Cost-optimal quantum error mitigation based on universal cost bound

Department of Applied Physics, The University of Tokyo Kento Tsubouchi, Takahiro Sagawa, and Nobuyuki Yoshioka

arXiv:2208.09385

Outline

Introduction: quantum error mitigation (QEM)

- What is the theoretical limit of the sampling cost of QEM?
- What is the cost-optimal QEM method?

Method: quantum estimation theory

Optimal measurement on noisy quantum systems

Main results: two lower bounds on the sampling cost of QEM

- First bound: exponential growth with circuit depth
- Second bound: exponential growth with circuit depth and qubit count

Application: cost-optimal QEM method

Rescaling the measurement results is cost-optimal

Outline

- Introduction: quantum error mitigation (QEM)
 - What is the theoretical limit of the sampling cost of QEM?
 - What is the cost-optimal QEM method?
- Method: quantum estimation theory
 - Optimal measurement on noisy quantum systems
- Main results: two lower bounds on the sampling cost of QEM
 - First bound: exponential growth with circuit depth
 - Second bound: exponential growth with circuit depth and qubit count
- Application: cost-optimal QEM method
 - Rescaling the measurement results is cost-optimal

Noise of quantum computers



Countermeasure 1: quantum error correction (QEC)

By encoding multiple physical qubits into a single logical qubit, errors are detected and **corrected** during the computation.



Logical qubit Error correction

Single logical qubit on superconducting hardware



Google Quantum AI, 2023.

Requires a lot of qubits

 \Rightarrow Can we combat noise more hardware-friendly?

Countermeasure 2: quantum error mitigation (QEM)

The effect of errors is mitigated by increasing the number of samples and performing appropriate post-processing.



No qubit overhead

 \Rightarrow Can be performed on existing hardware.

Methods of QEM : Zero-noise extrapolation

Mitigates errors by increasing the noise rate and extrapolating noisy expectation values Y. Li et al., 2017. K. Temme et al., 2017. Expectation Increasing noise rate in quantum circuits U value /ideal Pulse time: τ $\langle \hat{X} \rangle_{\rm est}$ $\hat{\rho}_0$ U \mathcal{M}_2 Error rate : ε_0 $\langle \hat{X} \rangle(\varepsilon_0)$ $\langle \hat{X} \rangle (\lambda \varepsilon_0)$ Pulse time: $\lambda \tau$ Error rate : $\lambda \varepsilon_0$ A. Kandala et al., 2019. $\lambda \varepsilon_0$ Error rate \mathcal{E}_0 Increase in the sampling cost Increase in the variance

Methods of QEM : Zero-noise extrapolation

Performing Zero-noise extrapolation in experiments



Methods of QEM : Probabilistic error cancellation

Mitigates errors by virtually implementing the inverse of the noise channel K. Temme et al., 2017. $\mathcal{E}^{-1} - \sum_{\alpha \in \mathcal{C}_{i}} \left[\mathcal{C}_{i} : \text{ implementable operation} \right]$



 $\mathcal{E}^{-1} = \sum_{i} q_i \mathcal{C}_i \quad \begin{pmatrix} \mathcal{C}_i : \text{ implementable operation} \\ q_i : \text{ quasi-probability distribution} \end{pmatrix}$ $= \gamma \sum_{i} p_i \operatorname{sgn}(q_i) \mathcal{C}_i \quad \begin{pmatrix} \gamma = \Sigma_i |q_i| \\ p_i = |q_i|/\gamma \end{pmatrix}$ Increase in the variance $\langle \hat{X} \rangle_{\text{est}} = \gamma \sum_{i} p_i \operatorname{sgn}(q_i) \operatorname{tr}[\mathcal{C}_i \circ \mathcal{E}(\hat{\rho}) \hat{X}]$ $= \operatorname{tr}[\mathcal{E}^{-1} \circ \mathcal{E}(\hat{\rho})\hat{X}]$ $= \operatorname{tr}[\hat{\rho}\hat{X}]$

Obtain unbiased estimator with sampling overhead γ^2

Methods of QEM : Probabilistic error cancellation

Performing Zero-noise extrapolation in experiments



Comparing QEM methods

	Zero-noise extrapolation	Probabilistic error cancellation	Best QEM
Idea	$ \langle \hat{X} \rangle_{\text{est}} \langle \hat{X} \rangle_{\text{ideal}} \\ \langle \hat{X} \rangle_{(\mathcal{E}_0)} \\ \langle \hat{X} \rangle (\lambda \mathcal{E}_0) \\ \varepsilon_0 \\ \lambda \mathcal{E}_0 $	$\widehat{p_0} - u \xrightarrow{\varepsilon^{-1}} \underbrace{\mathcal{K}_i}_{\varepsilon^{-1}} \xrightarrow{\mathcal{K}_i}_{i} \xrightarrow{\mathcal{K}_i}_{\varepsilon^{-1}} \underbrace{\mathcal{K}_i}_{\varepsilon^{-1}} \xrightarrow{\mathcal{K}_i}_{\varepsilon^{-1}} \xrightarrow{\mathcal{K}_i}_{\varepsilon^{$	
Hardware constraint	Boosting noise rate	Learning noise	Easily implementable
Bias (Accuracy)	Large bias for large ε_0	Unbiased	Unbiased
Sampling cost	$\left(\sum_{m=1}^{M} \left \prod_{k \neq m} \frac{\lambda_k}{\lambda_k - \lambda_m} \right \right)^2$	$\gamma^{2\# ext{noise}}$	Low

Sampling cost of QEM

The expected sampling overhead : $2^{O(\#noise)} \sim 2^{O(\#gate)}$



Theoretical limits of QEM

What is the theoretical limit of the sampling cost of unbiased QEM?

- Exponential growth with circuit depth
 - For general circuit

 $N \ge O(\gamma^{2L})$ L: circuit depth $\gamma > 1$:Noise-dependent constants Exponential growth with qubit count • For scrambling circuit $\mathbb{E}_{U}[N] \ge O(\nu^{nL})$ n: qubit count $\nu > 1: \text{Noise-dependent constants}$

What is the cost-optimal QEM method achieving the limit?

Rescaling the measurement result is cost-optimal. $\left< \hat{X} \right>_{\rm est} = (1-p)^{-k_{\rm mean}L} \left< \hat{X} \right>_{\rm noisy}$

Outline

Introduction: quantum error mitigation (QEM)

- What is the theoretical limit of the sampling cost of QEM?
- What is the cost-optimal QEM method?

Method: quantum estimation theory

Optimal measurement on noisy quantum systems

Main results: two lower bounds on the sampling cost of QEM

- First bound: exponential growth with circuit depth
- Second bound: exponential growth with circuit depth and qubit count

Application: cost-optimal QEM method

• Rescaling the measurement results is cost-optimal

Quantum estimation theory

Characterize the accuracy and cost of parameter estimation.



Quantum Cramér-Rao bound: C. W. Helstrom, 1969.

 The accuracy limit of the estimator is bounded by the amount of information in the quantum state.

$$\underbrace{\operatorname{Var}(f(\boldsymbol{\theta})_{\text{est}})}_{\text{Accuracy}} \geq \frac{1}{N} (\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}))^{T} J^{-1} (\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}))$$
Quantum Fisher information matrix

How much cost do we need to estimate the expectation value of an observable X from the measurement of the noisy state? Y. Watanabe et al., 2010.



Parametrize quantum state using
generalized Bloch vector $\boldsymbol{\theta} \in \mathbb{R}^{4^{n}-1}$
$\hat{\rho}(\boldsymbol{\theta}) = \frac{1}{2^n} \hat{I} + \frac{1}{2^{(n+1)/2}} \boldsymbol{\theta} \cdot \hat{\boldsymbol{P}}$
$\mathcal{E}(\hat{\rho}(\boldsymbol{\theta})) = \frac{1}{2^n}\hat{I} + \frac{1}{2^{(n+1)/2}}(\boldsymbol{A\boldsymbol{\theta}} + \boldsymbol{c})\cdot\hat{\boldsymbol{P}}$
<i>n</i> : qubit count \widehat{P} : Array of nontrivial Pauli operators $A_{ii} = 2^{-n} \operatorname{tr}[\widehat{P}_i \mathcal{E}(\widehat{P}_i)]$: Unital part of the PTM of \mathcal{E}
$c_i = 2^{(1-3n)/2} \text{tr}[\hat{P}_i \mathcal{E}(\hat{I})] : \text{Non-unital effect of } \mathcal{E}$

How much cost do we need to estimate the expectation value of an observable X from the measurement of the noisy state? Y. Watanabe et al., 2010.





How much cost do we need to estimate the expectation value of an observable X from the measurement of the noisy state? Y. Watanabe et al., 2010.



SLD Quantum Fisher information matrix of $\mathcal{E}(\hat{\rho}(\boldsymbol{\theta}))$:

$$J(\mathcal{E}(\hat{\rho}(\boldsymbol{\theta}))) = A^T C_{\mathcal{E}(\hat{\rho})}^{-1} A$$
$$\left(\lesssim ||A||^2 I \right)$$

 $A_{ij} = 2^{-n} \operatorname{tr} \left[\hat{P}_i \mathcal{E} \left(\hat{P}_j \right) \right] : \text{Unital part of the PTM of } \mathcal{E}$ $C_{\mathcal{E}(\hat{\rho})} : \text{Covariance matrix of } \hat{P} \text{ for } \mathcal{E}(\hat{\rho})$

How much cost do we need to estimate the expectation value of an observable X from the measurement of the noisy state? Y. Watanabe et al., 2010.



Quantum Cramér-Rao bound $\operatorname{Var}(f(\boldsymbol{\theta})_{\text{est}}) \geq \frac{1}{N} (\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}))^T J^{-1} (\nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}))$ $Var(\langle \hat{X} \rangle_{\text{est}}) \geq \frac{1}{N} \left(\text{tr}[\mathcal{E}(\hat{\rho})\hat{Y}^2] - \text{tr}[\mathcal{E}(\hat{\rho})\hat{Y}]^2 \right)$ $\left[\begin{array}{c} \hat{Y} = (\mathcal{E}^{-1})^{\dagger}(\hat{X}) \\ \operatorname{tr}[\mathcal{E}(\hat{\rho})\hat{Y}] = \operatorname{tr}[\hat{\rho}\hat{X}] \end{array} \right]$ Measuring $\mathcal{E}(\rho)$ with \hat{Y} is cost-optimal

Outline

Introduction: quantum error mitigation (QEM)

- What is the theoretical limit of the sampling cost of QEM?
- What is the cost-optimal QEM method?

Method: quantum estimation theory

• Optimal measurement on noisy quantum systems

Main results: two lower bounds on the sampling cost of QEM

- First bound: exponential growth with circuit depth
- Second bound: exponential growth with circuit depth and qubit count

Application: cost-optimal QEM method

Rescaling the measurement results is cost-optimal

Problem setup

Quantum circuit affected by noise:



What are the lower bounds on the sampling cost of estimating

 $\langle \hat{X} \rangle_{\text{ideal}}$ without bias through QEM?

Main results: two lower bounds on the sampling cost of QEM



Roadmap for obtaining the lower bound



③ Bound QFIM of the virtual circuit $J(\mathcal{E}'(\hat{\rho}(\boldsymbol{\theta}))) = A'^T C_{(\mathcal{E}'(\hat{\rho}))}^{-1} A'$ $\lesssim ||A'||^2 I$ $\lesssim ||A||^{2L} I$

② Define QFIM of the virtual circuit

$$\hat{\rho}(\boldsymbol{\theta}) = \frac{1}{2^n} \hat{I} + \frac{1}{2^{(n+1)/2}} \boldsymbol{\theta} \cdot \hat{\boldsymbol{P}}$$

$$J\left(\Pi_{m=1}^N \mathcal{E}'_m\left(\hat{\rho}(\boldsymbol{\theta}) \otimes |0^{aL}\rangle \langle 0^{aL}|\right)\right)$$

④ Obtain the lower bound $N \ge O(\gamma^{2L})$ for general circuit $N \ge O(\nu^{nL})$ for scrambling circuit

① Mapping QEM to virtual circuit

The virtual quantum circuit that gives an equivalent representation of the QEM methods:
Depth 7



① Mapping QEM to virtual circuit

The virtual quantum circuit that gives an equivalent representation of the QEM methods:



Probabilistic error cancellation



① Mapping QEM to virtual circuit

The virtual quantum circuit that gives an equivalent representation of the QEM methods:
Death 7



■ Number of copies *N*: sampling cost of QEM.

② Defining QFIM of the virtual circuit

Parameterize the noiseless quantum state with the generalized Bloch vector. G. Kimura, 2003.

$$\hat{\rho}_{0} - u_{1} - \cdots - u_{L} = \hat{\rho}(\theta)$$

$$= \frac{1}{2^{n}}\hat{I} + \frac{1}{2^{(n+1)/2}}\theta \cdot \hat{P}$$

$$n : \text{Number of qubits}$$

$$\hat{P} : \text{Array of nontrivial Pauli operators}$$

$$\theta \in \mathbb{R}^{4^{n}-1} : \text{Generalized Bloch vector}$$

② Defining QFIM of the virtual circuit

Rewrite the virtual circuit as in the following way.



② Defining QFIM of the virtual circuit

Define the QFIM of the state before the measurement of the virtual circuit.



Quantum Fisher information matrix $J(\bigotimes_{m=1}^{N} \mathcal{E}'_{m}(\hat{\rho}(\boldsymbol{\theta}) \otimes |0^{aL}\rangle\langle 0^{aL}|))$

③ Bounding QFIM of the virtual circuit

Simplify QFIM using additivity and convexity of QFIM.



 $J(\bigotimes_{m=1}^{N} \mathcal{E}'_{m}(\hat{\rho}(\boldsymbol{\theta}) \otimes |0^{aL}\rangle \langle 0^{aL}|))$

③ Bounding QFIM of the virtual circuit

Analyze $J(\mathcal{E}'(\hat{\rho}(\boldsymbol{\theta})))$.



From the result of Y. Watanabe et al., 2010,

$$\begin{split} J(\mathcal{E}'(\hat{\rho}(\boldsymbol{\theta}))) &= A'^T C_{(\mathcal{E}'(\hat{\rho}))}^{-1} A' \\ &\lesssim ||A'||^2 I \\ &\lesssim ||A||^{2L} I \end{split} \begin{cases} \mathcal{C}_{\mathcal{E}'(\hat{\rho})} : \text{Covariance matrix of } \hat{P} \text{ for } \mathcal{E}'(\hat{\rho}) \\ A' : \text{Unital part of the PTM of } \mathcal{E}' \\ A : \text{Unital part of the PTM of } \mathcal{E} \end{cases} \end{split}$$

④ Obtaining the lower bound for general circuit

Bound on QFIM *J* of the virtual circuit

$$J \leq O(N\gamma^{-2L})$$
Quantum Cramér-Rao bound

$$\operatorname{Var}(f(\theta)_{est}) \geq \frac{1}{N} (\nabla_{\theta} f(\theta))^T J^{-1} (\nabla_{\theta} f(\theta))$$

Exponential growth with circuit depth

For general circuit

 $N \geq O(\gamma^{2L})$

L: circuit depth $\gamma > 1$:Noise-dependent constants



 $\gamma = ||A||^{-1}$ Minimal degree of shrinkage of the generalized Bloch sphere

Applying general bound to some noise models

Global depolarizing noise

• Correlated uniform noise: $\theta \mapsto (1-p)\theta$

•
$$\gamma = (1-p)^{-1}$$

 $N \gtrsim \left(\frac{1}{1-p}\right)^{2L}$

Local depolarizing noise

(1-p)

Noise acting independently on each qubit

•
$$\gamma = (1-p)^{-1}$$

 $N \gtrsim \left(\frac{1}{1-p}\right)^{2L}$





Problem with the general bound



Bound for general circuit

• Includes the case where all \mathcal{U}_i are single-qubit gate

Expected lower bound under entanglement: $2^{O(\#noise)} = 2^{O(nL)}$

Second bound: exponential growth with qubit count

Scrambling circuit

- Noise : local noise $\mathcal{E}^{\otimes n}$
- Gates : drawn from unitary 2-design



• QFIM J of the virtual circuit $J \leq NA'^T C_{\mathcal{E}'(\hat{\rho})}^{-1} A'$



- Exponential growth with qubit count
 - For scrambling circuit

 $\mathbb{E}_U[N] \ge O(\nu^{nL})$

n: qubit count v > 1:Noise-depending constants

Lower bound under local randomness

Even when gates are not drawn from unitary 2-design, we can still observe the exponential growth with n from numerics.



$$\hat{\rho}(\theta) \boxed{u_{L}^{\dagger} + \dots - u_{1}^{\dagger} + u_{1}} \underbrace{\varepsilon}_{\varepsilon} \\ \text{effective noise} \underbrace{\varepsilon}_{\varepsilon} \\ \underbrace{\varepsilon}_{\varepsilon} \\$$

- Pauli transfer matrix A' of effective noise \mathcal{E}'
 - singular values: $(1-p)^{kL}$
- Convergence to $k = k_{\text{mean}} = O(n)$

$$\Rightarrow ||A'|| = 2^{-O(nL)}$$

$$N \ge 2^{O(nL)}$$

Outline

Introduction: quantum error mitigation (QEM)

- What is the theoretical limit of the sampling cost of QEM?
- What is the cost-optimal QEM method?

Method: quantum estimation theory

Optimal measurement on noisy quantum systems

Main results: two lower bounds on the sampling cost of QEM

- First bound: exponential growth with circuit depth
- Second bound: exponential growth with circuit depth and qubit count

Application: cost-optimal QEM method

Rescaling the measurement results is cost-optimal

Cost-optimal QEM for Global depolarizing noise



Convergence to global depolarizing noise

- Randomness causes noise to converge to global depolarizing noise.
 - Our numerics supports the convergence.



A. M. Dalzell et al., 2022.

Cost-optimal QEM for scrambling circuit

 10^{0}

10

0

10

40

50

Depth L

30

20

60

70

80

90 100 110



 10^{0}

0

40

Depth L

10 20 30

39/40

Cost-optimal

50 60 70 80 90 100 110

Summary of today's talk

What is the theoretical limit of the sampling cost of unbiased QEM?

- Exponential growth with circuit depth
 - For general circuit

 $N \ge O(\gamma^{2L})$ L: circuit depth $\gamma > 1:$ Noise-dependent constants Exponential growth with qubit count • For scrambling circuit $\mathbb{E}_{U}[N] \ge O(\nu^{nL})$ n: qubit count $\nu > 1: \text{Noise-dependent constants}$

What is the cost-optimal QEM method achieving the limit?

Rescaling the measurement result is cost-optimal $\langle X\rangle_{\rm est} = (1-p)^{-k_{\rm mean}L} \langle X\rangle_{\rm noisy}$

Future direction

What is the tradeoff between bias and sampling costs in QEM ?

- Our work : cost bound for unbiased QEM.
- C.f.) Cost bound for biased QEM R. Takagi et al., 2022. E. Quek et al., 2022

What is the cost-optimal QEM method for structured circuit ?

- Is the rescaling method still effective for structured circuit?
- If not, what other method will be useful?
- Can we unify QEC and QEM ?
 - What is the optimal countermeasure in this regime?

Summary of today's talk

What is the theoretical limit of the sampling cost of unbiased QEM?

- Exponential growth with circuit depth
 - For general circuit

 $N \ge O(\gamma^{2L})$ L: circuit depth $\gamma > 1:$ Noise-dependent constants Exponential growth with qubit count • For scrambling circuit $\mathbb{E}_{U}[N] \ge O(\nu^{nL})$ n: qubit count $\nu > 1: \text{Noise-dependent constants}$

What is the cost-optimal QEM method achieving the limit?

Rescaling the measurement result is cost-optimal $\langle X\rangle_{\rm est} = (1-p)^{-k_{\rm mean}L} \langle X\rangle_{\rm noisy}$