



HPCA 2025 Tutorial

Topic 4. QuFEM: Fast and Accurate Quantum Readout Calibration Using the Finite Element Method







Speaker: Kaiwen Zhou

College of Computer Science and Technology Zhejiang University (ZJU)

https://janusq.github.io/HPCA_2025_Tutorial/

Presenter





Kaiwen Zhou kaiwenzhou@zju.edu.cn

Kaiwen Zhou is a first-year PhD student at the College of Computer Science, Zhejiang University. He is interested in quantum computer architecture and high-performance computing. He is currently working on designing a QLDPC decoding accelerator.

Outline of Presentation





Background and challenges

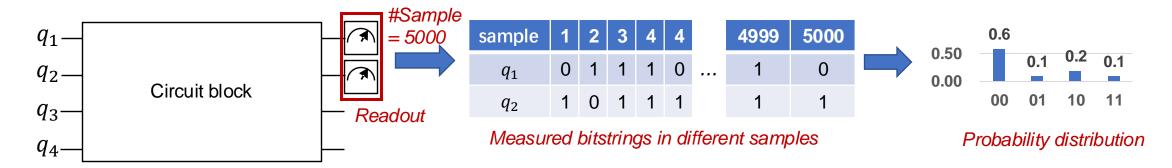
- Overview of QuFEM
- QuFEM characterization and calibration
- Experiment
- API of QuFEM

Background

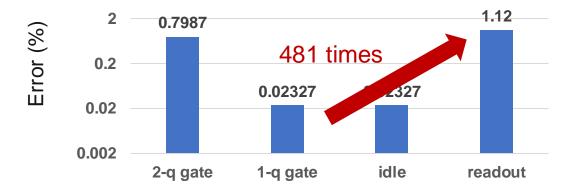




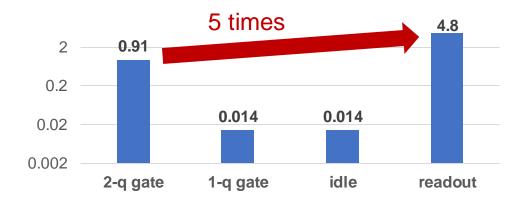
Quantum readout is an operation to read the information from quantum bits to classical bits.



Readout error is significant on current quantum hardware.



Noise on 127-qubit IBM Sherbrooke quantum device



Noise on 10-qubit Tianmu quantum device

Background





Implementation of readout on superconducting qubits

(2) Pulse takes the information out. *FFT* (1) Pulse in Readout line -1200 **Qubits share the** readout line 1800 I XY_1 XY_2

Source of readout error

from 1 to 0



Relaxation error

from 0 to 1



Excitation error



from 1/0 to 0/1

Crosstalk



e.g. Das, et al. JigSaw: Boosting Fidelity of NISQ Programs via Measurement Subsetting. MICRO 2021

State-dependent Readout Error





Readout errors vary in different combinations of measured qubits due to crosstalk.

Crosstalk has different frequencies when Q2 is measured 0, 1 or not measured

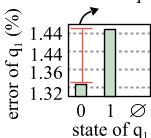
Example of state-dependent and readoutdependent noises on the IBMQ Perth quantum device.

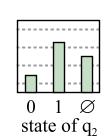


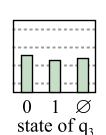


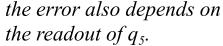


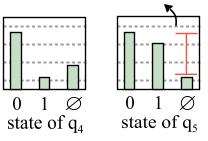
the error of q_1 depends on its own operation.











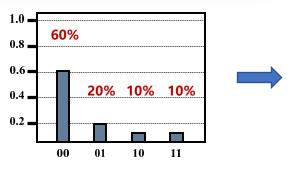
The readout output of qubits has correlations similar to the entanglement, making the calibration difficult.

Basic Matrix-based readout calibration





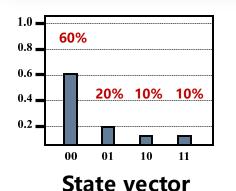
Ideal readout



0.6 0.2 0.1

State vector Ideal distribution (ideal program output)

Readout with noise

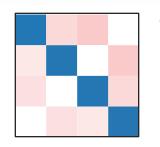




0.5 0.1 0.09 0.31

Noisy distribution (noisy program output)

Matrix-based readout error calibration



Noise

matrix

0.5 0.1 0.09 0.31

Noisy

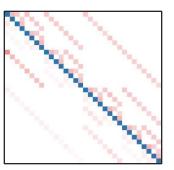
distribution

=

0.2 0.1 0.1

Calibrated distribution

0.6



The size exponentially increases!

Calibration matrix of a 5-qubit readout

 $2^5 \times 2^5$

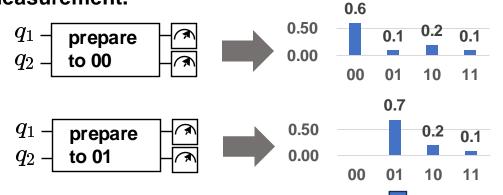
Basic Matrix-based readout calibration





Step 1. Matrix characterization

Prepares qubits to different basis states and apply measurement.



Fill in a noise matrix.

$$M = \begin{bmatrix} 0.6 & 0.1 & 0.2 & 0.1 \\ 0 & 0.7 & 0.2 & 0.6 & 0 \\ 0 & 0.1 & 0.1 & 0.8 \end{bmatrix}$$

Inverse the noise matrix

$$M^{-1} =$$

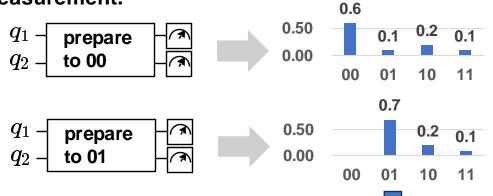
Basic Matrix-based Readout Calibration





Step 1. Matrix characterization

Prepares qubits to different basis states and apply measurement.



Fill in a noise matrix.

$$M = \begin{bmatrix} 0.6 & 0.1 & 0.2 & 0.1 \\ 0 & 0.7 & & & \\ 0.2 & 0.2 & 0.6 & 0 \\ 0 & 0.1 & 0.1 & 0.8 \end{bmatrix}$$

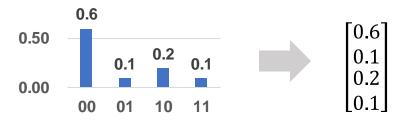
Inverse the noise matrix

$$M^{-1} =$$

Calibration matrix

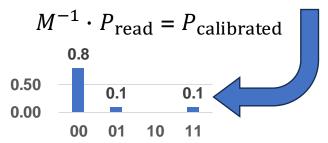
Step 2. Calibration for any input

Represent the measured distribution as a vector.



Apply matrix-vector multiplication.

$$\begin{bmatrix} 0.6 & 0.1 & 0.2 & 0.1 \\ 0 & 0.7 & 0.2 & 0.1 \\ 0.2 & 0.2 & 0.6 & 0 \\ 0 & 0.1 & 0.1 & 0.8 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 0.6 \\ 0.1 \\ 0.2 \\ 0.1 \end{bmatrix} = \begin{bmatrix} 0.8 \\ 0.1 \\ 0 \\ 0.1 \end{bmatrix}$$



Complexity Analysis



Step 1. Matrix characterization

Prepares qubits to different basis states and apply measurement.

 2^N circuits are executed to measure qubits on all basis states.

Fill in a noise matrix.

The size of the noise matrix is $2^N \times 2^N$.

Inverse the noise matrix

Calcauting the inverse has $O(4^N)$ complexity.

Step 2. Calibration for any input

Represent the measured distribution as a vector.

The transformation has linear complexity.

Apply matrix-vector multiplication.

The multiplication has $O(4^N)$ complexity.

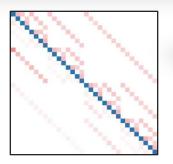
8.8 TB and 10 hours for a 32-qubit calibration on a server with AMD EPYC 2.25GHz 64-core CPUs

Limitations of Current Methods





IBU (Google Science 2021) Realizing topologically ordered states on a quantum processor.



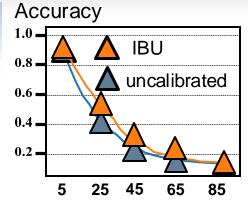
Crosstalk makes the matrix not simple tensor-product result.



Real calibration matrix

Single-qubit matrix

Use tensor product of a series of single-qubit metamatrices

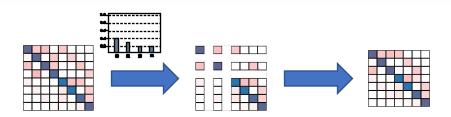


Fail to calibrate on 80qubit readout output

#qubit

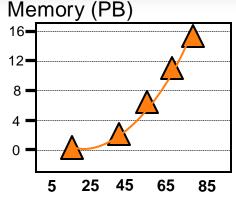
Fast but not accurate: ignore the qubit interactions.

M3 (IBM PRA 2021): Scalable mitigation of measurement errors on quantum computers



Before pruning Pruning based on After pruning program output

Use a sparsity-aware method to prune on the matrix under a threshold of Hamming distance



Require 16PB to calibrate a 85-qubit result. (4 times the Fugaku supercomputer)

qubit

Accurate but not fast: many matrix elements cannot be ignored

Outline of Presentation





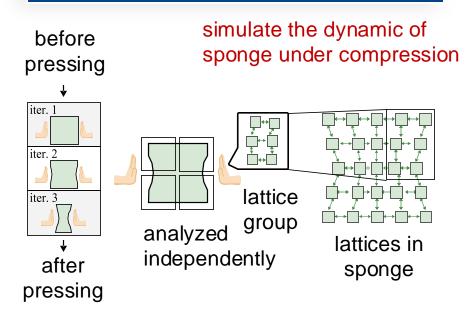
- Background and challenges
- Overview of QuFEM
- QuFEM characterization and calibration
- Experiment
- API of QuFEM

Calibration based on Finite Element method (FEM)





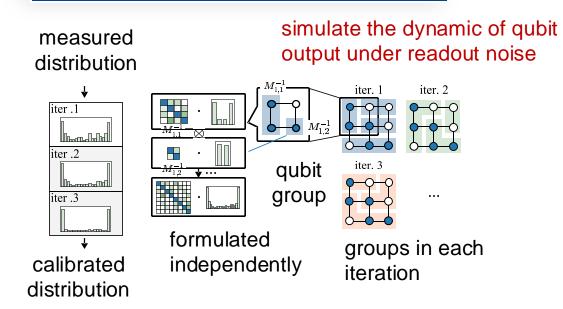
Classical Finite Element Method



- ① partitions the sponge into lattices
- ② analyzes the state of each lattice independently
- 3 simulate the interaction
- ④ update the state of sponge



Quantum Finite Element Method



- ① partitions qubits into groups
- ② analyze the noise in each group independently
- 3 simulate the interaction
- 4 update the calibration result of qubits

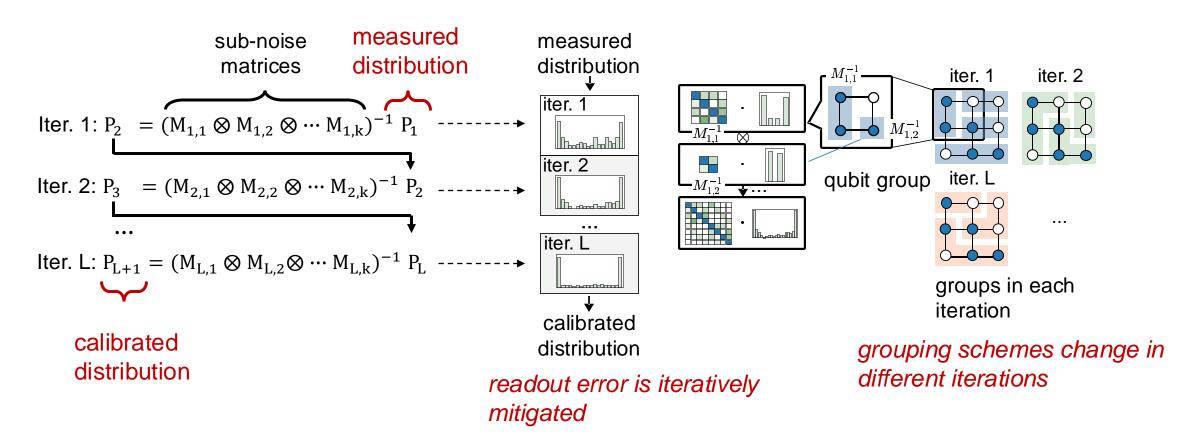
A divide-and-conquer strategy to calibrate measured distribution

Calibration formulation





QuFEM reformulates the calibration as an iterative process with a series of sub-noise matrices.



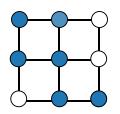
- Reason for fast: adopt finite element method
- Reason for accurate: dynamically generate noise matrices for different measured qubits



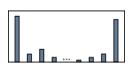


Input:

measured qubits

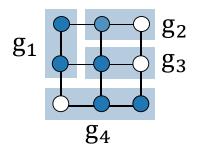


measured distribution

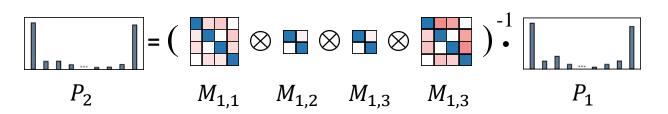


Iteration 1:

grouping scheme

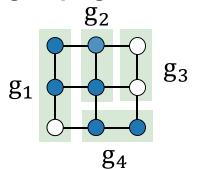


formulation

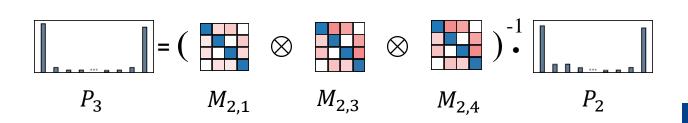


Iteration 2:

grouping scheme



formulation

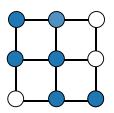






Input:

measured qubits



measured distribution

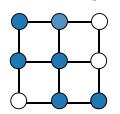




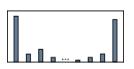


Input:

measured qubits

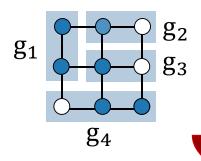


measured distribution

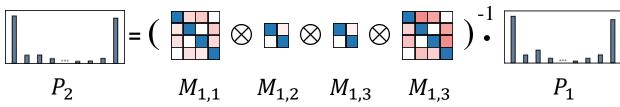


Iteration 1:

grouping scheme



formulation





Matrices are generated according to the measured qubits.

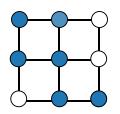
Since crosstalk varies in different combinations of measured qubits



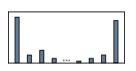


Input:

measured qubits

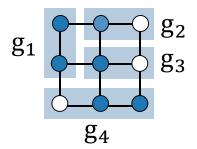


measured distribution

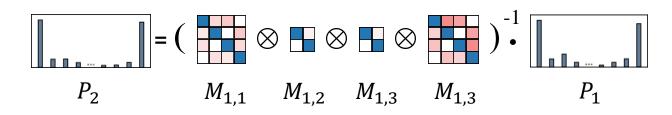


Iteration 1:

grouping scheme

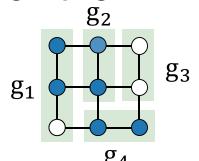


formulation

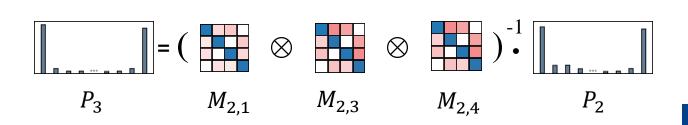


Iteration 2:

grouping scheme



formulation



Outline of Presentation





- Background and challenges
- Overview of QuFEM
- QuFEM characterization and calibration
- Experiment
- API of QuFEM

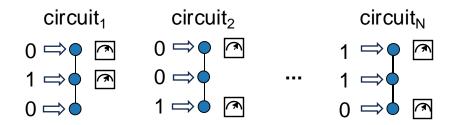




Technique 1: determine the grouping scheme

Data collection

Run benchmarking circuits.



Possible states of a qubit in a benchmarking circuit:

- 0: qubit is set 0 and measured
- 1: qubit is set 1 and measured
- 2: qubit is set 0 or 1 and not measured

Not all qubits are measured to maximize the variety.

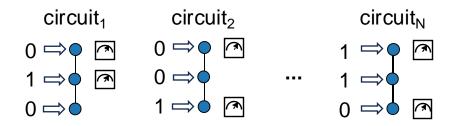




Technique 1: determine the grouping scheme

Data collection

Run benchmarking circuits.



Possible states of a qubit in a benchmarking circuit:

- 1: qubit is set 0 and measured
- 2: qubit is set 1 and measured
- 3: qubit is set 0 or 1 and not measured

Not all qubits are measured to maximize the variety.

Qubit partition

Characterize the **interaction** from one qubit to another qubit under different states:

$$interact(q_i. state = x \rightarrow q_i. state = x)$$

=
$$|P(q_j. error = 1 | C1, C2) - P(q_j. error = 1 | C2)|$$

error rate of q_i under C1, C2

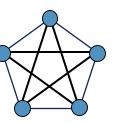
average error rate of qzi

C1:
$$q_i$$
. state = x , C2: q_i . state = y

C2:
$$q_i$$
. state = y

Construct weighted graph





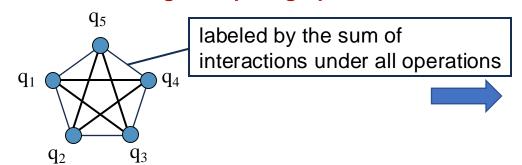




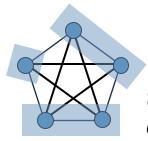
Technique 1: determine the grouping scheme

Qubit partition

Construct a weighted qubit graph:



Partitions with a **MAX-CUT solver**:



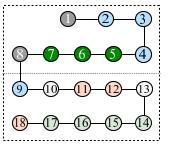
try to comprehensively capture the interactions between qubits

An Example

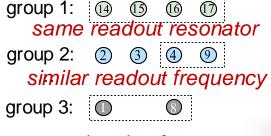
Prior knowledge of hardware helps grouping

Readout resonator 1

Readout resonator 2



18-qubit topology



overlapping frequency shift region

- demonstrated in the results from other quantum devices
- can be used as prior knowledge to facilitate the partition.





Technique 2: sub-noise matrix generation

Perform matrix-vector multiplication

Iter. i:
$$P_{i+1} = (M_{i,1} \otimes M_{i,2} \otimes \cdots M_{i,k})^{-1} P_i$$

Matrix generation

Noise matrix formulates the transformation probability from the ideal state to measured state.

							set state				
							00	01	10	11	
[0.6	0	0.1	0]			00	0.6	0	0.1	0	
$\begin{bmatrix} 0.6 \\ 0.1 \\ 0.2 \end{bmatrix}$	0.7	0.2	0.1	=		read	01	0.1	0.7	0.2	0.1
0.2	$0.2 \\ 0.1$		$\begin{bmatrix} 0.1 \\ 0.8 \end{bmatrix}$		read state	01	0.2	0.2	0.6	0.1	
Loiz	011	011	0.0]			11	0.1	0.1	0.1	0.8	





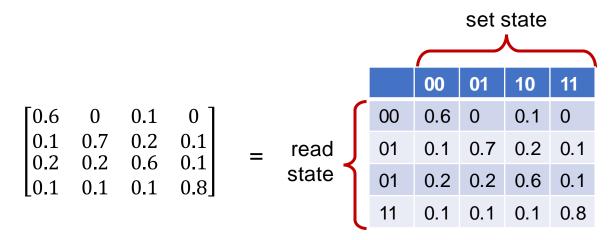
Technique 2: sub-noise matrix generation

Perform matrix-vector multiplication

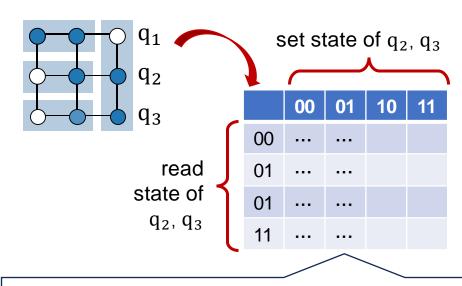
Iter. i:
$$P_{i+1} = (M_{i,1} \otimes M_{i,2} \otimes \cdots M_{i,k})^{-1} P_i$$

Matrix generation

Noise matrix formulates the transformation probability from the ideal state to measured state.



Sub-noise matrices of QuFEM formulates the transformation probability of states inside the qubit groups.



$$M[x][y] = P({q_2, q_3}. read = x | {q_2, q_3}. set = y, q_1 = 2)$$

Transformation probability when q_1 is not measured

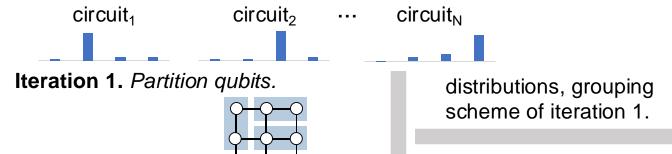
Put all together





Characterization

Iteration 1. Run benchmarking circuits.



Iteration 1. Calibrate.

Iteration 2. Partition qubits.



Iteration 2. Calibrate.

distributions, grouping scheme of iteration 2.

Calibration

Input. measured qubits measured distribution





Iteration 1. Generate sub-noise matrices.











Iteration 1. Calibrate.

Iteration 2. Generate sub-noise matrices.







Iteration 2. Calibrate.

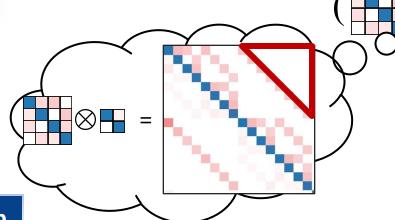


Sparse Tensor-Product Engine









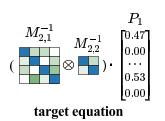
A large number of sparse intermediate vectors is generated in the tensor-product.

Implementation

Use a key-value table to store sparse vector

Calculate the tensor product

Aggregate the tensorproduct result



х	prob.	1		value	2		value
$P_1(000)$	0.47	→	00	0.50	\otimes	0	0.99
$P_1(011)$	0.53		01	-0.02		1	0.01
		•	10	0.01			
			11	0.01			3

 $\begin{array}{c|c} 9 & \cdot 0.47 = \\ \hline 0 & 000 & 0.47 \\ \hline 0 & 001 & 0.67 \\ \hline 0 & 010 & -0.67 \\ \hline 0 & 010 & 0.67 \\ \hline 0 &$

	value	4	X	prob.
000	0.49	$ \rightarrow\rangle$	$P_2(000)$	0.48
001	0.01	<u> </u>	$P_2(001)$	6×10 ⁻³
010	-0.01		$P_2(010)$	6×10 ⁻³
100	0.01		$P_2(011)$	0.50
101	10-4		$P_2(111)$	6×10 ⁻³
110	10-4			

Prune values < threshold (e.g., 10

For each basis states

- ① calculate the matrix-vector multiplication
- ② calculate the tensor-product
- 3 prune intermediate values
- (4) sum intermediate values to obtain output.

Compute the tensor-product of other basis states

Outline of Presentation





- Background and challenges
- Overview of QuFEM
- QuFEM characterization and calibration
- Experiment
- API of QuFEM

Experiment



Setup

Platform	#Qubits	1-q fidelity	2-q fidelity	Instructions
Overfor	136	94.6±3.1%	94.6±3.0%	ID,RX,RY,RZ,H,CX
Quafu	18	95.9±1.3%	95.9±1.3%	ID,RX,RY,RZ,H,CX
Rigetti	79	99.5±1.1%	90.0±6.4%	CPHASE,XY
Self-developed	36	99.9±0.1%	98.7±0.8%	U3,CZ
IBMQ	7	99.9±0.1%	99.2±0.1%	CX,ID,RZ,SX,X

Evaluated hardware

IBU: KJ Satzinger, et al. Realizing topologically ordered states on a quantum processor. Science 2021

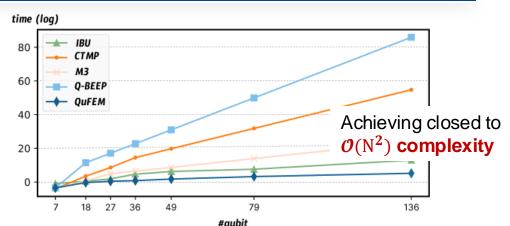
CTMP: Sergey, et al. Mitigating measurement errors in multiqubit experiments. PRA 2021.

M3: Paul D Nation, et al. Scalable mitigation of measurement errors on quantum computers. PRX Quantum 2021.

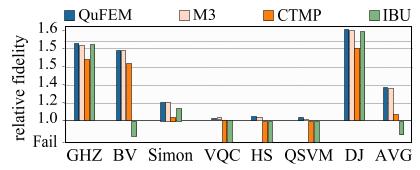
Q-BEEP: Nathan Wiebe, et al. QBEEP: Quantum Bayesian error mitigation employing Poisson modeling over the hamming spectrum. ISCA 2023.

Baselines

Result



QuFEM reduces the calibration time of the 136-qubit program output from 119.44 hours (IBU) to 169.65 seconds (119.44 \times reduction).



QuFEM shows an average improvement in relative fidelity of 1.003×, 1.2×, and 1.4× compared to M3, CTMP, and IBU, respectively.

Conclusion





- 1. Limitations of prior matrix-based calibration methods: slow and inaccurate
- 2. Finite element method: a divide and conquer strategy
- 3. Detailed techniques to partition qubits and generate noise matrix
- 4. Sparse tensor product engine to speed up the computation

before pressing iter. 2 iter. 3 after pressing after pressing Classical Finite Element Method before pressing lattice group analyzed independently sponge pressing

measured distribution iter .1 iter .2 iter .3 calibrated calibrated distribution formulated independently groups in each iteration

Outline of Presentation





- Background and challenges
- Overview of QuFEM
- QuFEM characterization and calibration
- Experiment
- API of QuFEM

API of QuFEM





File:

- JanusQ/examples/ipynb/4_1_readout_calibration_simulator.ipynb
- JanusQ/examples/ipynb/4_2_readout_calibration_realqc.ipynb
- https://janusq.github.io/tutorials/demo/4_1_readout_calibration_ simulator
- https://janusq.github.io/tutorials/demo/4_2_readout_calibration_realqc

```
from janusq.calibration.readout_mitigation.qufem import Mitigator
from janusq.calibration.readout_mitigation.qufem.tools import npformat_to_statuscnt

construct mitigator

mitigator = Mitigator(n_qubits=8, n_iters=2)
scores = mitigator.init(benchmark_circuits_and_results, group_size=2,
multi_process=False, draw_grouping=True)

calibrate output of
GHZ circuit

n_qubits = 4
outout_ideal = { '1'*n_qubits: 0.5, '0'*n_qubits: 0.5 }
output_fem = mitigator.mitigate(ghz_error[0], [_ for _ in range(n_qubits) ], cho = 1)
output_fem = npformat_to_statuscnt(output_fem)
```



Thanks for listening

QuFEM: Fast and Accurate Quantum Readout Calibration Using the Finite Element Method

Siwei Tan, Liqiang Lu*, Hanyu Zhang, Jia Yu, Congliang Lang, Yongheng Shang, Xinkui Zhao, Mingshuai Chen, Yun Liang, and Jianwei Yin*