Robustness of Classical Shadows under Gate-Dependent Noise via Readout Error Deconvolution

Guedong Park

IRC NextQuantum and Department of Physics and Astronomy, Seoul National University, Seoul, 08826, Korea

Index

1. Preliminaries

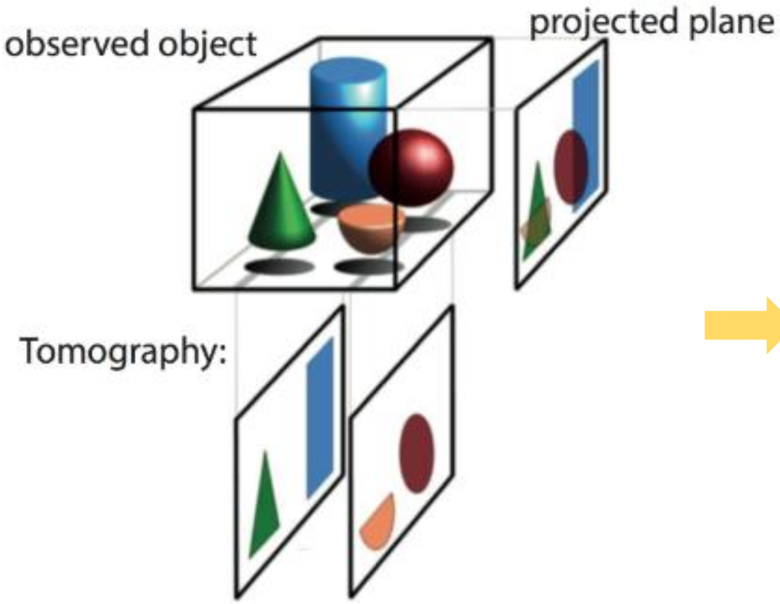
2. Read-out error of Clifford measurement

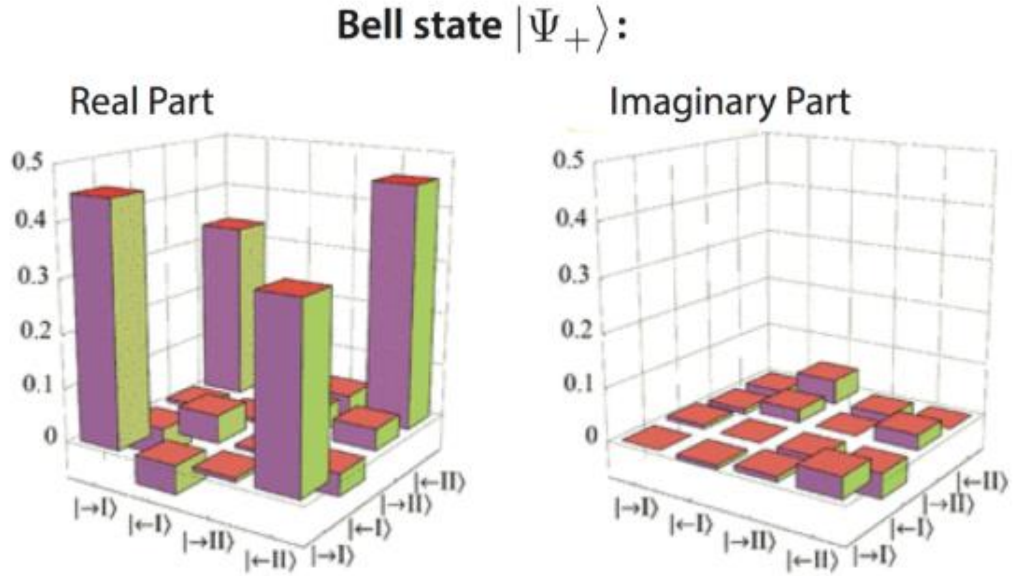
3-1. (Result) Robust classical shadow under read-out error
3-2. (Result) Graph state benchmarking and high-order stability

4. Summary

Preliminaries

[1] J. Acharya, arXiv:2502.18170 (2025)

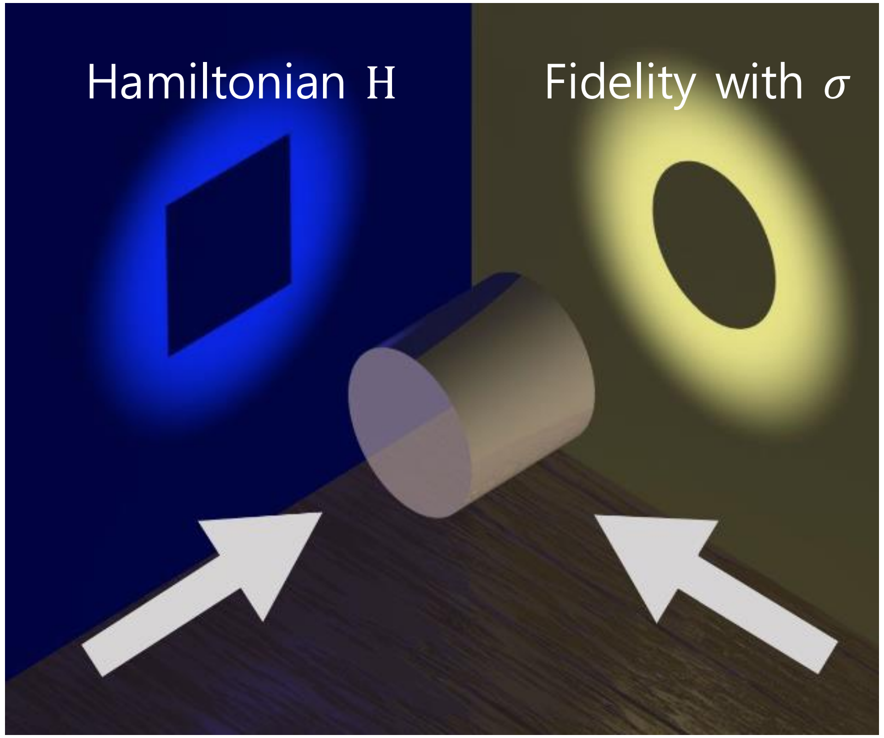




Quantum tomography is a task to estimate the density matrix structure of unknown input state

It generally requires exponentially many sampling copies ($O(4^n)$)[1] by the number of qubits n.

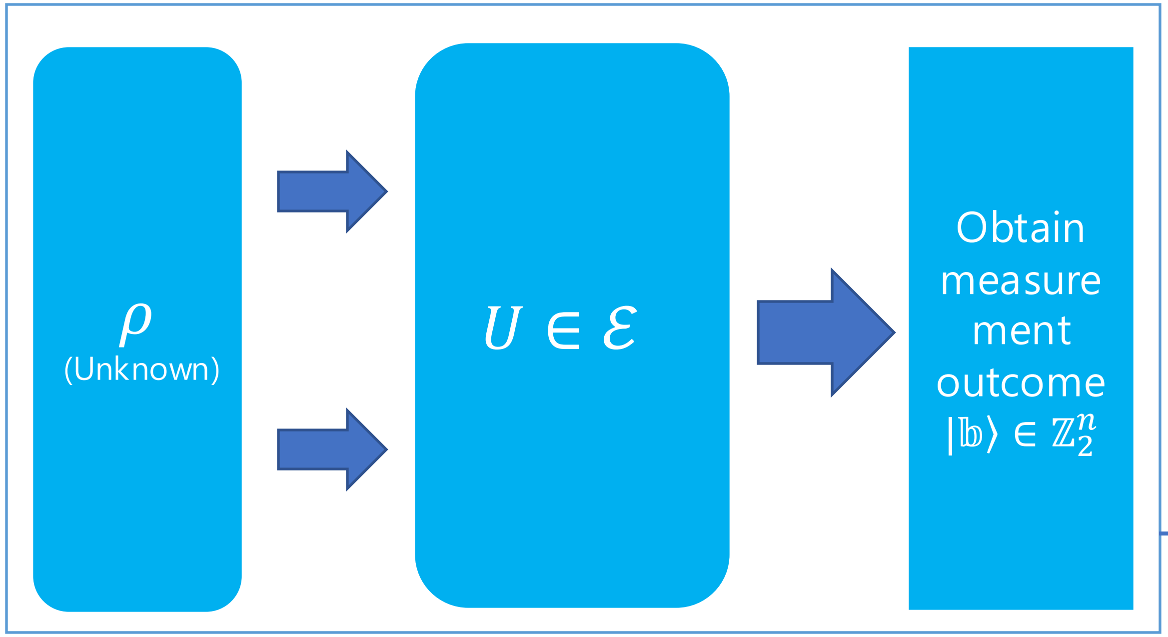
Shadow tomography[1] seeks a method to estimate the designated M number of physical properties of the unknown quantum states, hence reducing the required sampling complexity.



For the linear property, the task is to estimate $\mathrm{tr}(\rho O)$ for the given observable O, not the whole ρ.

[1] S. Aaronson, arXiv:1711.01053 (2017)

<Shadow tomography> ($\mathcal{E}$: unitary ensemble)



$U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger$

snap shot

$$\mathcal{M}(\rho) = \mathbb{E}_{U\in\mathcal{E}} \sum_{\mathbb{b}} \langle\mathbb{b}|U\rho U^\dagger|\mathbb{b}\rangle \cdot U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger$$

Measurement channel

$$\mathrm{tr}(O\rho) = \mathbb{E}_{U\in\mathcal{E}} \sum_{\mathbb{b}} \langle\mathbb{b}|U\rho U^\dagger|\mathbb{b}\rangle tr\{O\mathcal{M}^{-1}(U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger)\}$$

$\mathcal{M}^{-1}(U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger)$: classical shadow, $\mathcal{M}^{-1}$ exists for IC POVM.

$$\mathrm{tr}(O\rho) = \mathbb{E}_{U\in\mathcal{E}} \sum_{\mathbb{b}} \langle \mathbb{b}|U\rho U^\dagger|\mathbb{b}\rangle tr\{O\mathcal{M}^{-1}(U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger)\}$$

$\mathcal{M}^{-1}(U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger)$: classical shadow.

<Algorithm> (Sampling copies, $\mathrm{N} = RK$)

(1) We randomly choose $U \in \mathcal{E}$.

(2) Enact U to ρ

(3) Measure with computational basis to obtain $|\mathbb{b}\rangle \in \mathbb{Z}_2^n$

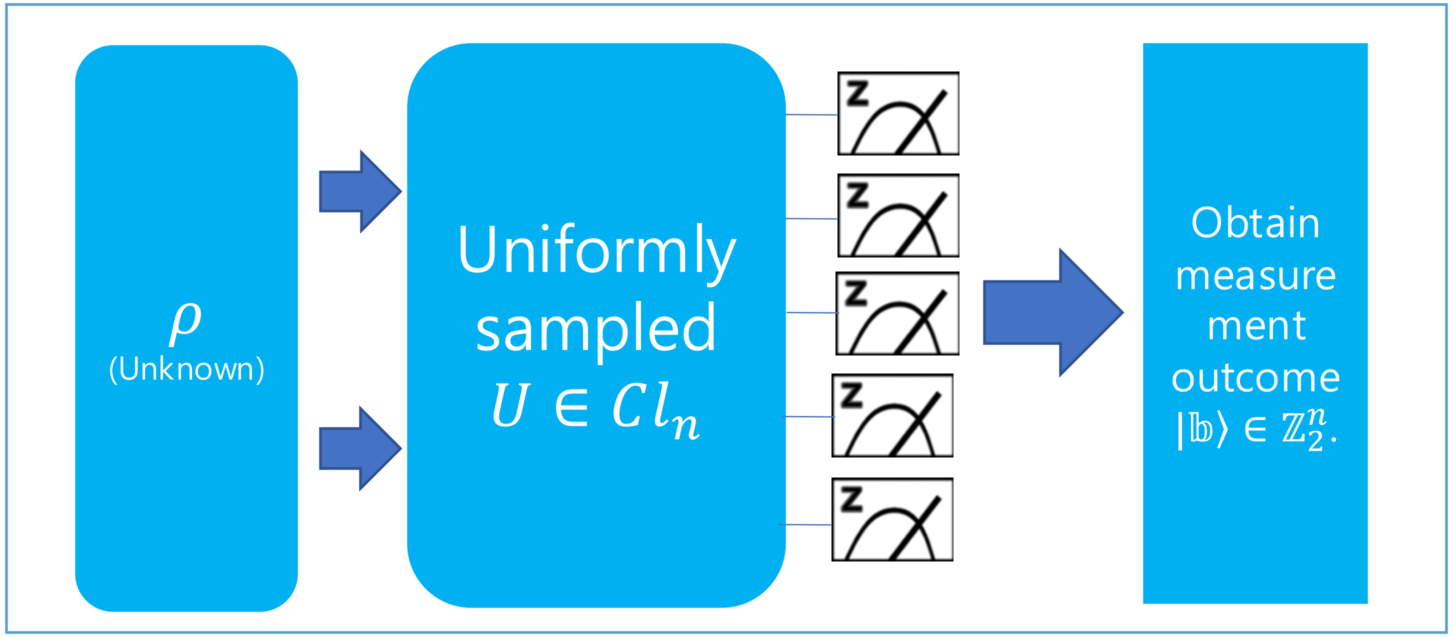
(4) Take the estimator $m_i = tr\{O\mathcal{M}^{-1}(U|\mathbb{b}\rangle\langle\mathbb{b}|U^\dagger)\}$

(5) Repeat (1)~(4) to get $m_1, m_2, \dots m_K$ and set $\hat{O}_j = \frac{1}{K}\sum_{i=1}^{K} m_i$

(6) Repeat (5) R times and conclude $\hat{O} = median\{\hat{O}_1, \hat{O}_2, \dots, \hat{O}_R\}$.

A representative shadow tomography is random Clifford tomography[1].

$$\mathcal{M}^{-1}(O) = (2^n + 1)O - \mathrm{tr}(O)I$$



Estimator m is,

$$m = (2^n + 1)\langle \mathbf{b}|UOU^\dagger|\mathbf{b}\rangle - \mathrm{tr}(O)$$

Shadow norm

$$\|O\|_{sh}^2 \le \|O_0^2\|_\infty \le \mathcal{O}(\mathrm{tr}(O_0^2))$$

$$O_0 \equiv O - \frac{\mathrm{tr}(O)}{2^n} I$$

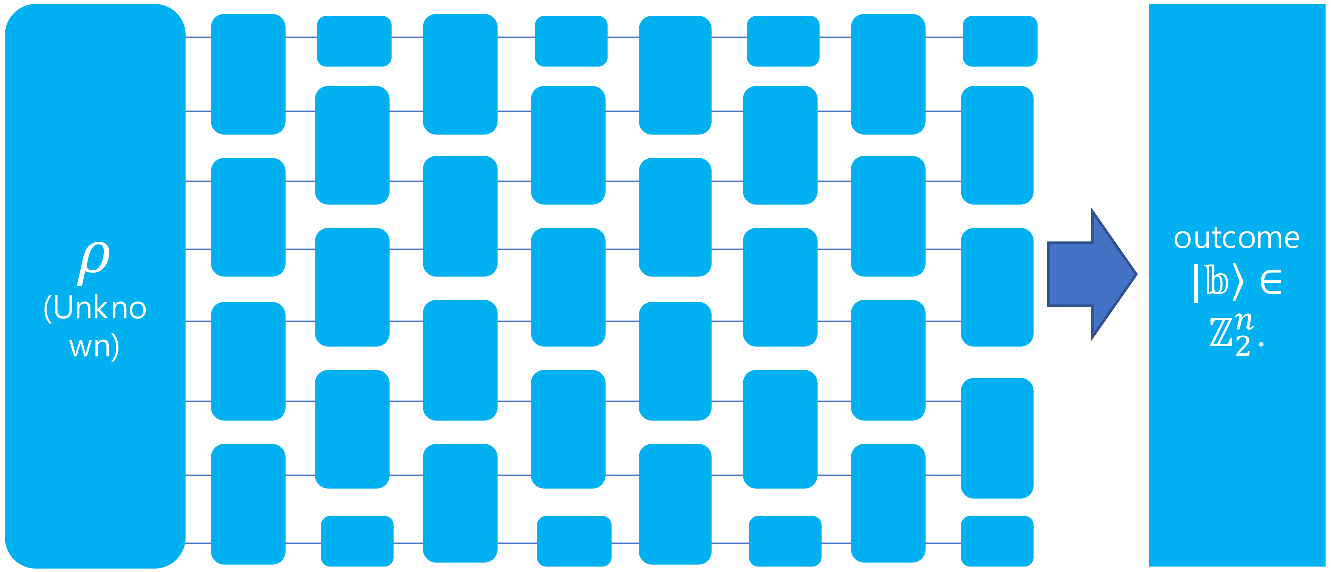
Sampling copies

$$N \le \mathcal{O}\left(\frac{\|O\|_{sh}^2}{\epsilon^2}\log(\delta_f^{-1})\right)$$

[1] Huang,H-Y et al. Nat. Phys 16 (10),1050-1057 (2020)

Random Clifford tomography needs $O(n)$ −depth neighboring Clifford blocks

However, $O(\log(n))$-depth Clifford blocks are sufficient[1] for k-local observable estimation



log(n)-depth blocks

$\forall P \in \mathcal{P}_n$ d: circuit depth

$$\mathcal{M}(P) = t_{P,d} P \;\Rightarrow\; \mathcal{M}^{-1}(P) = \frac{1}{t_{P,d}} P$$

$$t_{P,d} = \mathrm{Prob}_{V \in \mathrm{Cl}_{n,d}}(VPV^\dagger \in \mathcal{Z}) > 0$$

$$\|P\|_{\mathrm{sh}}^2 = t_{P,d}^{-2} \lesssim \mathcal{O}(3^{|P|})$$

[1] C. Bertoni et al., Phys. Rev. Lett. 133,020602

However, random Clifford sampling circuit is very noisy in current setups.

$$m = (2^n + 1)\langle \mathbf{b}|UOU^\dagger|\mathbf{b}\rangle - \mathrm{tr}(O)$$

Small noise in U may rise much bigger bias of the estimation

Recently, there have been many researches about Error mitigation for the classical shadow[1,2].

⟨ ⟩uantum
the open journal for quantum science

PAPERS PERSPECTIVES

Classical Shadows With Noise

Dax Enshan Koh[1,2] and Sabee Grewal[2,3]

[1]Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, #16-16 Connexis, Singapore 138632, Singapore
[2]Zapata Computing, Inc., 100 Federal Street, 20th Floor, Boston, Massachusetts 02110, USA
[3]Department of Computer Science, The University of Texas at Austin, Austin, TX 78712, USA

Published: 2022-08-16, volume 6, page 776
Eprint: arXiv:2011.11580v2
Doi: https://doi.org/10.22331/q-2022-08-16-776
Citation: Quantum 6, 776 (2022).

PRX Quantum

Highlights Recent Accepted Collections Authors Referees About Scope

OPEN ACCESS

Robust Shadow Estimation

Senrui Chen[1,2,3,†], Wenjun Yu[1,†], Pei Zeng[1,3,*], and Steven T. Flammia[4]

Show more

PRX Quantum 2, 030348 – Published 22 September, 2021

DOI: https://doi.org/10.1103/PRXQuantum.2.030348

[1] D. E. Koh et al., Quantum 6 776 (2022)
[2] S. Chen et al., PRX Quantum 2, 030348 (2021)

<Previous works>


PRX Quantum

Highlights Recent Accepted Collections Authors Referees About Scope

OPEN ACCESS

Robust Shadow Estimation

Senrui Chen[1,2,3,†], Wenjun Yu[1,†], Pei Zeng[1,3,*], and Steven T. Flammia[4]

Show more

PRX Quantum **2**, 030348 – **Published 22 September, 2021**

DOI: https://doi.org/10.1103/PRXQuantum.2.030348


nature communications

Explore content About the journal Publish with us

nature > nature communications > articles > article

Article | Open access | Published: 26 March 2025

Demonstration of robust and efficient quantum property learning with shallow shadows

Hong-Ye Hu, Andi Gu, Swarnadeep Majumder, Hang Ren, Yipei Zhang, Derek S. Wang, Yi-Zhuang You, Zlatko Minev, Susanne F. Yelin & Alireza Seif

Nature Communications **16**, Article number: 2943 (2025) | Cite this article

8519 Accesses | **5** Citations | **63** Altmetric | Metrics


PRX Quantum

Highlights Recent Accepted Collections Authors Referees About Scope Editorial Team RSS

OPEN ACCESS

Quantum Error Mitigated Classical Shadows

Hamza Jnane[1,2,†], Jonathan Steinberg[3,4,†], Zhenyu Cai[1,2], H. Chau Nguyen[3], and Bálint Koczor[5,1,2,*]

Show more

PRX Quantum **5**, 010324 – **Published 9 February, 2024**

DOI: https://doi.org/10.1103/PRXQuantum.5.010324

Randomized benchmarking of Pauli transfer matrix (PTM)[1,2,3]

Probabilistic error cancellation (PEC)[4]>

Efficient noise benchmarking for shallow circuit is known[2].

Full random Clifford noise benchmarking can be efficient, but assumes gate-independent noise[1].
PEC exponentially increases by the gate count[4].

[1] S. Chen et al., PRX Quantum 2, 030348 (2021)
[2] H. T. Hu et al., Nat. Comm. 16, 2943 (2025)
[3] K. Bu et al., npj Quantum 10, 6 (2024)
[4] PRX Quantum 5, 010324 (2024)

<Problem>

We assume single qubit unitary is free. Noise is gate-dependent.

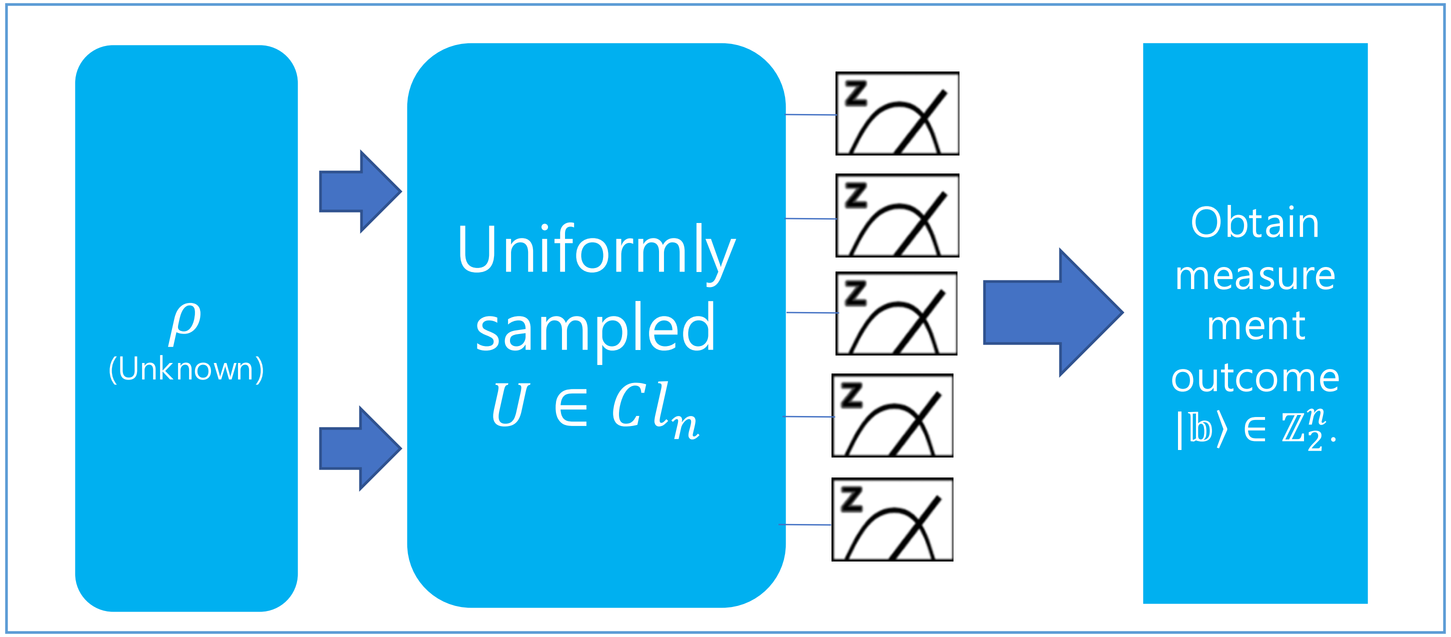
(1) Sample-improved unbiased mitigation scheme of noisy (shallow or full) Clifford shadow?

(2) High-order stability of Clifford shadow under unknown noise? [1]

$$|\mathbb{E}[\hat{o}] - \langle O \rangle| \leq \min\{\|O\|_2, \|O\|_{\mathrm{st}}\} \max_{a \in \mathbb{F}_2^{2n}} |1 - \bar{\lambda}_a|^{w(>1)?}$$

[1] R. Brieger Phys. Rev. Lett. 134, 090801 (2025)

Read-out error of Clifford measurement



$\mathcal{N}_U$: (Unital) Noise channel after unitary U.

Effective noise: Read-out error.

$$P^{(U)}_{\mathbf{b},\mathbf{a}} \equiv \langle \mathbf{b}|\mathcal{N}_U(|\mathbf{a}\rangle\langle\mathbf{a}|)\mathbf{b}\rangle$$

(Bi-stochastic)

$$p^{(U)}_{\mathbf{b}|\mathbf{a}} \equiv P^{(U)}_{\mathbf{b},\mathbf{a}}$$ Read-out errors[1]

$g^{(U)}$: Ideal measurement distribution, $\mu^{(U)}$: Noisy measurement distribution

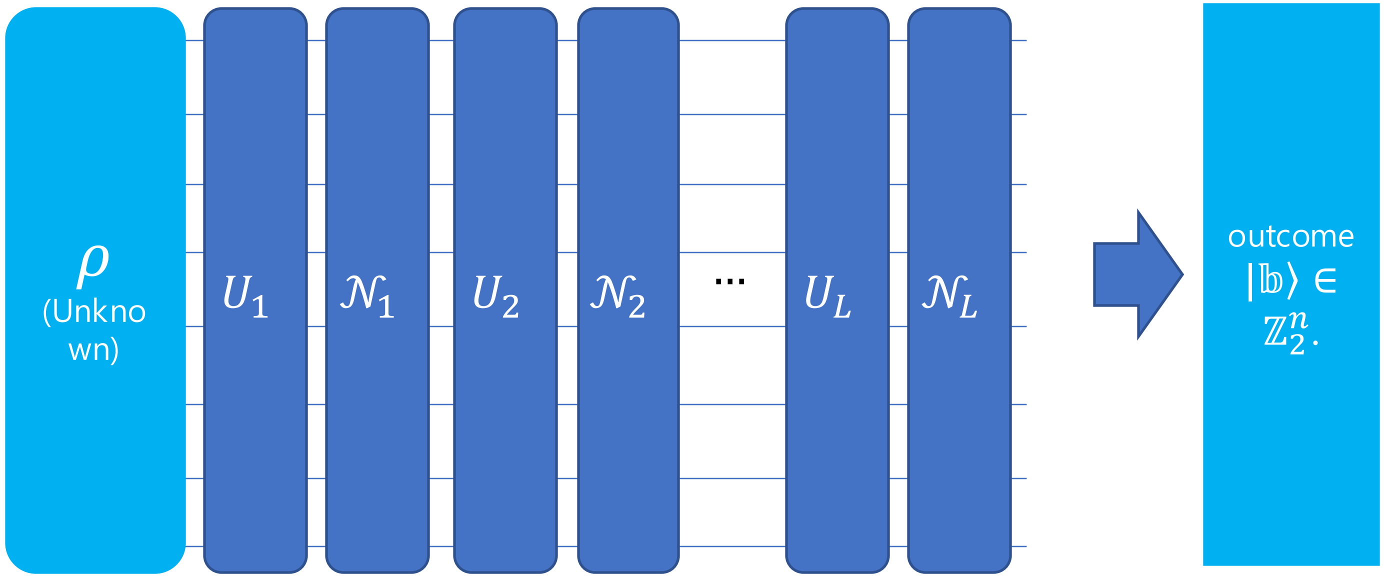
$$g^{(U)} = P^{(U)}\mu \;\Rightarrow\; g^{(U)}_{\mathbf{a}} = \sum_{\mathbf{b}} p^{(U)}_{\mathbf{a}|\mathbf{b}}\mu^{(U)}_{\mathbf{b}}$$

[1] S. Bravyi et al., Phys. Rev. A 103, 042605 (2021)
[2] P. D. Nation et al., PRX Quantum 2, 040326 (2021)

<Pauli noise>

$$\mathcal{N}_l(\cdot) = \sum_{\mathbf{a}\in\mathbb{Z}_2^{2n}} p_{\mathbf{a}}^{(l)} T_{\mathbf{a}}(\cdot) T_{\mathbf{a}}$$

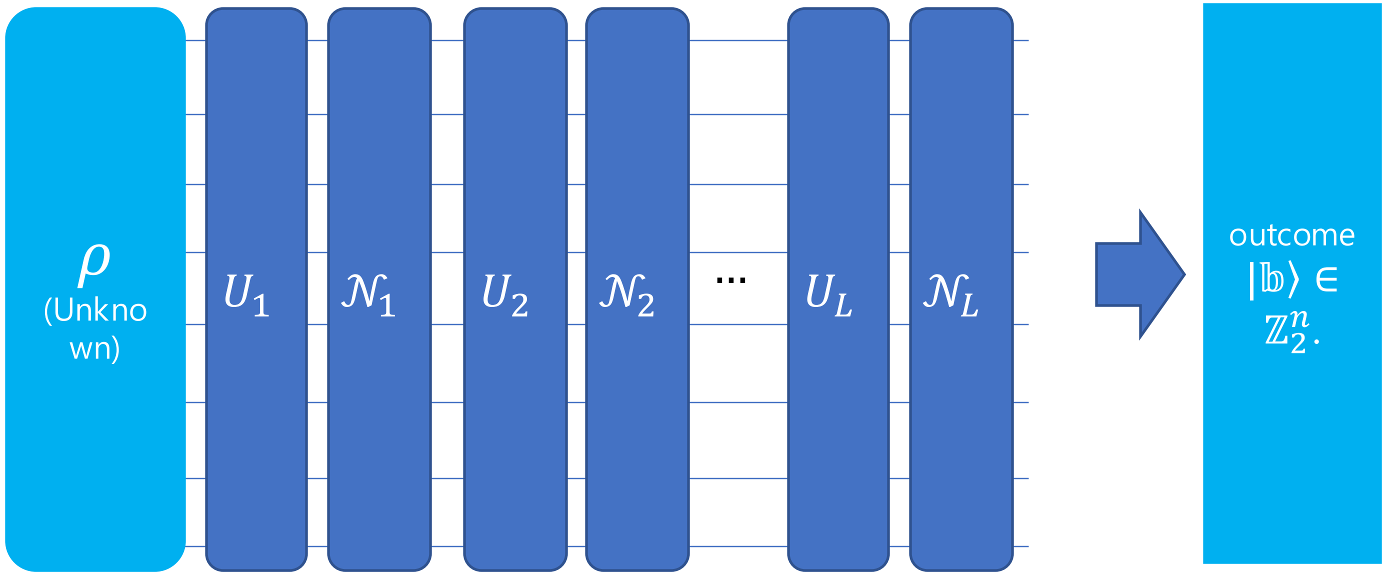
$$T_{\mathbf{a}} \equiv \bigotimes_{i=1}^{n} i^{a_{ix}a_{iz}} X^{a_{ix}} Z^{a_{iz}}$$



Each Clifford layer undergoes Pauli noise with probability $p^{(l)}$.

$$U_l T_{\mathbf{a}} U_l^{\dagger} \sim T_{S_l(\mathbf{a})} \quad S_l \in Sp(2n, \mathbb{Z}_2)$$

We omit the upper-script (U) for convenience



$$\mathcal{N}_l(\cdot) = \sum_{\mathbf{a}\in\mathbb{Z}_2^{2n}} p_{\mathbf{a}}^{(l)} T_{\mathbf{a}}(\cdot)T_{\mathbf{a}}$$

$$T_{\mathbf{a}} \equiv \bigotimes_{i=1}^{n} i^{a_{ix}a_{iz}} X^{a_{ix}} Z^{a_{iz}}$$

$$\mu_{\mathbf{b}} = \sum_{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_L\in\mathbb{Z}_2^{2n}} p_{\mathbf{a}_1}^{(1)}\cdots p_{\mathbf{a}_L}^{(L)} \langle \mathbf{b}| \left(\prod_{l=1}^{L} T_{S_l(\mathbf{a}_l)}\right) \rho \left(\prod_{l=1}^{L} T_{S_l(\mathbf{a}_l)}\right) |\mathbf{b}\rangle$$

$$= \sum_{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_L\in\mathbb{Z}_2^{2n}} p_{\mathbf{a}_1}^{(1)}\cdots p_{\mathbf{a}_L}^{(L)} \langle \mathbf{b} + \left(\sum_{l=1}^{L} S_l(\mathbf{a}_l)_x\right) |\rho| \left(\sum_{l=1}^{L} S_l(\mathbf{a}_l)_x\right) + \mathbf{b}\rangle$$

$$= \sum_{\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_L\in\mathbb{Z}_2^{2n}} p_{\mathbf{a}_1}^{(1)}\cdots p_{\mathbf{a}_L}^{(L)} p_{\mathbf{b}+(\sum_{l=1}^{L} S_l(\mathbf{a}_l)_x)},$$

$$p_{\mathbf{b}(\in\mathbb{Z}_2^n)}^{(U)} \equiv \sum_{\mathbf{a}_1,\ldots,\mathbf{a}_L\in\mathbb{Z}_2^{2n}} p_{S_1^{-1}(\mathbf{a}_1)}^{(1)}\cdots p_{S_L^{-1}(\mathbf{a}_L)}^{(L)} \delta_{\mathbf{b},\mathbf{a}_{1x}+\cdots+\mathbf{a}_{Lx}}$$

➡ Read-out error function
(Samplable!)

$(\forall S_l \in Sp(2n,\mathbb{Z}_2^n))$

Ideal!!

$$\mu_{\mathbf{b}} = \sum_{\mathbf{c}\in\mathbb{Z}_2^n} p_{\mathbf{c}} g_{\mathbf{b}+\mathbf{c}} = \sum_{\mathbf{c}\in\mathbb{Z}_2^n} p_{\mathbf{b}+\mathbf{c}} g_{\mathbf{c}} \equiv (g*p)_{\mathbf{b}}.$$

How to sample by g?

<Solution 1: Approximation>

$$K(p)_{\mathbf{a},\mathbf{b}} \equiv \begin{cases} 1-p_{\mathbf{0}} & (\mathbf{a}=\mathbf{b}) \\ -p_{\mathbf{a}+\mathbf{b}} & (\mathbf{a}\neq\mathbf{b}), \end{cases}, \quad P(p)_{\mathbf{a},\mathbf{b}} \equiv p_{\mathbf{a}+\mathbf{b}}. \quad \Rightarrow \quad P(p)g = (I-K(p))g = \mu \;\; (K \geq 0)$$

(Matrix equation)

$$g = (I-K(p))^{-1}\mu = \sum_{l=0}^{\infty} K^l(p)\mu.$$

$$= \sum_{l=0}^{\infty} \left(K^l(p) - K^{l+1}(p)\right) g$$

$$= \sum_{l=0}^{w-1} c_l P^{l+1}(p) g + \sum_{l=w}^{\infty} (K^l(p) - K^{l+1}(p)) g$$

$$\Rightarrow \therefore \;\; g = \sum_{l=0}^{w-1} c_l P^{l+1}(p) g + \mathcal{O}((2\delta)^w) \quad (\forall c_l \leq \mathcal{O}(l^l))$$

$$= \sum_{l=0}^{w-1} c_l p^{*l} * \mu + \mathcal{O}((2\delta)^w)$$

Samplable!

$$(\delta \equiv 1 - p_{\mathbf{0}})$$

$$(p^{*0} = \mathbf{1} \equiv (1,0,0,\ldots,0),\; p^{*1} = p,\; p^{*2} = p * p, \ldots)$$

<Solution 2: (unbiased) Walsh-Hadamard transform>

$$\widehat{\mu}_{\mathbf{a}} \equiv \sum_{\mathbf{b}} \mu_{\mathbf{a}} (-1)^{\mathbf{a}\cdot\mathbf{b}}$$:Walsh-Hadamard (WH) transform (or Fourier transform)

<*Product rule>

$$\mu = p * g \;\Rightarrow\; \widehat{\mu} = \widehat{p} \cdot \widehat{g}.$$ $\cdot$: element-wise product

$$\Rightarrow \widehat{g} = \widehat{p}^{-1} \cdot \widehat{\mu}$$

$$\therefore \quad g = \frac{1}{2^n} \widehat{\widehat{p}^{-1} \cdot \widehat{\mu}}. \quad \left(\because \frac{1}{2^n} \widehat{\widehat{p}} = p\right)$$

The solution is well-defined, i.e. $\hat{p}$ has no zero element.

(Result) Robust classical shadow under read-out error

<Full random Clifford>

Recall $g^{(V)} = \frac{1}{2^n}\widehat{\left(\frac{\widehat{\mu^{(V)}}}{\widehat{p^{(V)}}}\right)}$

$\mathcal{M}^{-1}(O) = (2^n+1)O - \mathrm{tr}(O)I$

$$\langle O\rangle = \mathbb{E}_{U\sim \mathrm{Cl}_n}\sum_{\mathbf{b}} g_{\mathbf{b}}\left((2^n+1)\langle \mathbf{b}|UOU^\dagger|\mathbf{b}\rangle - \mathrm{tr}(O)\right) =$$

$$\Rightarrow \mathbb{E}_{V\sim \mathrm{Cl}_n}\frac{2^n+1}{2^n}\sum_{\mathbf{c}\in\mathbb{Z}_2^n}\frac{1}{\widehat{p}_{\mathbf{c}}^{(V)}}\sum_{\mathbf{b},\mathbf{k}\in\mathbb{Z}_2^n}\mu_{\mathbf{b}}^{(V)}(-1)^{\mathbf{c}\cdot(\mathbf{b}+\mathbf{k})}\langle \mathbf{k}|VOV^\dagger|\mathbf{k}\rangle - \mathrm{tr}(O)\,.$$

$$\Rightarrow \widehat{\langle O\rangle} = \frac{2^n+1}{2^n}\sum_{\mathbf{k},\mathbf{c}\in\mathbb{Z}_2^n}\frac{(-1)^{(\mathbf{b}+\mathbf{k})\cdot\mathbf{c}}}{\widehat{p}_{\mathbf{c}}^{(V)}}\langle \mathbf{k}|VOV^\dagger|\mathbf{k}\rangle - \mathrm{tr}(O)$$

estimator

$$\Rightarrow \widehat{\langle O\rangle} = \frac{2^n+1}{2^n}\widehat{p}^{(V)-1}Z^{\mathbf{b}}H\mathbb{O}^{(V)T} - \mathrm{tr}(O) \quad (H_{\mathbf{a},\mathbf{b}} = (-1)^{\mathbf{a}\cdot\mathbf{b}})$$

$$(Z^{\mathbf{b}} = diag((-1)^{\mathbf{b}\cdot\mathbf{c}})) \quad (\mathbb{O}_{\mathbf{k}}^{(V)} = \langle \mathbf{k}|VOV^\dagger|\mathbf{k}\rangle)$$

Algorithm : (1) We sample $\mathbf{b} \sim \mu$ (noisy distribution)

(2) Take the esitmator $\widehat{\langle O \rangle} = \frac{2^n+1}{2^n} \widehat{p}^{(V)-1} Z^{\mathbf{b}} H \mathbb{O}^{(V)T} - \mathrm{tr}(O)$. $(H_{\mathbf{a},\mathbf{b}} = (-1)^{\mathbf{a}\cdot\mathbf{b}})$

<**Theorem**: Sampling efficiency for full random Clifford>

$$\|O\|_{\mathrm{sh}}^2 \leq \mathbb{E}(\widehat{O_0}^2) \leq \min_{V,\mathbf{c}} (\widehat{p}_{\mathbf{c}}^{(V)})^{-2} \min\left\{\mathcal{F}(\mathrm{Cl}_n)\mathcal{O}(\mathrm{tr}(O_0^2)),\ \mathcal{O}(\mathrm{tr}(O_0^2) + \|O\|_{\mathrm{st}}^2)\right\}$$

Pure case: $\|O\|_{sh}^2 \leq \|O_0^2\|_\infty \leq \mathcal{O}(\mathrm{tr}(O_0^2))$

$\mathcal{F}(\mathrm{Cl}_n) \equiv \sum_{\mathbf{b}} \max_V p_{\mathbf{b}}^{(V)}$ ➡ (1) Noise has a low-fluctuation : sampling efficient

(2) This factor can be ignored for low-magic observables

Can we calculate the shadow $\widehat{\langle O \rangle} = \frac{2^n+1}{2^n} \widehat{p}^{(V)-1} Z^{\mathbf{b}} H \mathbb{O}^{(V)T} - \mathrm{tr}(O)$??

It normally takes $\mathcal{O}(2^n \mathrm{poly}(n))$....

However, there is a case we can calculate the shadow efficiently

Example:Fidelity estimation with stabilizer state

An arbitrary stabilizer state has a following expression (standard form),

$$|\phi\rangle = \frac{1}{\sqrt{|A|}} \sum_{\mathbf{x}\in A} i^{\mathbf{u}_\phi \cdot \mathbf{x}} (-1)^{Q_\phi(\mathbf{x})} |\mathbf{x} + \mathbf{v}_\phi\rangle$$

Q_ϕ: Quadratic function, u_ϕ, v_ϕ: binary vector

We can regard that $V^\dagger |\phi\rangle = |\phi\rangle$, $VOV^\dagger = |\phi\rangle\langle\phi|$

$$\sum_{\mathbf{k}} (-1)^{\mathbf{c}\cdot\mathbf{k}} \langle \mathbf{k}|V^\dagger O V|\mathbf{k}\rangle$$

$$= \frac{1}{|A|} \sum_{\mathbf{k},\mathbf{x},\mathbf{y}} (-1)^{\mathbf{c}\cdot\mathbf{k}} \delta_{\mathbf{x},\mathbf{k}+\mathbf{v}_\phi} \delta_{\mathbf{y},\mathbf{k}+\mathbf{v}_\phi} \delta_{\mathbf{k}+\mathbf{v}_\phi, A} i^{\mathbf{u}_\phi \cdot (\mathbf{x}-\mathbf{y})} \boxed{= \frac{1}{|A|} \sum_{\mathbf{k}\in A+\mathbf{v}_\phi} (-1)^{\mathbf{c}\cdot\mathbf{k}}}$$

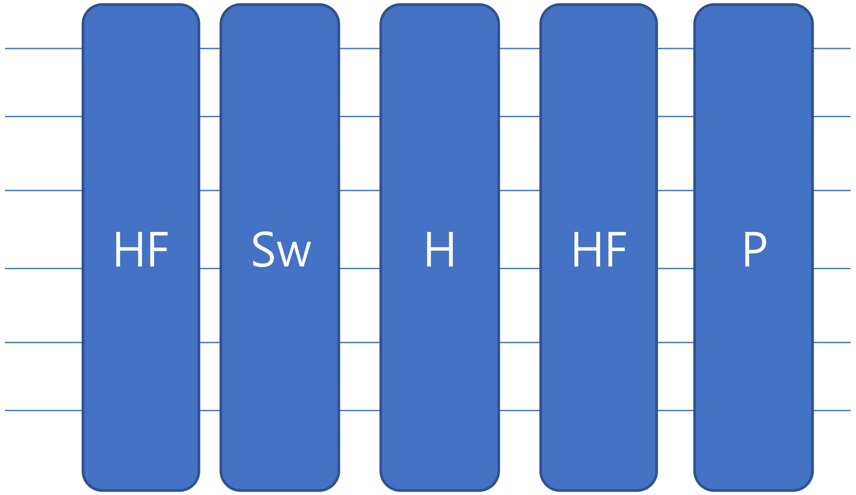
$A + \mathbf{v}_\phi \equiv \{\mathbf{a} + \mathbf{v}_\phi | \mathbf{a} \in A\}$ Affine subspace

Therefore,

[1] S. Bravyi et al., IEEE Trans. Inf. Theor. 67 (7) 4546-4563 (2021)

$$\widehat{\langle O\rangle} = \frac{2^n+1}{2^n|A|}\sum_{\mathbf{c}\in\mathbb{Z}_2^n}\sum_{\mathbf{k}\in A}\frac{1}{\widehat{p_{\mathbf{c}}^{(V)}}}(-1)^{\mathbf{v}_\phi\cdot\mathbf{c}+\mathbf{k}\cdot\mathbf{c}} = \frac{2^n+1}{2^n}\sum_{\mathbf{c}\in A^\perp}\frac{(-1)^{\mathbf{c}\cdot\mathbf{v}_\phi}}{\widehat{p_{\mathbf{c}}^{(V)}}}. \qquad (\mathbb{Z}_2^n = A\oplus A^\perp)$$

Single-depth H[1]



The gate sequence of random Clifford unitary

(HF:Hadamard-Free, Sw: Swapping, P: Pauli)

Total time: time to calculate $\hat{p}_c^{(V)} \times 2^{\dim(A^\perp)}$.

$\dim(A^\perp)$= the number of H-vacancy in the randomly sampled Clifford unitary

$$p_H(q) \geq \prod_{i=0}^{q-1}\frac{2^{n-i}}{2^{n-i}+1} > \left(\frac{2^{n-q}}{2^{n-q}+1}\right)^q$$

(Mallow distribution)

We set $q = n - c\log(n)$ ➡ $p_H(q) \geq (1-n^{-c})^{n-c\log(n)} \simeq 1 - n^{-c+1}$

$\therefore$ with failure prob. δ_{Hf}, $N < \delta_{Hf} M^{-1} n^{c+1}$ copies are calculated in poly-time (M: number of target stabilizer state)

(ex)

M=100 number of stabilizer states are targets,
c=5, δ_{Hf}=0.01, we can hold $N \ll n^6 \cdot 0.0001$ ($\sim 1562500\ for\ n = 50$) copies.

Total time: time to calculate $\hat{p}_c^{(V)} \times O(n^5)$.

$\Downarrow$

? (Next problem)

(Sol) it needs $O(n^2)$-time.

$$\because \widehat{p}_{\mathbf{c}}^{(V)} = \prod_{g=1}^{G} \widehat{p}_{\mathbf{c}}^{(g)}$$ (each can be computed in constant-time, G: gate-count)

$\Uparrow$

$$p^{(U)}_{\mathbf{b}(\in \mathbb{Z}_2^n)} \equiv \sum_{\mathbf{a}_1,\ldots,\mathbf{a}_L \in \mathbb{Z}_2^{2n}} p^{(1)}_{S_1^{-1}(\mathbf{a}_1)} \cdots p^{(L)}_{S_L^{-1}(\mathbf{a}_L)} \delta_{\mathbf{b},\mathbf{a}_{1x}+\cdots+\mathbf{a}_{Lx}}$$

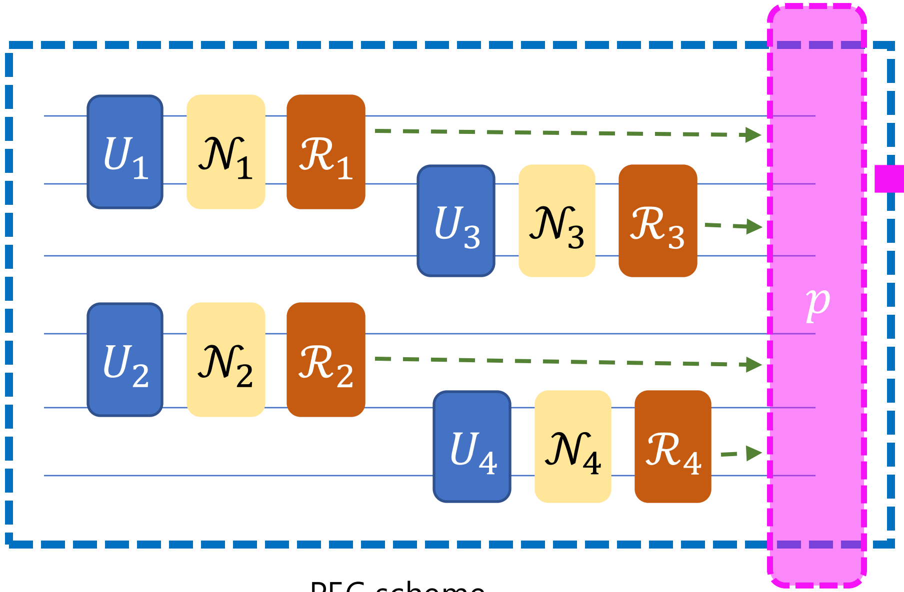
5-qubit GHZ state fidelity estimation,

Full random Clifford tomography with depol-rate 0.05

(target value=1)

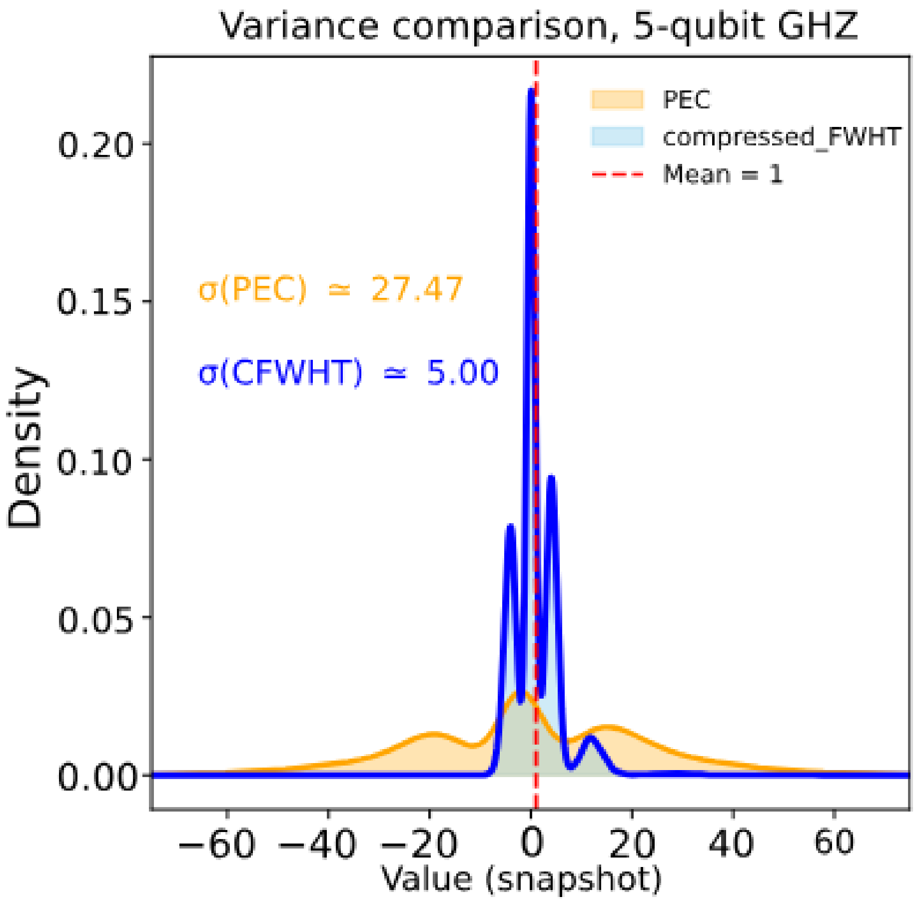
PEC: gate-by-gate probabilistic error cancellation

CFWHT: Our method



PEC scheme

Compressed& Marginalized



$$\mathcal{R}_i(\cdot) = \sum_{\mathbf{a}} q_{\mathbf{a}} T_{\mathbf{a}}(\cdot) T_{\mathbf{a}} \quad (\|\mathcal{R}_i\| = \|q\|_1)$$

$$(\text{negativity})\ \|\mathcal{R}\| = \prod_i \|\mathcal{R}_i\| \simeq \mathcal{O}(C^G)\ (C > 1)$$

$$\min_{V,\mathbf{c}} (\widehat{p}_{\mathbf{c}}^{(V)})^{-1} \lesssim (\text{compressed}) \|\mathcal{R}\|_1 \ll \prod_i \|\mathcal{R}_i\|$$ 'Sub-multiplicativity'

<Shallow (d-depth) Clifford shadow>

We recall that $\mathcal{M}(P) = t_{P,d}P \Rightarrow \mathcal{M}^{-1}(P) = \frac{1}{t_{P,d}}P$

$$\text{for some } t_{P,d} = \mathrm{Prob}_{V\in \mathrm{Cl}_{n,d}}(VT_{\mathbf{a}}V^\dagger \in \mathcal{Z}) > 0$$

$$\widehat{\langle T_{\mathbf{a}}\rangle}_{LS} = \frac{1}{4^n}\sum_{\mathbf{c},\mathbf{k}} \frac{(-1)^{\mathbf{c}\cdot(\mathbf{b}+\mathbf{k})}}{\widehat{p}^{(V)}_{\mathbf{c}}} \sum_{\mathbf{l}} (-1)^{\mathbf{k}\cdot\mathbf{l}} \frac{\mathrm{tr}\left(VT_{\mathbf{a}}V^\dagger Z^{\mathbf{l}}\right)}{t_{V^\dagger Z^{\mathbf{l}}V,d}} \qquad \left(T_{\mathbf{a}} \equiv \bigotimes_{i=1}^{n} i^{a_{ix}a_{iz}} X^{a_{ix}} Z^{a_{iz}}\right)$$

$$= \frac{1}{2^n}\sum_{\mathbf{c}} \frac{(-1)^{\mathbf{c}\cdot\mathbf{b}}}{\widehat{p}^{(V)}_{\mathbf{c}}} \frac{\mathrm{tr}\left(VT_{\mathbf{a}}V^\dagger Z^{\mathbf{c}}\right)}{t_{V^\dagger Z^{\mathbf{c}}V,d}}$$

$$= \sum_{\mathbf{c}} \frac{(-1)^{\mathbf{c}\cdot\mathbf{b}}}{\widehat{p}^{(V)}_{\mathbf{c}}} \frac{(-1)^{P_V(\mathbf{a})}\delta_{S_V(\mathbf{a})_z,\mathbf{c}}}{t_{V^\dagger Z^{\mathbf{c}}V,d}} \delta_{S_V(\mathbf{a})_x,\mathbf{0}}$$

$$\boxed{= \frac{(-1)^{S_V(\mathbf{a})_z\cdot\mathbf{b}+P_V(\mathbf{a})}}{\widehat{p}^{(V)}_{S_V(\mathbf{a})_z} t_{V^\dagger Z^{S_V(\mathbf{a})_z}V,d}} \delta_{S_V(\mathbf{a})_x,\mathbf{0}},}$$

➡ $\mathbb{b}$ is sampled by μ. $\hat{p}_{S_V(a)_z}$ is calculated efficiently ➡ Always efficient!!

<**Theorem**: Sampling efficiency for d-depth shallow Clifford shadow>

Given a Pauli observable $O = T_{\mathbf{a}}$,

$\mathbb{E}(O_0^2) \leq \min_{V,\mathbf{c}} (\widehat{p}_{\mathbf{c}}^{(V)})^{-2} t_{T_{\mathbf{a}},d}^{-1}$, where, $t_{P,d} = \mathrm{Prob}_{V \in \mathrm{Cl}_{n,d}}(V T_{\mathbf{a}} V^{\dagger} \in \mathcal{Z}) > 0$

For the pure case,

$\mathbb{E}(O_0^2) \leq t_{T_{\mathbf{a}},d}^{-1}$.

<Tighter analysis from Ref.[1]... for depolarizing noise>

$$\mathbb{E}(O_0^2) \leq \min_{V,\mathbf{c}} (\widehat{p}_{\mathbf{c}}^{(V)})^{-2} t_{T_{\mathbf{a}},d}^{-1} \leq 3^{(k+d)\left\{\frac{3}{4}+t^{-\frac{3}{2}}\left(\frac{4}{5}\right)^{2t}\right\}}(1+2\delta)^2 \qquad (k \equiv \mathrm{weight}(T_{\mathbf{a}}), \delta \ll 1)$$

,which is tighter than $3^{(k+d)\left\{\frac{3}{4}+t^{-\frac{3}{2}}\left(\frac{4}{5}\right)^{2t}\right\}+(k+d)\delta}$ [1]

[1] H. Y. Hu et al., Nat. Comm. 16, 2943 (2025)

<Challenges>

(1) How to boost the shadow calculation speed in general?

$$\widehat{\langle O \rangle} = \frac{2^n+1}{2^n}\widehat{p}^{(V)-1} Z^{\mathbf{b}} H \mathbb{O}^{(V)T} - \mathrm{tr}(O)??$$

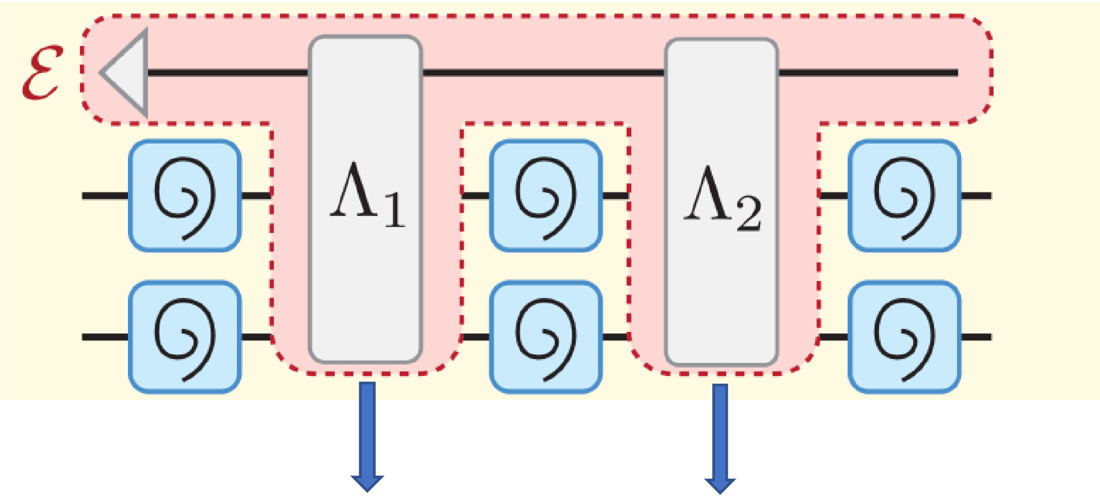
(2) How do we reduce the sampling when the noise fluctuation of random Clifford is large?

$$\|O\|_{\mathrm{sh}}^2 \leq \mathbb{E}(\widehat{O_0}^2) \leq \min_{V,\mathbf{c}} (\widehat{p}_{\mathbf{c}}^{(V)})^{-2} \min\left\{\mathcal{F}(\mathrm{Cl}_n)\mathcal{O}(\mathrm{tr}\big(O_0^2\big)),\ \mathcal{O}(\mathrm{tr}\big(O_0^2\big) + \|O\|_{\mathrm{st}}^2)\right\}$$

(Result) Graph state benchmarking and high-order stability

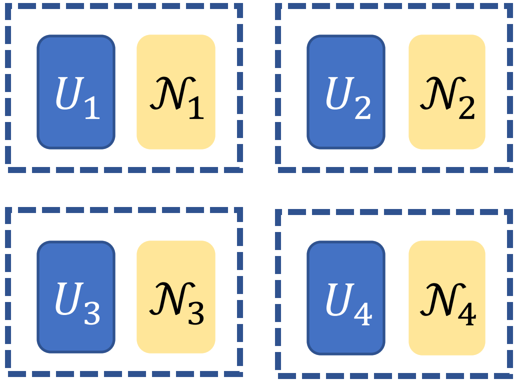
[1] S. Endo et al., Phys. Rev. X 8, 031027 (2018)
[2] S. Chen et al., Nat. Comm 14, 52 (2023)

What if the noise is arbitrary and unknown?

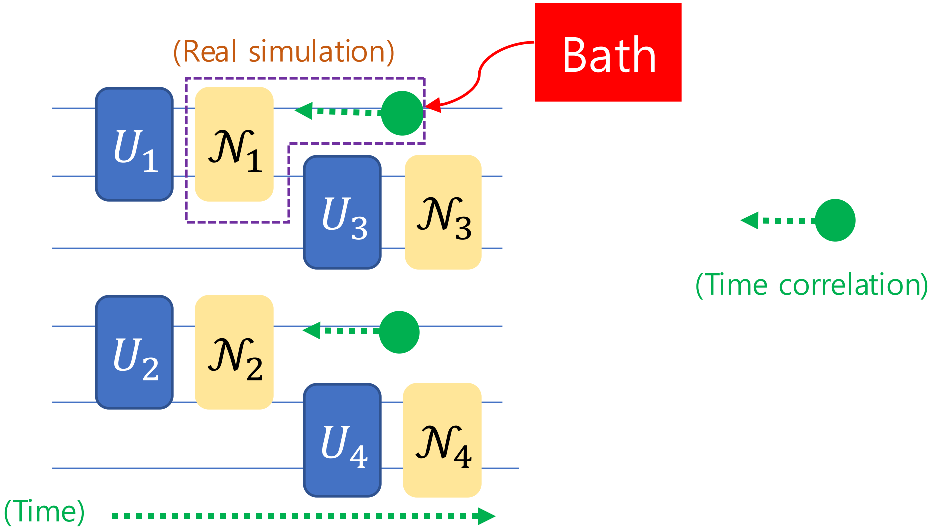


Even after twirled, additional noise occurs during the interaction with the environment[1]

(Noise learning)



Each (twirled)noise is learned in a restricted space[1,2]



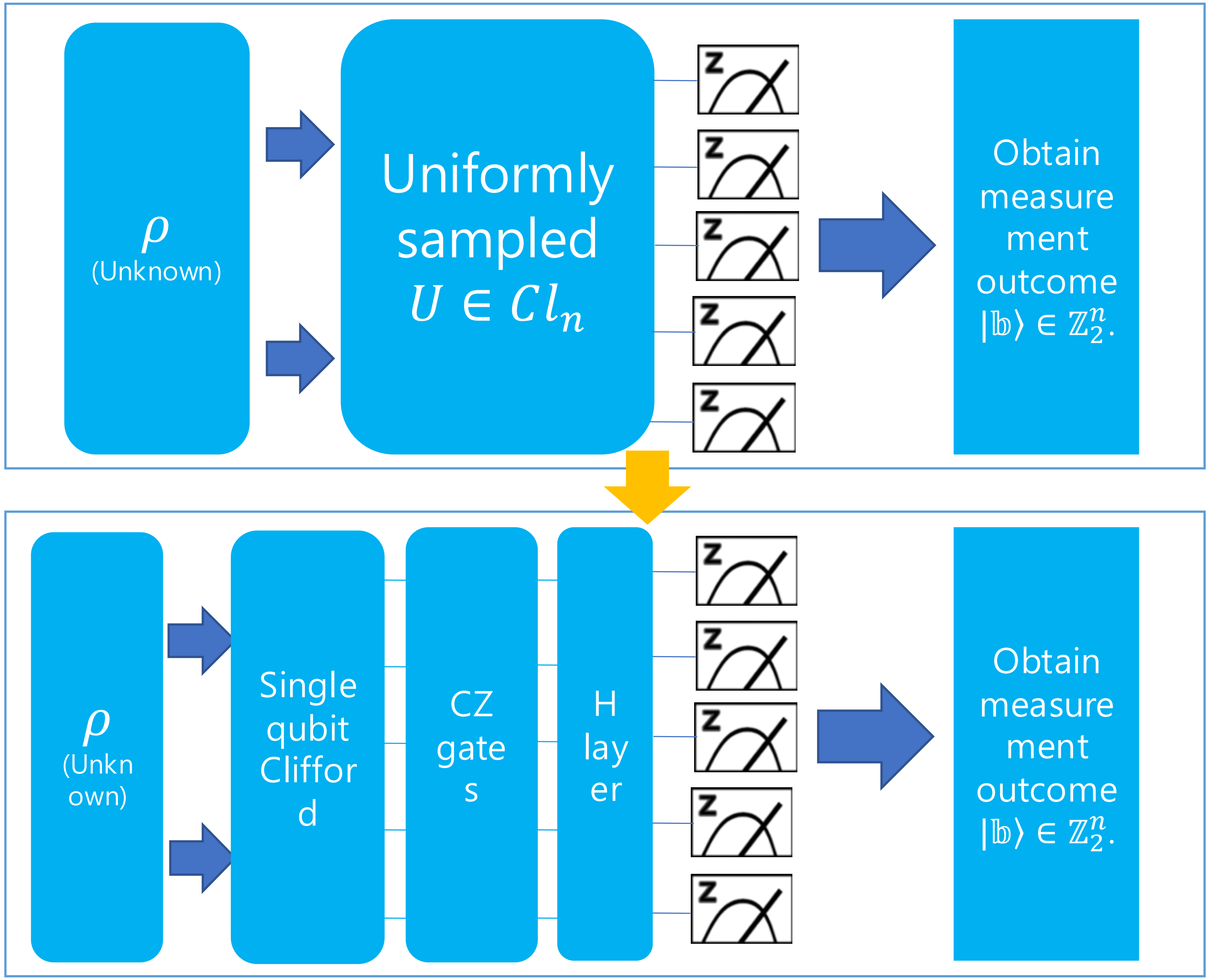
[1] H. T. Hu et al., Nat. Comm. 16, 2943 (2025)

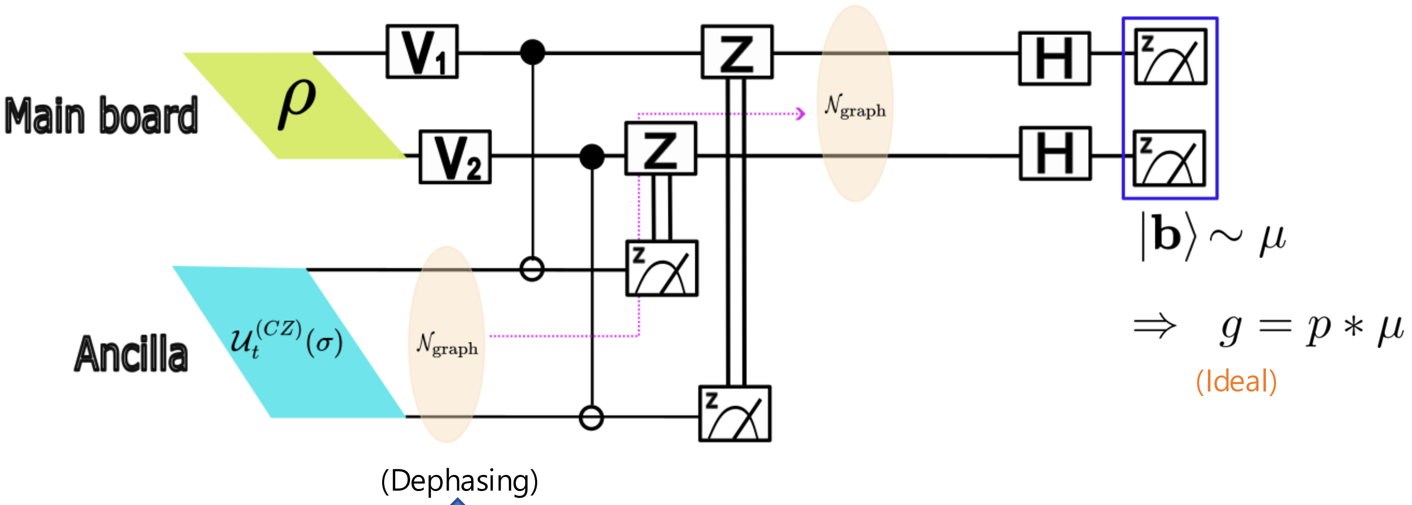
<Problem>

Can we learn the arbitrary noise more efficiently?

<Solution>

We use the graph state measurement





Given $|\psi = G(V,E)\rangle = \prod_{A \in E} CZ_A \left|+\right\rangle^{\otimes n}, \quad \sigma = \mathcal{N}(|\psi\rangle\langle\psi|),$

Random Pauli operation

$$\mathcal{N}_t(\sigma) = \frac{1}{2^n} \sum_{\mathbf{a} \in \mathbb{Z}_2^n} X_\psi^{\mathbf{a}} \sigma X_\psi^{\mathbf{a}} = \sum_{\mathbf{a} \in \mathbb{Z}_2^n} p_{\mathbf{a}} Z^{\mathbf{a}} |\psi\rangle\langle\psi| Z^{\mathbf{a}} = \mathcal{N}_{\text{graph}}(|\psi\rangle\langle\psi|)$$

$$X_{\mathbf{a}}^{\psi} = \prod_{A \in E} CZ_A X^{\mathbf{a}} \prod_{A \in E} CZ_A \in \mathcal{P}_n$$

"Noise tailoring"[1]

[1] G. Park et al., arXiv:2503.12870 (2025)

$$g = \sum_{l=0}^{w-1} c_l p^{*l} * \mu + \mathcal{O}((2\delta)^w) \quad (\forall c_l \leq \mathcal{O}(l^l)) \qquad g_{\text{approx}} = \sum_{l=0}^{w-1} c_l p^{*l} * \mu$$

$$\Rightarrow \langle O \rangle_{\text{approx}} = \sum_{l=0}^{w-1} c_l \sum_{\mathbf{b} \in \mathbb{Z}_2^n} (p^{*l} * \mu)(\mathbf{b})(2^n+1) \langle \mathbf{b} | U O U^\dagger | \mathbf{b} \rangle - \mathrm{tr}(O)$$

<Algorithm>

(1) We sample $\mathbf{b}_1, \mathbf{p}_1, \ldots, \mathbf{p}_{w-1}$ (i.i.d) Following μ, p

(2) Set $\mathbf{b}_2 = \mathbf{b}_1 + \mathbf{p}_1,\ \mathbf{b}_3 = \mathbf{b}_2 + \mathbf{p}_2, \ldots, \mathbf{b}_{w-1} = \mathbf{b}_{w-2} + \mathbf{p}_{w-1}$

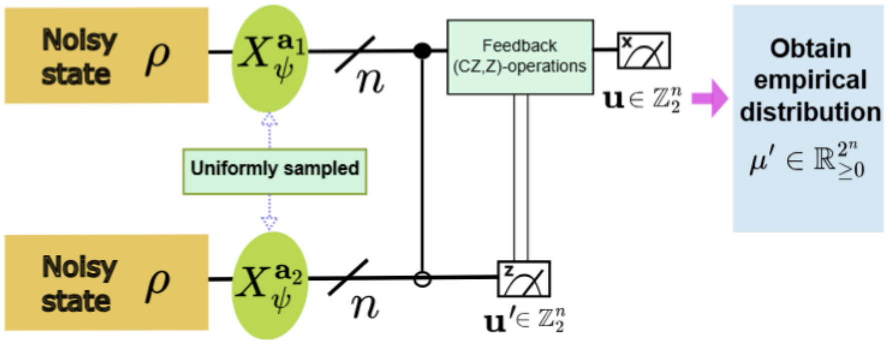
(3) For each $0 \leq l \leq w-1$, calculate $m_i^{(l)} = c_l (2^n+1) \langle \mathbf{b} | U \rho U^\dagger | \mathbf{b} \rangle$

(4) For each $0 \leq l \leq w-1$, calculate $m_i = \sum_{l=0}^{w-1} m_i^{(l)} - \mathrm{tr}(O)$

(5) We repeat this procedure (1)~(4) N times and take $m = \frac{1}{N} \sum_{i=1}^{N} m_i.$

Problem: How to sample $\mathbf{p}$?

We can sample following $p * p$



<Theorem: Estimating p from the convolution powers> [1]

Given the approximation order $(w, s) \in \mathbb{N}^2$ and $\delta < \frac{1}{3w}$,

$$p = \sum_{l=0}^{\mathcal{O}(w^2, ws)} d_l^{(w,s)} (p * p)^{*l} + \mathcal{O}\left(\left(\frac{3w\delta}{2} \right)^{w+s} + \frac{(w\delta)^w}{w^{\frac{w}{2}}} \right). \qquad d_l^{(w,s)} \leq \mathcal{O}(l^l)$$

(ex) (w,s)=(2,0)

$$p = \frac{3}{2}(p * p) - \frac{1}{2}(p * p)^{*1} + \mathcal{O}(5\delta^2) = \frac{3}{2}\mu - \frac{1}{2}\mu^{*1} + \mathcal{O}(5\delta^2)$$

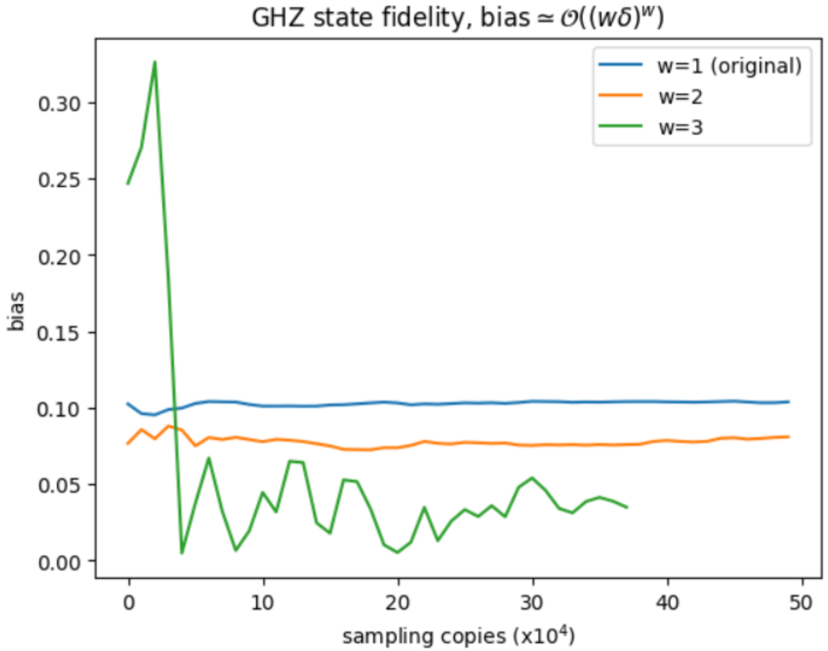
(Samplable)

[1] G. Park et al., arXiv:2503.12870 (2025)

$$\therefore \quad g = \sum_{k=0}^{\mathcal{O}(w^3, w^2 s)} c_k (p * p)^{*k} + \mathcal{O}\left((2\delta)^w + w^w \left(\frac{3w\delta}{2} \right)^{w+s} \right)$$

$$\Rightarrow \langle O \rangle_{\text{approx}} = \sum_{l=0}^{\mathcal{O}(w^3, w^2 s)} c_k \sum_{\mathbf{b} \in \mathbb{Z}_2^n} (p * p)^{*k}(\mathbf{b})(2^n + 1) \langle \mathbf{b} | U O U^\dagger | \mathbf{b} \rangle - \mathrm{tr}(O)$$

(Samplable)



Bias (y_axis) of fidelity estimation between pure 2-qubit GHZ states.

Single qubit depol-noise with 0.05 rate for the sampled two-qubit gates

Target value: 1

<Challenges for the unknown noise case>

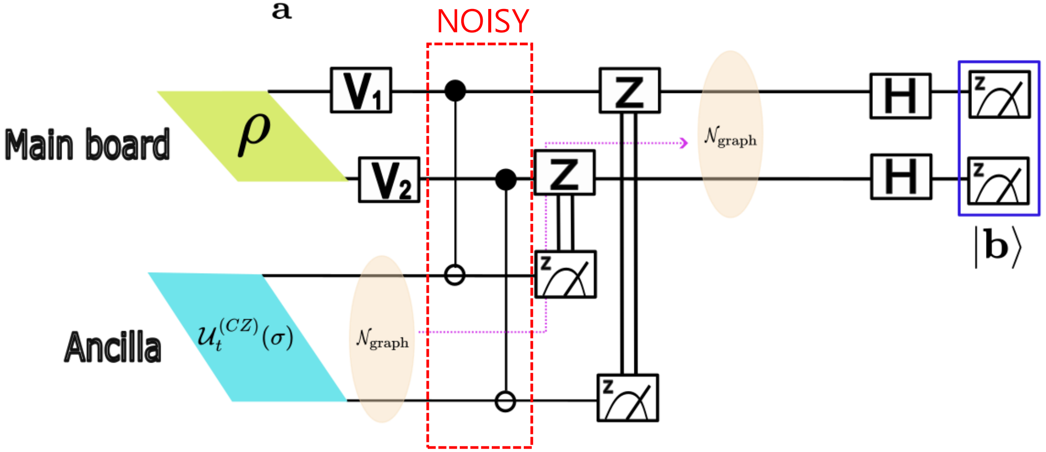
1. δ should be lower than $\frac{1}{3w^2}$ for the valid approximation. Can we increase the upper bound of threshold?

2. Can we improve the upper bound of bias?

$$\mathcal{O}\left((2\delta)^w + w^w\left(\frac{3w\delta}{2}\right)^{w+s}\right)$$

3. Unbiased estimator for sparse noise? $(\widehat{p} = \sqrt{\widehat{p * p}},\ \widehat{p * p}_{\mathbf{b}} = \sum_{\mathbf{a}} (p * p)_{\mathbf{a}} (-1)^{\mathbf{a}\cdot\mathbf{b}})$

4. How to manage the noisy transversal CNOTs?



Summary

1. We demonstrated a robust shadow tomography scheme under the gate dependent noise.

2. The noise channel in random Clifford measurement shrinks to read out error convolution with the ideal measurement distribution.

3. For Pauli noise case, the above fact offers a tighter estimation variance (sampling complexity) of noise mitigation via WH transform, compared to the conventional PEC shadow.

4. When the noise is arbitrary and unknown, we can transform the Clifford measurement to graph state measurement, to tailor the noise into dephased form.

Then the tailored dephasing noise p can be learned by the approximation with the convolution power series of $(p * p)$ hence applying to mitigate the shadow.

Thank you very much