Quantum Wednesday

`arXiv:2201.10672`

Nate Stemen

Nov 16, 2022

Today's Paper

Efficiently improving the performance of noisy quantum computers

Samuele Ferracin,[1,2,*] Akel Hashim,[3,4,*] Jean-Loup Ville,[3,†] Ravi Naik,[3,4] Arnaud Carignan-Dugas,[1] Hammam Qassim,[1] Alexis Morvan,[3,4,‡] David I. Santiago,[3,4] Irfan Siddiqi,[3,4,5] and Joel J. Wallman[1,2]

[1] *Keysight Technologies Canada, Kanata, ON K2K 2W5, Canada*
[2] *Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*
[3] *Quantum Nanoelectronics Laboratory, Dept. of Physics, University of California at Berkeley, Berkeley, CA 94720, USA*
[4] *Computational Research Division, Lawrence Berkeley National Lab, Berkeley, CA 94720*
[5] *Materials Sciences Division, Lawrence Berkeley National Lab, Berkeley, CA 94720, USA*

(Dated: July 20, 2022)

https://arxiv.org/abs/2201.10672

Main Ideas

- Noiseless Output Extrapolation (NOX)
- Pauli Error Cancellation (PEC)

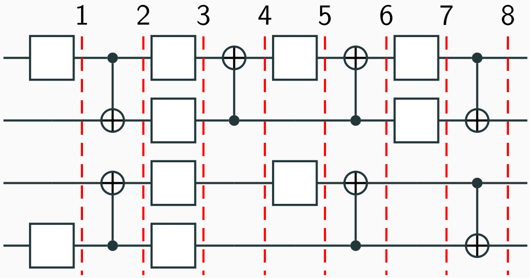
Main Ideas

- Noiseless Output Extrapolation (NOX)
- Pauli Error Cancellation (PEC)

Question

How are these different from Zero-Noise Extrapolation (ZNE) and Probabilistic Error Cancellation (PEC)?

Cycles

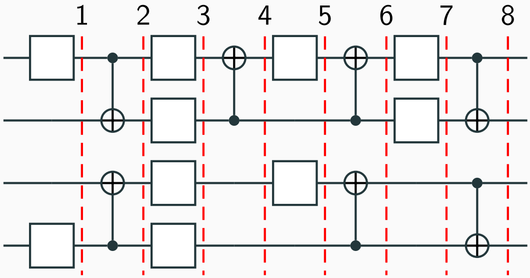


- $\mathcal{E}_i$: single qubit cycles
- $\mathcal{H}_i$: Clifford two-qubit cycles

$$\mathcal{C} = \mathcal{E}_{m+1}\mathcal{H}_m\mathcal{E}_m \cdots \mathcal{E}_2\mathcal{H}_1\mathcal{E}_1 \tag{1}$$

$$=: \mathcal{E}_{m+1}\left(\bigcirc_{j=1}^{m} \mathcal{H}_j\mathcal{E}_j\right) \tag{2}$$

Cycles

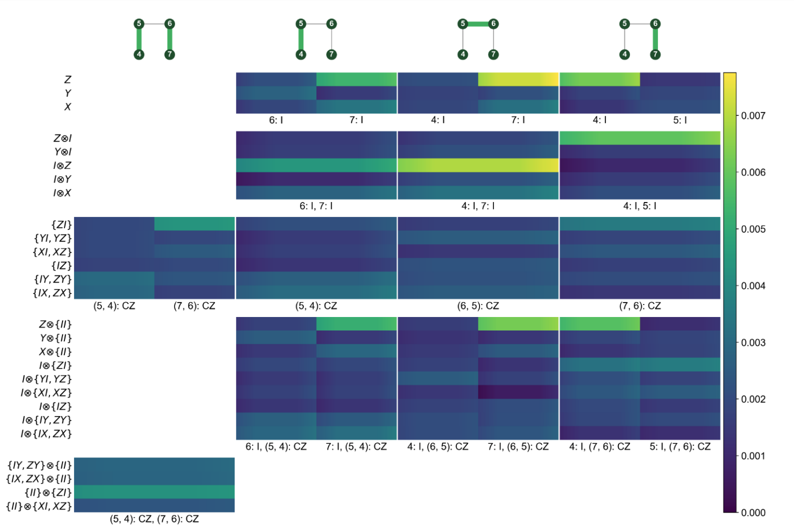


- $\mathcal{E}_i$: single qubit cycles
- $\mathcal{H}_i$: Clifford two-qubit cycles

$$\mathcal{C} = \mathcal{E}_{m+1}\mathcal{H}_m\mathcal{E}_m \cdots \mathcal{E}_2\mathcal{H}_1\mathcal{E}_1 \tag{1}$$

$$=: \mathcal{E}_{m+1}\left(\bigcirc_{j=1}^{m} \mathcal{H}_j\mathcal{E}_j\right) \tag{2}$$

Cycle Error Reconstruction[1] (CER)



[1] Hashim et al., "Randomized Compiling for Scalable Quantum Computing on a Noisy Superconducting Quantum Processor".

Assumptions:

1. Noise is Markovian and time stationary ($\mathcal{D}_{\mathcal{U}}\mathcal{U}$).
2. $\mathcal{D}_{\mathcal{U}} = \mathcal{D}$ when $\mathcal{U}$ consists solely of single qubit gates ($\mathcal{D}_{\mathcal{E}_i}\mathcal{E}_i = \mathcal{D}\mathcal{E}_i$ for all i).

Noise processes $\mathcal{D}_{\mathcal{U}}$ taken to be a Pauli channels by Randomized Compiling.[2]

$$\mathcal{D}_{\mathcal{U}}(\rho) = \sum_{k=0}^{4^n-1} \epsilon_k^{(\mathcal{U})}\mathcal{P}_k(\rho)$$

$$\mathcal{P}_k \in \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}^{\otimes n}$$

$\epsilon_k^{(\mathcal{U})}$: Pauli error rates

[2] Wallman and Emerson, "Noise tailoring for scalable quantum computation via randomized compiling".

Setup

Assumptions:

1. Noise is Markovian and time stationary ($\mathcal{D}_{\mathcal{U}}\mathcal{U}$).
2. $\mathcal{D}_{\mathcal{U}} = \mathcal{D}$ when $\mathcal{U}$ consists solely of single qubit gates ($\mathcal{D}_{\mathcal{E}_i}\mathcal{E}_i = \mathcal{D}\mathcal{E}_i$ for all i).

Noise processes $\mathcal{D}_{\mathcal{U}}$ taken to be a Pauli channels by Randomized Compiling.[2]

$$\mathcal{D}_{\mathcal{U}}(\rho) = \sum_{k=0}^{4^n-1} \epsilon_k^{(\mathcal{U})} \mathcal{P}_k(\rho)$$

$\mathcal{P}_k \in \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}^{\otimes n}$

$\epsilon_k^{(\mathcal{U})}$: Pauli error rates

[2] Wallman and Emerson, "Noise tailoring for scalable quantum computation via randomized compiling".

Assumptions:

1. Noise is Markovian and time stationary ($\mathcal{D}_{\mathcal{U}}\mathcal{U}$).
2. $\mathcal{D}_{\mathcal{U}} = \mathcal{D}$ when $\mathcal{U}$ consists solely of single qubit gates ($\mathcal{D}_{\mathcal{E}_i}\mathcal{E}_i = \mathcal{D}\mathcal{E}_i$ for all i).

Noise processes $\mathcal{D}_{\mathcal{U}}$ taken to be a Pauli channels by Randomized Compiling.[2]

$$\mathcal{D}_{\mathcal{U}}(\rho) = \sum_{k=0}^{4^n-1} \epsilon_k^{(\mathcal{U})} \mathcal{P}_k(\rho)$$

$\mathcal{P}_k \in \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}^{\otimes n}$

$\epsilon_k^{(\mathcal{U})}$: Pauli error rates

[2] Wallman and Emerson, "Noise tailoring for scalable quantum computation via randomized compiling".

Setup

Assumptions:

1. Noise is Markovian and time stationary ($\mathcal{D}_{\mathcal{U}}\mathcal{U}$).
2. $\mathcal{D}_{\mathcal{U}} = \mathcal{D}$ when $\mathcal{U}$ consists solely of single qubit gates ($\mathcal{D}_{\mathcal{E}_i}\mathcal{E}_i = \mathcal{D}\mathcal{E}_i$ for all i).

Noise processes $\mathcal{D}_{\mathcal{U}}$ taken to be a Pauli channels by Randomized Compiling.[2]

$$\mathcal{D}_{\mathcal{U}}(\rho) = \sum_{k=0}^{4^n-1} \epsilon_k^{(\mathcal{U})}\mathcal{P}_k(\rho)$$

$\mathcal{P}_k \in \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}^{\otimes n}$

$\epsilon_k^{(\mathcal{U})}$: Pauli error rates

[2] Wallman and Emerson, "Noise tailoring for scalable quantum computation via randomized compiling".

Setup

Assumptions:

1. Noise is Markovian and time stationary ($\mathcal{D}_{\mathcal{U}}\mathcal{U}$).
2. $\mathcal{D}_{\mathcal{U}} = \mathcal{D}$ when $\mathcal{U}$ consists solely of single qubit gates ($\mathcal{D}_{\mathcal{E}_i}\mathcal{E}_i = \mathcal{D}\mathcal{E}_i$ for all i).

Noise processes $\mathcal{D}_{\mathcal{U}}$ taken to be a Pauli channels by Randomized Compiling.[2]

$$\mathcal{D}_{\mathcal{U}}(\rho) = \sum_{k=0}^{4^n-1} \epsilon_k^{(\mathcal{U})} \mathcal{P}_k(\rho)$$

$$\mathcal{P}_k \in \{\mathcal{I}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}\}^{\otimes n}$$

$\epsilon_k^{(\mathcal{U})}$: Pauli error rates

[2] Wallman and Emerson, "Noise tailoring for scalable quantum computation via randomized compiling".

Noiseless Output Extrapolation

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}_{\mathcal{H}_i}^{\alpha}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
  1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
  2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle:
  $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots \left(\mathcal{D}_{\mathcal{H}_j}\right)^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\text{NOX}}(O) := E_{\text{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\text{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

Noiseless Output Extrapolation

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}^{\alpha}_{\mathcal{H}_i}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
  1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
  2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle:
  $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots \left(\mathcal{D}_{\mathcal{H}_j}\right)^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\text{NOX}}(O) := E_{\text{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\text{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

Noiseless Output Extrapolation

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}^{\alpha}_{\mathcal{H}_i}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
  1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
  2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle:
  $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots \left(\mathcal{D}_{\mathcal{H}_j}\right)^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\text{NOX}}(O) := E_{\text{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\text{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

Noiseless Output Extrapolation

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}^{\alpha}_{\mathcal{H}_i}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
  1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
  2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle: $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots \left(\mathcal{D}_{\mathcal{H}_j}\right)^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\mathsf{NOX}}(O) := E_{\mathsf{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\mathsf{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}^{\alpha}_{\mathcal{H}_i}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
  1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
  2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle:
  $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots (\mathcal{D}_{\mathcal{H}_j})^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\mathsf{NOX}}(O) := E_{\mathsf{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\mathsf{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

Noiseless Output Extrapolation

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}^{\alpha}_{\mathcal{H}_i}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
    1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
    2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle:
  $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots \left(\mathcal{D}_{\mathcal{H}_j}\right)^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\mathsf{NOX}}(O) := E_{\mathsf{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\mathsf{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

Noiseless Output Extrapolation

- Need a way to scale the noise: $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i \to \mathcal{D}^{\alpha}_{\mathcal{H}_i}\mathcal{H}_i$
- ZNE uses pulse stretching, or unitary folding[3]: $\mathcal{H}_i(\mathcal{H}_i\mathcal{H}_i^{-1})^{\alpha}$
  1. $\mathcal{H}_i$ and $\mathcal{H}_i^{-1}$ have identical noise models
  2. $\mathcal{D}_{\mathcal{H}_i}\mathcal{H}_i = \mathcal{H}_i\mathcal{D}_{\mathcal{H}_i}$
- CER can detect when these assumptions apply
- Otherwise, scale noise from single cycle:
  $\mathcal{C}'_{j,\alpha} = \mathcal{E}_{m+1}\mathcal{D}_{\mathcal{H}_m}\mathcal{H}_m\mathcal{E}_m \cdots \left(\mathcal{D}_{\mathcal{H}_j}\right)^{\alpha}\mathcal{H}_j \cdots \mathcal{D}_{\mathcal{H}_1}\mathcal{H}_1\mathcal{E}_1$

$$\langle O \rangle \approx E_{\mathsf{NOX}}(O) \coloneqq E_{\mathsf{noisy}}(O) + \sum_{j=1}^{m} \frac{E_{\mathsf{noisy}}(O) - E_{j,\alpha}(O)}{\alpha - 1} \tag{3}$$

[3] This is referred to as "identity insertion" in the paper.

Pauli Error Cancellation

- Like Prob. EC, start with quasiprobability distribution:
$C_{\text{tot}} \sum_{i=1}^{L} s_i q_i \,\mathrm{tr}\Big(O\widetilde{\mathcal{U}}_i(\rho_{\text{in}})\Big) = \mathrm{tr}\big[O\mathcal{U}(\rho_{\text{in}})\big] + \delta$

- $\widetilde{\mathcal{U}}_i$: implementable operations
- $s_k = \pm 1$ depending on number of Pauli cycles
- r_k results from PEC circuits

$$C_{\text{tot}} = \prod_{j=1}^{m} \frac{1}{\left(\epsilon_0^{(\mathcal{H}_j)}\right)^2 - \sum_{k=1}^{4^n-1} \left(\epsilon_k^{(\mathcal{H}_j)}\right)^2} \tag{4}$$

$$\langle O \rangle \approx E_{\text{PEC}}(O) = C_{\text{tot}} \sum_{k=1}^{N} \frac{s_k r_k}{N} \tag{5}$$

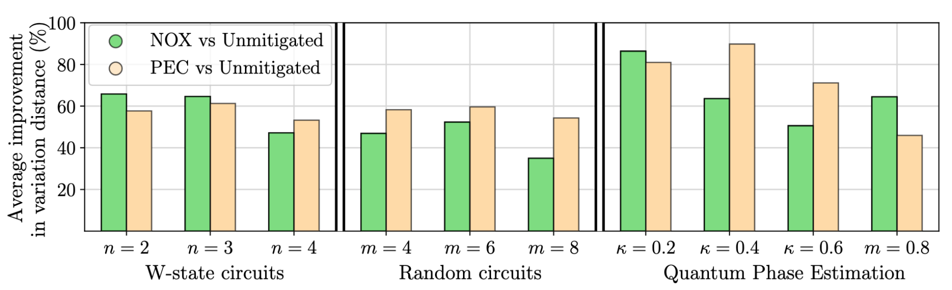
Pauli Error Cancellation

- Like Prob. EC, start with quasiprobability distribution:
  $C_{\text{tot}} \sum_{i=1}^{L} s_i q_i \operatorname{tr}\left(O\widetilde{\mathcal{U}}_i(\rho_{\text{in}})\right) = \operatorname{tr}\left[O\mathcal{U}(\rho_{\text{in}})\right] + \delta$

- $\widetilde{\mathcal{U}}_i$: implementable operations
- $s_k = \pm 1$ depending on number of Pauli cycles
- r_k results from PEC circuits

$$C_{\text{tot}} = \prod_{j=1}^{m} \frac{1}{\left(\epsilon_0^{(\mathcal{H}_j)}\right)^2 - \sum_{k=1}^{4^n-1} \left(\epsilon_k^{(\mathcal{H}_j)}\right)^2} \tag{4}$$

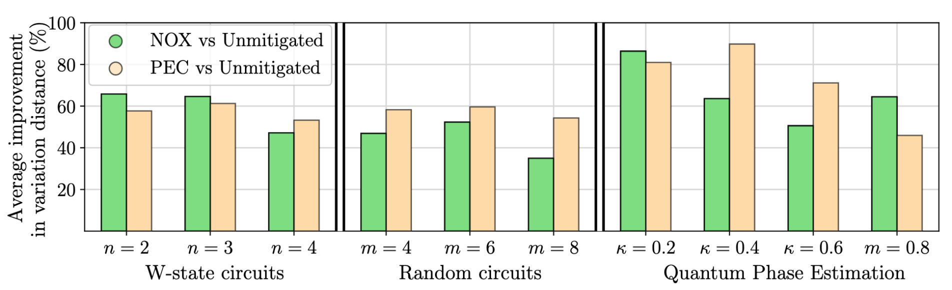
$$\langle O \rangle \approx E_{\text{PEC}}(O) = C_{\text{tot}} \sum_{k=1}^{N} \frac{s_k r_k}{N} \tag{5}$$



	PEC	NOX	Unmit.
Runtime	$\frac{1}{(1-n\varepsilon)^{2m}}$	m^3	1
Bias	$\mathcal{O}\left(mn^2\varepsilon^2\right)+\delta_{\text{rec}}$	$\mathcal{O}\left(m^2n^2\varepsilon^2\right)+\delta_{\text{rec}}$	$\mathcal{O}(mn\varepsilon)$

- m circuit depth
- $n\varepsilon$ cycle error rate
- n number of qubits
- δ_{rec} accuracy of noise reconstruction

Summary



	PEC	NOX	Unmit.
Runtime	$\frac{1}{(1-n\varepsilon)^{2m}}$	m^3	1
Bias	$\mathcal{O}\left(mn^2\varepsilon^2\right)+\delta_{\text{rec}}$	$\mathcal{O}\left(m^2n^2\varepsilon^2\right)+\delta_{\text{rec}}$	$\mathcal{O}(mn\varepsilon)$

- m circuit depth
- $n\varepsilon$ cycle error rate
- n number of qubits
- δ_{rec} accuracy of noise reconstruction

References

- Ferracin, Samuele et al. "Efficiently improving the performance of noisy quantum computers". In: (Jan. 2022). arXiv: 2201.10672 [quant-ph].
- Hashim, Akel et al. "Randomized Compiling for Scalable Quantum Computing on a Noisy Superconducting Quantum Processor". In: *Phys. Rev. X* 11 (4 Nov. 2021), p. 041039. DOI: 10.1103/PhysRevX.11.041039.
- Wallman, Joel J. and Joseph Emerson. "Noise tailoring for scalable quantum computation via randomized compiling". In: *Phys. Rev. A* 94 (5 Nov. 2016), p. 052325. DOI: 10.1103/PhysRevA.94.052325.

Thank you!