Fundamental limits of quantum error mitigation

Ryuji Takagi
NTU Singapore

Joint work with S. Endo (NTT), S. Minagawa (Nagoya), M. Gu (NTU)

npj Quantum Inf. 8, 114 (2022)

NISQ devices

NISQ = Noisy Intermediate-Scale Quantum

Focus on manipulating O(10) – O(100) qubits.

Let's NOT dive into a complicated error correction and see if we can do **anything useful** with it.

Obstacle: noise, of course!!

Error mitigation: run a given NISQ device many times + classical postprocessing

- Trade space resource with time resource (<u>sampling cost</u>)

Zoo of error-mitigation protocols

Various error-mitigation protocols proposed so far.

- Probabilistic error cancellation [Buscemi et al., ’13] [Temme et al., PRL ’17]
- Noise extrapolation [Temme et al., PRL ’17] [Li, Benjamin, PRX ’17]
- Virtual distillation [Koczor, PRX ’21] [Huggins, PRX ’21]
- Symmetry verification [Bonet-Monroig et al., PRA ’18]
- …

What is the ultimate performance shared by all error-mitigation protocols?

General theory for quantum error mitigation?

General quantum error mitigation

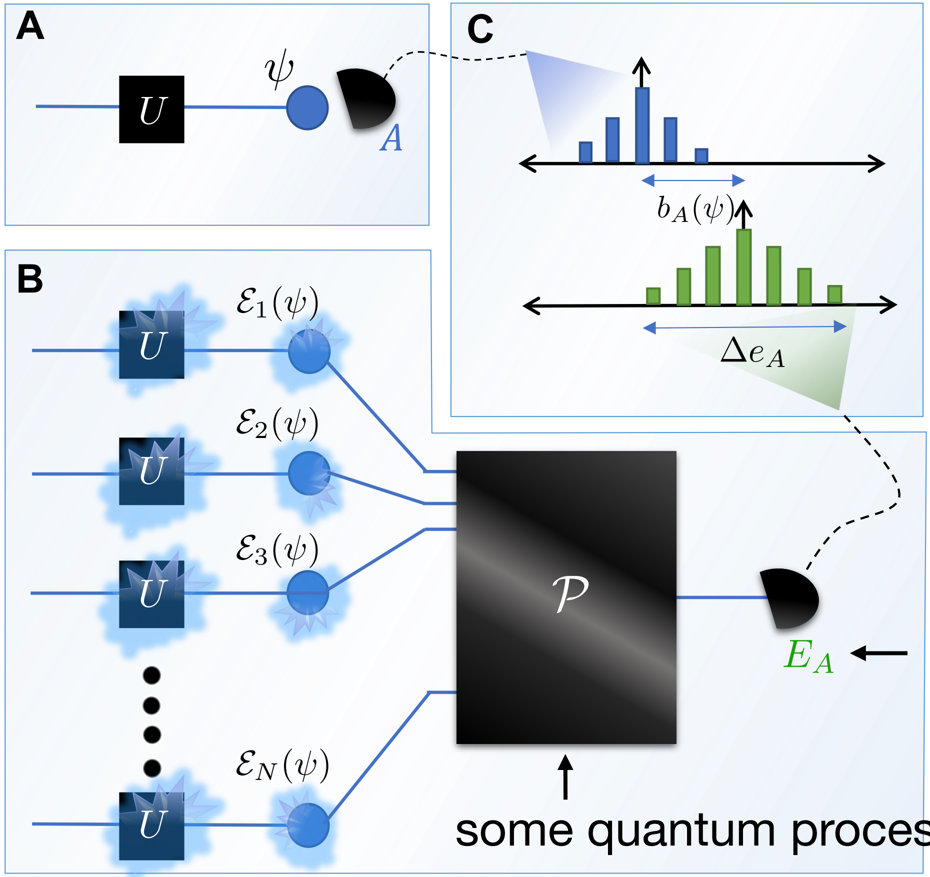
Ideal:

$\psi \longrightarrow \mathrm{Tr}(A\psi)$

Real:

$\mathcal{E}_1(\psi), \ldots, \mathcal{E}_N(\psi)$

$\longrightarrow \mathrm{Tr}(A\psi)$



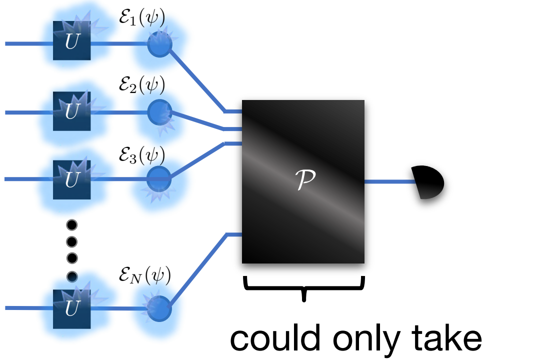
$b_A(\psi)$: bias

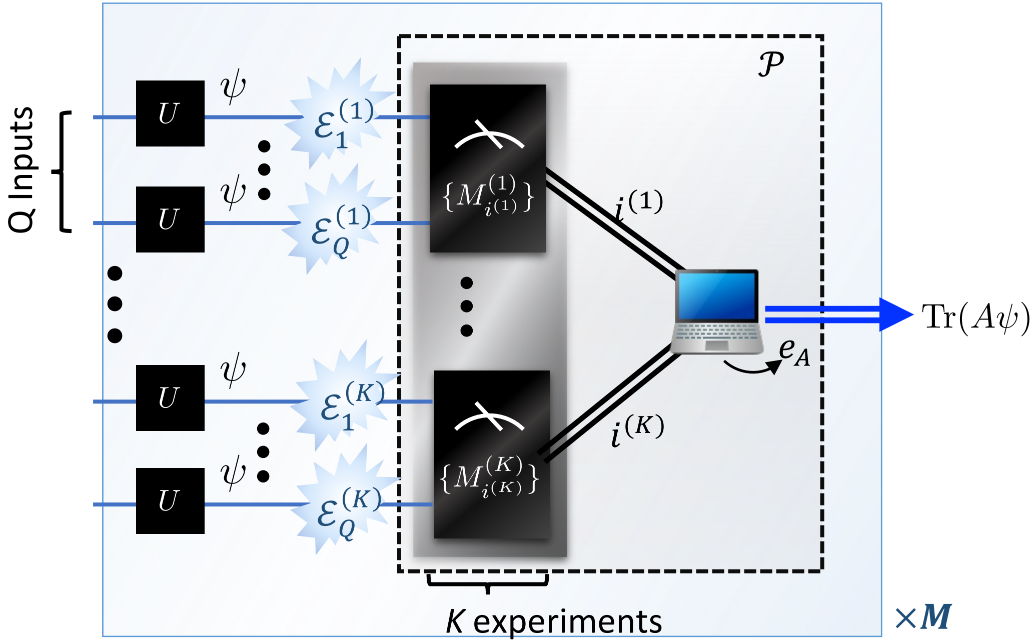
Δe_A : spread

$E_A \longleftarrow$ estimate of $\mathrm{Tr}(A\psi)$ for each round

$\uparrow$ some quantum process

General quantum error mitigation



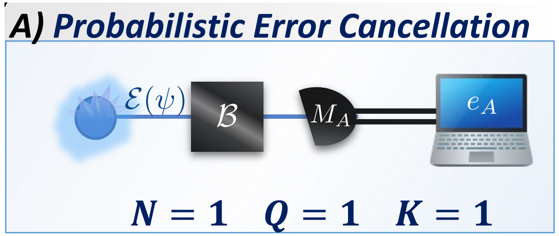


(Q,K)-error mitigation

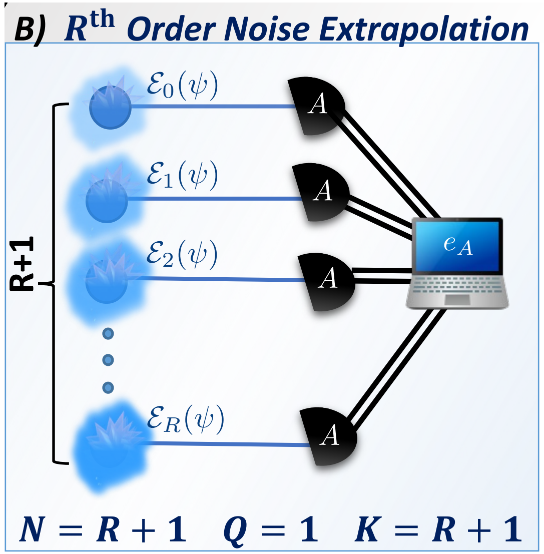
Repeat *M* rounds to get estimate $e_A(i^{(1)},\ldots,i^{(K)}) = e_{A,m}$ for each *m*=1,…,*M*

$\bar{E}_A^{(M)} = \frac{1}{M}\sum_{m=1}^{M} e_{A,m}$: estimate of $\mathrm{Tr}(A\psi)$

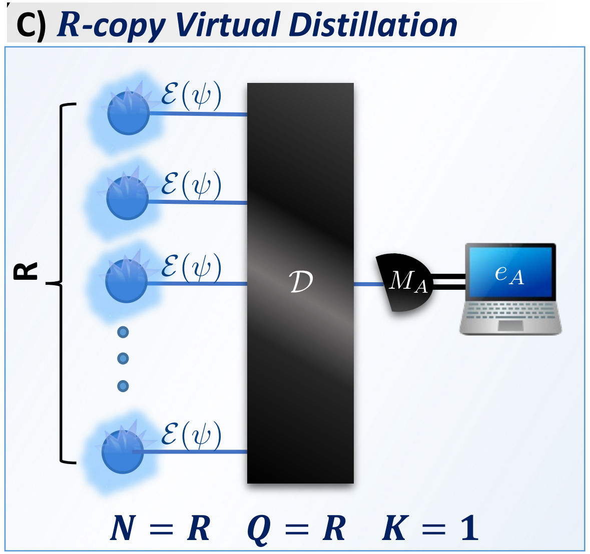
Examples



[Temme et al., PRL '17]



[Temme et al., PRL '17] [Li, Benjamin, PRX '17]



[Koczor, PRX '21] [Huggins, PRX '21]

Quantifying performance

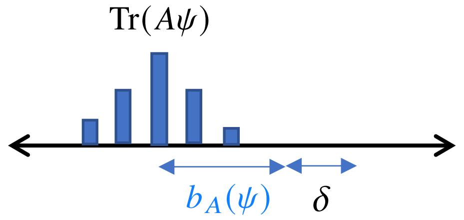
- Bias $b_A(\psi) = \langle E_A \rangle - \mathrm{Tr}(A\psi)$ E_A : random variable for each estimate
- Statistical Error $|\bar{E}_A^{(M)} - \langle E_A \rangle|$ $\bar{E}_A^{(M)} = \frac{1}{M}\sum_{m=1}^{M} e_{A,m}$: estimate of $\mathrm{Tr}(A\psi)$

How many round number *M* ensures the statistical error δ ?

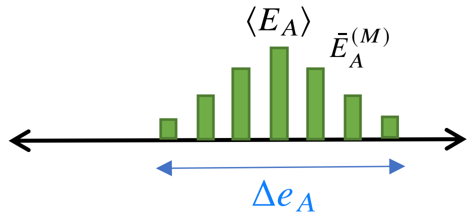
- Hoeffding inequality

M samples ensure error $|\bar{E}_A^{(M)} - \langle E_A \rangle| < \delta$ with probability $1-\varepsilon$ when

$$M > \frac{2\,\Delta e_A^2\,\log(2/\varepsilon)}{\delta^2}$$



- Maximum bias $b_{\max} = \max_{-\mathbb{I}/2 \le A \le \mathbb{I}/2} \max_{\psi} b_A(\psi)$
- Maximum spread $\Delta e_{\max} = \max_{-\mathbb{I}/2 \le A \le \mathbb{I}/2} \Delta e_A$



Lower bounds for maximum spread

What is the unavoidable maximum spread for a given maximum bias?

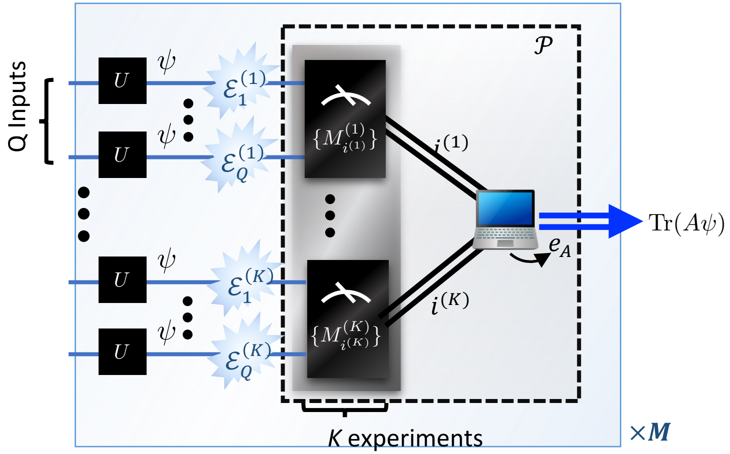
Maximum bias $b_{\max} = \max_{-\mathbb{I}/2 \le A \le \mathbb{I}/2} \max_{\psi} b_A(\psi)$

Maximum spread $\Delta e_{\max} = \max_{-\mathbb{I}/2 \le A \le \mathbb{I}/2} \Delta e_A$

For an arbitrary (*Q*,*K*)-mitigation protocol,

$$\Delta e_{\max} \ge \max_{\psi,\phi} \frac{D_{\mathrm{tr}}(\psi,\phi) - 2b_{\max}}{D_{\mathrm{LM}}\left(\tilde{\psi}_Q^{(K)}, \tilde{\phi}_Q^{(K)}\right)}$$

$$\tilde{\psi}_Q^{(K)} := \otimes_{k=1}^{K} \otimes_{q=1}^{Q} \left[\mathcal{E}_q^{(k)}(\psi)\right] \quad \tilde{\phi}_Q^{(K)} := \otimes_{k=1}^{K} \otimes_{q=1}^{Q} \left[\mathcal{E}_q^{(k)}(\phi)\right]$$



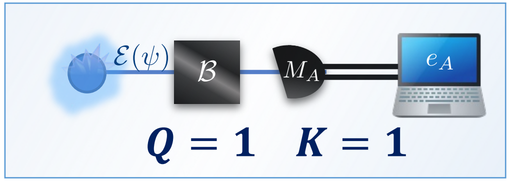
$D_{\mathrm{tr}}(\rho,\sigma)$: trace distance (= distinguishability using any measurement)

$D_{\mathrm{LM}}(\rho,\sigma)$: local distinguishability (= distinguishability using local measurements)

Connecting error-mitigation performance with information-theoretic measures.

Application: protocol benchmarking

Probabilistic error cancellation: (1,1)-protocol. (Q=K=1)



$\mathcal{E}^{-1} = \sum_j c_j \mathcal{B}_j,\ c_j \in \mathbb{R}$ $\qquad \gamma := \sum_j |c_j|$

$\gamma_{\rm opt}$: minimum γ over all feasible $\{\mathcal{B}_j\}_j$

$\Delta e_{\max}^{\rm PEC} = \gamma_{\rm opt}$ characterizes the sampling cost for this protocol. [Takagi, PRR '21] [Regula, Takagi, Gu, Quantum '21]

e.g.) Local dephasing noise $\mathcal{Z}_\epsilon^{\otimes n}$

$$\mathcal{Z}_\epsilon(\rho) = (1-\epsilon)\rho + \epsilon Z\rho Z$$

$$\Delta e_{\max}^{\rm PEC} = \gamma_{\rm opt} = \frac{1}{(1-2\epsilon)^n}$$

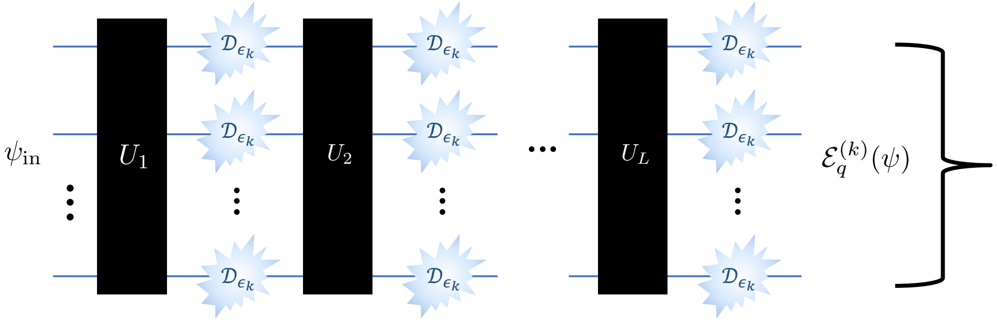
Our bound implies that any unbiased (1,1)-protocol satisfies

$$\Delta e_{\max} \geq \max_{\psi,\phi} \frac{D_{\rm tr}(\psi,\phi)}{D_{\rm tr}(\mathcal{Z}_\epsilon^{\otimes n}(\psi), \mathcal{Z}_\epsilon^{\otimes n}(\phi))} \geq \frac{1}{(1-2\epsilon)^n}$$

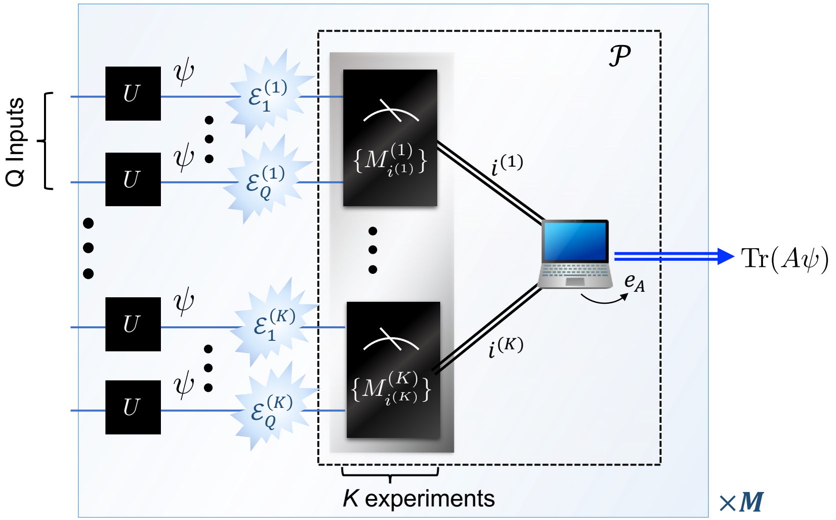
Lower bound achieved!

Probabilistic error cancellation is optimal for this noise.

Application: error-mitigating layered circuits



$\mathcal{D}_{\epsilon_k}(\rho) = (1-\epsilon_k)\rho + \epsilon_k \mathbb{I}/2$: local depolarizing noise



For an arbitrary (*Q*,*K*)-mitigation protocol,

$$\Delta e_{\max} \geq \frac{1-2b_{\max}}{\sqrt{2\ln 2}\sqrt{nQK}} \left(\frac{1}{1-\epsilon_{\min}}\right)^L$$

$\epsilon_{\min} := \min_k \epsilon_k$

Exponential error spread w.r.t. circuit depth *L*!

Summary

- Introduced a general framework for quantum error mitigation.
- Established a general lower bound for the maximum spread.
- Showed the optimality of probabilistic error cancellation for the local dephasing noise in terms of the maximum spread.
- Showed that mitigating errors in layered circuits requires exponentially large maximum spread.
- Sampling lower bounds? See our follow-up work "Universal sample lower bounds for quantum error mitigation", arXiv:2208.09178