Title: How to learn Pauli noise over a gate set

Speakers: Senrui Chen

Collection/Series: Quantum Information

Subject: Quantum Information

Date: December 11, 2024 - 11:00 AM

URL: https://pirsa.org/24120035

Abstract:

Understanding quantum noise is an essential step towards building practical quantum information processing systems. Pauli noise is a useful model widely applied in quantum benchmarking, quantum error mitigation, and quantum error correction. Despite previous research, the problem of how to learn a Pauli noise model self-consistently, completely, and efficiently has remained open. In this talk, I will introduce a framework of gate-set Pauli noise learning that aims at addressing this problem. The framework treats initialization, measurement, and a set of quantum gates to suffer from unknown Pauli noise channels, which are allowed to have customized locality constraints. The goal is to learn all the Pauli noise channels using only those noisy operations. I will first introduce a theory on the "learnability" of Pauli noise model, i.e., what information is fundamentally identifiable within the model and what is not. This is established using tools from algebraic graph theory and ideas from gate set tomography; I will then discuss a sample-efficient procedure to learn all learnable information of a Paul noise model to any desired precision; Finally, I will demonstrate how to apply our theoretic framework for concrete practical gate set and noise assumptions, and discuss the potential impact on quantum error mitigation and other applications.

How to learn Pauli noise over a gate set

Senrui Chen (U of Chicago)

Talk at Perimeter Institute, Dec 2024

Based on

[1] SC*, Y Liu*, M Otten, A Seif, B Fefferman, L Jiang. Nat. Commum. 14, 52 (2023)

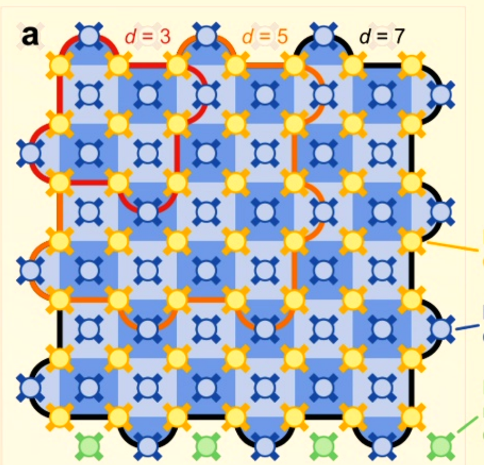
[2] SC, Z Zhang, L Jiang, S Flammia. arXiv: 2410.03906 (2024)

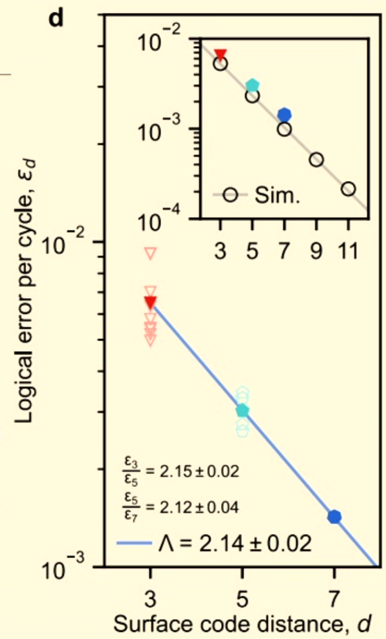
nature
https://doi.org/10.1038/s41586-024-08449-y
Accelerated Article Preview

Quantum error correction below the surface code threshold

Received: 24 August 2024

Google Quantum AI and Collaborators



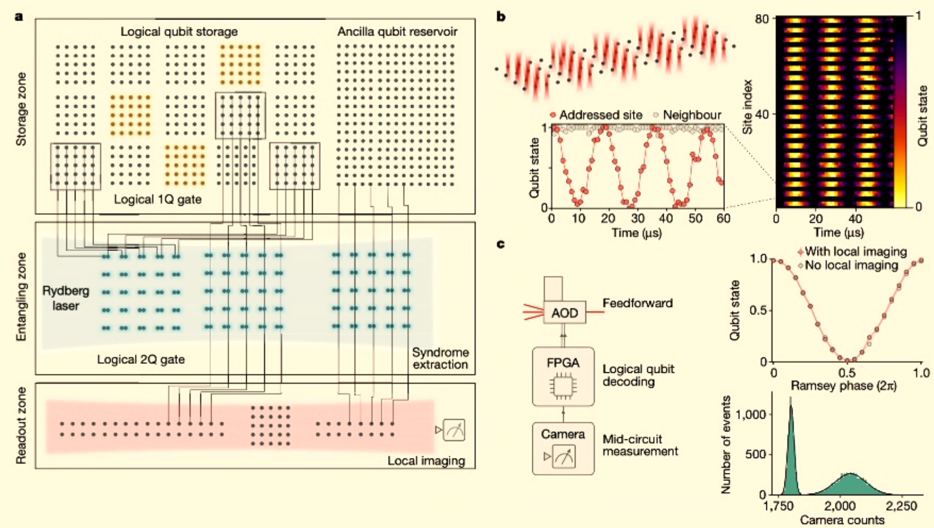


Article

Logical quantum processor based on reconfigurable atom arrays

https://doi.org/10.1038/s41586-023-06927-3
Received: 21 October 2023
Accepted: 1 December 2023
Published online: 6 December 2023

Dolev Bluvstein[1], Simon J. Evered[1], Alexandra A. Geim[1], Sophie H. Li[1], Hengyun Zhou[1,2], Tom Manovitz[1], Sepehr Ebadi[1], Madelyn Cain[1], Marcin Kalinowski[1], Dominik Hangleiter[3], J. Pablo Bonilla Ataides[1], Nishad Maskara[1], Iris Cong[1], Xun Gao[1], Pedro Sales Rodriguez[2], Thomas Karolyshyn[2], Giulia Semeghini[4], Michael J. Gullans[3], Markus Greiner[1], Vladan Vuletić[5] & Mikhail D. Lukin[1]

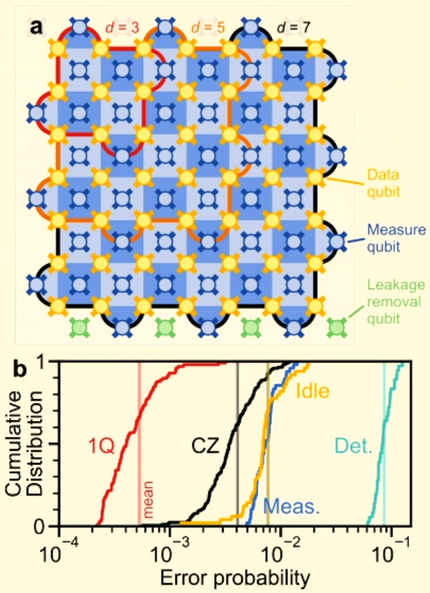


Noise remains the major challenges in building a quantum computer!

1

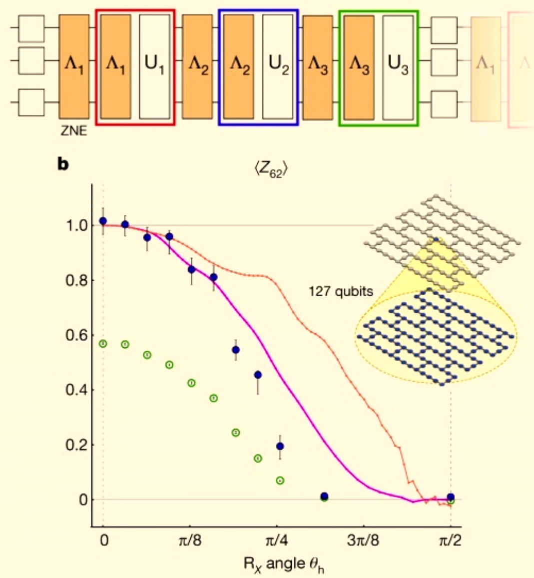
Why noise characterization?

Hardware calibration & improvement



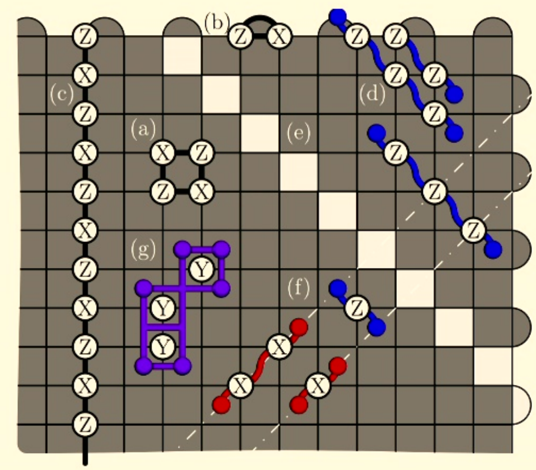
Google Quantum AI, Nature 2024

Learning-based quantum error mitigation (QEM)



Y Kim et al., IBM Quantum, Nature 2023

Improved designs for quantum error correction (QEC)



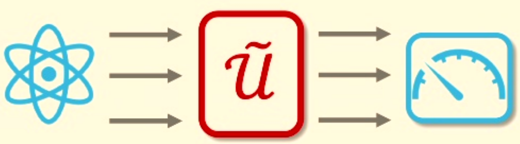
J.P. Bonilla Ataides et al., Nat. Commun. 2021

Learning noise in quantum circuits

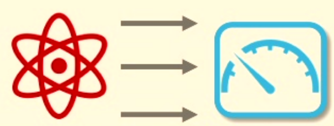
- Components of quantum circuits:
  (Ignore leakage, non-Markovianity, time-fluctuation, …)

State preparation (SP) Gate Measurement (M)

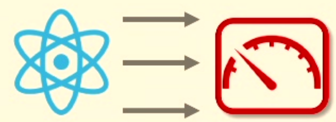
- Black-box learning method?



Process tomography



State tomography



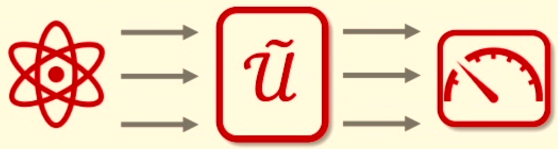
Detector tomography

3

Learning noise in quantum circuits

- Components of quantum circuits:
  (Ignore leakage, non-Markovianity, time-fluctuation, …)

- Black-box learning method?



ibm_marrakesh, 24-12-10

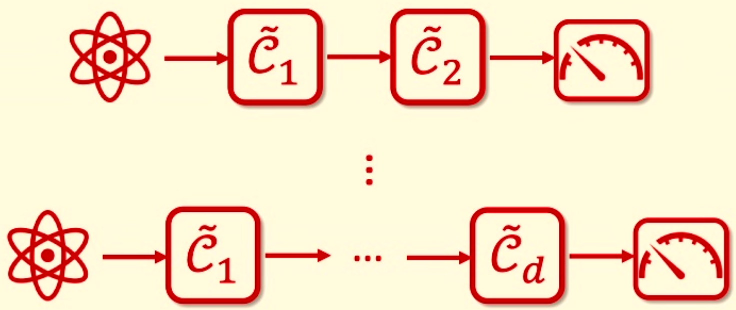
Median CZ error:	2.242e-3
Median SX error:	2.774e-4
Median Readout error:	1.610e-2

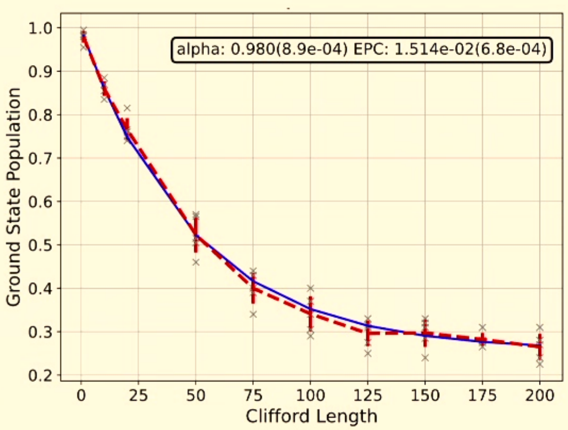
  - All are noisy! SPAM noise non-negligible in practice.

- Questions: How to learn about the noise self-consistently?
  - i.e., learning noise using noisy operations themselves

4

Randomized benchmarking (RB)



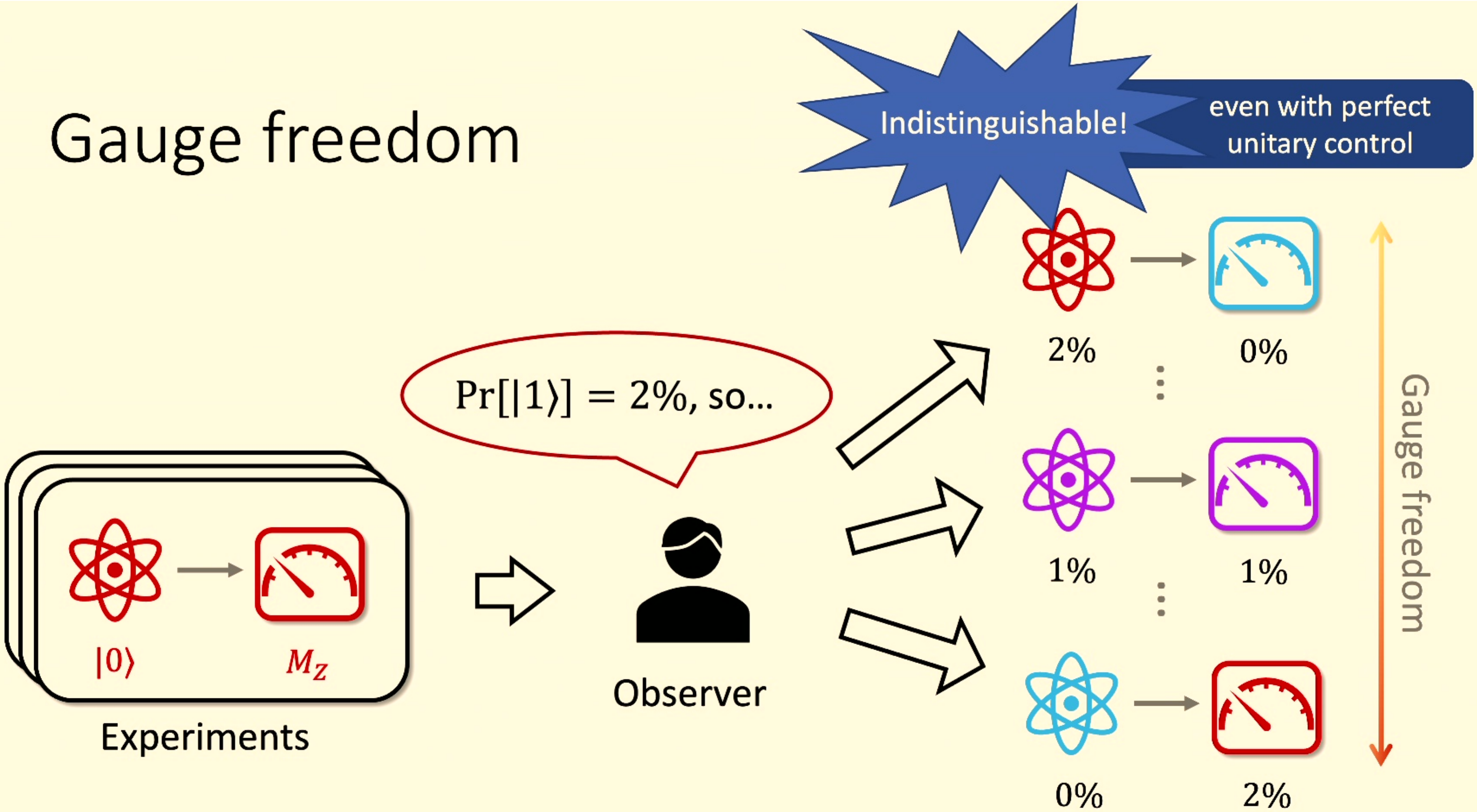


Randomized Benchmarking, Qiskit Textbook

- Idea: Separate gate noise from SPAM noise via gate repetition and fitting
- Learning gate noise SPAM-robustly assuming gate-independent noise of Clifford group[1]
- In practice, difficulty arises with **gate-dependent** noise

[1] Kimmel et al., PRX (2014)

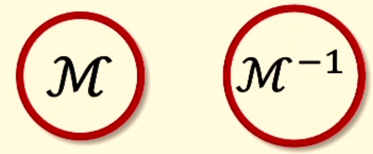
Gauge freedom

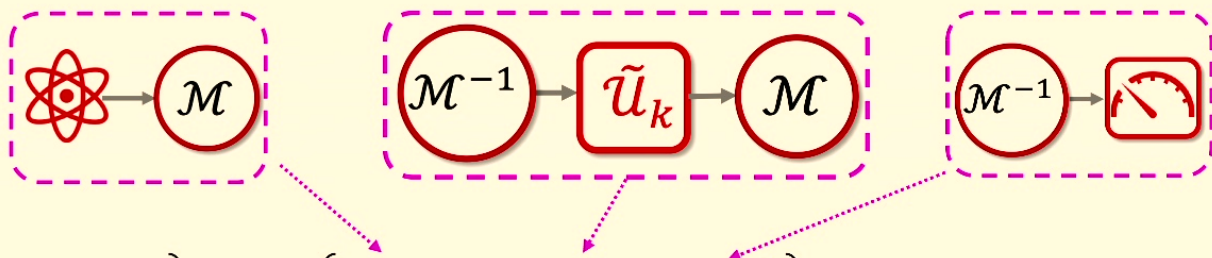


*Assume depolarizing noise (for simplicity)

Gauge transform

- Noisy operation set (SPAM+gate): 

- Gauge transform map (invertible): 

- Gauge transform: 

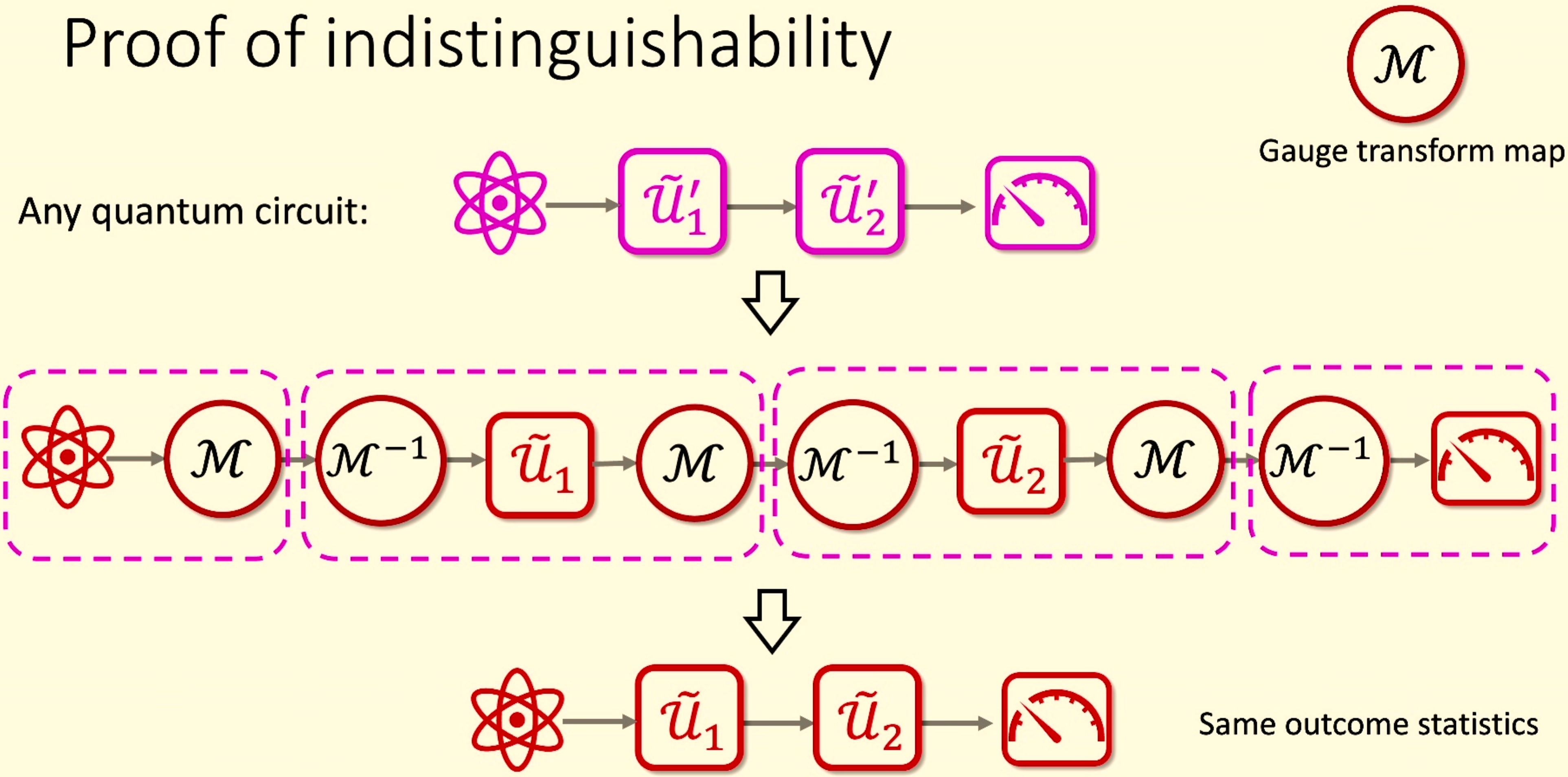
- Claim: 

and 

are **indistinguishable**

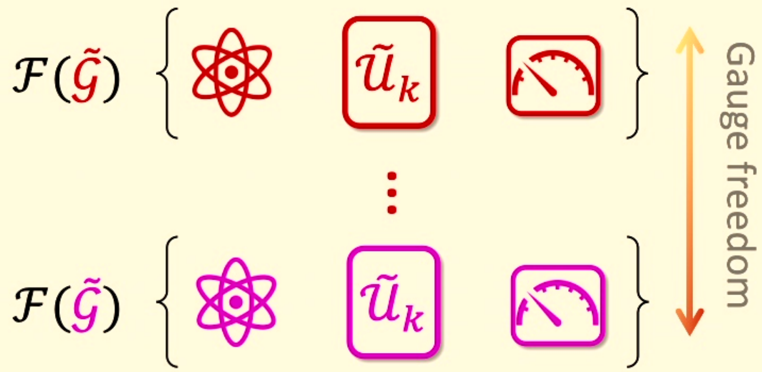
See, e.g., "Gate set tomography". E Nielsen et al., Quantum 5, 557 (2021).

Proof of indistinguishability



8

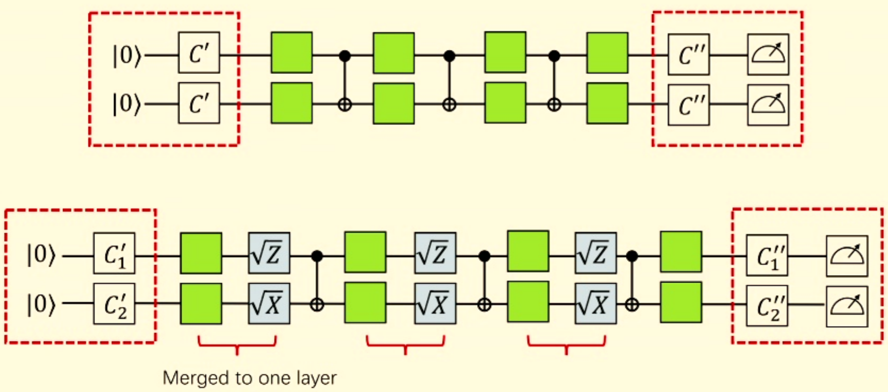
Gauge and noise learnability

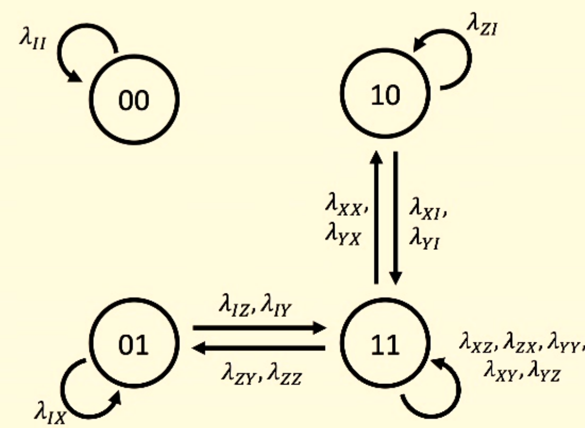


- Gauge transform relates a manifold of indistinguishable noise realizations
  - Restriction: All noisy components still need to be physical and obey model assumptions.
- Consider a function $\mathcal{F}$ of the noise parameters (e.g., SPAM fidelity, gate fidelity)
  - If $\mathcal{F}$ is **learnable**, then $\mathcal{F}$ must be invariant under any gauge transformations
  - In contrast, unlearnable function can only be determined up to gauge freedom
- Can we characterize the “learnability” of a given noise model?
  - Which functions are learnable? How to construct learning experiments?

Summary of contributions

- Characterization of the learnability for Pauli noise model via graph theory
- Learning algorithm that is self-consistent, complete, and efficient.
- Experimental results justifying the practical relevance of noise unlearnability.
- Applications to error mitigation





[SC*, Y Liu*, et al., Nat. Commum. 14, 52 (2023)], [SC et al., arXiv: 2410.03906 (2024)]

Pauli channels

- n-qubit Pauli channel:

$$\Lambda(\rho) = \sum_{a \in \mathrm{P}^n} p_a P_a \rho P_a = \sum_{b \in \mathrm{P}^n} \lambda_b P_b \mathrm{Tr}(P_b \rho)/2^n$$

  - $\mathrm{P}^n = \{I, X, Y, Z\}^{\otimes n}$ - n-qubit Pauli group
  - $\{p_a\}_a$ – Pauli error rates
  - $\{\lambda_b\}_b$ – Pauli fidelities or Pauli eigenvalues, $\Lambda(P_b) = \lambda_b P_b$

Example:

	I	X	Y	Z
p (error rate)	0.7	0	0	0.3
λ (eigenvalue)	1	0.4	0.4	1

E.g., Phase-flip noise channel

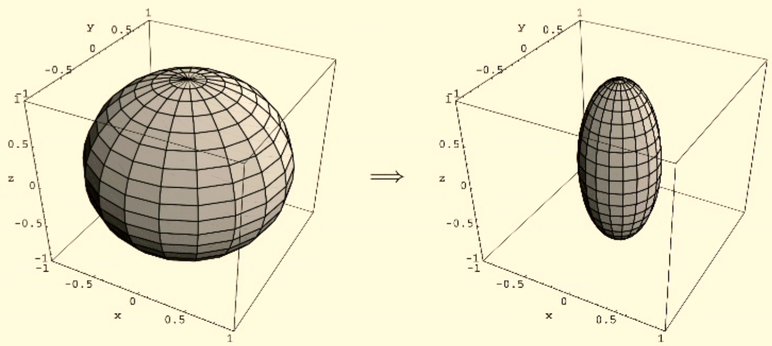
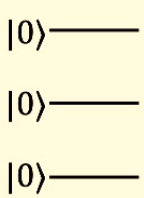


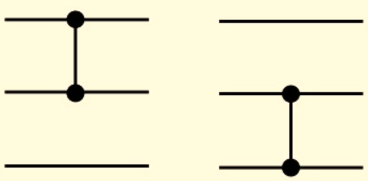
Fig. 8.9, Nielsen & Chuang

Pauli noise model

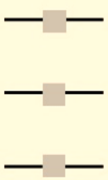
- Consider an n-qubit system with the following gadgets:



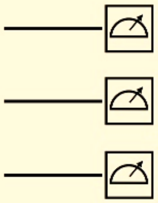
Initialization to $|0\rangle^{\otimes n}$



Layer of multi-qubit Clifford gates $\mathcal{G}$ from a set $\mathfrak{G}$



Layer of arbitrary 1-qubit gates $\otimes_j \mathcal{U}_j$

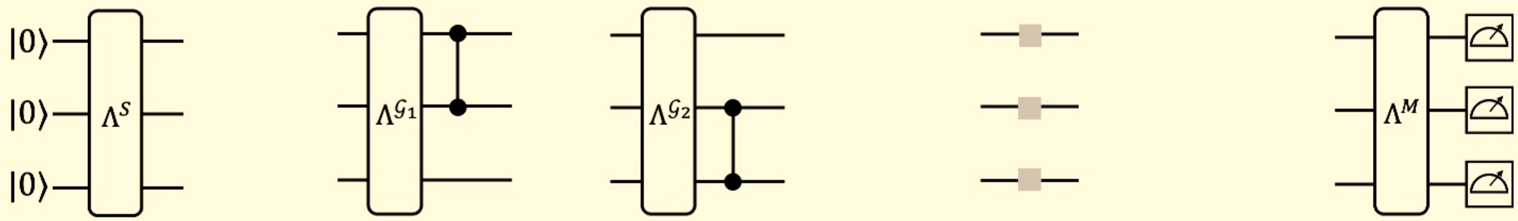


Computational-basis measurement $\{E_j\}_j$

13

Pauli noise model

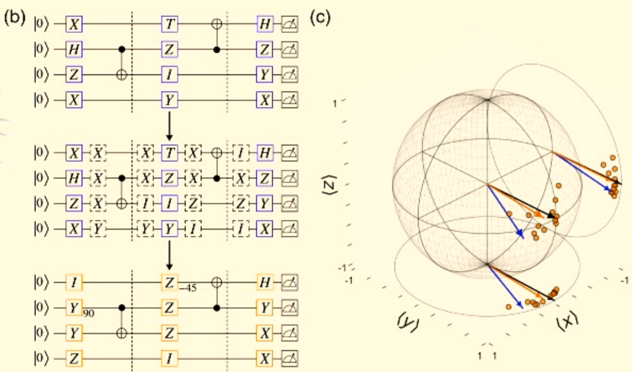
- Consider an n-qubit system with the following gadgets:



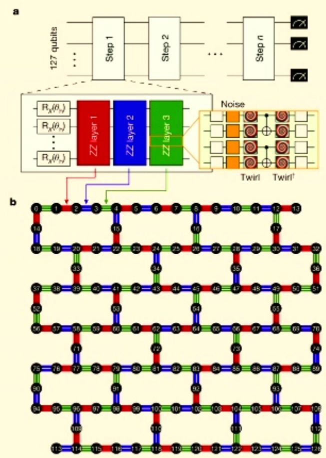
$\tilde{\rho} = \Lambda^S(|0\rangle\langle 0|^{\otimes n})$ $\quad$ $\tilde{\mathcal{G}} = \mathcal{G} \circ \Lambda^{\mathcal{G}}$ $\quad$ $\otimes_j \mathcal{U}_j$ noise negligible $\quad$ $\tilde{E}_j = \Lambda^M(E_j)$

- Pauli noise model assume (ignore leakage, non-Markovianity, …)
    1) Assume 1q gate has negligible noise
    2) Assume multi-qubit Clifford gates and SPAM has Pauli noise channels
- The Pauli fidelity for $\Lambda^{S/M/\mathcal{G}}$ is denoted by $\lambda_a^{S/M/\mathcal{G}}$
    - WLOG, assume λ_a^S or λ_a^M only depend on the support of a

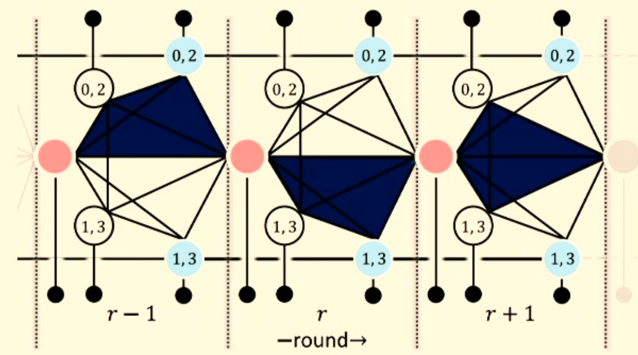
Why Pauli Noise Model



[Wallman and Emerson PRA 2016]
[A. Hashim et al., PRX 2021]



[Y. Kim et al., IBM, Nature 2023]



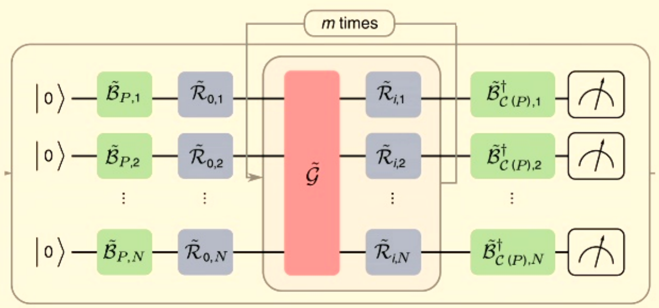
[E Chen et al., IBM, PRL. 2022]

- Generic noise can be twirled into Pauli channel via **randomized compiling**, given sufficiently good 1q gates.
- State-of-the-art **quantum error mitigation** techniques based on Pauli noise model
- Knowledge of Pauli noise rates useful for decoder optimizer in **QEC**.

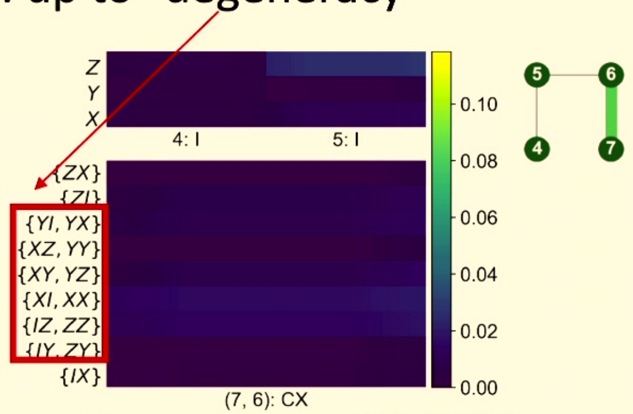
15

Pauli noise learning

- Start with a simpler problem: learn $\{\lambda_a^{\mathcal{G}}\}$ for every Clifford gate $\mathcal{G} \in \mathfrak{G}$
- **Cycle benchmarking** (CB) solves this problem … up to "degeneracy"



[A Erhard et al., Nat. Commum. (2019)]



[Akel Hashim et al., PRX (2021)]

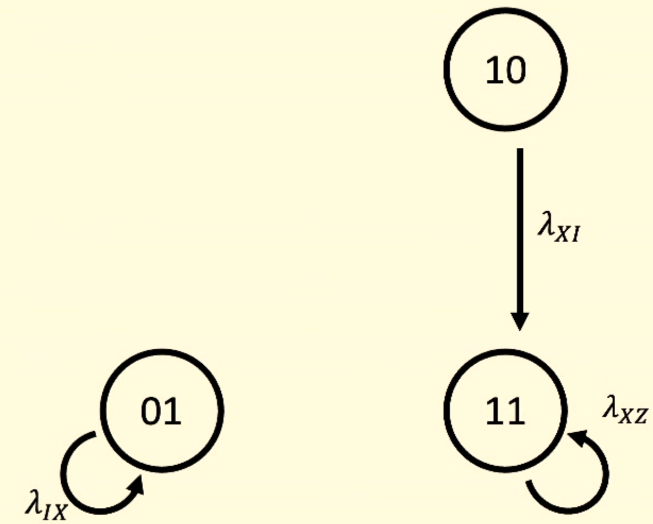
- We prove such degeneracy is a manifestation of **gauge freedom**
- Our results: Learnability of Pauli noise & learning algorithms

Graph for Pauli noise

- Define the pattern of a Pauli:
  - pt: $\mathrm{P}^n \mapsto \{0,1\}^n$ via $\{X, Y, Z\} \mapsto 1, I \mapsto 0$
  - E.g., $XYIIZ \mapsto 11001$
- Let's put all Pauli fidelities $\lambda_a^{\mathcal{G}}$ on a graph:
  - Nodes: All non-trivial Pauli pattern $\{0,1\}^n \backslash 0^n$
  - Edges: $\mathrm{pt}(P) \rightarrow \mathrm{pt}(\mathcal{G}(P))$ for all P and $\mathcal{G} \in \mathfrak{G}$
- Example: $\mathfrak{G} = \{\mathrm{CNOT}\}$

$IX \mapsto IX \quad XZ \mapsto YY \quad XI \mapsto XX$

$10 \mapsto 10 \quad 11 \mapsto 11 \quad 10 \mapsto 11$

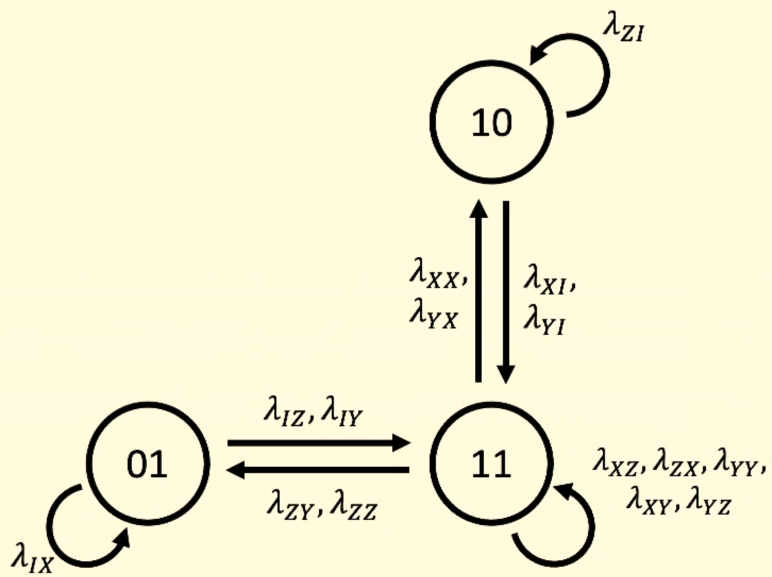


$n = 2, \mathcal{G} = \mathrm{CNOT}$

[SC*, Y. Liu* et al., Nature Communication (2023)]

Graph for Pauli noise

- Define the pattern of a Pauli:
  - pt: $\mathrm{P}^n \mapsto \{0,1\}^n$ via $\{X,Y,Z\} \mapsto 1, I \mapsto 0$
  - E.g., $XYIIZ \mapsto 11001$
- Let's put all Pauli fidelities $\lambda_a^{\mathcal{G}}$ on a graph:
  - Nodes: All non-trivial Pauli pattern $\{0,1\}^n \backslash 0^n$
  - Edges: $\mathrm{pt}(P) \to \mathrm{pt}(\mathcal{G}(P))$ for all P and $\mathcal{G} \in \mathfrak{G}$
- Example: $\mathfrak{G} = \{\mathrm{CNOT}\}$
  - Multiple edges only drawn once



$n = 2, \mathcal{G} = \mathrm{CNOT}$
Pattern Transform Graph (PTG)

[SC*, Y. Liu* et al., Nature Communication (2023)]

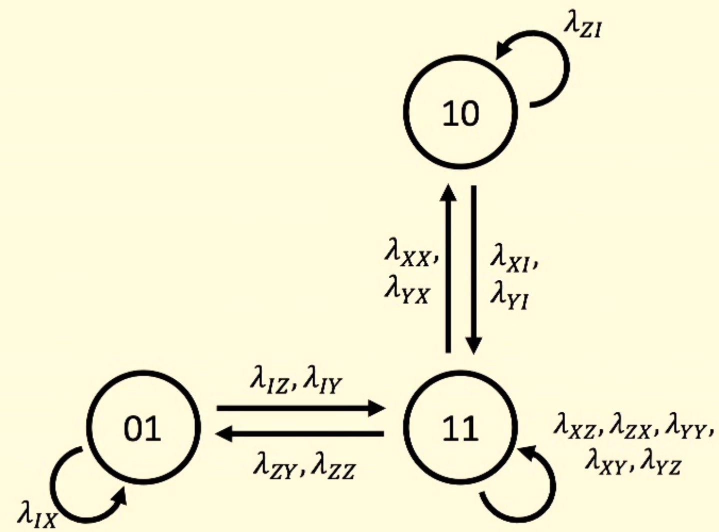
Learnability of Pauli noise

- Theorem:

1) Any $\lambda_P^{\mathcal{G}}$ is learnable $\Leftrightarrow$ $\lambda_P^{\mathcal{G}}$ is on a self-loop of PTG.
   - That is, $\mathrm{pt}(P) = \mathrm{pt}(\mathcal{G}(P))$.
   - CNOT: λ_{IX} is learnable, λ_{XX} or λ_{XI} is unlearnable.

2) Any product form $\lambda_{P_1}^{\mathcal{G}} \lambda_{P_2}^{\mathcal{G}} \cdots \lambda_{P_m}^{\mathcal{G}}$ is learnable $\Leftrightarrow$ $(\lambda_{P_1}^{\mathcal{G}}, \lambda_{P_2}^{\mathcal{G}}, \cdots, \lambda_{P_m}^{\mathcal{G}})$ forms a **cycle** of PTM.
   - CNOT: The product $\lambda_{XX}\lambda_{XI}$ is learnable

- Rigorous statement needs algebraic graph theory



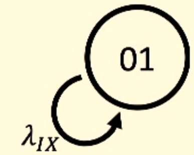
$n = 2$, CNOT

Pattern Transform Graph (PTG)

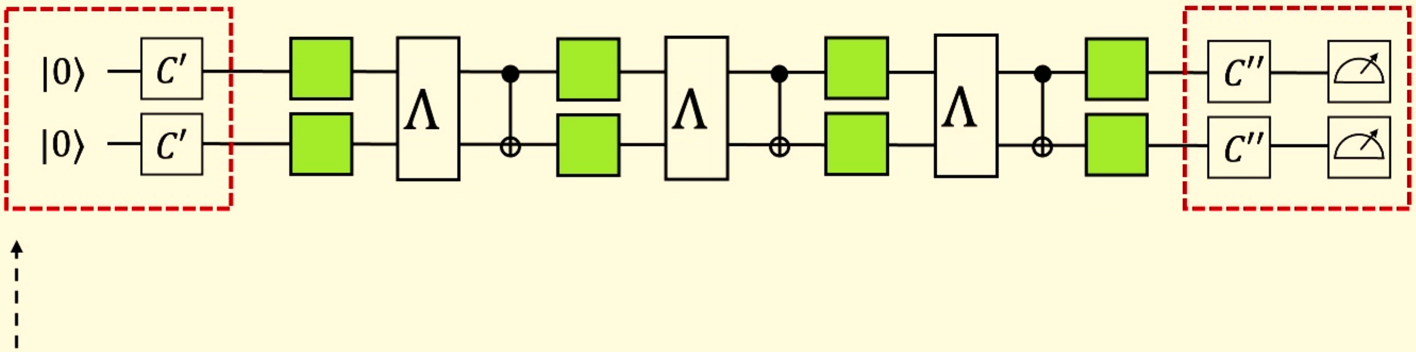
[SC*, Y. Liu* et al., Nature Communication (2023)]

Proof Sketch – Learnable part

1. For $\mathcal{G}(P_a) = P_a$, CB gives SPAM-robust estimate for $\lambda_a^{\mathcal{G}}$



$\mathrm{CNOT}(IX) = IX$



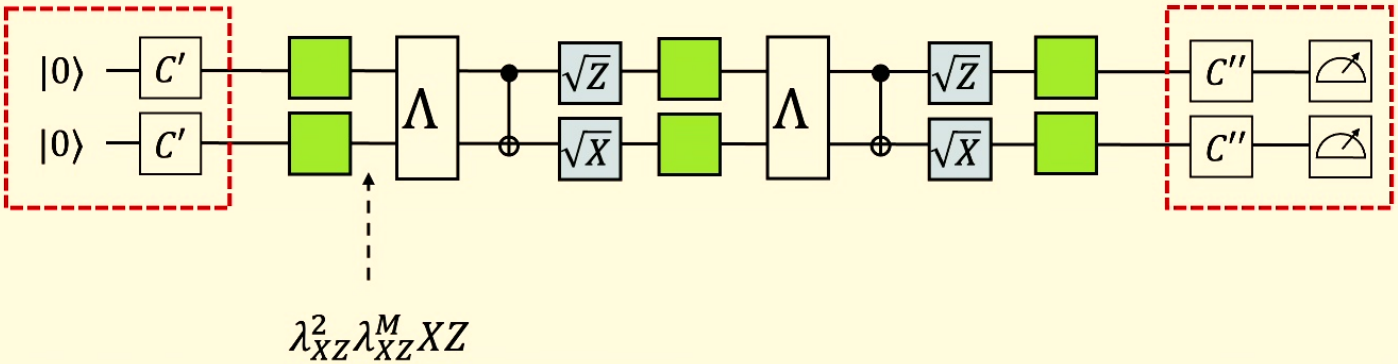
$\mathbb{E}_{\tilde{\rho}_0}\langle\widetilde{IX}\rangle = \lambda_{IX}^3 \lambda_{IX}^S \lambda_{IX}^M$

$\lambda_{IX}^S \coloneqq \mathrm{Tr}(\tilde{\rho}_0\, IX)$ Pauli fidelity of state preparation noise

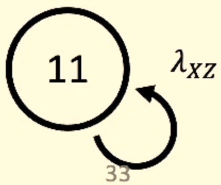
λ_{IX}^M: Pauli Fidelity for measurement noise

Proof Sketch – Learnable part

2. If $\mathcal{G}(P_a), P_a$ have same pattern, there is 1q Clifford C such that $(C \circ \mathcal{G})(P_a) = P_a$



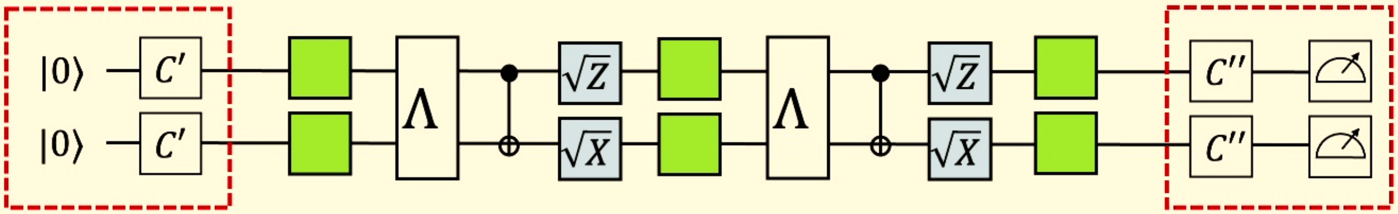
$\mathrm{CNOT}(XZ) = YY;$
$\sqrt{Z} \otimes \sqrt{X}(YY) = XZ$



33

Proof Sketch – Learnable part

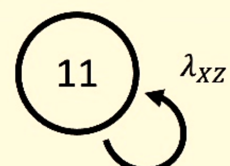
2. If $\mathcal{G}(P_a), P_a$ have same pattern, there is 1q Clifford C such that $(C \circ \mathcal{G})(P_a) = P_a$



Expectation value for depth t: $F_{XZ}(t) = \lambda_{XZ}^{t} \lambda_{XZ}^{S} \lambda_{XZ}^{M}$

gate noise SPAM noise

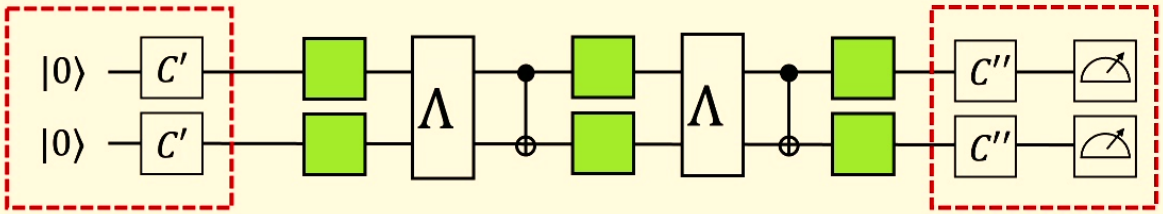
$\mathrm{CNOT}(XZ) = YY;$
$\sqrt{Z} \otimes \sqrt{X}(YY) = XZ$

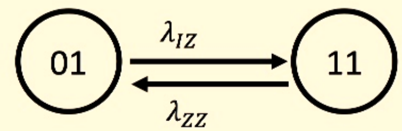


34

Proof Sketch – Learnable part

3. Products on any cycle on pattern transform graph can be estimated via CB.





Expectation value for depth $2t$: $F_{YY}(2t) = \lambda_{ZZ}^{t}\lambda_{IZ}^{t}\lambda_{ZZ}^{S}\lambda_{ZZ}^{M}$

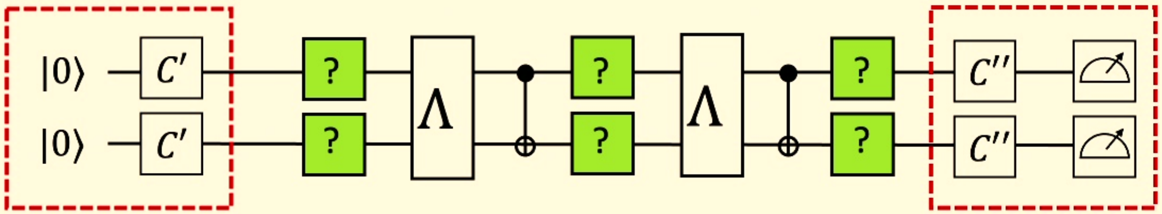
gate noise SPAM noise

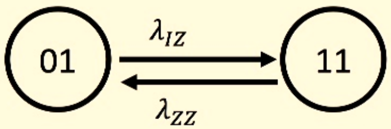
Fitting $\mathrm{F}(t)$ different t gives SPAM-robust estimate for $\sqrt{\lambda_{ZZ}\lambda_{IZ}}$

40

Break the degeneracy?

Can we turn ZZ back to IZ using single-qubit gates, so as to break degeneracy?

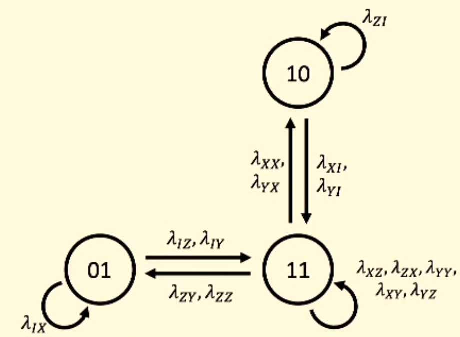




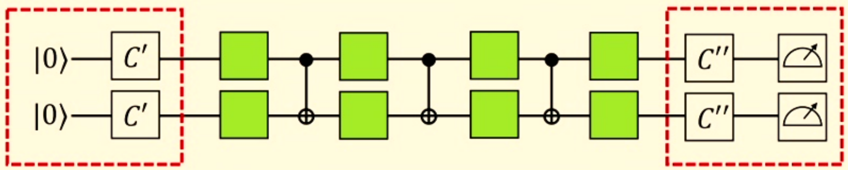
No, because single-qubit gates preserve Pauli patterns

(i.e., never transform between I and {X,Y,Z})

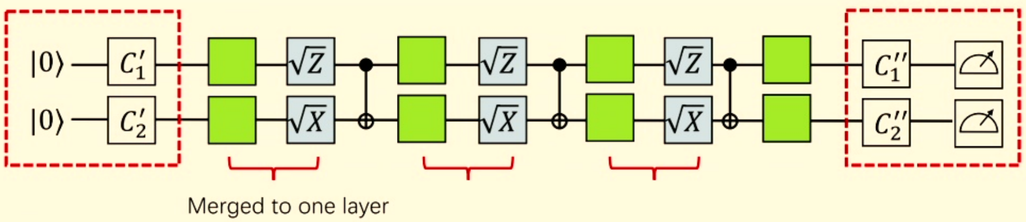
Proof Sketch – Learnable part



1. For $\mathcal{G}(P_a) = P_a$, CB gives SPAM-robust estimate for $\lambda_a^{\mathcal{G}}$



2. If $\mathcal{G}(P_a), P_a$ have same pattern, there is 1q Clifford C such that $(C \circ \mathcal{G})(P_a) = P_a$



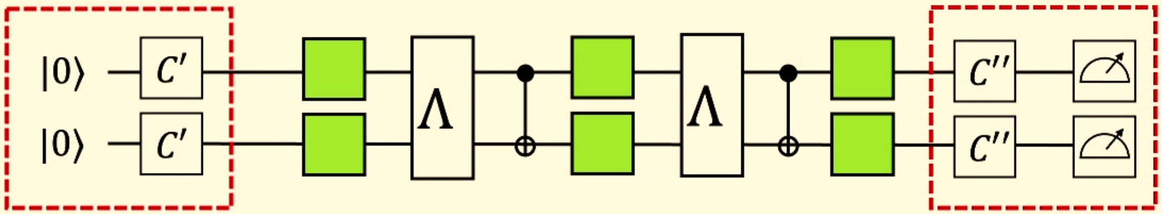
3. Finally, any cycle on pattern transform graph can be estimated via generalized CB.

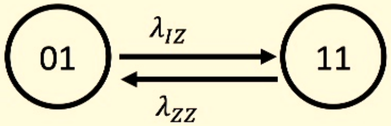
How to prove these are **all** the learnable functions?

42

Proof Sketch – Learnable part

3. Products on any cycle on pattern transform graph can be estimated via CB.





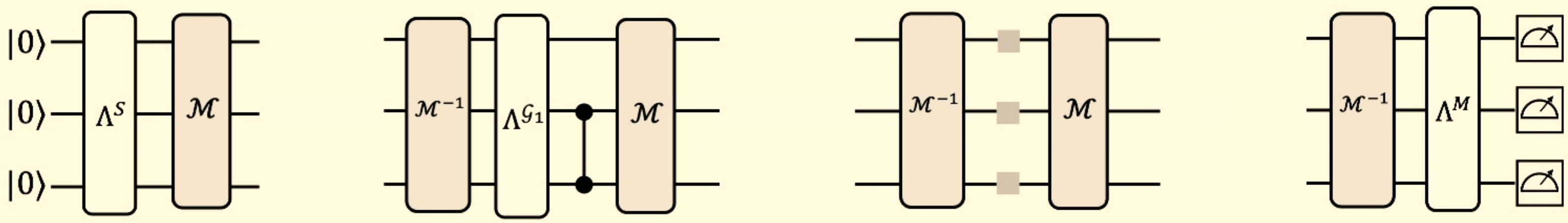
Expectation value for depth $2t$: $F_{YY}(2t) = \lambda_{ZZ}^{t}\lambda_{IZ}^{t}\lambda_{ZZ}^{S}\lambda_{ZZ}^{M}$

gate noise SPAM noise

Fitting $\mathrm{F}(t)$ different t gives SPAM-robust estimate for $\sqrt{\lambda_{ZZ}\lambda_{IZ}}$

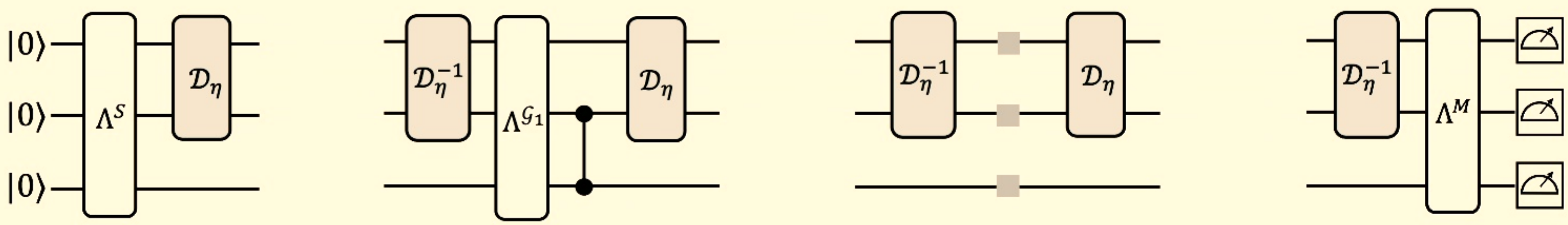
Gauge freedom for Pauli noise

- Recall the following **gauge transformation** (defined by an invertible map $\mathcal{M}$)



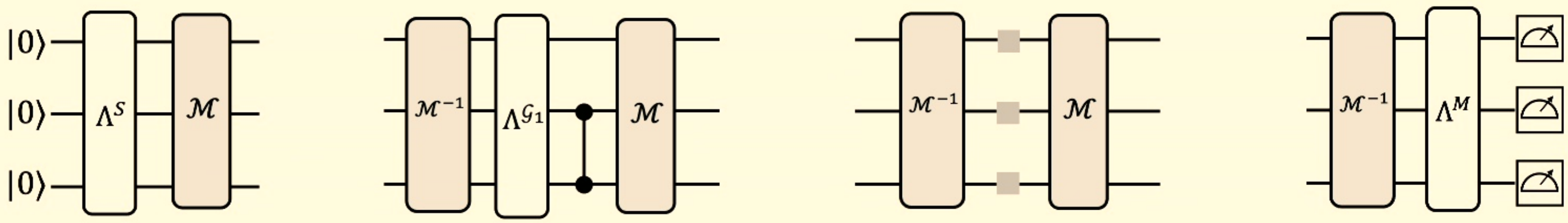
$\tilde{\rho} \mapsto \mathcal{M}(\tilde{\rho})$ $\quad$ $\tilde{\mathcal{G}} \mapsto \mathcal{M} \circ \tilde{\mathcal{G}} \circ \mathcal{M}^{-1}$ $\quad$ $\bigotimes_i \mathcal{U}_i \mapsto \mathcal{M} \circ \bigotimes_i \mathcal{U}_i \circ \mathcal{M}^{-1}$ $\quad$ $\tilde{E}_j \mapsto (\mathcal{M}^{-1})^*(E_j)$

- Let $\mathcal{M}$ be a depolarizing map $\mathcal{D}_\eta(\rho) = \eta\, I/d + (1-\eta)\,\rho$ on some subset of qubits



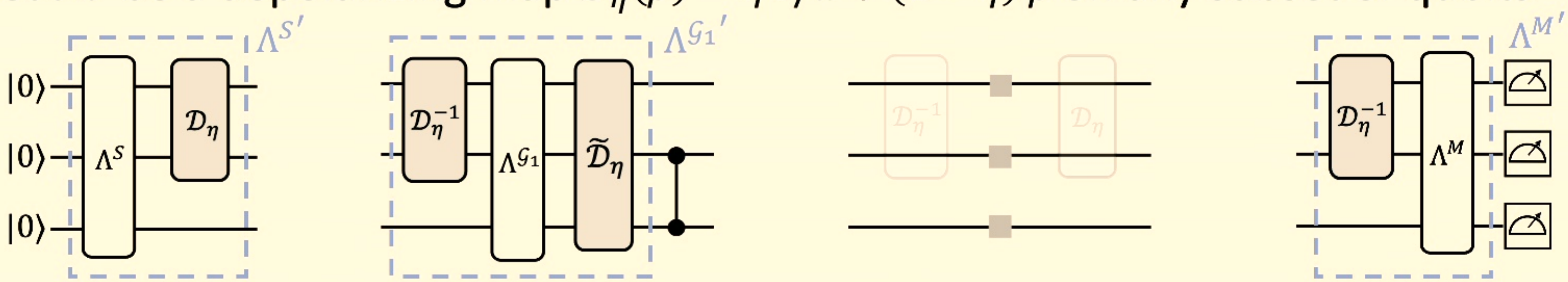
Gauge freedom for Pauli noise

- Recall the following **gauge transformation** (defined by an invertible map $\mathcal{M}$)



$\tilde{\rho} \mapsto \mathcal{M}(\tilde{\rho})$ $\quad$ $\tilde{\mathcal{G}} \mapsto \mathcal{M} \circ \tilde{\mathcal{G}} \circ \mathcal{M}^{-1}$ $\quad$ $\otimes_i \mathcal{U}_i \mapsto \mathcal{M} \circ \otimes_i \mathcal{U}_i \circ \mathcal{M}^{-1}$ $\quad$ $\tilde{E}_j \mapsto (\mathcal{M}^{-1})^*(E_j)$

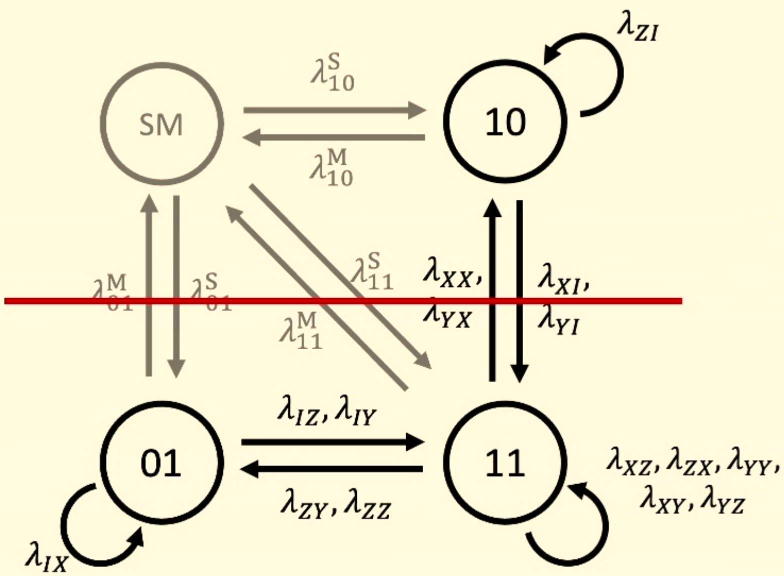
- Let $\mathcal{M}$ be a depolarizing map $\mathcal{D}_\eta(\rho) = \eta\, I/d + (1-\eta)\,\rho$ on any subset of qubits



- Lemma: Subsystem depolarizing gauge (SDG) is a valid **gauge transformation** for Pauli noise

Proof Sketch – Unlearnable part

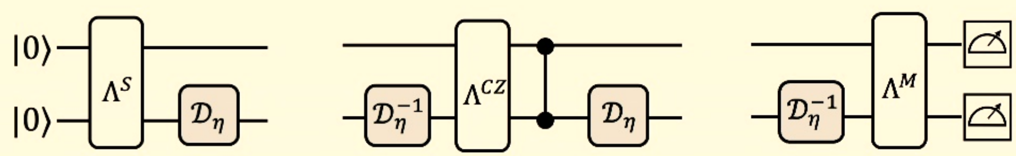
- Any cut defines a **gauge transformation** of the Pauli noise model.



- Add a root node to encode SPAM params
  - WLOG, $\lambda_a^{S/M}$ only depends on $\mathrm{pt}(a)$
- For edges crossing the cut, multiply the fidelity by η or η^{-1} depending on its direction:

$$\lambda_{XX}, \lambda_{YX}, \lambda^S_{11}, \lambda^S_{01} *= \eta;$$
$$\lambda_{XI}, \lambda_{YI}, \lambda^M_{11}, \lambda^M_{01} /= \eta$$

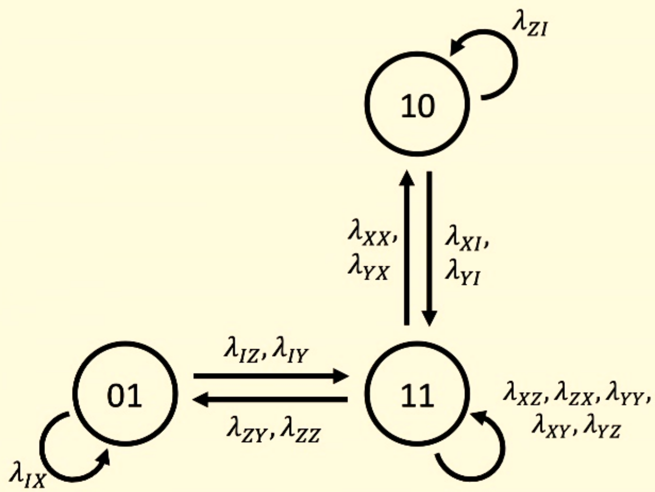
- This corresponding to the following SDG:



Thus, provably indistinguishable by any experiments

- Learnable function must be orthogonal to all gauge transformation
- Now that Gauge space = Cut space ⊥ Cycle space, learnable functions must be inside cycle space

Full example of CNOT

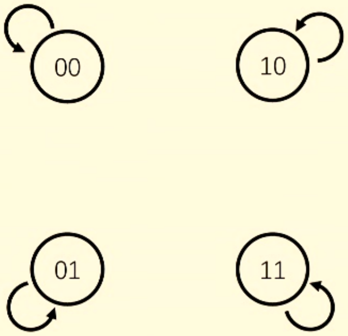


Pattern transform graph for CNOT

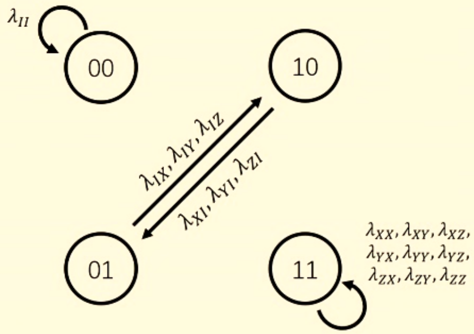
Gate	CNOT
(a) Cycle basis	$l_{II}, l_{ZI}, l_{IX}, l_{ZX}, l_{XZ}, l_{YY}, l_{XY}, l_{YZ}$, $l_{IZ}+l_{ZZ}, l_{IY}+l_{ZY}, l_{IZ}+l_{ZY}$, $l_{XI}+l_{XX}, l_{YI}+l_{YX}, l_{XI}+l_{YX}$
(b) Learnable Pauli fidelities	$\lambda_{II}, \lambda_{ZI}, \lambda_{IX}, \lambda_{ZX}, \lambda_{XZ}, \lambda_{YY}, \lambda_{XY}, \lambda_{YZ}$, $\lambda_{IZ}\cdot\lambda_{ZZ}, \lambda_{IY}\cdot\lambda_{ZY}, \lambda_{IZ}\cdot\lambda_{ZY}$, $\lambda_{XI}\cdot\lambda_{XX}, \lambda_{YI}\cdot\lambda_{YX}, \lambda_{XI}\cdot\lambda_{YX}$
(c) Learnable Pauli errors (1st order approx.)	$p_{II}, p_{ZI}, p_{IX}, p_{ZX}, p_{XZ}, p_{YY}, p_{XY}, p_{YZ}$, $p_{IZ}+p_{ZZ}, p_{IY}+p_{ZY}, p_{IZ}+p_{ZY}$, $p_{XI}+p_{XX}, p_{YI}+p_{YX}, p_{XI}+p_{YX}$
(d) Unlearnable degrees of freedom	$\lambda_{XI}, \lambda_{IZ}$ (e.g.)

Num of unlearnable parameters = dim of cut space

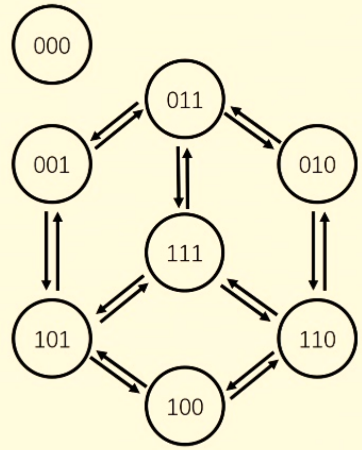
More examples



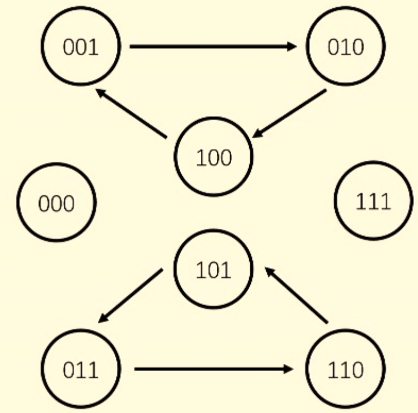
Identity



SWAP



$\{CNOT_{12}, CNOT_{23}, CNOT_{31}\}$



$CIRC_3$

Number of qubits n	Gate set $\mathfrak{G}$	Number of params	Unlearnable degrees
2	I	16	0
2	CNOT	16	2
2	SWAP	16	1
3	$\{\mathrm{CNOT}_{12}, \mathrm{CNOT}_{23}, \mathrm{CNOT}_{31}\}$	192	6
3	CIRC_3	64	4

Fact: Number of unlearnable DOFs is at most $2^n - 1$.

Rigorous statement

- Let's put the Pauli noise model into a linear space:
  1. Let $\boldsymbol{x}$ be a list of all log Pauli fidelities $x_a^{\mathcal{S}\backslash M\backslash \mathcal{G}} := -\log \lambda_a^{\mathcal{S}\backslash M\backslash \mathcal{G}}$
     - **Parameter space** X: The real linear space where $\boldsymbol{x}$ lives in
  2. Experiment $\boldsymbol{F}_{\mathcal{C}}$: $X \mapsto \mathbb{R}^{2^n}$ such that $\boldsymbol{F}_{\mathcal{C}}[j] = \mathrm{Tr}(\tilde{E}_j \tilde{\mathcal{C}}(\tilde{\rho}_0))$ for operation sequence $\mathcal{C}$*
  3. Learnable linear function $\boldsymbol{f} \in X^*$: $\boldsymbol{f}(\boldsymbol{x})$ can be determined by a set of experiments $\boldsymbol{F}_k(\boldsymbol{x})$
     - **Learnable space** L: subspace of all learnable functions
  4. Gauge vectors $\boldsymbol{\mathfrak{d}} \in X$: For any experiments $\boldsymbol{F}$ and $\boldsymbol{x} \in X$, $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{F}(\boldsymbol{x} + \boldsymbol{\mathfrak{d}})$.
     - **Gauge space** T: subspace of all gauge vectors
- Fact: $L \perp T$, i.e., $\boldsymbol{f}(\boldsymbol{\mathfrak{d}}) \equiv \boldsymbol{f} \cdot \boldsymbol{\mathfrak{d}} = 0$ for all learnable $\boldsymbol{f}$ and gauge $\boldsymbol{\mathfrak{d}}$.

* Extend the definitions of experiments to non-physical states for mathematical convenience

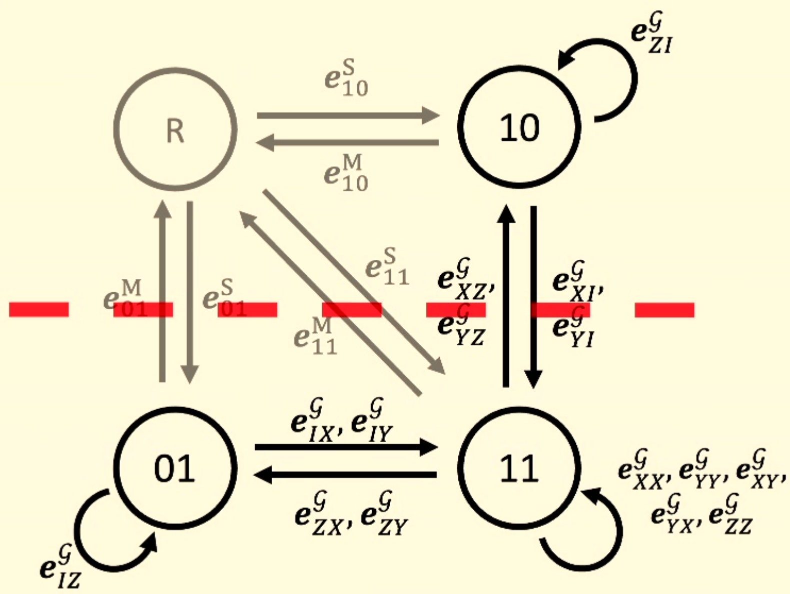
Linear space for graph

- Define linear spaces on pattern transform graph:

1. Edge space E: real linear space spanned by all edges $\{\boldsymbol{e_i}\}$
2. Cycle space Z: spanned by all **cycle vectors**
   - e.g.: $\boldsymbol{e}_{01}^{S} + \boldsymbol{e}_{IX}^{\mathcal{G}} + \boldsymbol{e}_{11}^{M}$
3. Cut space U: spanned by all cut vectors
   - e.g.: $\boldsymbol{e}_{01}^{S} + \boldsymbol{e}_{11}^{S} + \boldsymbol{e}_{XI}^{\mathcal{G}} + \boldsymbol{e}_{XI}^{\mathcal{G}} - \boldsymbol{e}_{01}^{M} - \boldsymbol{e}_{11}^{M} - \boldsymbol{e}_{XZ}^{\mathcal{G}} - \boldsymbol{e}_{YZ}^{\mathcal{G}}$

- Fact: $E = Z \oplus U$ and $Z \perp U$

Pattern: pt$(XIZYI) \mapsto 10110$



$\mathfrak{G} := \{\mathcal{G} = \mathrm{CZ}\}$

See e.g. *[Gleiss, Leydold, Stadler., Discuss. Math. Graph Theory 23, 241 (2003).]*

Learnability of Pauli noise

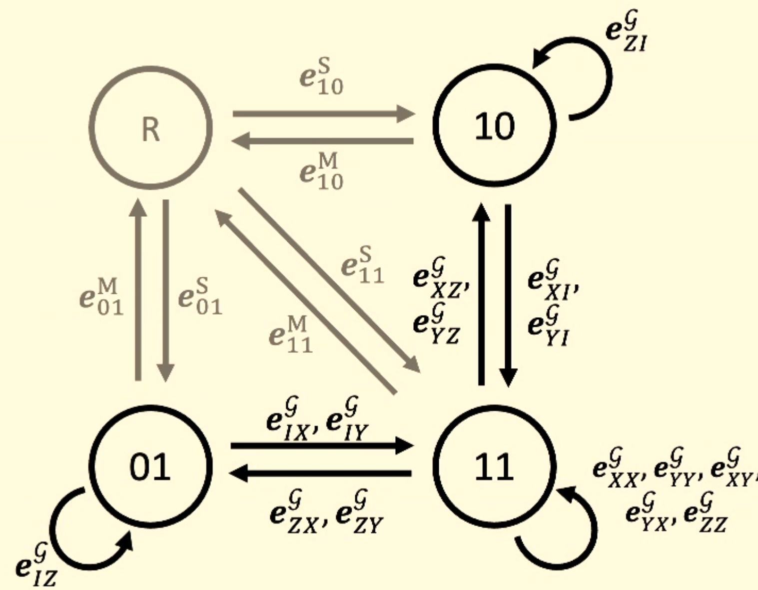
- Theorem:

$$\underset{\text{(edge)}}{E} = \underset{\text{(cycle)}}{Z} \oplus^{\perp} \underset{\text{(cut)}}{U}$$

$$\parallel \qquad \parallel \qquad \parallel$$

$$\underset{\text{(parameter)}}{X} = \underset{\text{(learnable)}}{L} \oplus^{\perp} \underset{\text{(gauge)}}{T}$$

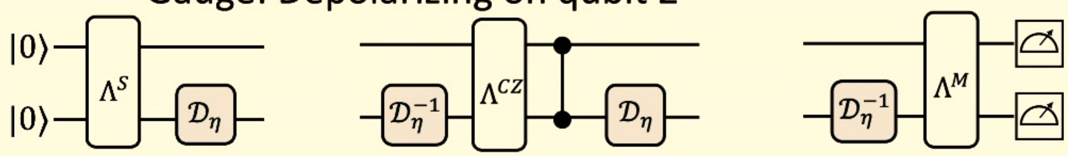
Note: $\oplus^{\perp}$ stands for orthogonal complement
Note 2: Identify X and X' using the standard basis $\{\boldsymbol{e}_i\}$

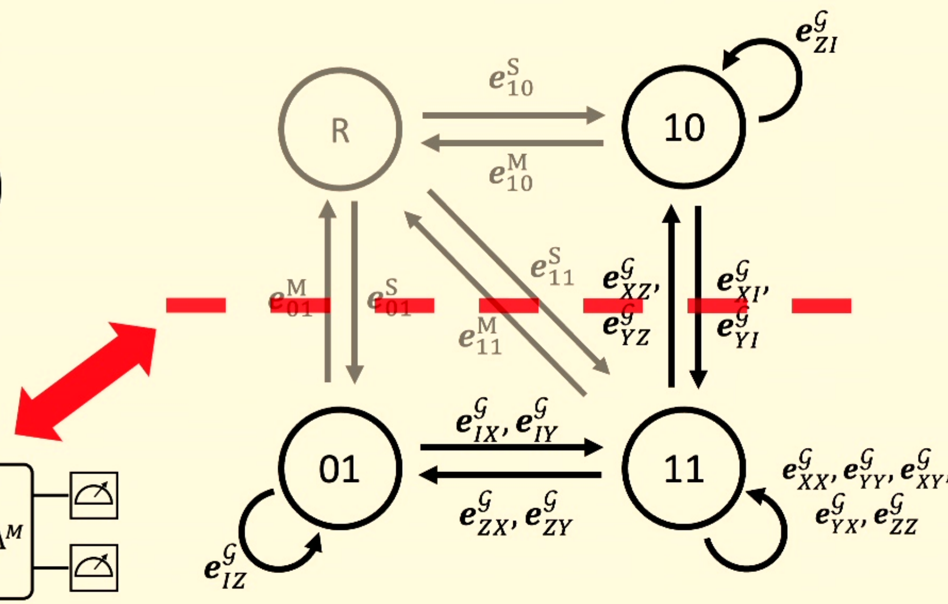


SC, Liu, et al. “The learnability of Pauli noise”. Nat. Commun. (2023).
SC, Z Zhang, L Jiang, S Flammia. arXiv: 2410.03906 (2024)

Key proof ideas (more rigorously)

- Any **rooted cycle** implies an experiment
  - Clifford gates sequence + Pauli measurements
  - $\mathrm{Tr}(\tilde{P}\tilde{\mathcal{C}}(\tilde{\rho}_0)) = \lambda^{S}_{\mathrm{pt}(a_1)}\lambda^{\mathcal{G}_1}_{a_1}\lambda^{\mathcal{G}_2}_{a_2}\cdots\lambda^{\mathcal{G}_M}_{a_M}\lambda^{M}_{\mathrm{pt}(a_M)}$
    $= \exp\left(-\left(x^{S}_{\mathrm{pt}(a_1)} + \cdots + x^{M}_{\mathrm{pt}(a_M)}\right)\right)$
  - Rooted cycles span cycle space
- Any cut vector is a gauge vector
  - Cut: {R, 10} / {01, 11} ⇔
  - Gauge: Depolarizing on qubit 2

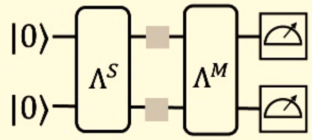


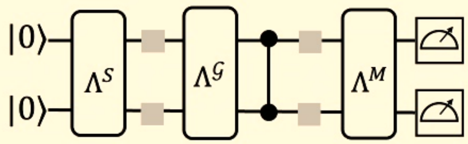


[1] Chen, Liu, et al. "The learnability of Pauli noise". Nat. Commun. (2023).

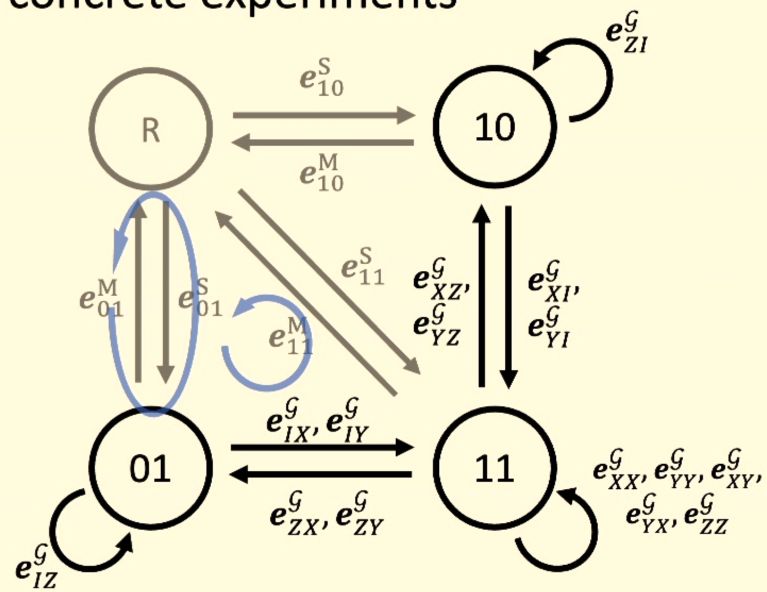
Learning algorithm

- Learnable space = Cycle space; Rooted cycles yield concrete experiments
- Find a rooted cycle basis and learn all of them.
- $\{e_t^S + e_t^M\} \quad \cup \quad \{e_{\mathrm{pt}(a)}^S + e_a^{\mathcal{G}} + e_{\mathrm{pt}(\mathcal{G}(a))}^M\}$



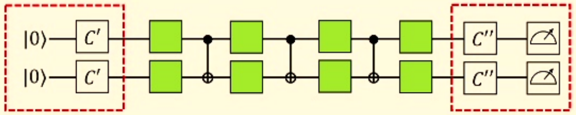


  - Learn $\mathcal{G} \in \mathfrak{G}$ one-by-one. No concatenation needed.



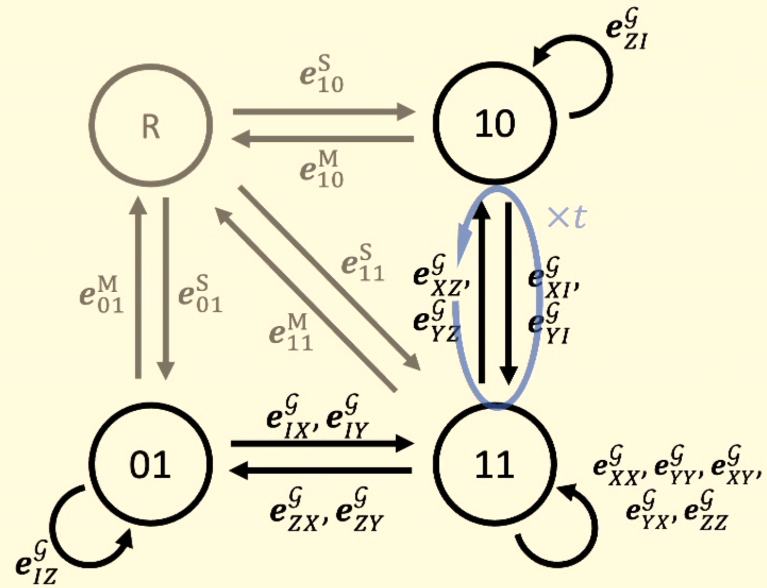
SC, Z Zhang, L Jiang, S Flammia. arXiv: 2410.03906 (2024)

Learning to relative precision



- One often hopes to learn noise parameters to relative precision
  - With a small number of measurements
- Key: Amplify certain noise parameters using concatenation $\lambda^{S}_{\mathrm{pt}(a_1)}(\lambda^{\mathcal{G}_1}_{a_1}\lambda^{\mathcal{G}_2}_{a_2}\cdots\lambda^{\mathcal{G}_M}_{a_M})^t\,\lambda^{M}_{\mathrm{pt}(a_M)}$
  $= \exp(-(\boldsymbol{x}_{SPAM} + \boldsymbol{x}^t_{\mathrm{cycles}})$
- Theorem (informally):

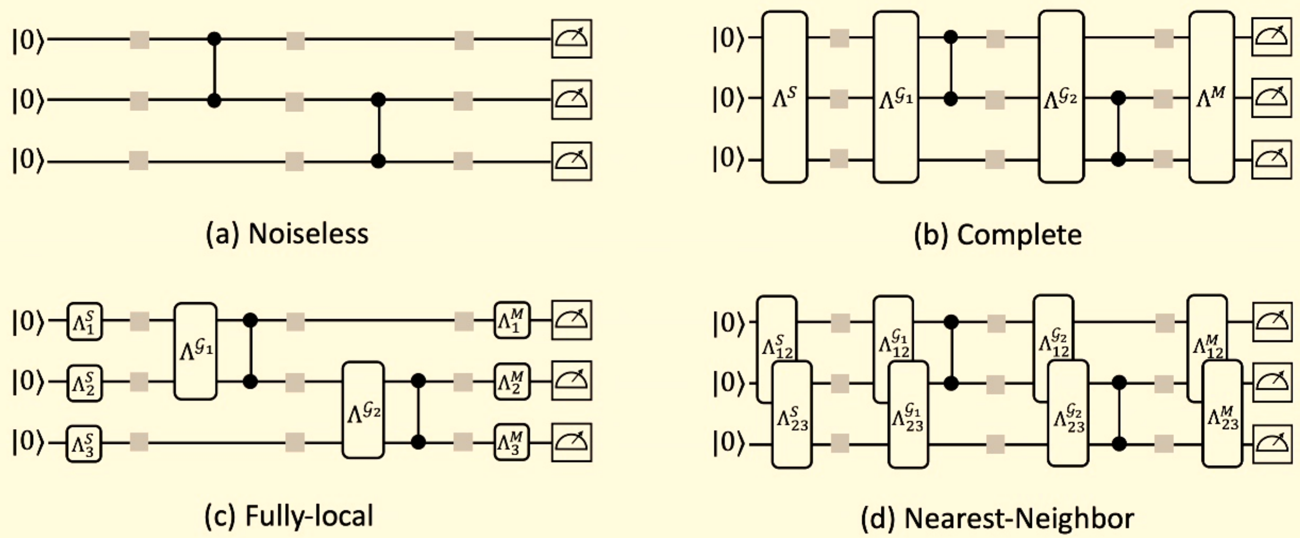
Any cycle consisting only of gate noise params can by amplified and learned via concatenation



57

SC, Z Zhang, L Jiang, S Flammia. arXiv: 2410.03906 (2024)

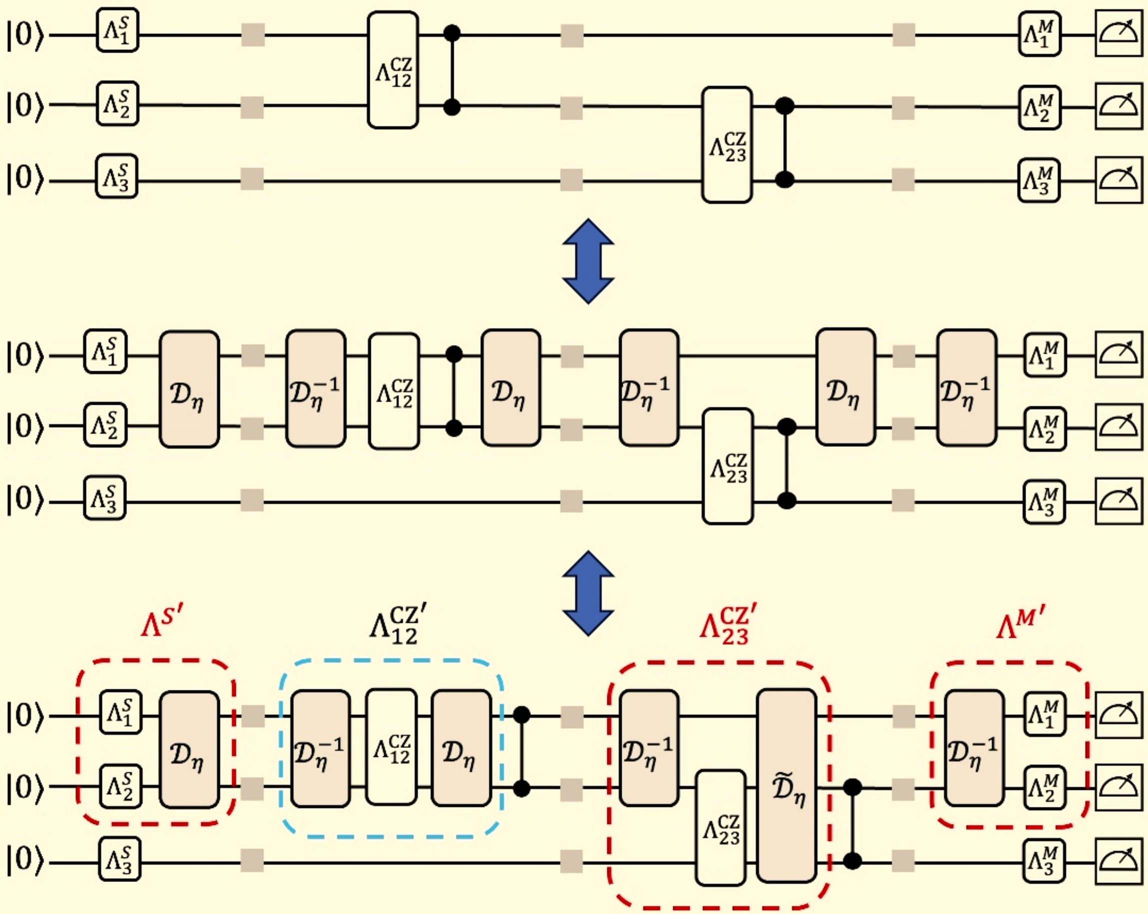
Reduced Pauli noise model



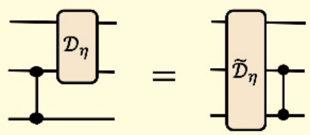
- A complete Pauli noise model contains **exponentially** many parameters
- In practice, additional noise assumptions (such as locality) are often introduced.
- We refer to such as **reduced** Pauli noise models.
- How to learn a reduced Pauli noise models?

58

Reduced model



- Certain gauge transformations become invalid due to additional locality assumptions.
- In this example, the transformed $\Lambda^{S'}, \Lambda_{23}^{CZ'}, \Lambda^{M'}$ are not within the fully-local ansatz. Thus, the gauge is invalid.



60

Definitions for reduced model

- Reduced noise model $(X_R, \mathcal{Q})$:
  1. **Reduced parameter space** X_R: where the reduced set of parameters $\boldsymbol{r}$ lives in
  2. **Embedding map** $\mathcal{Q}$: describes how $\boldsymbol{r}$ defines the Pauli channels. $\boldsymbol{x} = \mathcal{Q}(\boldsymbol{r})$

     We require $\mathcal{Q}$ to be linear and injective

- Similar Linear algebraic definitions:
  1. **Experiments** on reduced models $X_R \mapsto \mathbb{R}^{2^n}$ $\quad \boldsymbol{F}_R(\boldsymbol{r}) \coloneqq \boldsymbol{F}\,(\mathcal{Q}(\boldsymbol{r}))$
  2. **Reduced learnable function** $\boldsymbol{f} \in X_R^*$: can be determined by a set of experiments
     - Reduced learnable space L_R: subspace of all reduced learnable functions
  3. **Reduced gauge vectors** $\mathfrak{d} \in X_R$: For any experiments $\boldsymbol{F}_R$ and $\boldsymbol{r} \in X_R$, $\boldsymbol{F}_R(\boldsymbol{r}) = \boldsymbol{F}_R(\boldsymbol{r} + \mathfrak{d})$.
     - Reduced gauge space T_R: subspace of all reduced gauge vectors

Learnability of reduced model

- Theorem:
- $L_R = Q^{\mathrm{T}}(L) \equiv \{f(Q(\cdot)) \mid \forall f \in L\}$
  - Reduced learnable space are complete learnable space projected by Q^{T}
- $T_R = Q^{-1}(T) \equiv \{\mathfrak{d} \in X_R \mid Q(\mathfrak{d}) \in T\}$
  - Reduced gauge space are the preimage of complete gauge space via Q
  - More intuitively, $Q(T_R) = T \cap \mathrm{Im}Q$, a gauge is allowed iff it is in the image of embedding

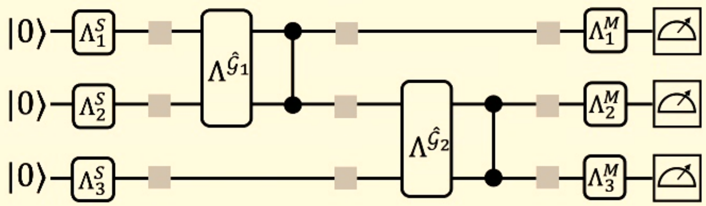
- Putting everything together:

$$
\begin{array}{ccccc}
\underset{\text{(edge)}}{E} & = & \underset{\text{(cycle)}}{Z} & \oplus^{\perp} & \underset{\text{(cut)}}{U} \\
\parallel & & \parallel & & \parallel \\
\underset{\text{(parameter)}}{X} & = & \underset{\text{(learnable)}}{L} & \oplus^{\perp} & \underset{\text{(gauge)}}{T} \\
\uparrow Q & & \downarrow Q^{\mathrm{T}} & & \downarrow Q^{-1} \\
\underset{\text{(reduced param)}}{X_R} & = & \underset{\text{(reduced learnable)}}{L_R} & \oplus^{\perp} & \underset{\text{(Reduced gauge)}}{T_R}
\end{array}
$$

- Now we have a linear algebra procedure to determine L_R and T_R.

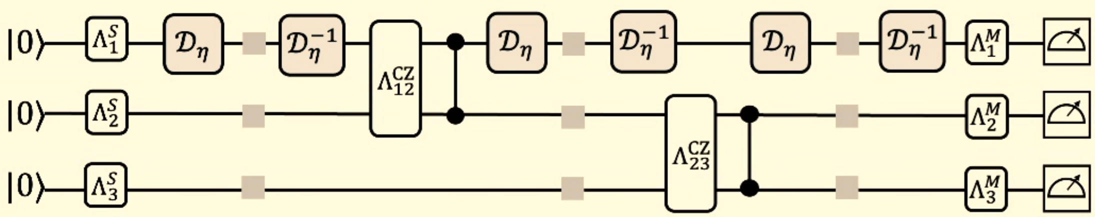
* Q^{T} denotes the conjugate map of Q. Here it can basically be understood as transposition.

Example: Fully-local model



- Noise ansatz:
- Local SPAM noise: Product of 1q noise
- Local Gate noise: within gate support

- **Thm**: The allowed gauge space $\mathcal{Q}(T_R)$ is spanned by the 1-qubit depolarizing gauges
  - Only n gauge params



- Efficient learning discussed in paper
  - Local experiments suffices, usually

SC, Z Zhang, L Jiang, S Flammia. arXiv: 2410.03906 (2024)

Summary

- We develop a framework of efficient self-consistent gate set Pauli noise learning

1. Characterization of learnable/gauge space for flexible gate set/noise locality
2. Efficient learning algorithms, additive vs. relative precision
3. Applications to concrete examples, experiments

- Outlook:

1. Graph-theoretical techniques beyond Pauli noise model
2. Including MCMs *[Zhang et al. 2406.02669, Hines et al. 2406.09299]*, extending to FTQC learning.
3. Applications to efficient self-consistent quantum error mitigation (on-going)
4. Optimal experiment design, fine-grained complexity analysis

80

What if we simply ignore SP noise?

- Got unphysical results!
- This implies the assumptions of perfect SP is NOT practical
- Violation of physical constraints can be used to lower bound SP noise $\sim 0.6\%$
  - Could also come from other imperfectness of assumptions