Quantum Error Mitigation for Sampling Algorithms

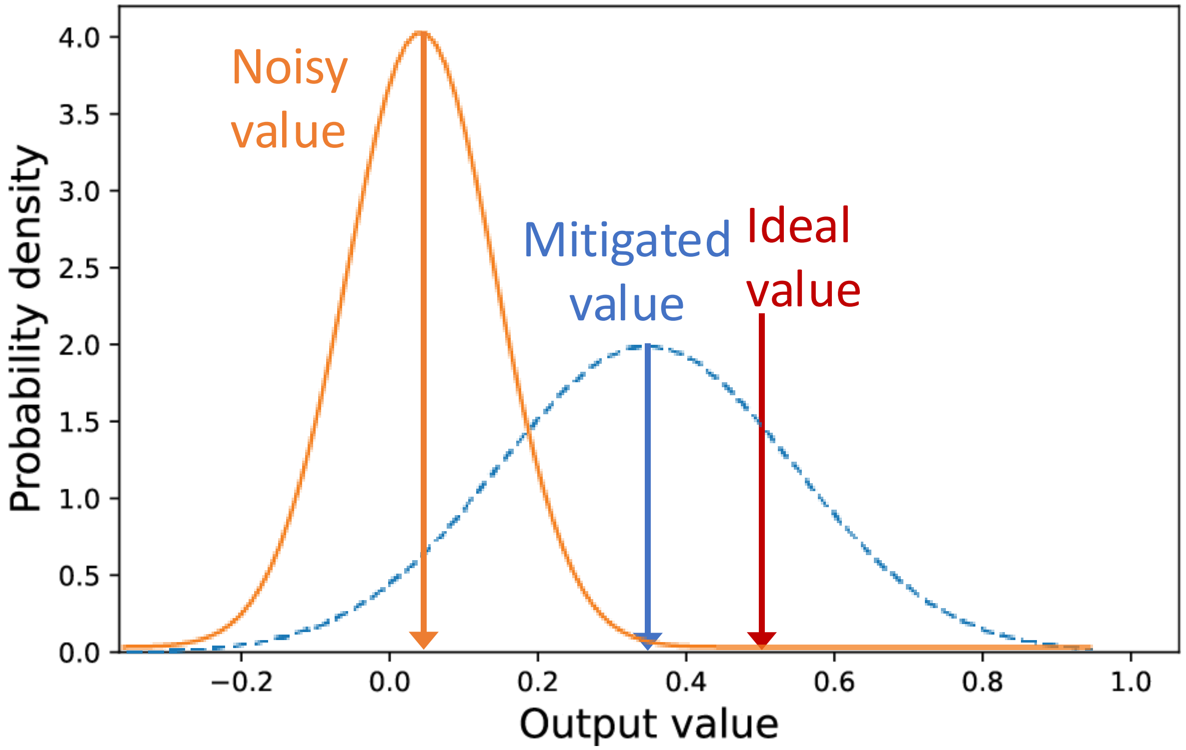
Zhenyu Cai, Joint work with Kecheng Liu

UNIVERSITY OF OXFORD
DEPARTMENT OF MATERIALS

DEPARTMENT OF ENGINEERING SCIENCE
UNIVERSITY OF OXFORD

Quantum Error Mitigation (QEM)

- Using additional circuit runs to reduce the bias in the expectation value via post-processing.



Cai *et al*, Quantum error mitigation, Rev. Mod. Phys. 95, 045005 (2023).

Not scalable $\neq$ Not useful

Applicability of QEM

- Cannot rely on QEM alone to deal with noise in large-scale computation.

- It is always useful in the finite error regime (average number of errors per circuit run $\lesssim 1$), extending the computational reach (i.e. allowing for larger and deeper computation) for noisy devices.

- This is especially relevant in the early fault-tolerant era, with non-negligible finite logical errors remaining after QEC.

Can QEM be applied to sampling-based algorithms like QPE?

QEM for Sampling Algorithm

- QEM use post-processing to combine the output from multiple noisy circuits to obtain the error-mitigated expectation values.
  - The effective damage from noise is only reduced for the entire ensemble of circuit runs.
  - The noise remains unchanged or even increases when zoom individual circuit runs.

- Sampling algorithms (e.g. quantum phase estimation): rely on accurate results for every circuit run, thus seems to be inherently incompatible with QEM (except for those uses post-selection).

Error-mitigated State

- QEM can also be viewed as trying to extract the error-mitigated "states" ρ_{em} out of the noisy circuit runs.

- The error-mitigated expectation value is given as $\mathrm{Tr}(O\rho_{em})$. (O is the observable of interests)

- The error-mitigated states ρ_{em} is obtained via linear combination of output states from different circuit configurations.

- This covers most mainstream QEM techniques.

Cai *et al*, arXiv:2110.05389.

Examples of Error-mitigated States

- Linear Zero-noise extrapolation (can be generalized to Richardson):

$$\rho_p = (1-p)\rho_0 + p\rho_{err} \quad \Rightarrow \quad \rho_0 = \rho_{em} \propto p_2\rho_{p_1} - p_1\rho_{p_2}$$

- Probabilistic error cancellation for bit-flip noise:

$$\rho = (1-p)\rho_0 + pX\rho_0X \quad \Rightarrow \quad \rho_0 = \rho_{em} \propto (1-p)\rho - pX\rho X$$

- Also applicable to other major QEM techniques like virtual purification.

Cai *et al*, arXiv:2110.05389.

QEM for Recovering Output Distribution

- Setting: Noiseless circuit gives the ideal output distribution $p_0(z)$ for binary strings z, but noise corrupts the output distribution to $p(z)$.
- Goal: obtain some error-mitigated distribution $p_{em}(\mathrm{z})$.

- Insight 1: The probability of obtaining a given output string z is simply the expectation value of the observable $\Pi_z = |z\rangle\langle z|$.
- i.e. the ideal and noisy output distributions are

$$p_0(\mathrm{z}) = \mathrm{Tr}(\Pi_z \rho_0), \qquad p(\mathrm{z}) = \mathrm{Tr}(\Pi_z \rho)$$

- The error-mitigated distribution is

$$p_{em}(\mathrm{z}) = \mathrm{Tr}(\Pi_z \rho_{em})$$

QEM for Recovering Output Distribution

- However, there are exponentially many observable Π_z!

- Insight 2: by running the circuit and measuring in the computational basis which output the string z' in a given run, we have actually obtained one sample for all $\{\Pi_z\}$ with
  - one sample of 1 for the Π_z with $z = z'$
  - one sample of 0 for the Π_z with $z \neq z'$

QEM for Recovering Output Distribution

- Combining the two insights: obtaining error-mitigated distributions $p_{em}(\mathrm{z}) = \mathrm{Tr}(\Pi_z \rho_{em})$ from error-mitigated states ρ_{em} is efficient (by measuring in the computation basis to obtain $\{\Pi_z\}$).

- Existing mainstream QEM techniques can be used to extract error-mitigated "states" ρ_{em}, thus can be straightforwardly extended to extract error-mitigated distributions.

PEC Example

- Probabilistic error cancellation for bit-flip noise:

$$\rho = (1-p)\rho_0 + pX\rho_0 X \quad \Rightarrow \quad \rho_0 = \rho_{em} = \frac{(1-p)\rho - pX\rho X}{1-2p}$$

- Implementation:

1. Sample ρ and $X\rho X$ with probability $(1-p)$ and p, respectively.
2. Measure in computation basis $\{Z_i\}$, post-process to obtain the set of observables $\{\Pi_z\}$.
3. If $X\rho X$ is sampled, attach minus sign to the output.
4. $p_{em}(\mathrm{z})$ is estimated by taking the average over all samples of Π_z and renormalise the result with the $(1-2p)^{-1}$ factor.

QEM for Recovering Output Distribution

Given the error-mitigated state as a linear combination of output states from different noisy circuit configurations.

1. Sample from the distribution of circuit configuration.
2. Measure in computation basis $\{Z_i\}$, post-process to obtain the set of observables $\{\Pi_z\}$.
3. Attaching minus sign to the output according to the circuit configuration or measurement results of additional observable.
4. Obtained one sample of $\{0, \pm 1\}$ for every Π_z in each run.
5. $p_{em}(\mathrm{z})$ is estimated by taking the average over all samples of Π_z, and multiply the result with a normalisation factor A.

Sampling overhead

- Let us consider the trivial observable I:

$$\hat{I} = \sum_z \widehat{\Pi}_z \quad \Rightarrow \quad \mathrm{Var}[\hat{I}] = \sum_z \mathrm{Var}[\widehat{\Pi}_z]$$

- i.e. the variance of estimating a single observable I is the same as the total variance of estimating the probability of all z, i.e. the entire probability distribution.

- For a given number of circuit runs, the total variance achieved for all entries in the entire estimated distribution is actually similar to the variance of one single observable.

How to sample from the QEM distribution?

- Without QEM, when measure z in a circuit run, we put one sample into the “bucket” corresponding to outcome z.

- With QEM, when measure z in a circuit run, there is also a additional sign associated with the circuit configuration we are running:
  - +ve sign: add one sample into the “bucket” corresponding to outcome z
  - -ve sign: remove one sample from the “bucket” corresponding to outcome z

How to sample from the QEM distribution?

- There can be negative number of samples! Esp. when the number of circuit run is small.

- When comes to interpretation of results, these negative number can effectively be treated as zero since any components below zero are entirely due to shot noise.

Application to Quantum Phase Estimation

- Considering using quantum phase estimation for obtaining ground state energy.

- Instead of trying to obtain the whole distribution, we are trying to obtain the smallest string from the output distribution.

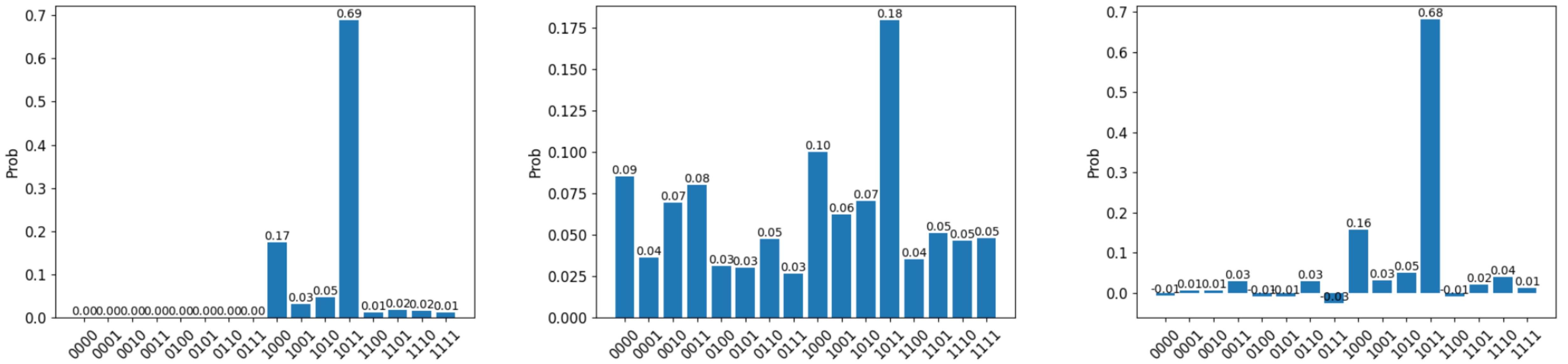
- Cannot simply output the smallest string from the estimated error-mitigated distribution, since shot noise can turn zero-probability entries to non-zero.

Application to Quantum Phase Estimation

- An additional step to test whether an entry is likely to be zero or not.
- Set a threshold probability $p_{\mathrm{th}}(z)$ for each entry such that
  - $\hat{p}_{\mathrm{em}}(z) \leq p_{\mathrm{th}}(z) \quad \Rightarrow \text{Accept null: } p_{\mathrm{em}}(z) = 0$
  - $\hat{p}_{\mathrm{em}}(z) > p_{\mathrm{th}}(z) \quad \Rightarrow \text{Accept null: } p_{\mathrm{em}}(z) > 0$

- $p_{\mathrm{th}}(z)$ can be set using:
  - Proportional to the sample standard deviation of the $\hat{p}_{\mathrm{em}}(z)$ estimator.
  - Known lower bound of the probability of the smallest string.

Numerical Simulation



(a) Ideal $p_0(z)$ (b) Noisy $p(z)$ (c) Mitigated $\hat{p}_{\mathrm{em}}(z)$, 10^6 circuit runs

- QPE with 4-bit precision
- Circuit error rate ~ 0.6
- 10^6 runs

- Total square errors reduced from 0.297 to 0.004
- Valid threshold: $0.03 < p_{\mathrm{th}} < 0.16$

Summary

- QEM can be used for recovering the output distribution and also sampling from it.
- Mitigating errors in the entire distribution is as cheap as one observable.
- Outlook:
  - Explicit analysis for more QEM techniques and more applications. Going
  - Beyond linear QEM.
  - Direct mitigation for a specific algorithm without estimating/sampling the distribution