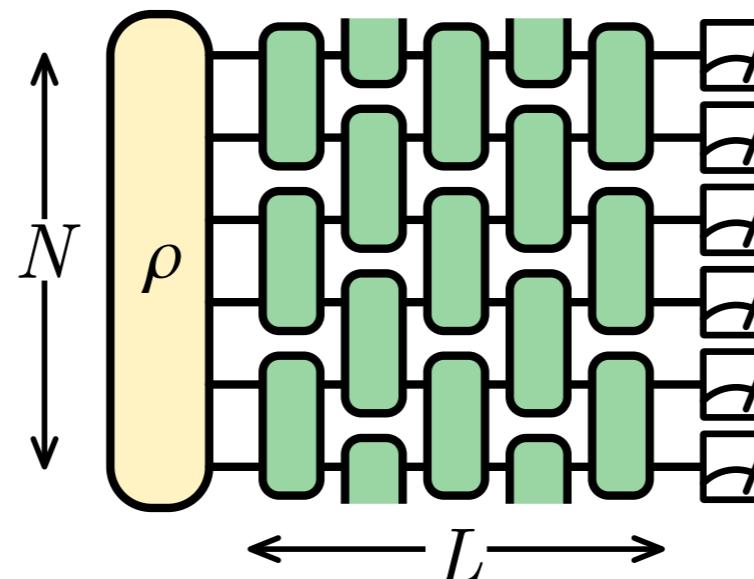
- Test states:

$$\rho_{\text{GHZ}} = |\text{GHZ}\rangle\langle\text{GHZ}| \quad |\text{GHZ}\rangle = \frac{1}{\sqrt{2}}(|00\cdots\rangle + |11\cdots\rangle)$$

$$\rho_{\text{ZXZ}} = |\text{ZXZ}\rangle\langle\text{ZXZ}| \quad Z_{i-1}X_iZ_{i+1}|\text{ZXZ}\rangle = |\text{ZXZ}\rangle$$

(Cluster state)

- Randomized measurement scheme:



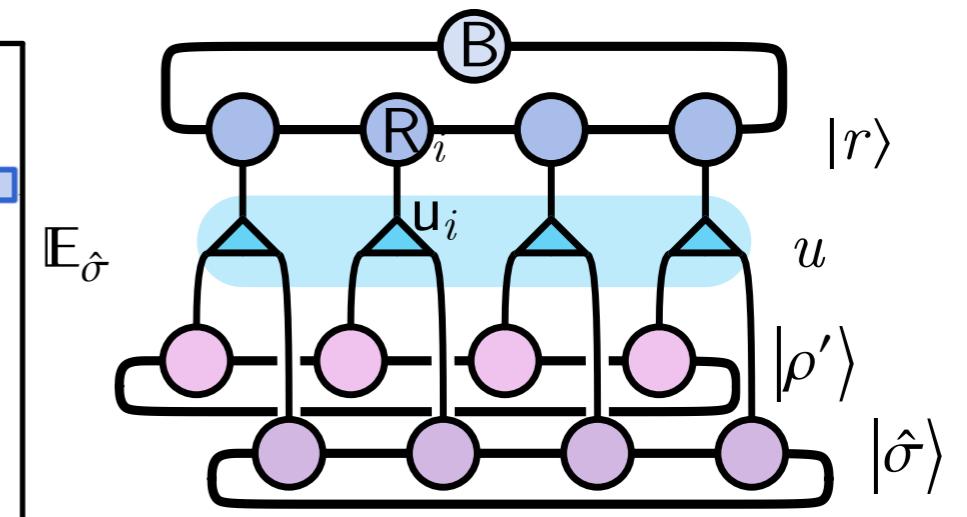
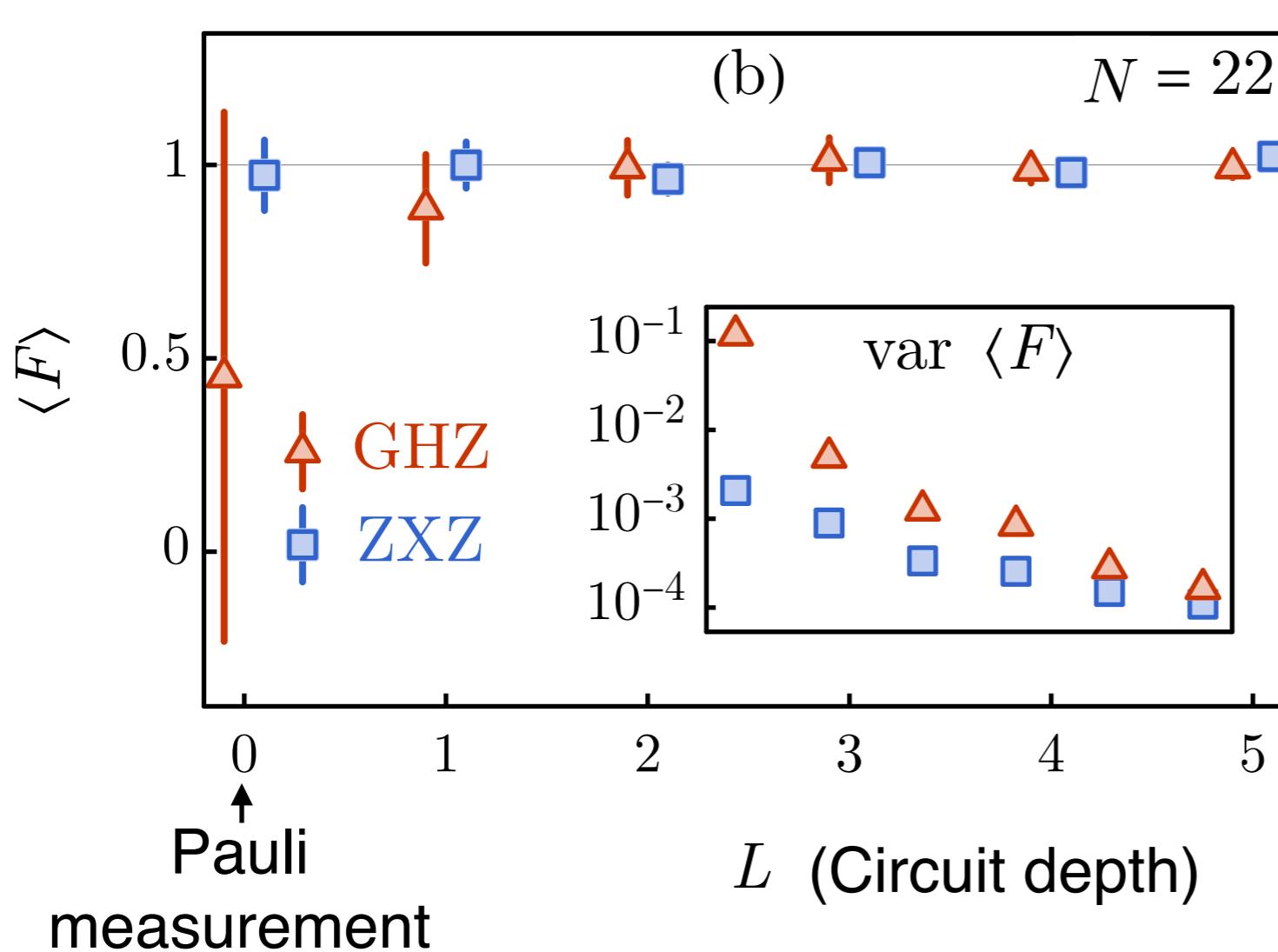
- Brick-wall arrangement of 2-qubit **random Clifford** gates
- Width (number of qubits): N
- Depth (number of layers): L

Fidelity Estimation

- Task 1: Estimate the fidelity (state overlap) between the reconstructed state and the original state

$$F(\rho, \rho') = \text{Tr}(\rho\rho')$$

(For pure states)

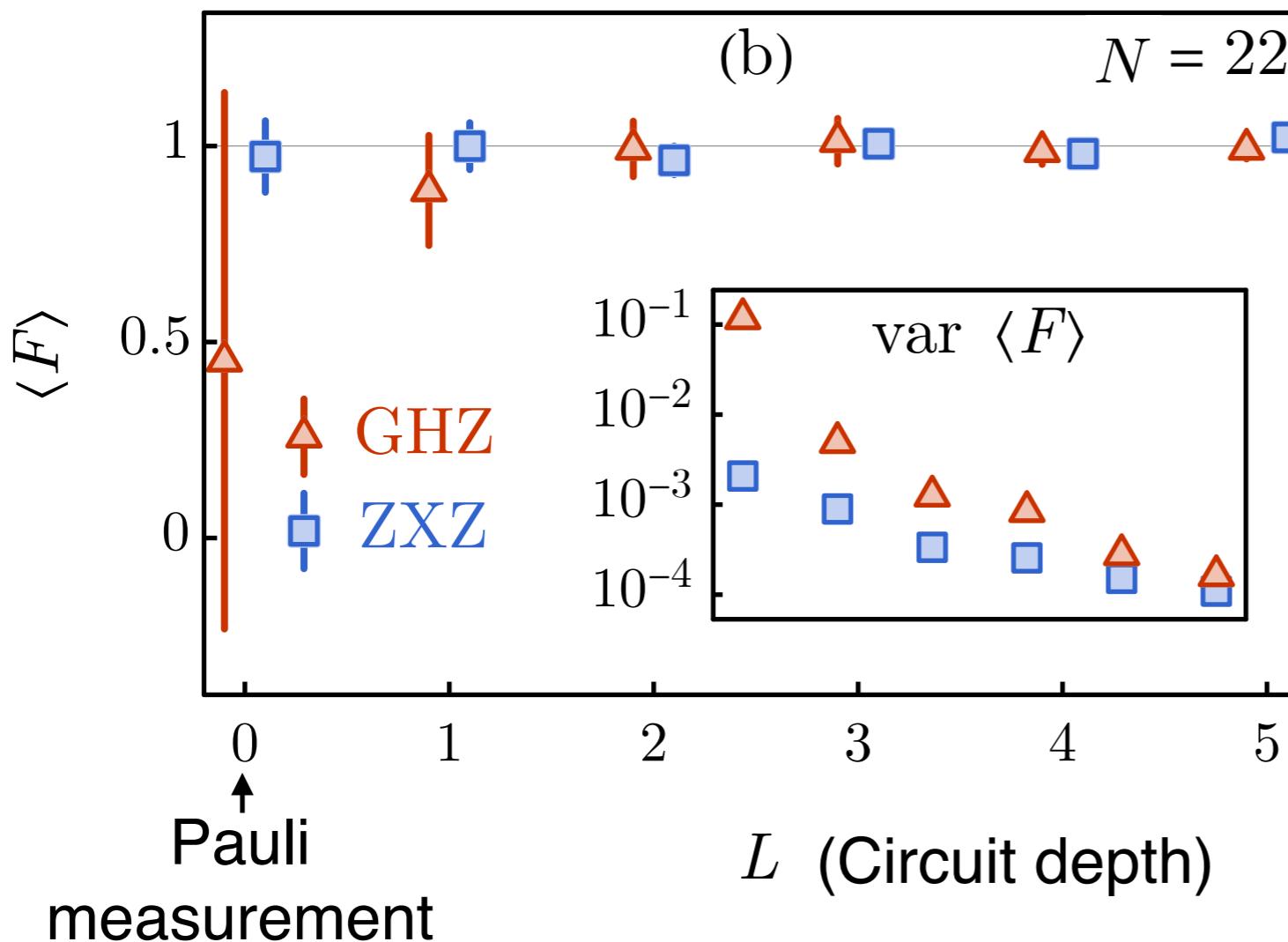


Efficient MPS-based
post-processing

Fidelity Estimation

- Task 1: Estimate the fidelity (state overlap) between the reconstructed state and the original state

$$F(\rho, \rho') = \text{Tr}(\rho\rho') \quad (\text{For pure states})$$

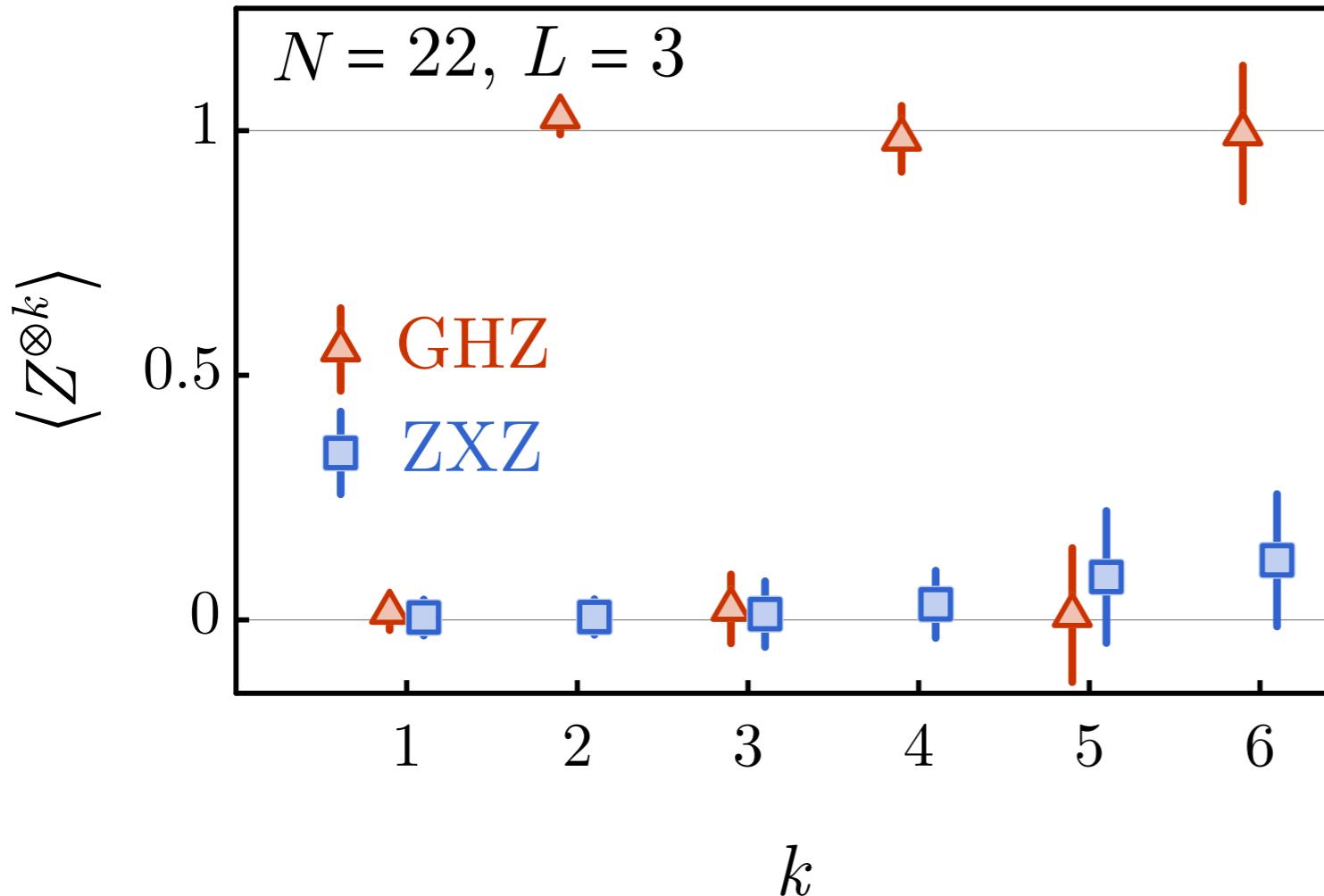


- Scalability (22 qubits)
- Reconstruction is unbiased for all circuit depth
- $M = 50k$ samples: variance is reduced for deeper circuits — quantum information scrambling helps!

Pauli Operator Estimation

- Task 2: Estimate the expectation value of a Pauli string operator

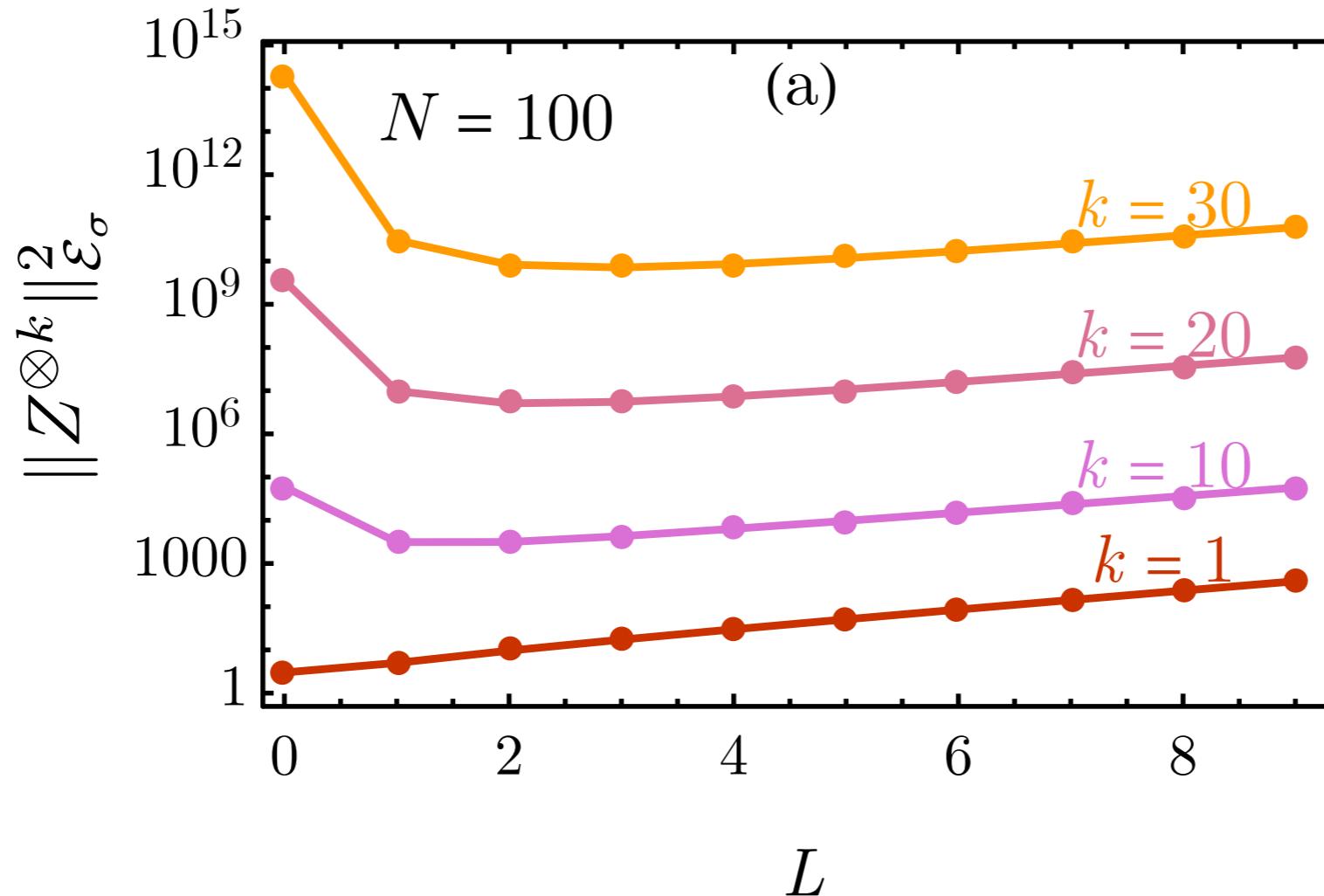
$$Z^{\otimes k} := \prod_{i=1}^k Z_i$$



- Estimations are unbiased (converge to the ground truth)
- Variance (shadow norm) increases with the weight (size) of the Pauli string
 - Note: For Pauli measurements ($L = 0$)
$$\|Z^{\otimes k}\|_{\mathcal{E}_\sigma}^2 \sim 3^k$$

Pauli Operator Estimation

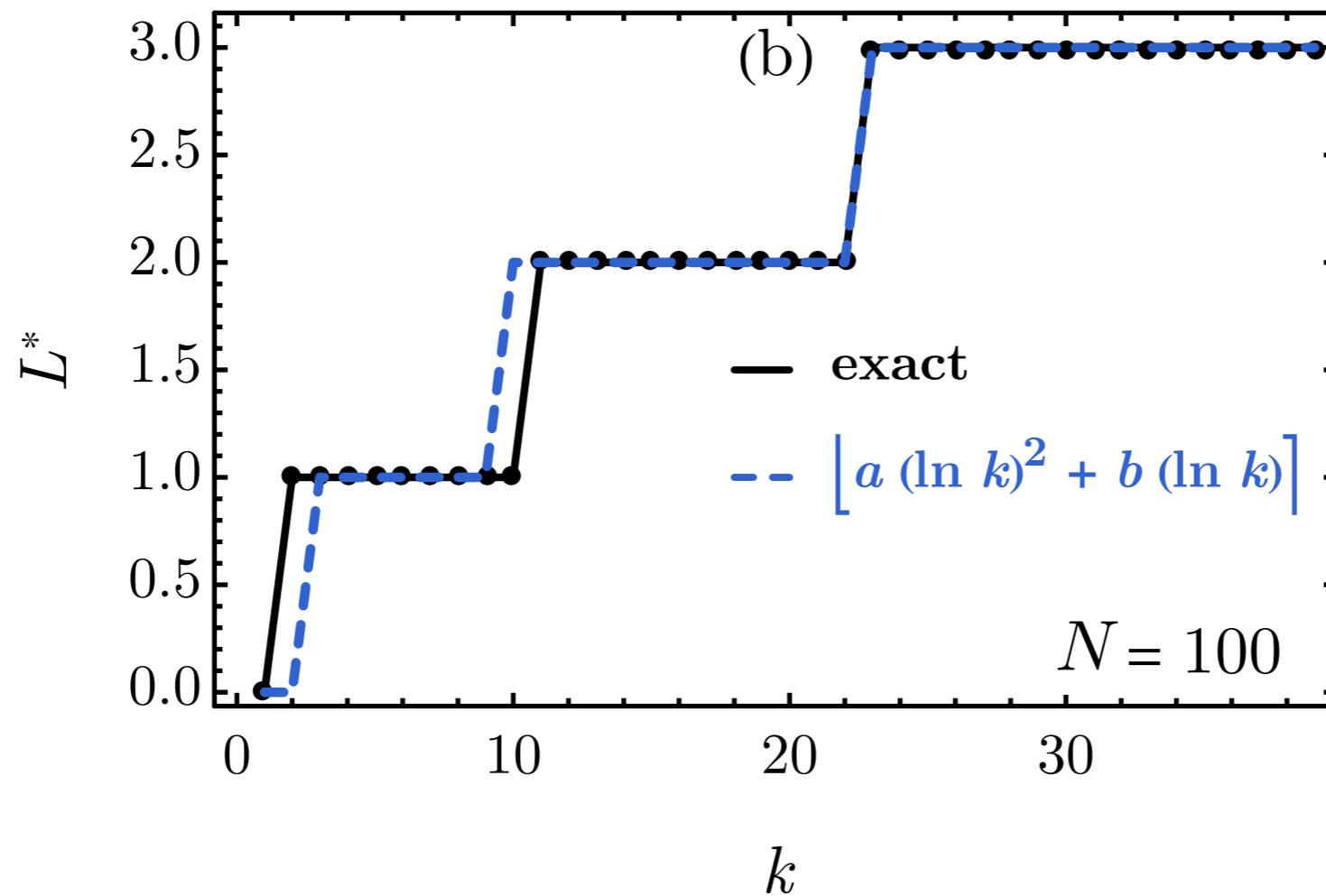
- Task 2: Estimate the expectation value of a Pauli string operator $Z^{\otimes k} := \prod_{i=1}^k Z_i$



- Given the size k of the Pauli string, there is an **optimal circuit depth** L^* minimizing the shadow norm (sample complexity)

Pauli Operator Estimation

- Task 2: Estimate the expectation value of a Pauli string operator $Z^{\otimes k} := \prod_{i=1}^k Z_i$

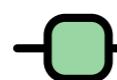


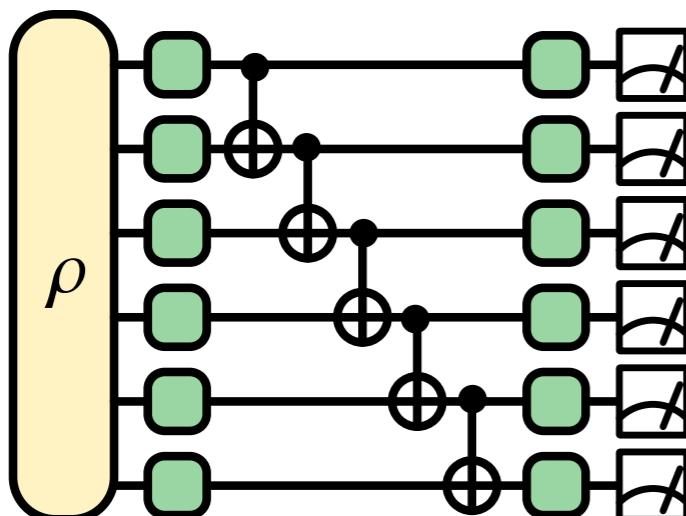
- The **optimal circuit depth** L^* scales with k poly-logarithmically

$$L^* \sim a(\ln k)^2 + b \ln k$$

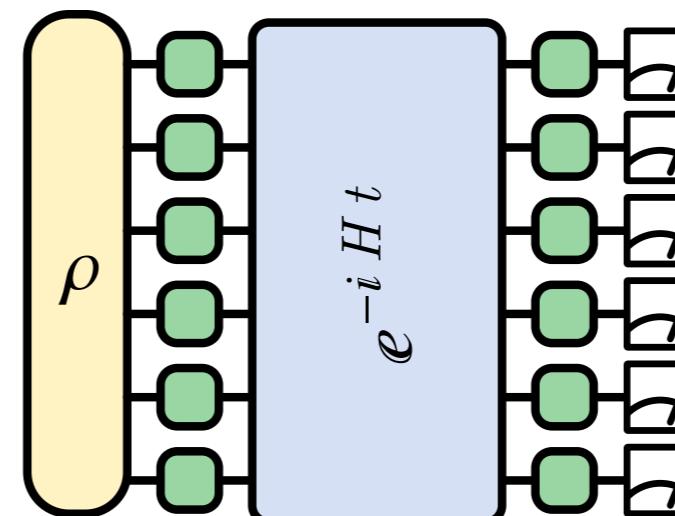
Classical Shadow Tomography on NISQ Devices

- Random Clifford gate may still be challenge for **noisy intermediate-scale quantum** (NISQ) devices
- Our approach enables very flexible design of randomized measurement protocol
 - Wrap any circuit / quantum dynamics with local scramblers
→ Locally scrambled unitary ensemble

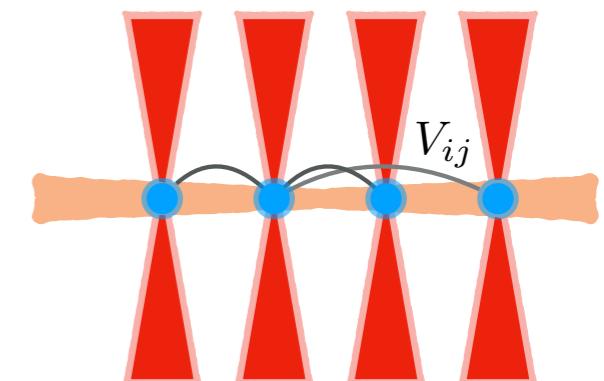
 = single-qubit random Clifford gate



Locally scrambled
CNOT circuit

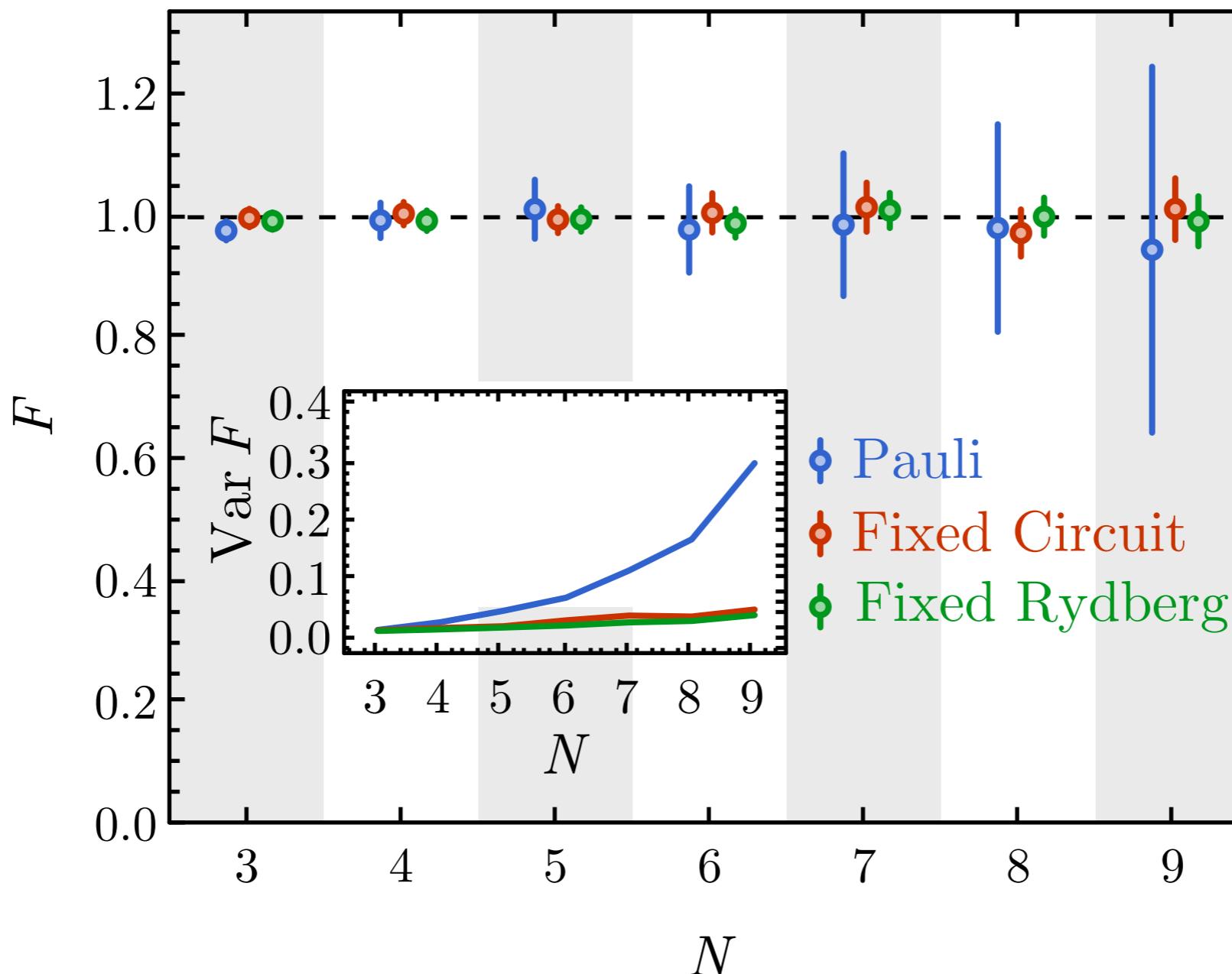


Locally scrambled
Hamiltonian dynamics



Classical Shadow Tomography on NISQ Devices

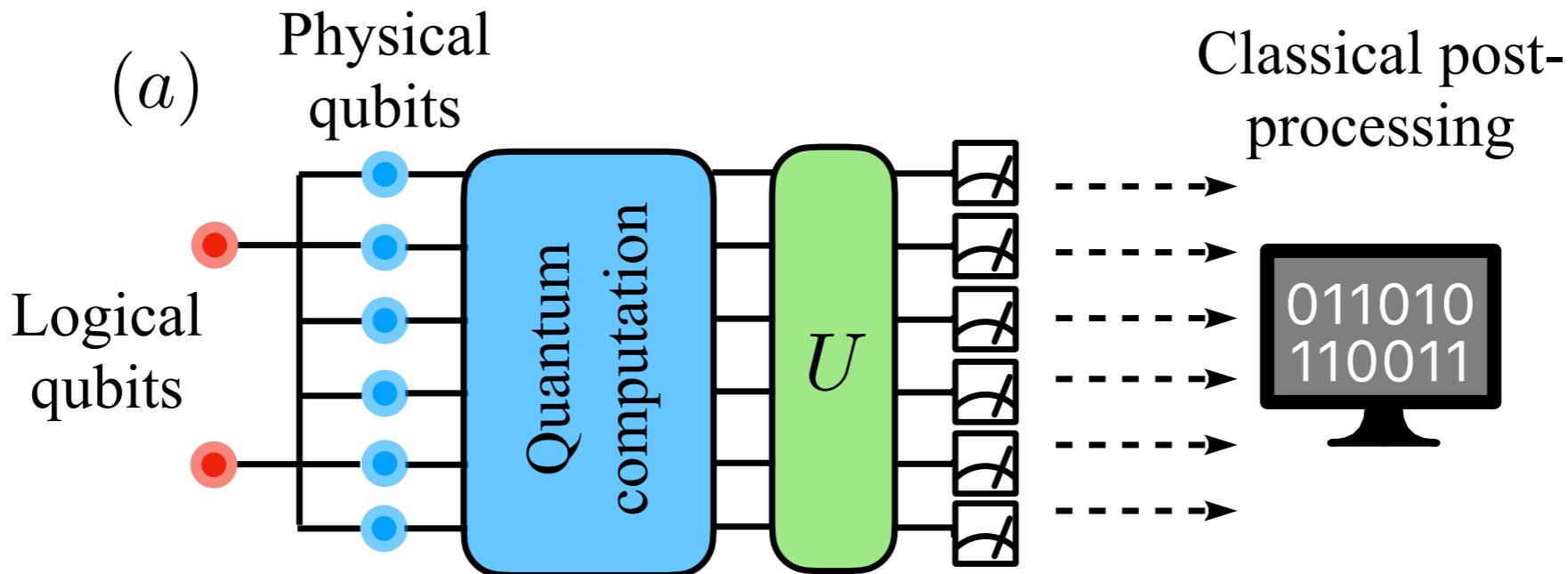
- They can just achieve the same effect in reducing the estimation variance as a shallow random circuit!



LSST for Quantum Error Mitigation

- Classical shadow tomography has many amazing applications, one example is to implement **code subspace projection in quantum error mitigation (QEM)**

Hu, LaRose, You, Rieffel, Wang (2022); Seif, Cian, Zhou, Chen, Jiang (2022)

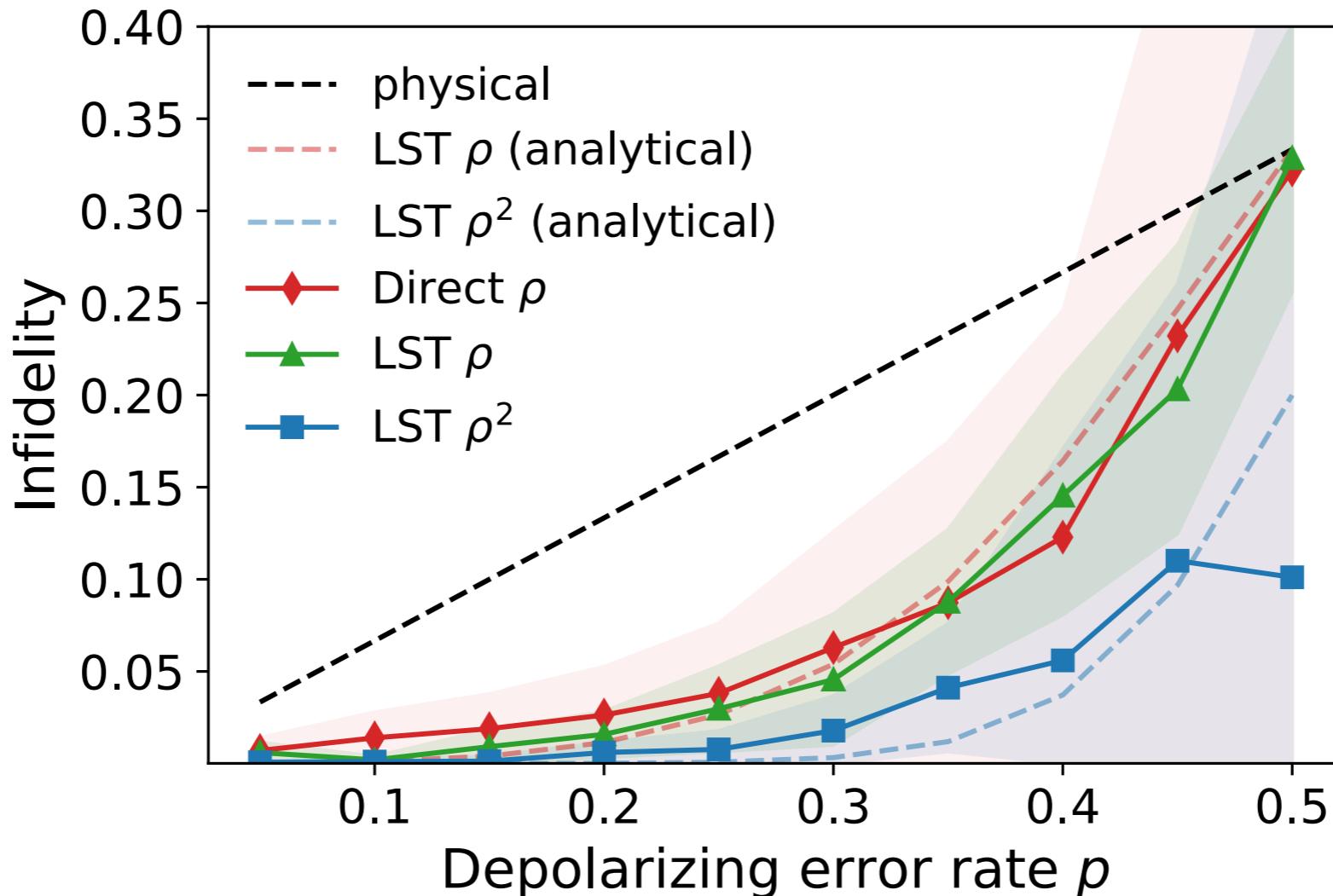


$$\langle O \rangle_{\text{QEM}} = \frac{\text{Tr}(\Pi \rho \Pi^\dagger O)}{\text{Tr}(\Pi \rho \Pi^\dagger)}$$

$$\text{Tr}(\Pi \rho \Pi^\dagger O) = \mathbb{E}_{\hat{\sigma} \in \mathcal{E}_{\sigma|\rho}} \sum_A r_A \text{Tr}(\hat{\sigma}_A \Pi^\dagger O \Pi)$$

LSST for Quantum Error Mitigation

- Assume each physical bit is subject to depolarization error
- QEM can reduce the infidelity when the error rate is small
- Logical shadow tomography (LST) demonstrates superior sample efficiency (small variance)



Summary

- Combining the quantum entanglement dynamics and classical shadow tomography, we introduce the **locally scrambled shadow tomography** (LSST), which is
 - **Scalable** (efficient classical post-processing)
 - **Flexible** (arbitrary circuit structure / quantum dynamics)
 - **NISQ friendly** (shallow circuits, simple gates, available devices)
- We expect our approach to have broad applications in many quantum information processing tasks (e.g. quantum error mitigation)

Thanks for your attention!