Machine learning for practical quantum error mitigation

Haoran Liao, Derek Wang, Iskandar Sitdikov, Ciro Salcedo, Alireza Seif, **Zlatko Minev (IBM*)**

Nature Machine Intelligence v. 6, p. 1478 (2024)

Acknowledgements: Brian Quanz, Patrick Rall, Oles Shtanko, Roland De Putter, Kunal Sharma, Barbara Jones, Sona Najafi, Thaddeus Pelligrini, Grace Harper, Vinay Tripathi, Antonio Mezzacapo, Christa Zoufal, Travis Scholten, Bryce Fuller, Swarnadeep Majumder, Sarah Sheldon, Youngseok Kim, Pedro Rivero, Will Bradbury, Nate Gruver, Minh Tran, Kristan Temme, and the broader IBM Quantum team

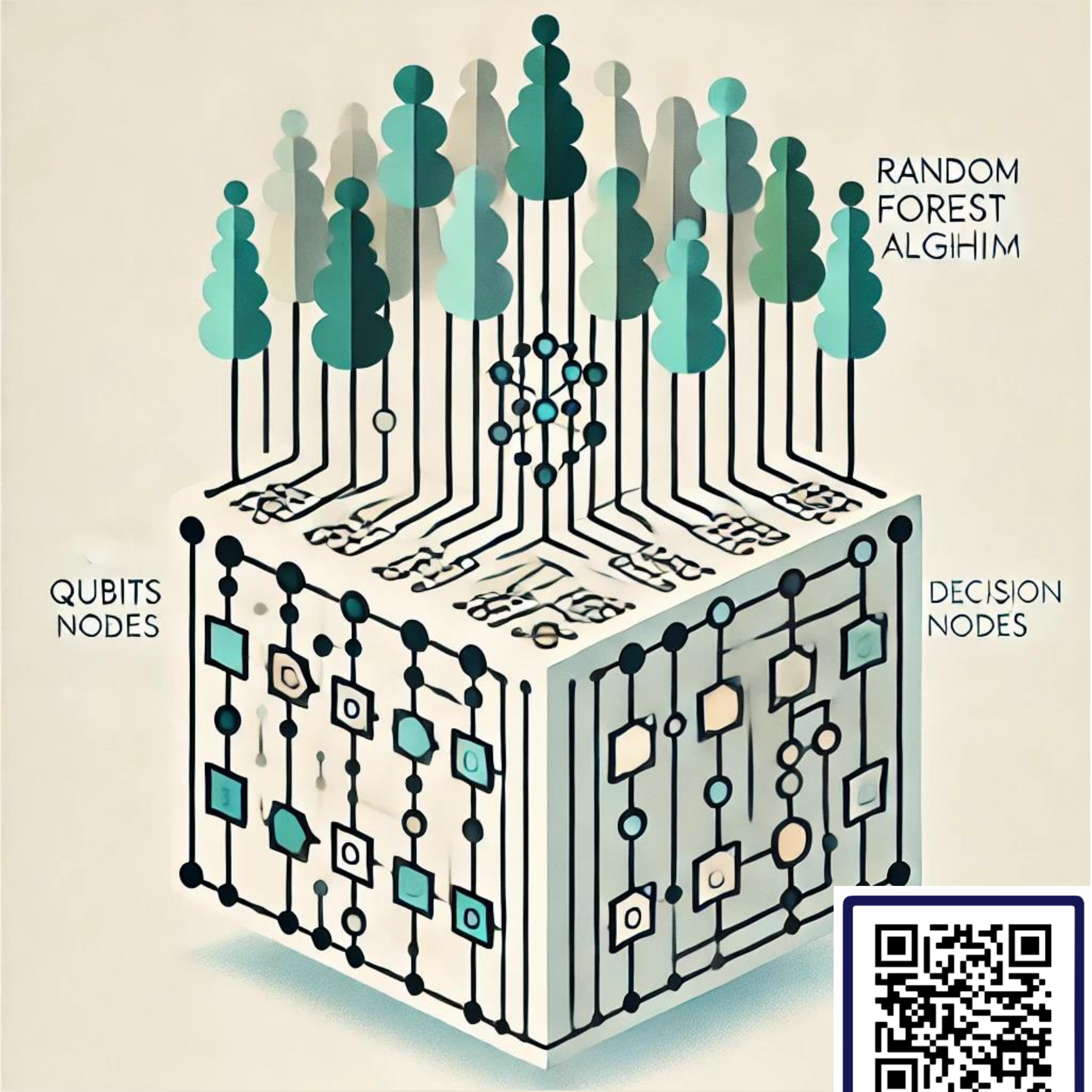


Illustration generated using AI (DALL·E)

IBM

Slides?

zlatko-minev.com/blog @zlatko_minev

Slides

Why?

Biggest challenge?

Please do share

Biggest challenge?

hardware development

decoherence

loss

stability

heat

Noise (Errors)

algo development

error correction overheads

high error rates

scalability

engineering

modularization

importance of N in NISQ

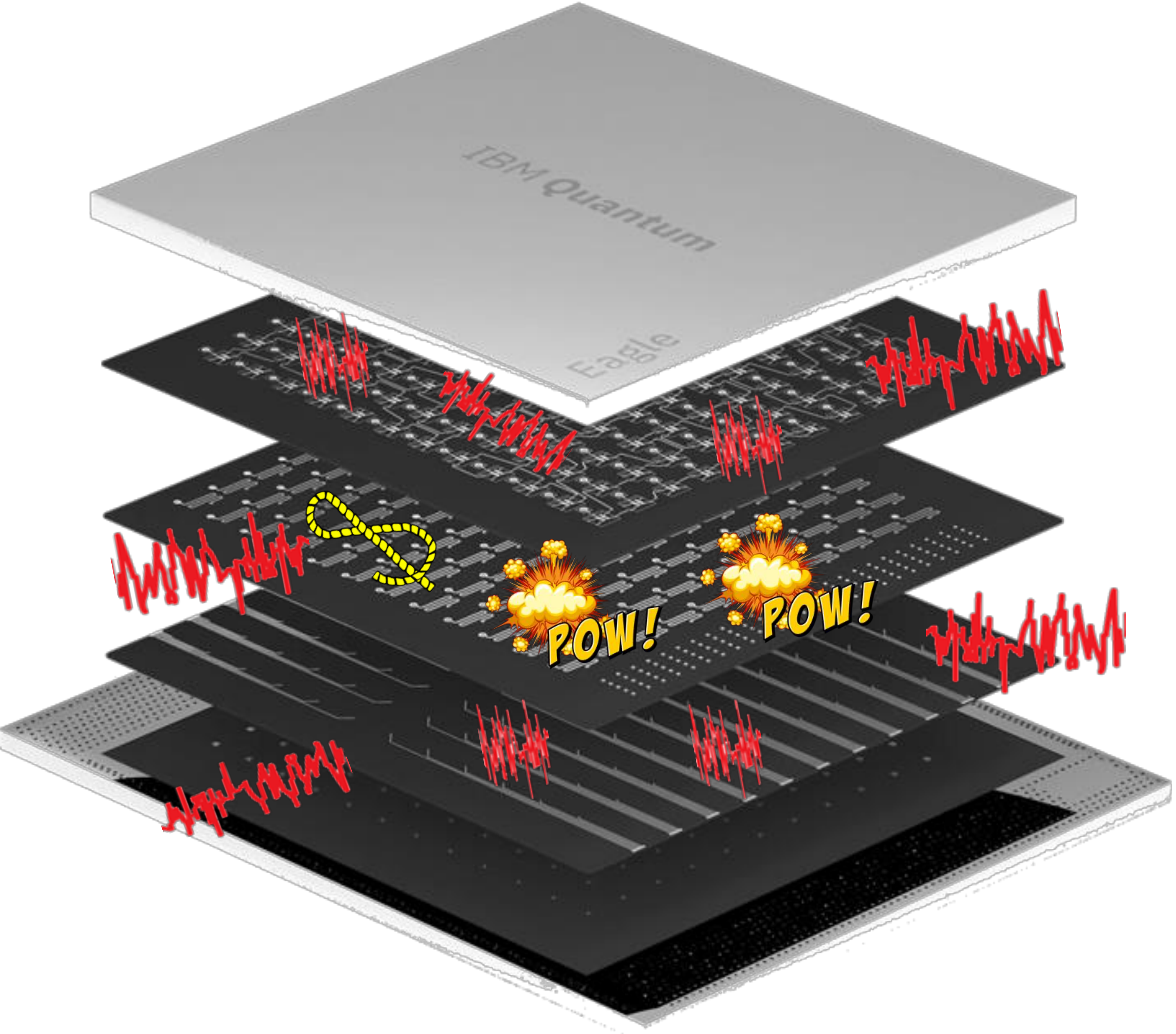
hype

need CS/EE talent

material quality

gravity

expectations



Noise
(Errors)

Zlatko Minev
IBM Quantum

How to deal with errors due to noise?

Monitor
Error occurs
Error detect

Quantum error correction

Shor, PRA (1995), ...

Monitor
Error anticipated
Tell signal detected

Catch and reverse

Minev, Nature (2019), ...

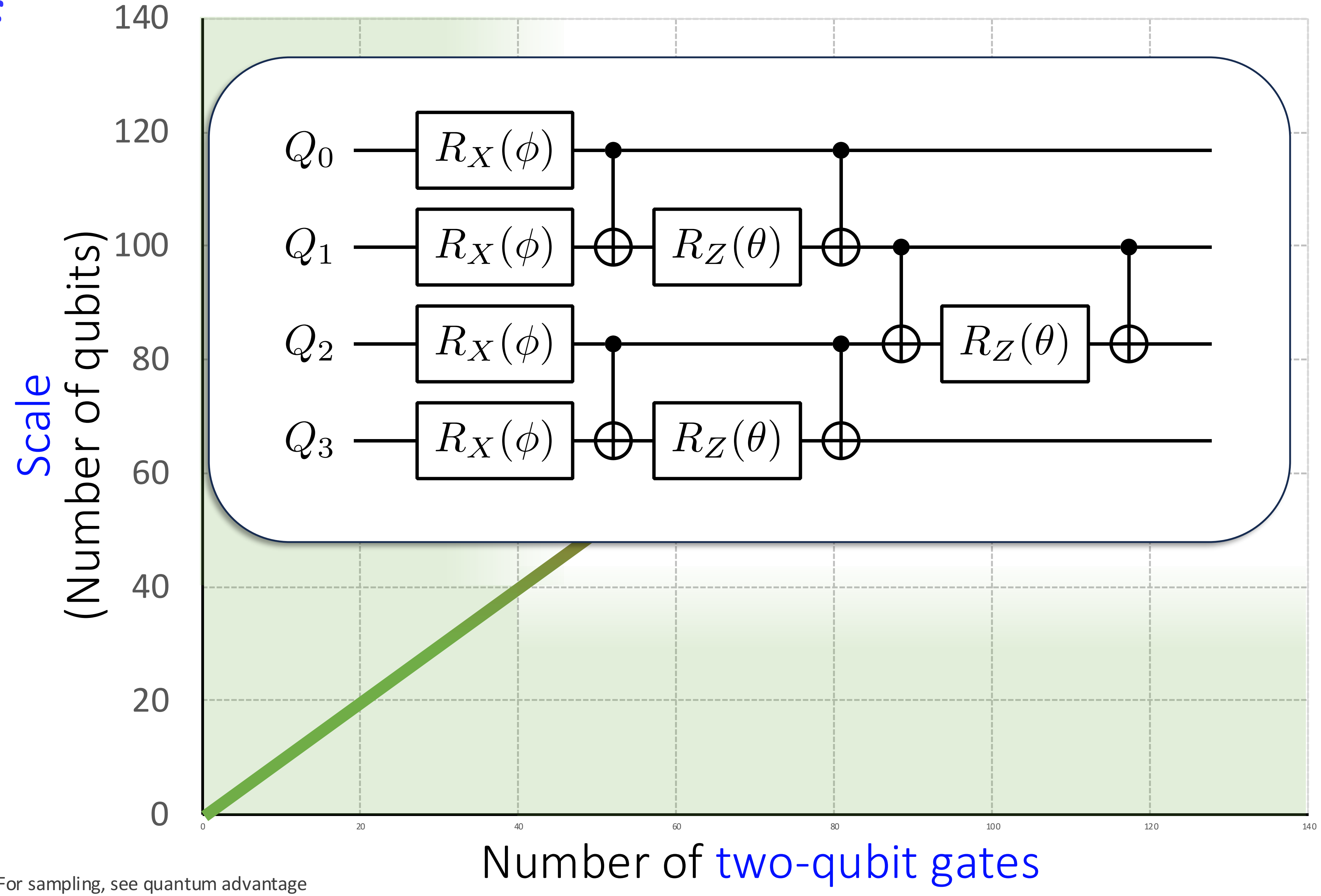
No monitor
Error occurs
Error undetected

Error mitigation

RMP 95, 045005 (2023)
... subject of today

Landscape of quantum circuits

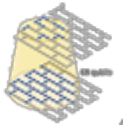
and the use of quantum error mitigation



Zlatko Minev
IBM Quantum

For expectation values. For sampling, see quantum advantage with shallow circuits, Bravyi, Gosset, Konig, and related

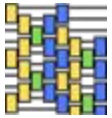
Some early utility-scale experiments



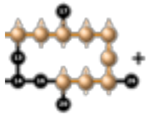
[0] Kim, Eddins, …, Temme, Kandala. Nature 618, 500–505 (2023)

[1] Yu, Zhao, Wei. arXiv: 2207.09994 (2022)

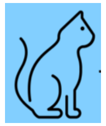
[2] Shtanko, Wang, Zhang, Harle, Seif, Movassagh, Minev. arXiv: 2307.07552 (2023)



[3] Farrell, Illa, Ciavarella, Savage. arXiv: 2308.04481 (2023)



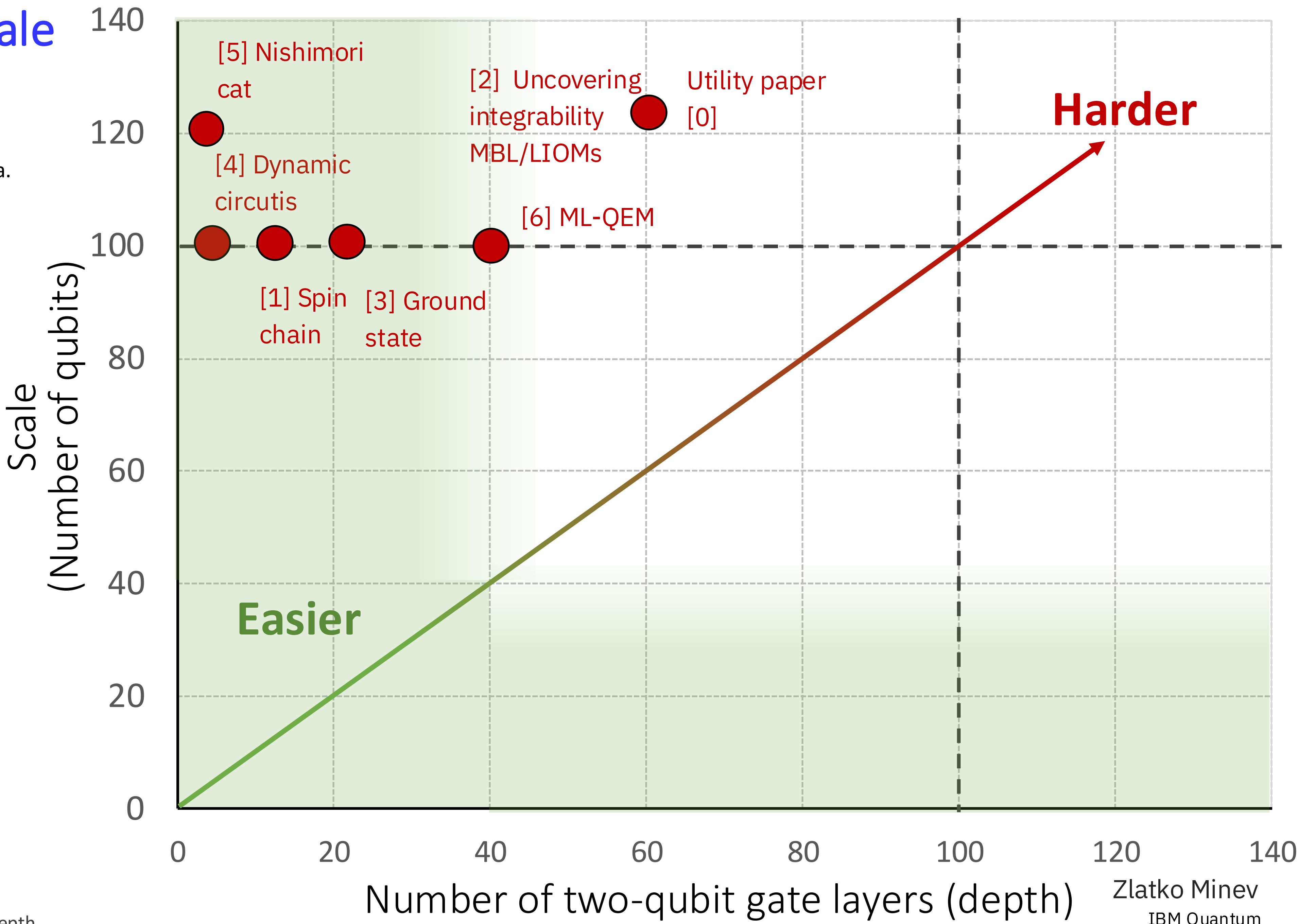
[4] Bäumer, Tripathi, …. Seif, Minev. arXiv: 2308.13065 (2023)



[5] Chen, Zhu, Verresen, Seif, Baümer, … Trebst, Kandala. arXiv: 2309.02863 (2023)

[6] Liao, Wang, Sitdikov, Salcedo, Seif, Minev. arXiv: 2308.13065 (2023)

… see next slide



On superconducting qubit platforms

Signal stops or decays more than 50% beyond this depth

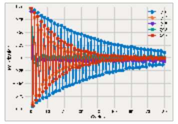
Zlatko Minev

IBM Quantum

Utility-scale demos: More recent examples

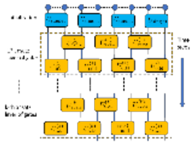
Evidence for the utility of quantum computing before fault tolerance
Nature, 618, 500 (2023)
127 qubits / 2880 CX gates

simulation



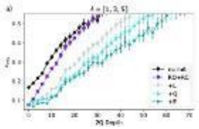
Characterizing quantum processors using discrete time crystals
arXiv:2301.07625
80 qubits / 7900 CX gates

simulation



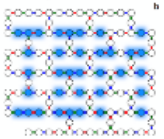
Simulating large-size quantum spin chains on cloud-based superconducting quantum computers
Phys. Rev. Research 5, 013183 (2023)
102 qubits / 3186 CX gates

simulation



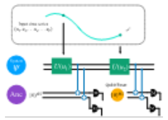
Best practices for quantum error mitigation with digital zero-noise extrapolation
arXiv:2307.05203
104 qubits / 3605 ECR gates

tools



Uncovering Local Integrability in Quantum Many-Body Dynamics
arXiv:2307.07552
124 qubits / 2641 CX gates

simulation

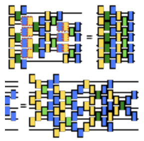


Quantum reservoir computing with repeated Measurements on superconducting devices
arXiv:2310.06706
120 qubits / 49470 ECR gates + meas.

QML

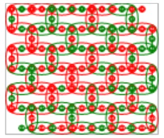
Realizing the Nishimori transition across the error threshold for constant-depth quantum circuits
arXiv:2309.02863
125 qubits / 429 gates + meas.

simulation



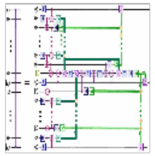
Scalable Circuits for Preparing Ground States on Digital Quantum Computers: The Schwinger Model Vacuum on 100 Qubits
PRX Quantum 5, 020315 (2024)
100 qubits / 788 CX gates

simulation



Scaling Whole-Chip QAOA for Higher-Order Ising Spin Glass Models on Heavy-Hex Graphs
arXiv:2312.00997
127 qubits / 420 CX gates

optimization

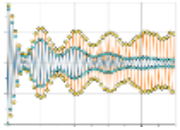


Efficient Long-Range Entanglement using Dynamic Circuits
arXiv:2308.13065
101 qubits / 504 ECR gates + meas

tools

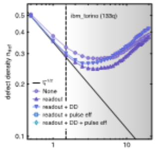
Quantum Simulations of Hadron Dynamics in the Schwinger Model using 112 Qubits
arXiv:2401.08044
112 qubits / 13,858 CZ gates

simulation



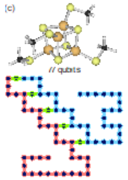
Unveiling clean two-dimensional discrete time quasicrystals on a digital quantum computer
arXiv:2403.16718
133 qubits / 15,000 CZ gates

simulation



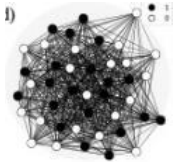
Benchmarking digital quantum simulations and optimization above hundreds of qubits using quantum critical dynamics
arXiv:2404.08053
133 qubits / 1440 CX gates

simulation



Chemistry Beyond Exact Solutions on a Quantum-Centric Supercomputer
arXiv:2405.05068
77 qubits / 3590 CZ gates

simulation



Towards a universal QAOA protocol: Evidence of quantum advantage in solving combinatorial optimization problems
arXiv:2405.09169
109 qubits / 21,200 ECR gates

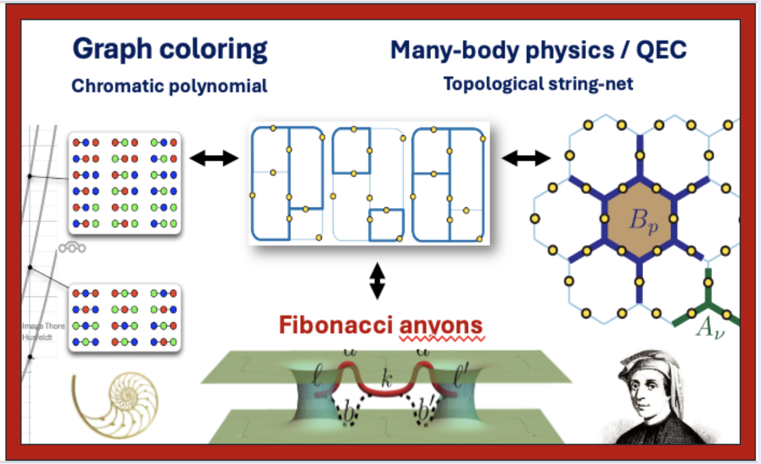
....

optimization

IBM Quantum | © 2024 IBM Corporation

Some related utility-scale & advantage experiments

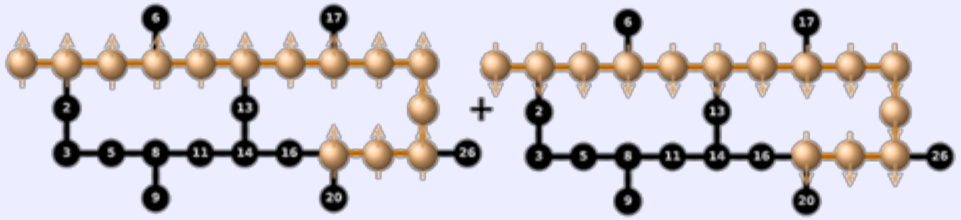
Towards classically-hard hard problems and universal topological quantum computation through Fibonacci string-net condensate



arXiv:2406.12820

Z. Minev, **K. Najafi,** S. Majumder, J. Wang, A. Stern, **Eun-Ah Kim,** C.-M. Jian, G. Zhu

Efficient Long-Range Entanglement using Dynamic Circuits

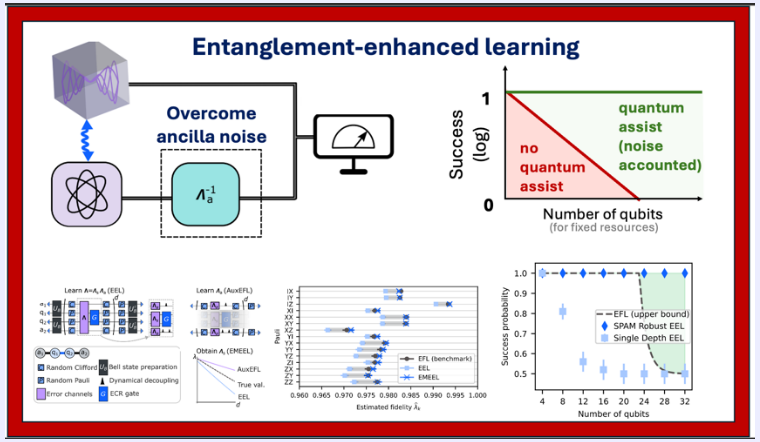


Gate teleportation over 100+ qubits

PRX Quantum 5, 030339

Bäumer, Tripathi, Wang, Rall, Chen, Majumder, Seif, Minev

Noise-Robust Quantum Learning Advantage at Scale



arXiv:2408.03376

A. Seif, S. Chen, S. Majumder, H. Liao, D. Wang, M. Malekakhlagh, A. Javadi-Abhari, L. Jiang, Z. Minev

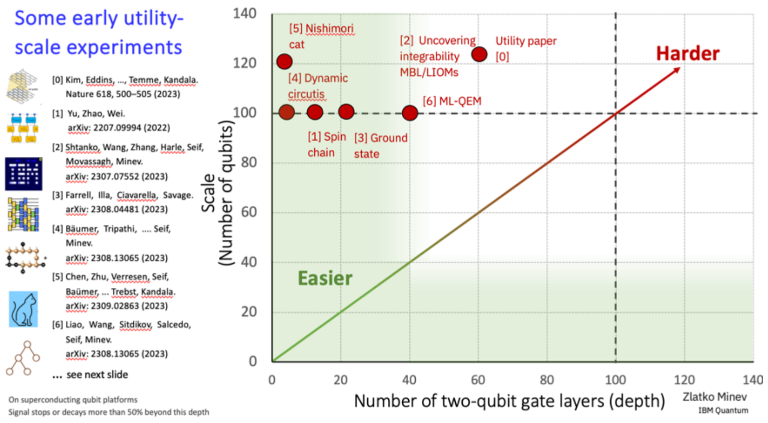
Demonstration of Robust and Efficient Quantum Property Learning with Shallow Shadows

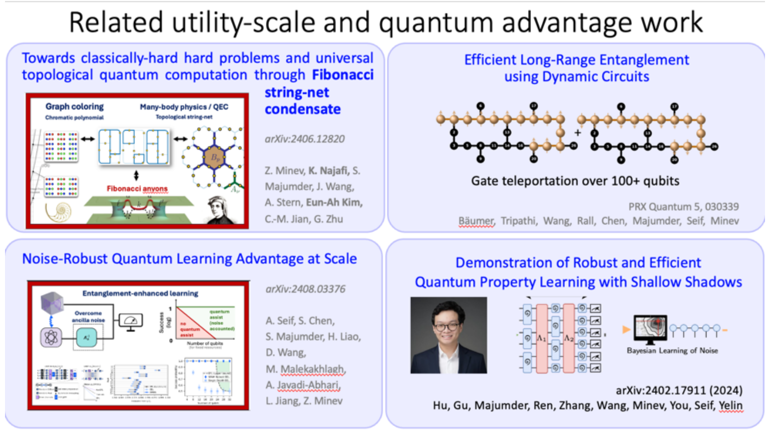
arXiv:2402.17911 (2024)

Hu, Gu, Majumder, Ren, Zhang, Wang, Minev, You, Seif, Yelin

These utility-scale results enabled by

quantum error mitigation

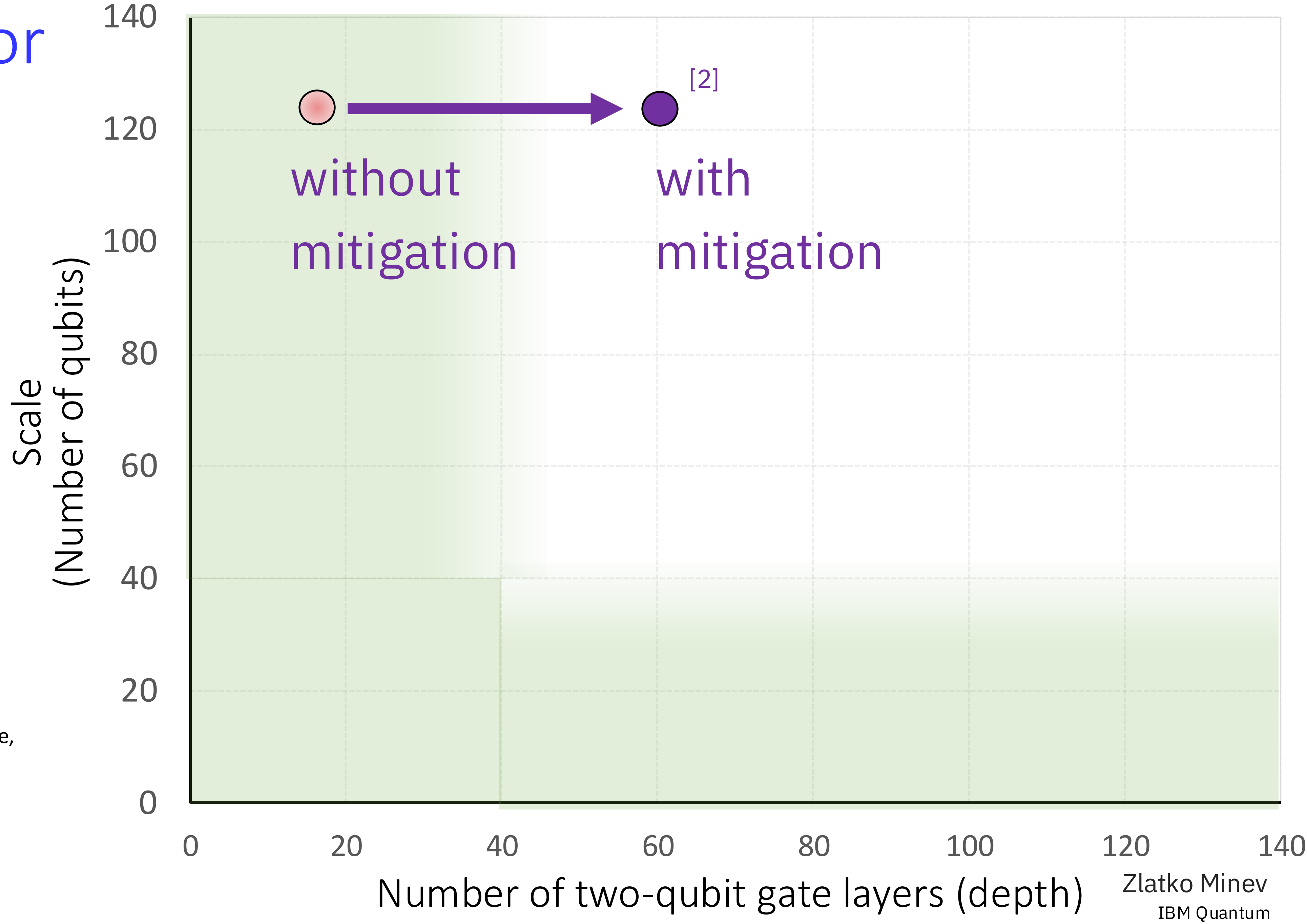




Quantum Error Mitigation

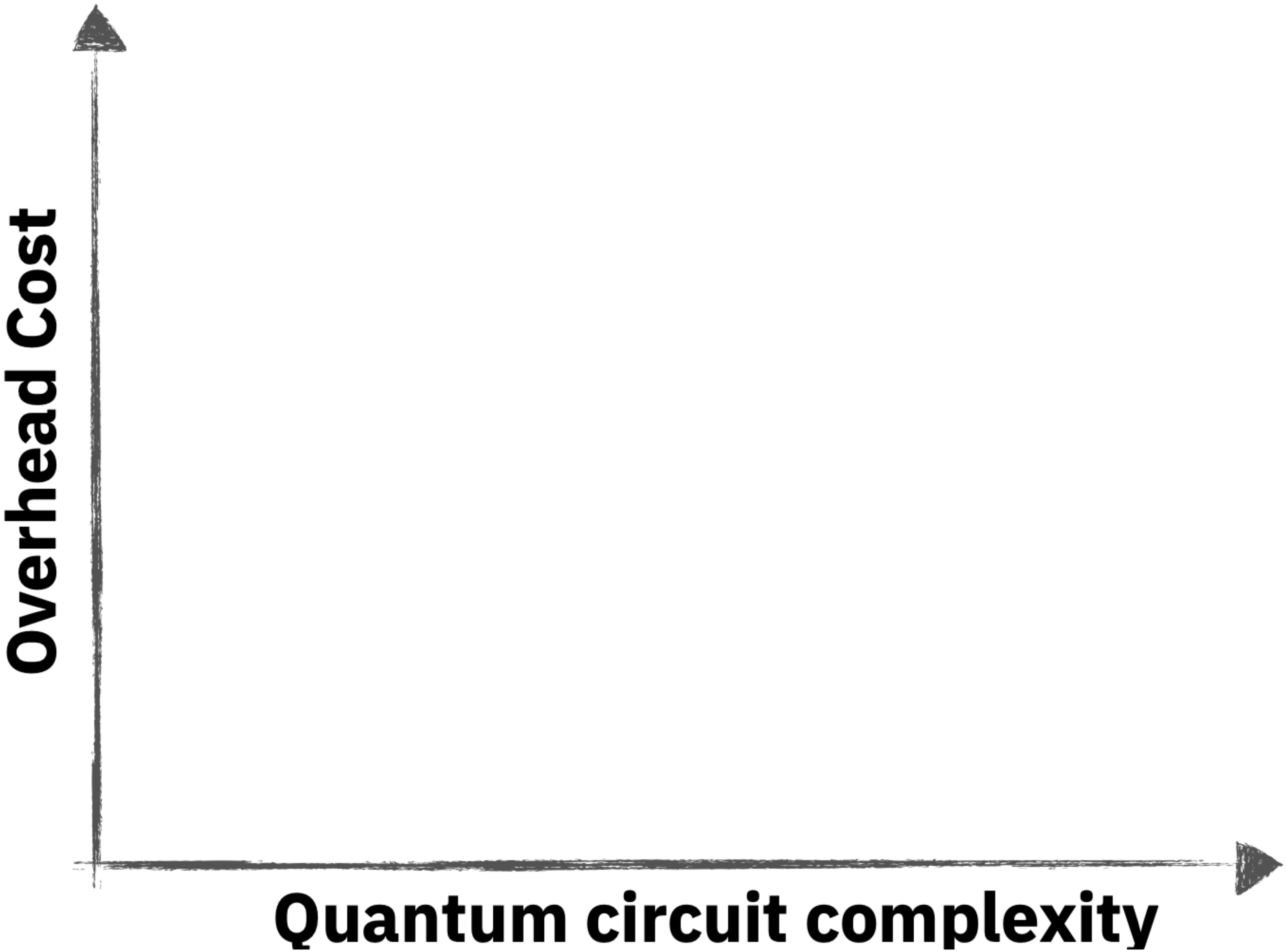
RMP 95, 045005 (2023)

Quantum error mitigation: example



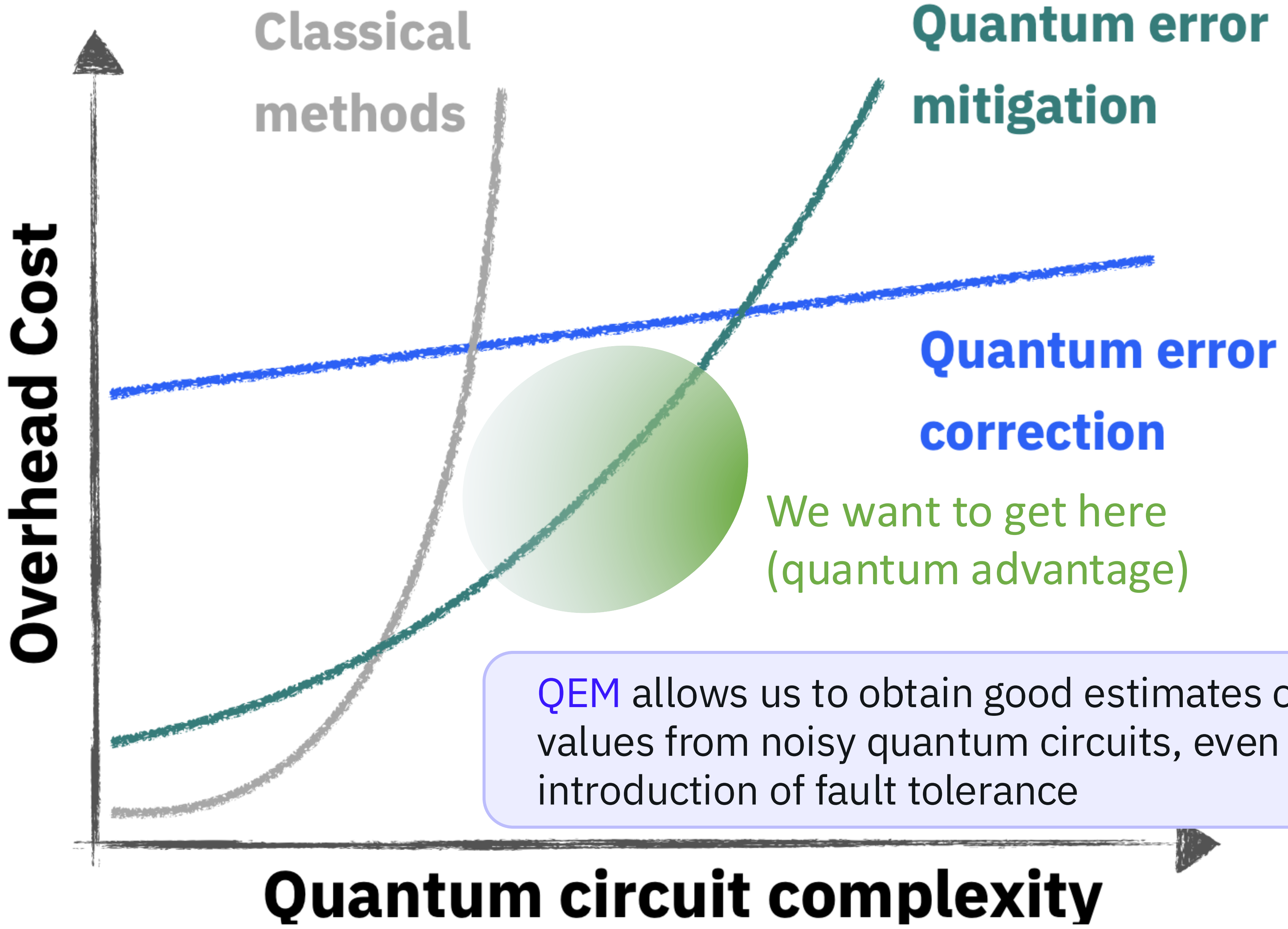
[2] Shtanko, Wang, Zhang, Harle, Seif, Movassagh, Minev. arXiv: 2307.07552 (2023)

Zlatko Minev
IBM Quantum



Zlatko Minev
IBM Quantum

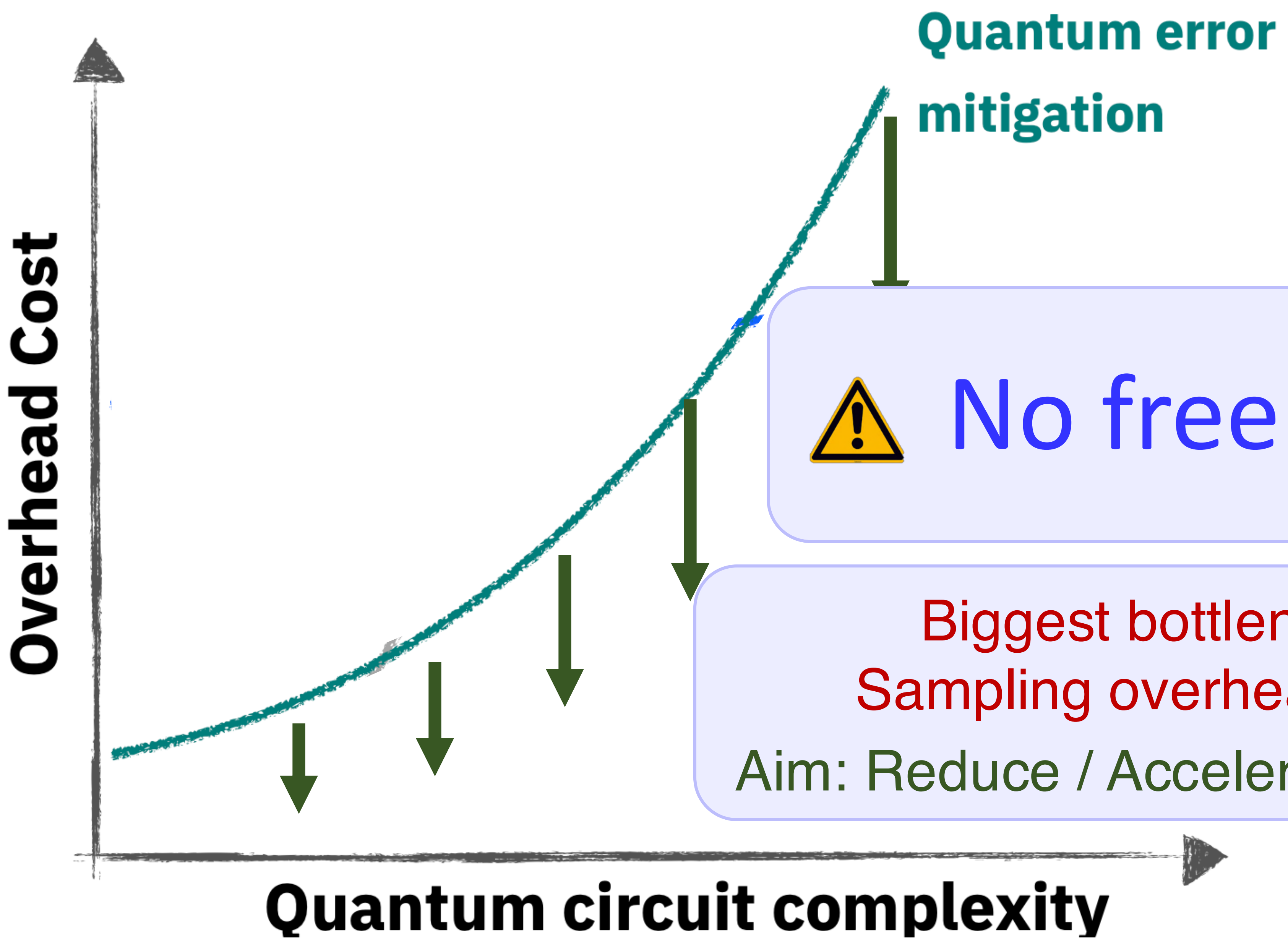
Quantum advantage: Quantum computation that delivers a significant, practical benefit beyond either brute force or approximate classical computing methods, calculating solutions in a way that is cheaper, faster or more accurate than all known classical alternatives.”
ibm.com/quantum/blog/gammabar-for-quantum-advantage



QEM allows us to obtain good estimates of expectation values from noisy quantum circuits, even before the introduction of fault tolerance

Zlatko Minev
IBM Quantum

"Quantum advantage: Quantum computation that delivers a significant, practical benefit beyond either brute force or approximate classical computing methods, calculating solutions in a way that is cheaper, faster or more accurate than all known classical alternatives."
ibm.com/quantum/blog/gammabar-for-quantum-advantage



⚠ No free lunch

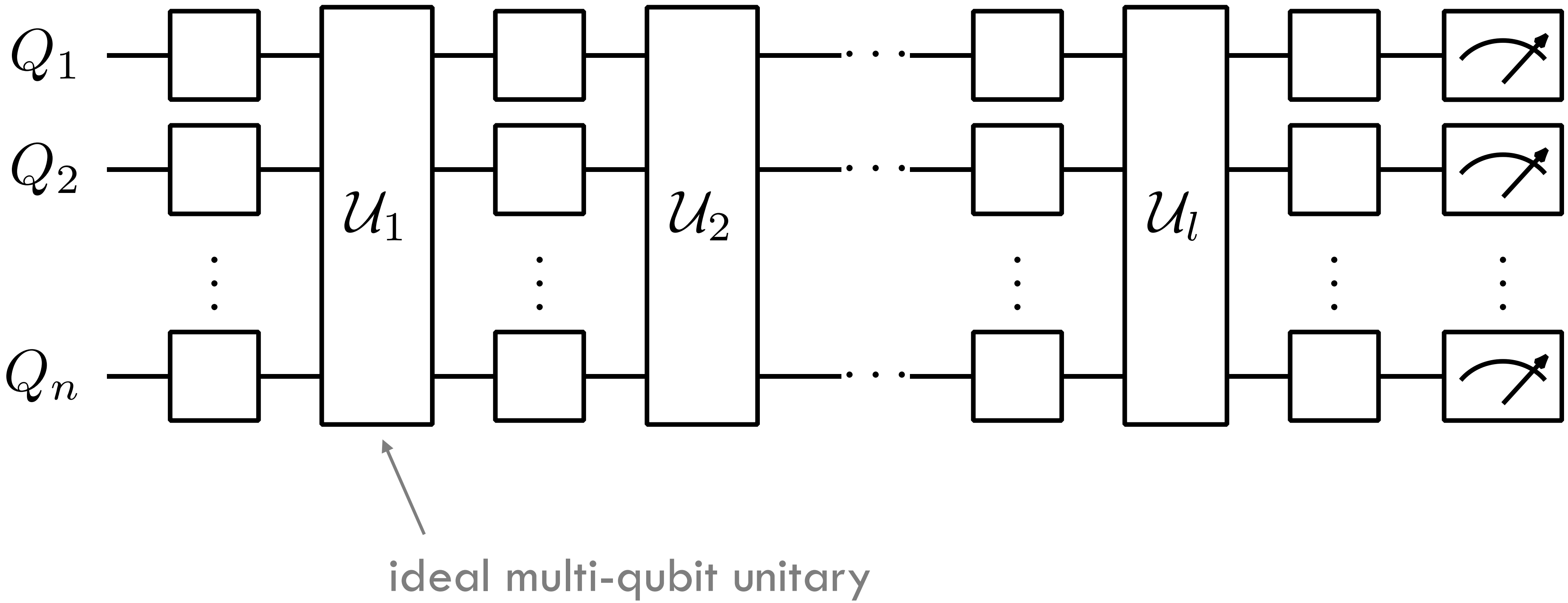
Biggest bottleneck:
Sampling overhead cost

Aim: Reduce / Accelerate with ML

Zlatko Minev
IBM Quantum

Tsubouchi, K., Sagawa, T. & Yoshioka, N. Universal cost bound of quantum error mitigation based on quantum estimation theory. PRL (2023).
Quek, Y. et al. Exponentially tighter bounds on limitations of quantum error mitigation. Nat. Phys. (2024).
Takagi, R., Tajima, H. & Gu, M. Universal sampling lower bounds for quantum error mitigation. PRL (2023).

Ideal (noise-free) quantum circuit

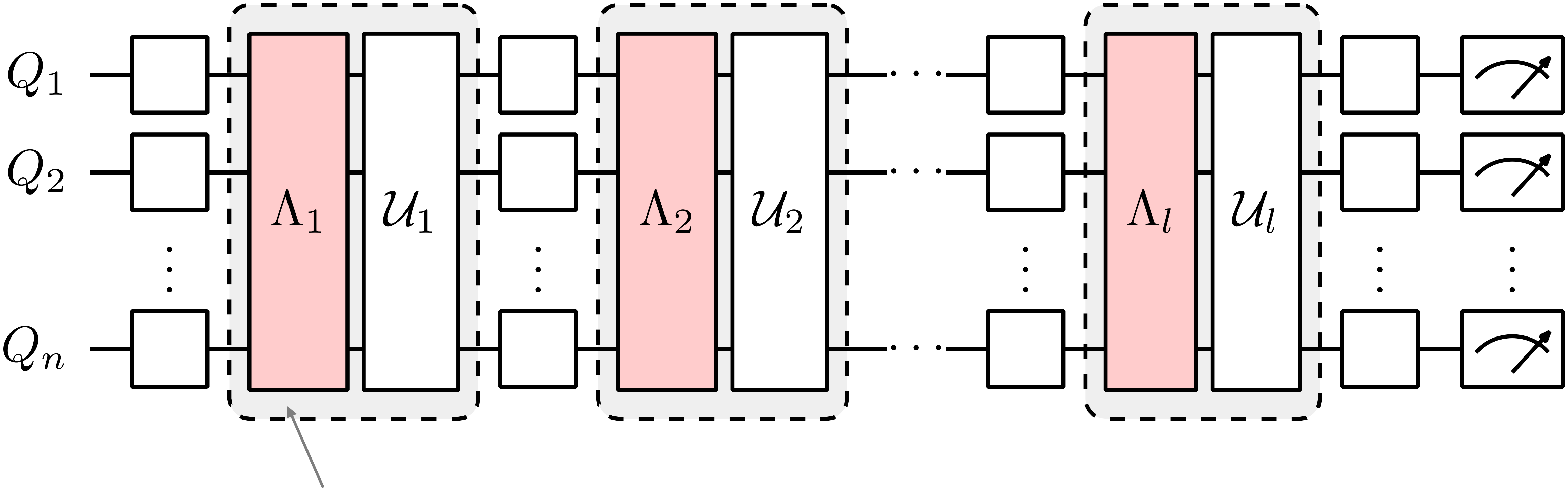


A circuit can be decomposed into a layer construction

Example: Trotterization of Ising model simulation

Zlatko Minev
IBM Quantum

Real (noisy) quantum circuit

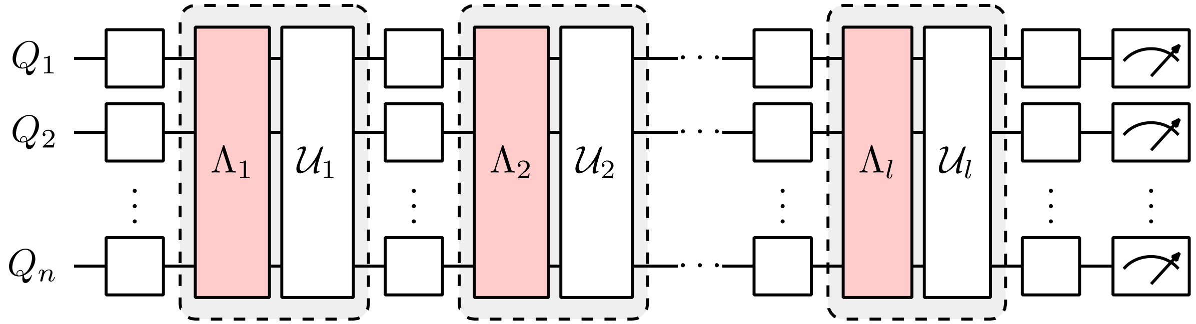


multi-qubit noise channel
inseparable from gate

completely positive and trace preserving (CPTP)
representable by a 4^n x 4^n matrix

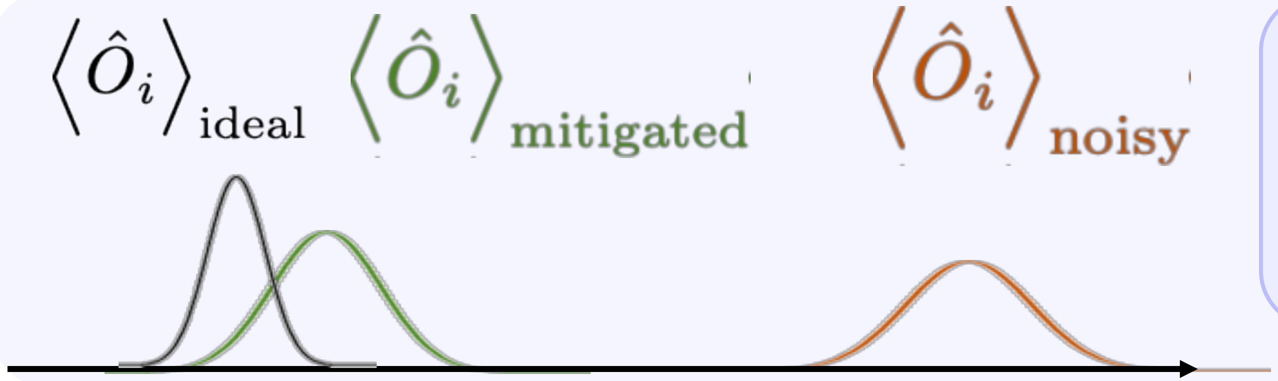
Zlatko Minev
IBM Quantum

Error mitigation: Wish list



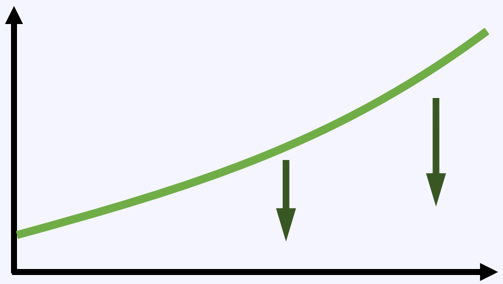
1. Mitigate *all* observables of interest

$$\left\{ \left\langle \hat{O}_1 \right\rangle, \left\langle \hat{O}_2 \right\rangle, \ldots \right\}$$



2. Precision with high probability

$$\Pr\left[\left| \left\langle \hat{O}_i \right\rangle_{\text{ideal}} - \left\langle \hat{O}_i \right\rangle_{\text{mitigated}} \right| \geq \epsilon \right] = \delta$$

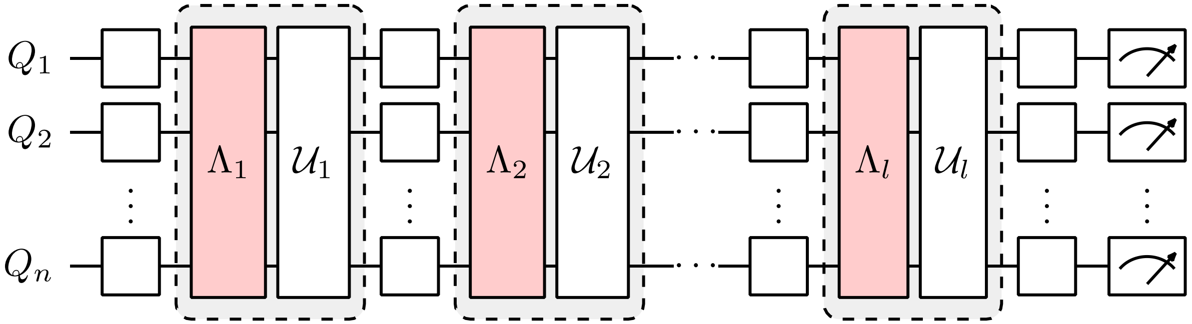


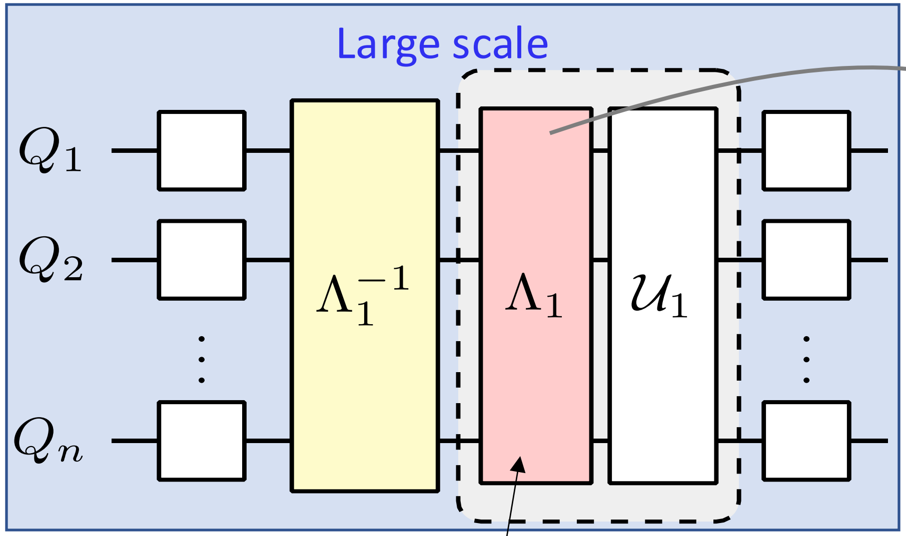
3. Minimize mitigation sampling overhead

$$\min N_{\text{mitigation}} / N_{\text{ideal}}$$

IBM Quantum

Error mitigation: Challenges





noise in full device

- cross-talk (quantum and classical), …
- correlated errors, context-dependence, …
- leakage, time instability, …

1. Beyond classical circuits

Paradoxical: How can classical ML learn handle these?

2. Quantum noise is exponentially complex

Quantum noise is exponentially complex and even simple noise mixed by complex unitary yields complex noise.

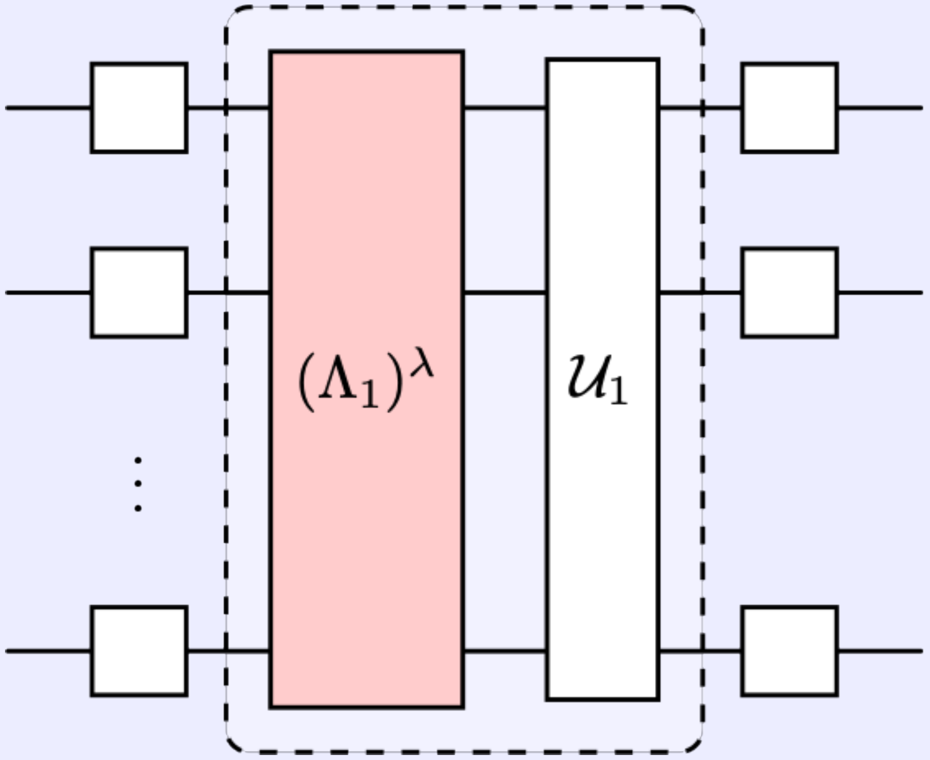
A. exp params — 10^{60} params for 50Q
B. noise param values — 10^{-2} - 10^{-5}
C. additive error sampling cost ($>10^2 - 10^{10}$)

3. Runtime cost

Number of circuits, construction and execution speed of circuits, platform speed, monetary expense, ...

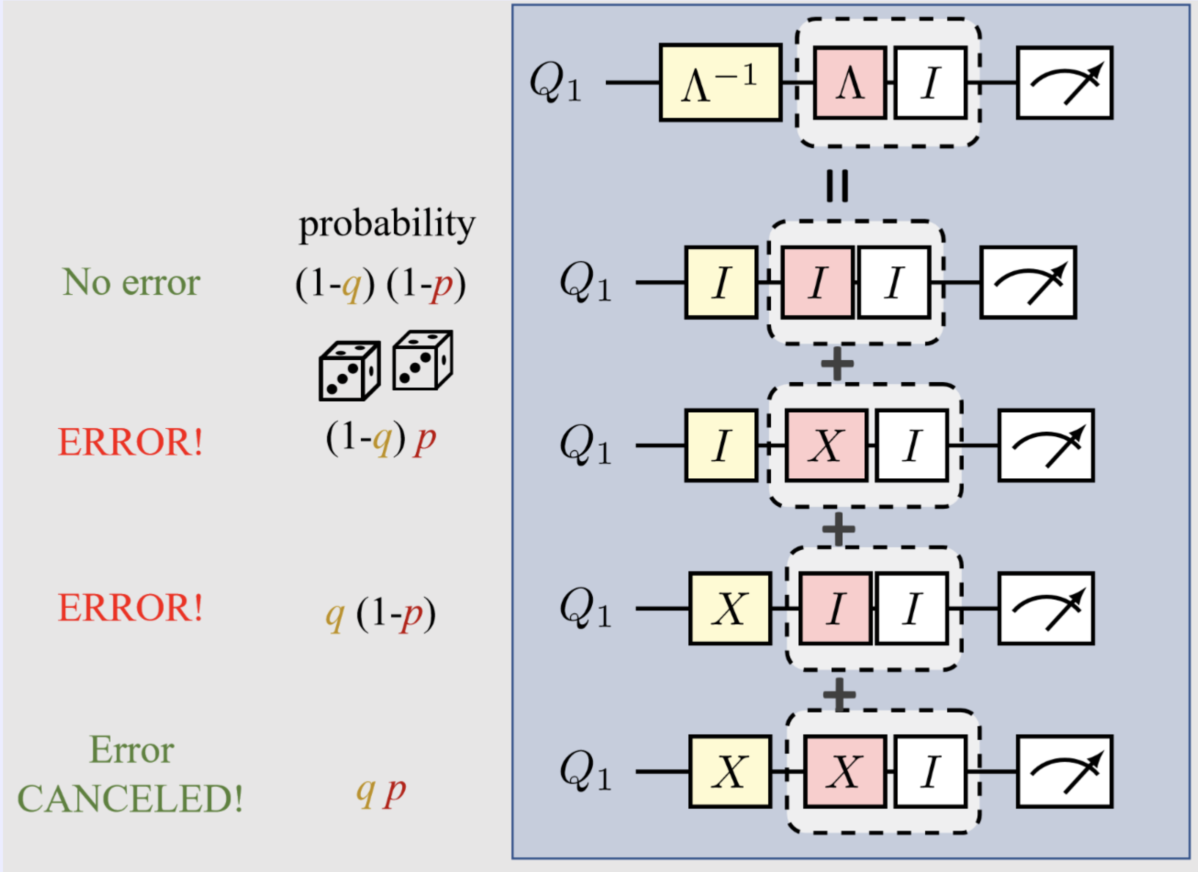
Two state-of-the-art physics-based QEM methods

Zero-noise extrapolation (ZNE)



K. Temme, S. Bravyi, and J. M. Gambetta. PRL (2017)
Li and Benjamin PRX (2017)
Note: PEA = ZNE + PEC combination

Probabilistic error cancellation (PEC)



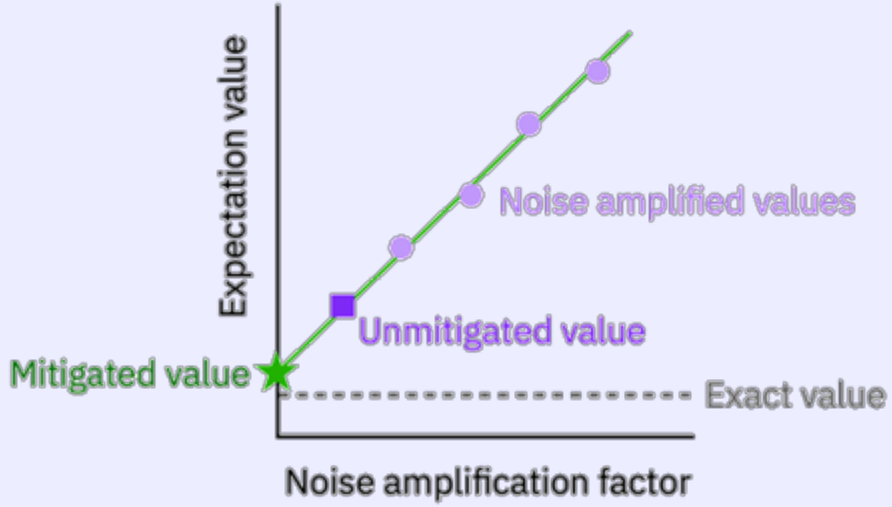
K. Temme, S. Bravyi, and J. M. Gambetta. PRL (2017)
E. Berg, Z. K. Minev, A. Kandala, and K. Temme. Nat. Phys. (2023)
Graphic: PEC tutorial zlatko-minev.com/blog/pec-talk

Physics-Based Methods

Zero-noise extrapolation (ZNE)

Pros: Relatively low overhead, no training, more straightforward to implement

Cons: Expectation values not bias-free in general

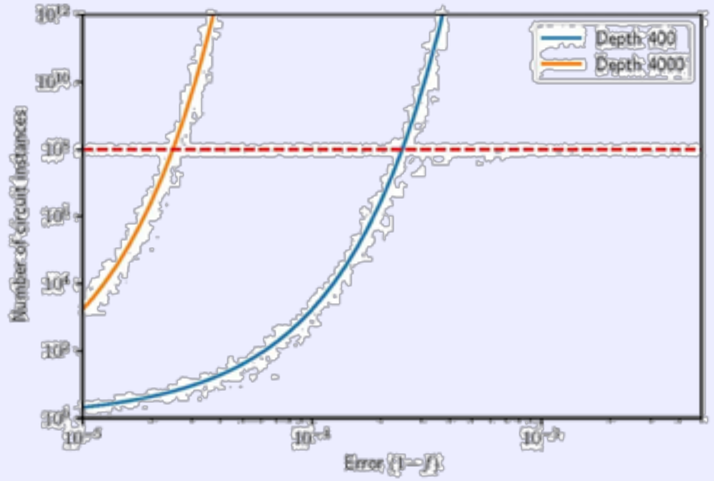


K. Temme, S. Bravyi, and J. M. Gambetta. Phys. Rev. Lett. 119, 180509 (2017)
Li and Benjamin PRX (2017)

Probabilistic error cancellation (PEC)

Pros: Unbiased expectation values

Cons: Fast and exponentially growing sampling overhead with noise. Restricted to layered circuits.



K. Temme, S. Bravyi, and J. M. Gambetta. Phys. Rev. Lett. 119, 180509 (2017)
E. Berg, Z. K. Minev, A. Kandala, and K. Temme. Nat. Phys. (2023)
Graphic: PEC tutorial zlatko-minev.com/blog/pec-talk

ML-QEM

Machine-learning quantum error mitigation

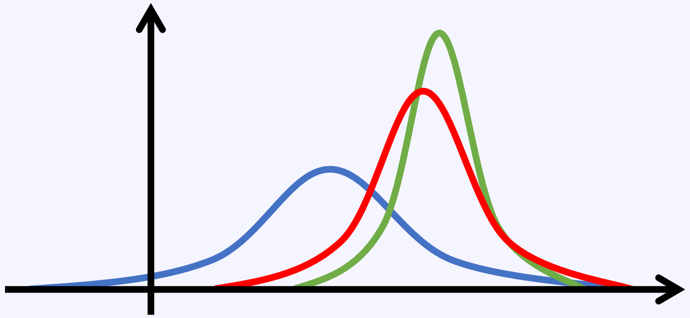
Nature Machine Intelligence v. 6, p. 1478 (2024)

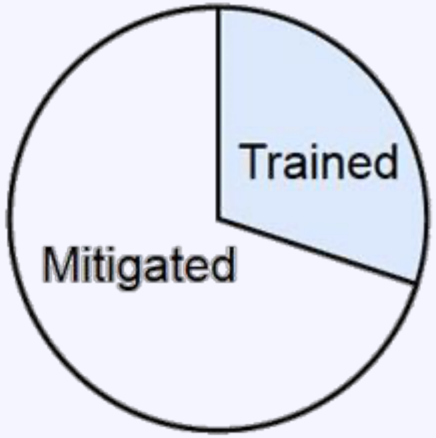
Machine learning accelerates error mitigation

ML-QEM reduces the cost of quantum error mitigation without sacrificing accuracy.

Pros: Arguably lowest runtime overhead, applies to multiple and generic expectation value, accelerates essentially any existing physics-based method ...

Cons: Could have biased expectation values, requires circuit-class specific training





Practical utility metrics

- Performance / precision
- Cost / efficiency
- Scalability
- Applicability

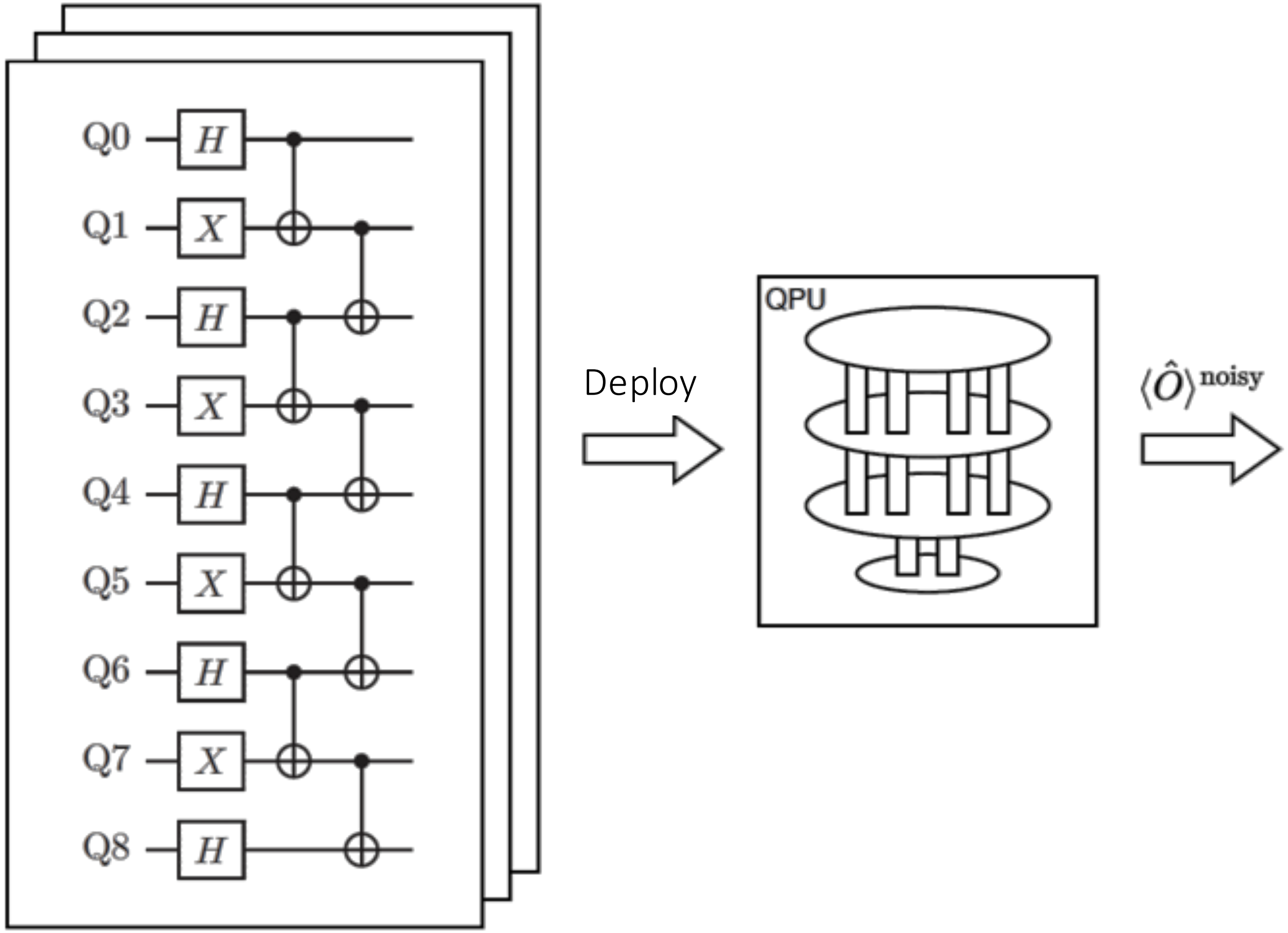
Disclaimer: No DeepSeek, ChatGPT, or Transformer moment here.

First proof of concept with very simple ideas.

Not be confused with the step of "learning device noise" in quasi-probability-based QEM approaches, e.g., PEC or probabilistic error amplification (PEA)**.

**Y. Kim, A. Eddins, S. Anand et al., Nature 618, 500–505 (2023)

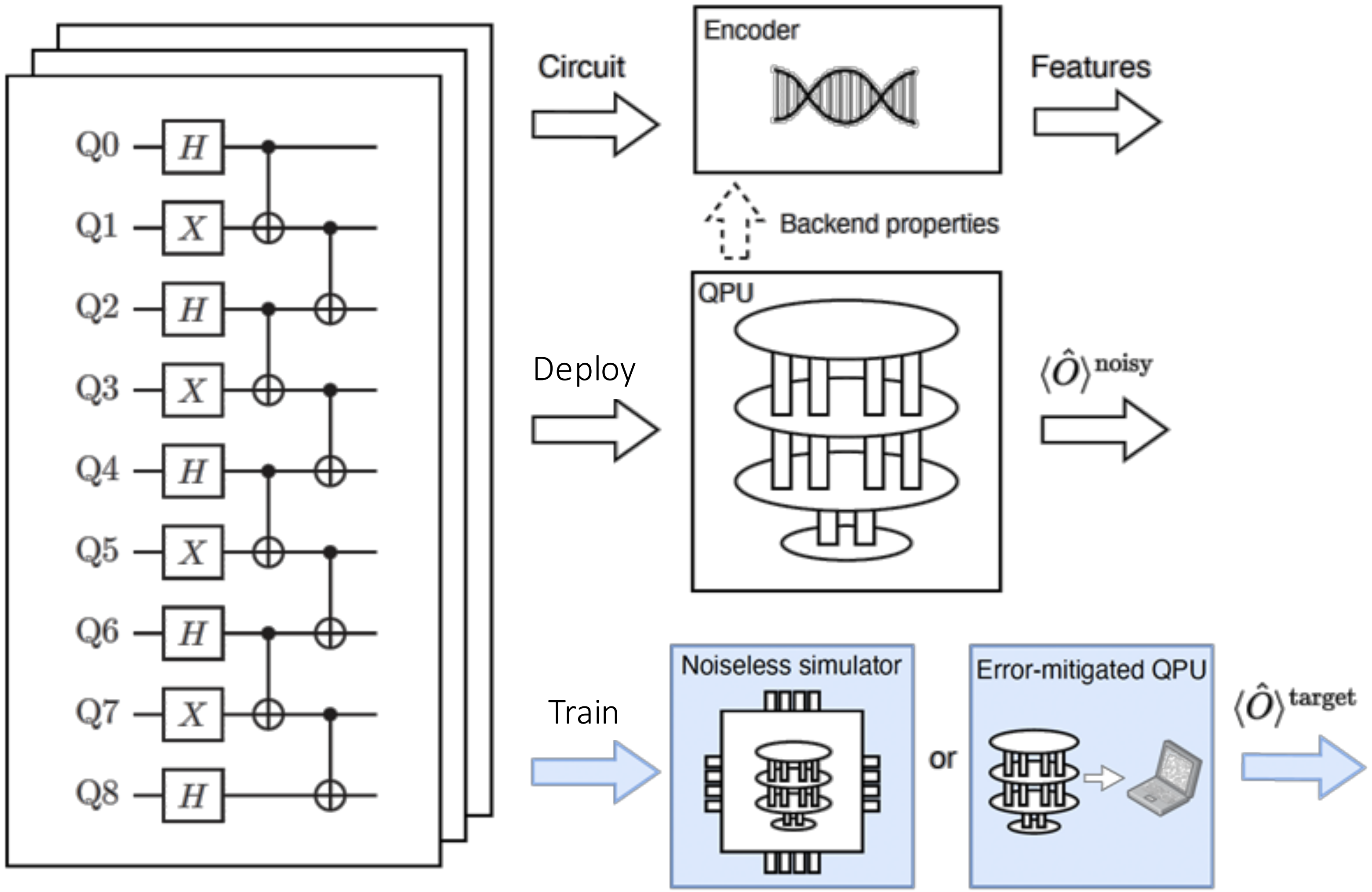
Normal execution of quantum circuit



Nature Machine Intelligence v. 6, p. 1478 (2024)

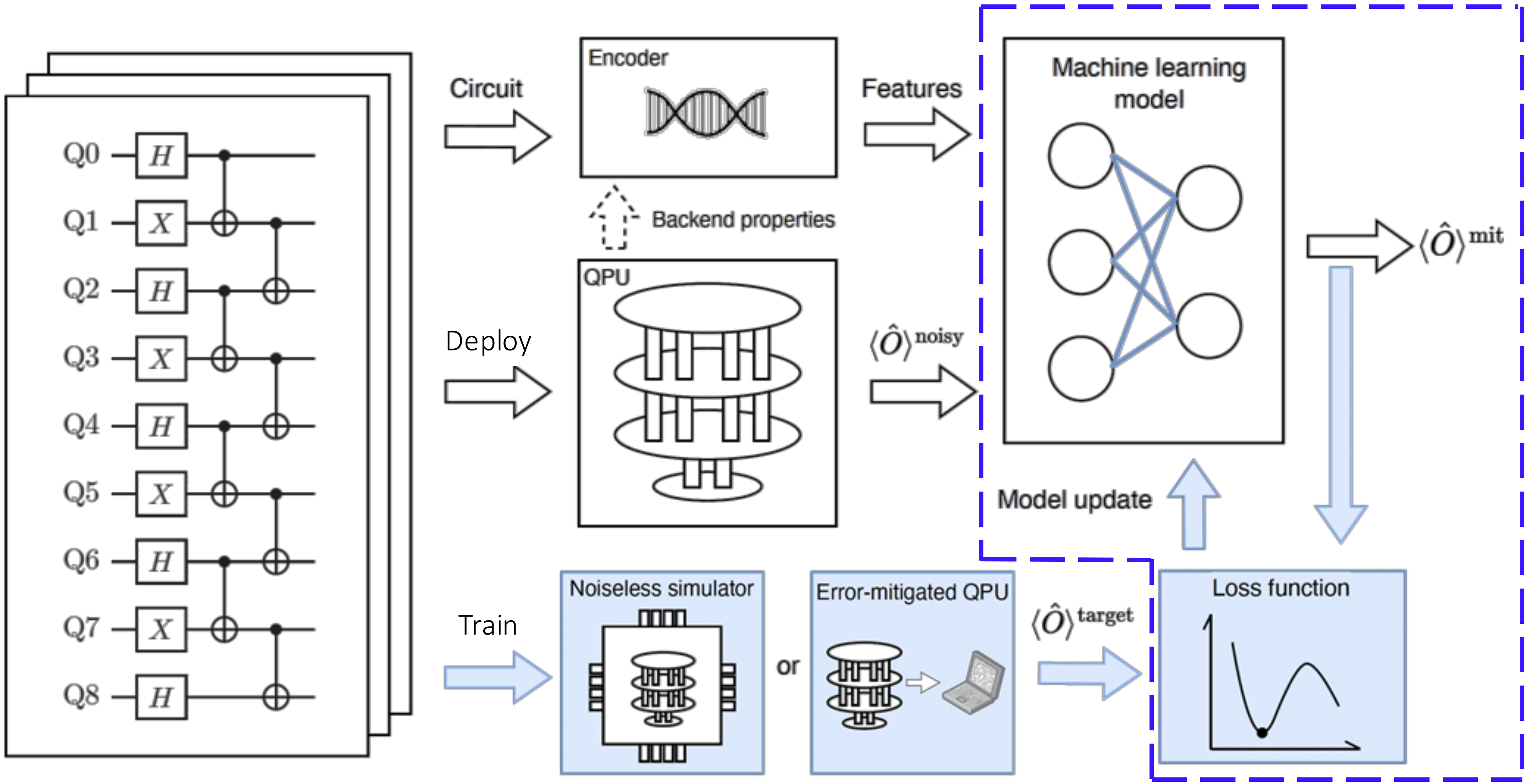
Workflow of ML-QEM

1. Generate training data



Nature Machine Intelligence v. 6, p. 1478 (2024)

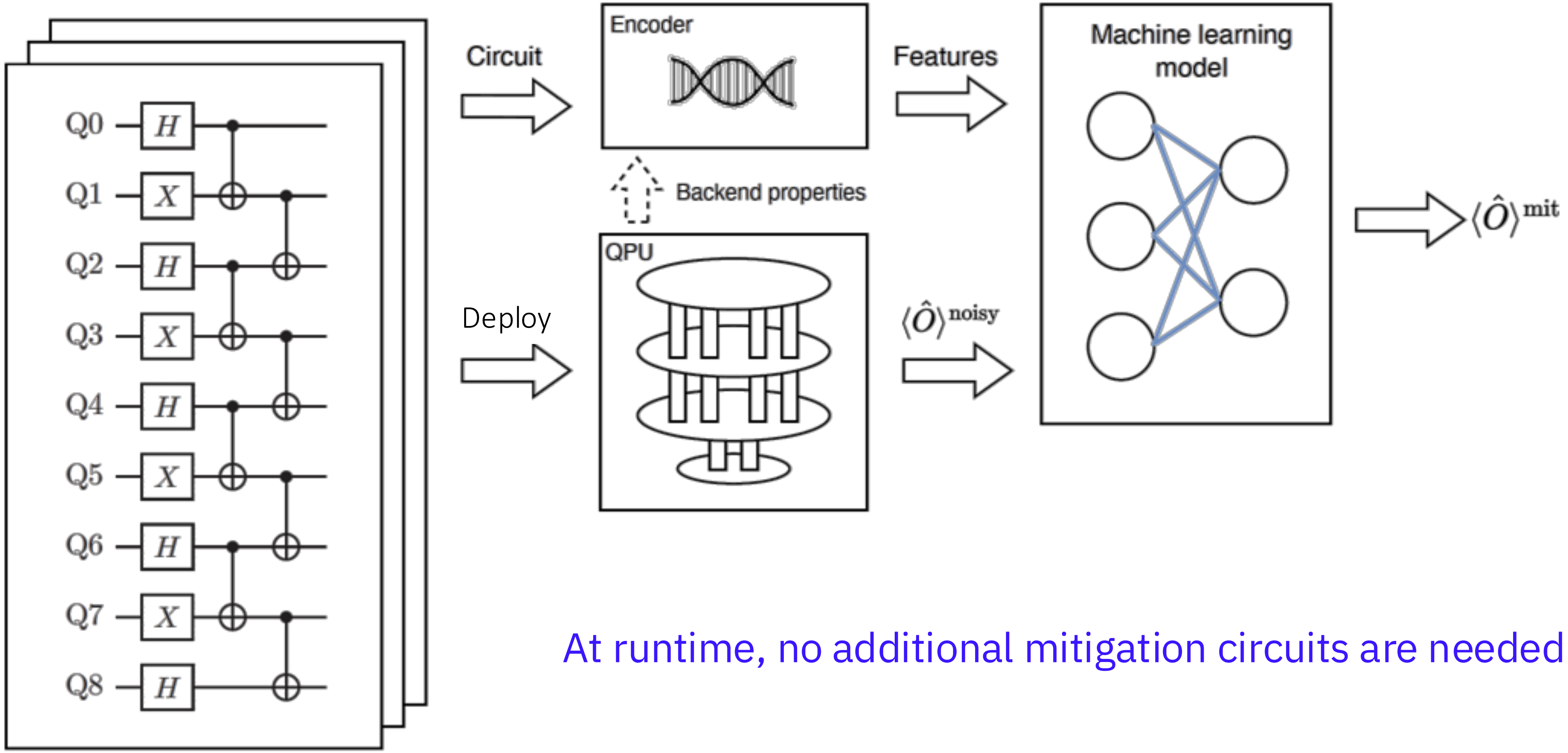
Workflow of ML-QEM: Train



Nature Machine Intelligence v. 6, p. 1478 (2024)

Workflow of ML-QEM: Deploy

3. Mitigate noisy expectation values at runtime



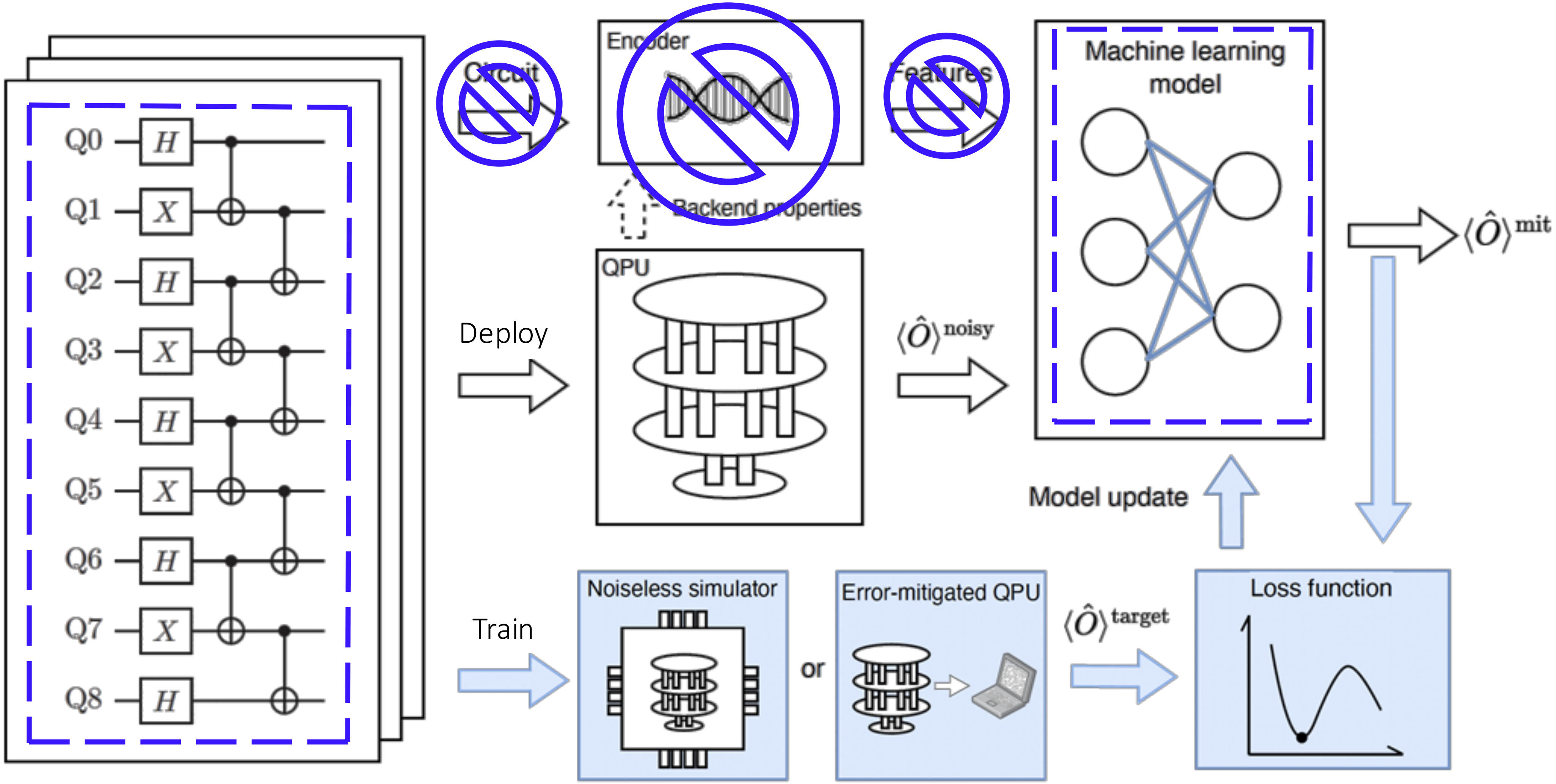
At runtime, no additional mitigation circuits are needed

Nature Machine Intelligence v. 6, p. 1478 (2024)

Relation to Clifford data regression (CDR)

Clifford data regression (CDR)* generates near-Clifford circuits constructed for only a single-expectation value for training

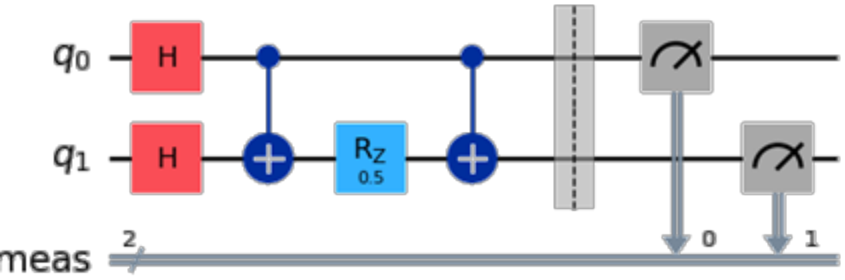
CDR uses only linear regression for mitigation



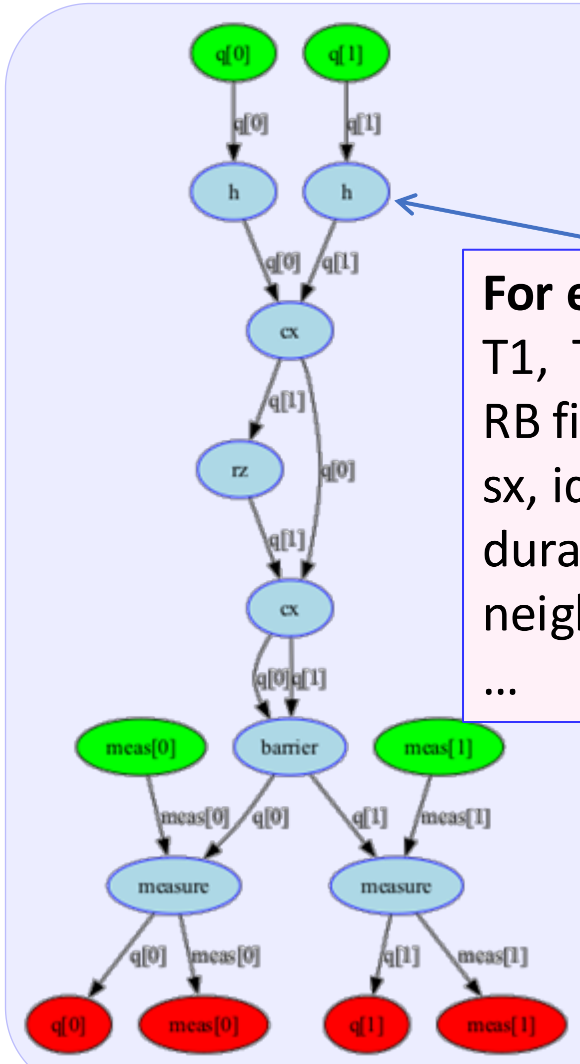
*P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Quantum 5, 592 (2021) and later variants

Encodings

Quantum circuit



Quantum circuit as a directed acyclic graph (DAG)



For each node
T1, T2E, T2R,
RB fidelity for
sx, id, … gate
duration,
neighbors,
…

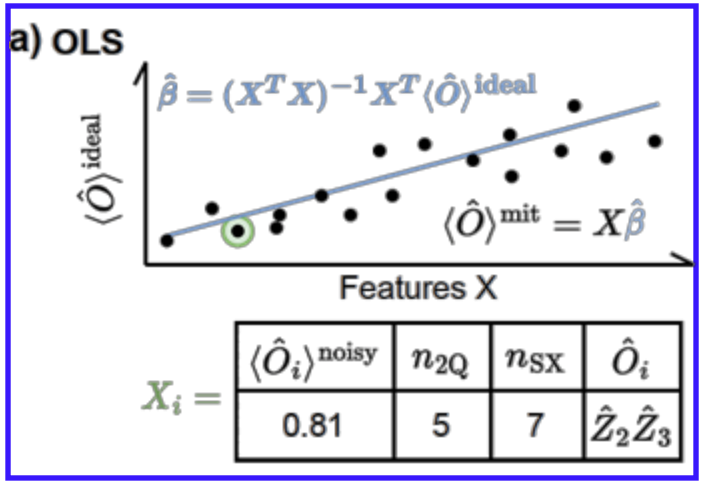
Feature vector

Circuit properties
number of qubits
number of cx gates
number of sx gates
number of id gates
number of rx gates
histogram of rotation angles
measurement basis
…

Device properties
T1 times
T2 times
readout error for each qubit
…

Benchmark performance of variety of ML models

Linear

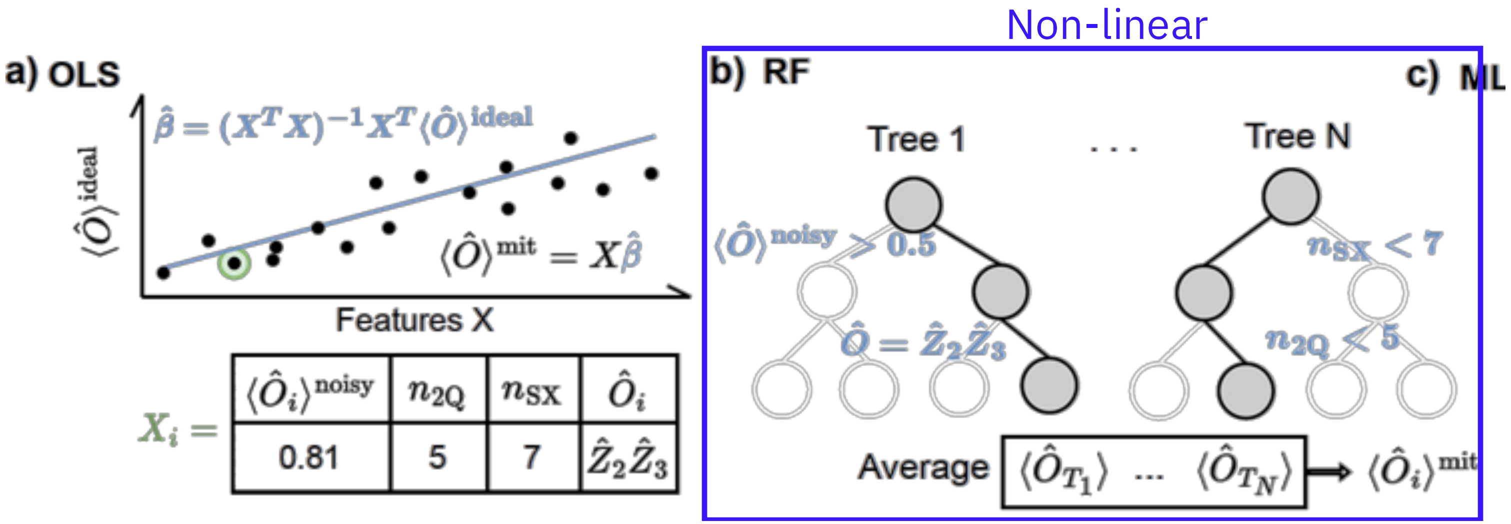


a) Ordinary least square (OLS)

Benchmark performance of variety of ML models

a) Ordinary least square (OLS)

b) Random forest regression (RF) [new]

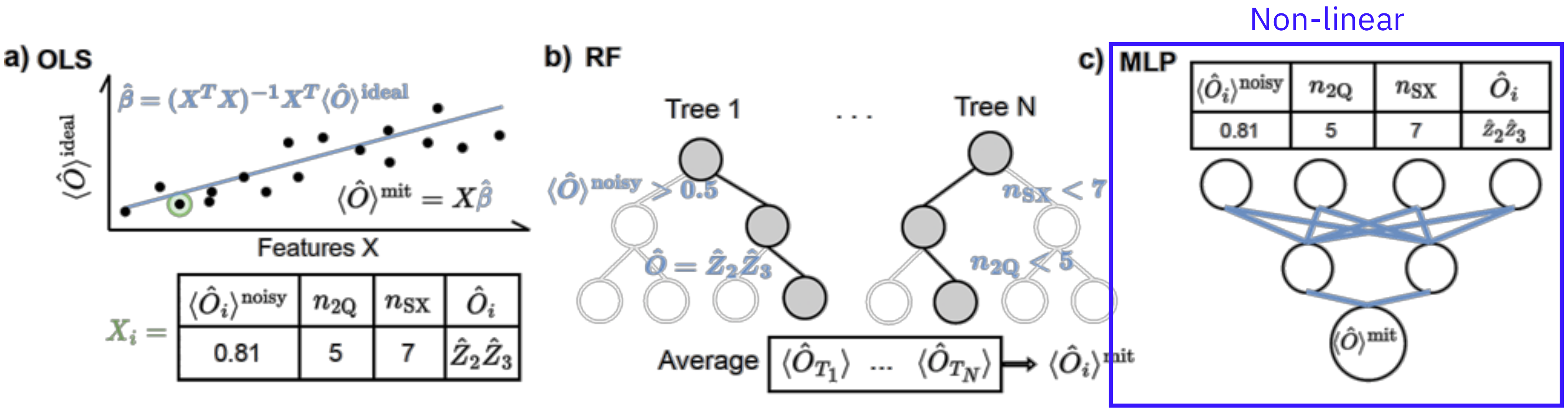


Benchmark performance of variety of ML models

a) Ordinary least square (OLS)

b) Random forest regression (RF)

c) Multi-layer perceptron (MLP)



Nature Machine Intelligence v. 6, p. 1478 (2024)

PyTorch

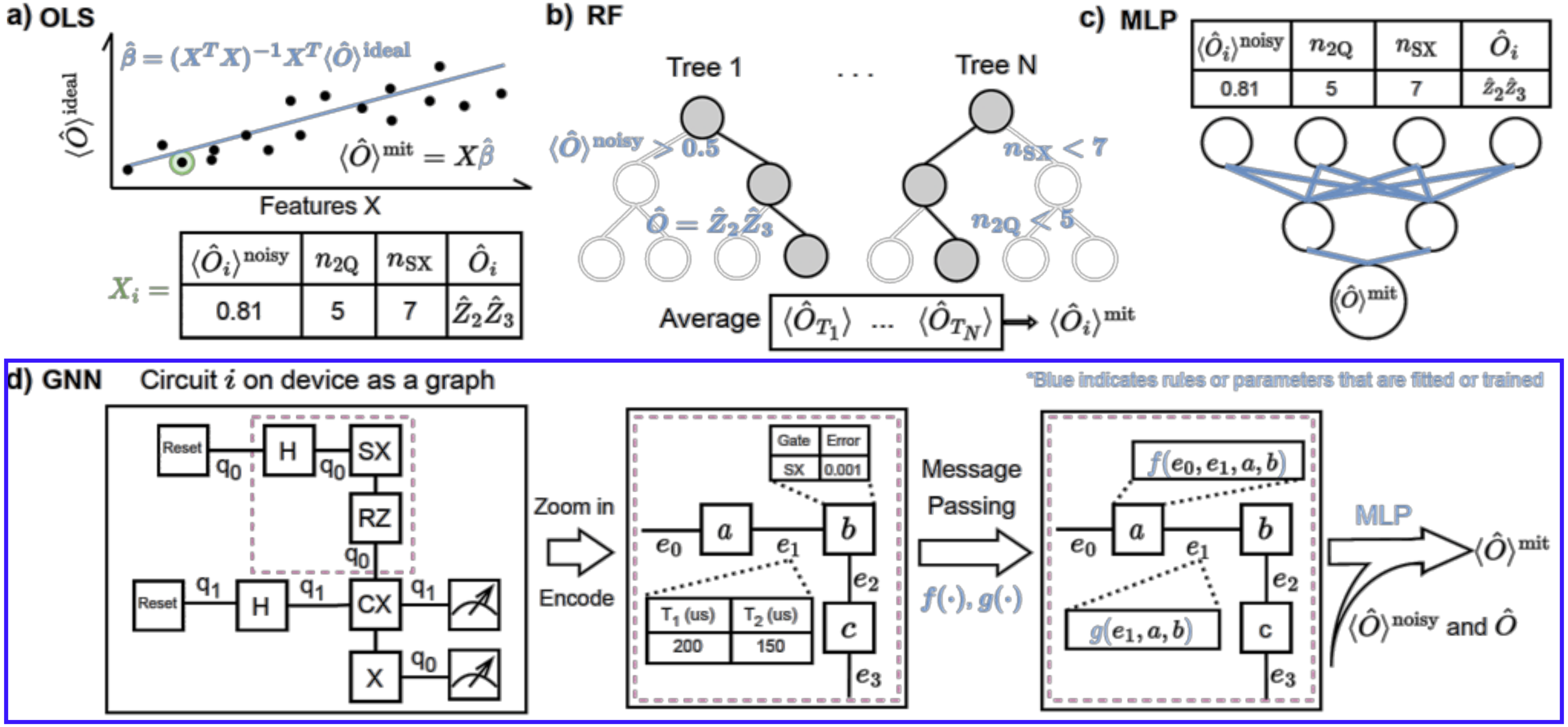
Benchmark performance of variety of ML models

a) Ordinary least square (OLS)

b) Random forest regression (RF) [new]

c) Multi-layer perceptron (MLP)

d) Graph neural network (GNN) [new]



Non-linear

Nature Machine Intelligence v. 6, p. 1478 (2024)

PyTorch

PyG (*PyTorch Geometric*)

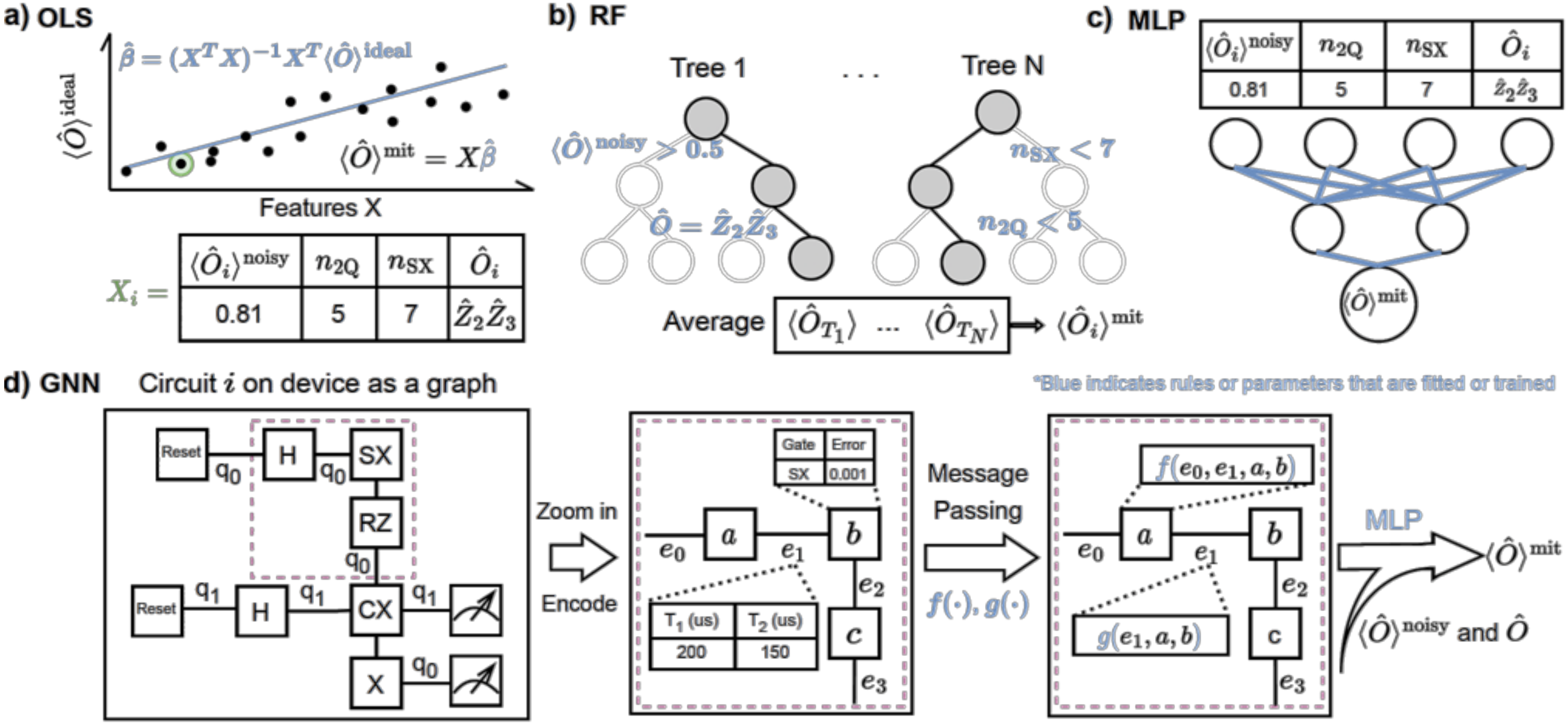
Benchmark performance of variety of ML models

a) Ordinary least square (OLS)

b) Random forest regression (RF) [new]

c) Multi-layer perceptron (MLP)

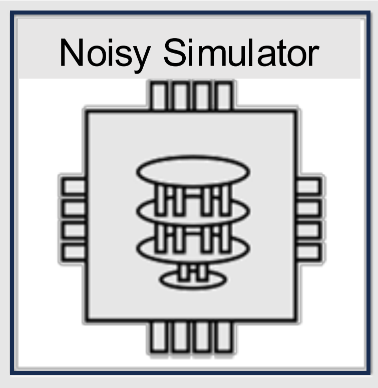
d) Graph neural network (GNN) [new]



Progression

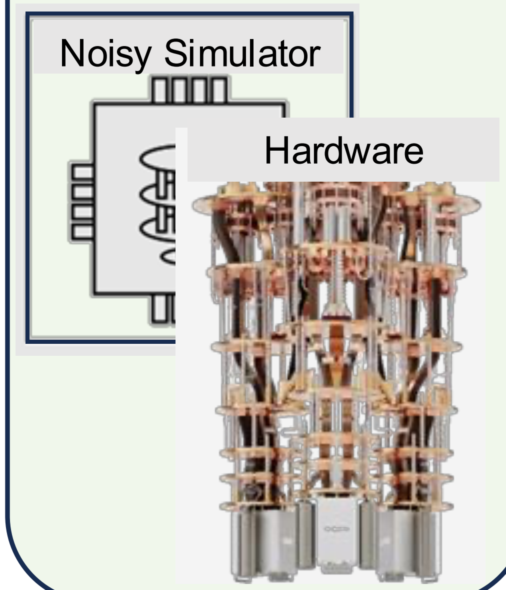
Random circuits

OLS, RF, MLP, GNN



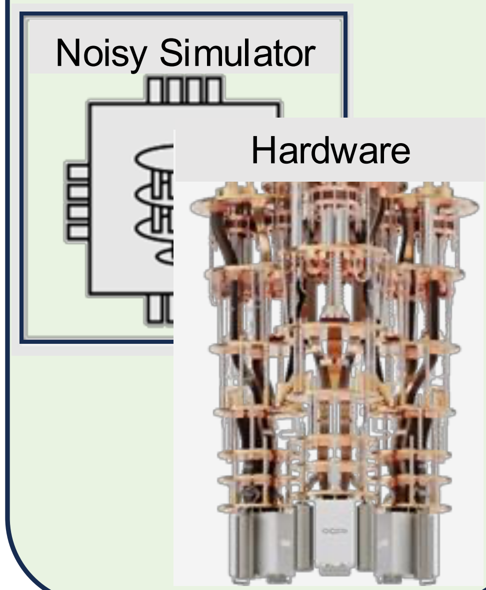
Structured circuits

Incoherent/coherent noise, readout error, ...



Use cases

Mitigating unseen Pauli observables, variational algorithms (e.g., VQE)



Scalability

Mimicry
Adaptability

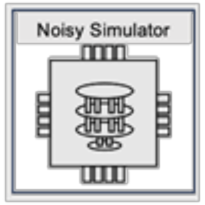
4 qubits
Classical sims available

100+ qubits
Classical sims not available

Nature Machine Intelligence v. 6, p. 1478 (2024)

Testing the waters

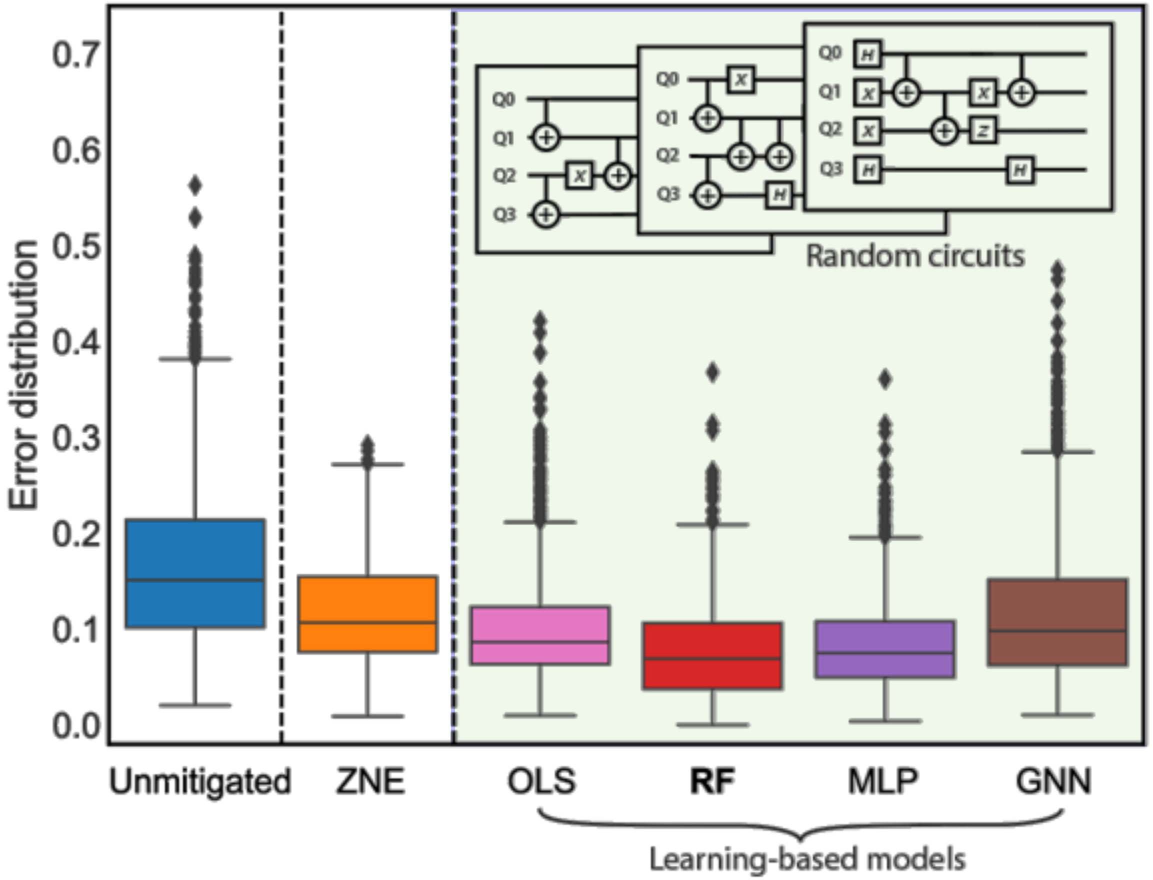
Start small: Performance on small random circuits

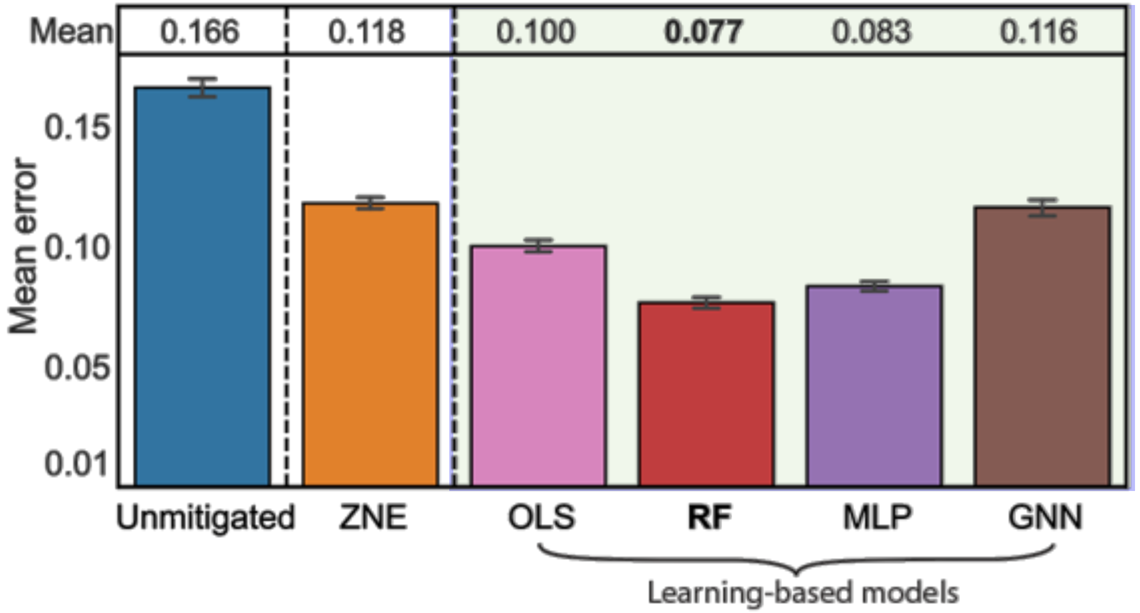


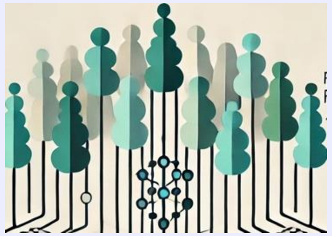
Circuits: 5,000 training and 2,000 benchmark random 4Q circuits, with 2Q gate depths from 0 to 18

Measure: All 2^4 single-Z observables

Execution: Simulation device-measured incoherent noise + SPAM errors. Ideal values computed.







ML-QEM outperforms digital ZNE

Random forest (RF) top performer

ZNE: Digital ZNE only, no readout error mitigation.

Performance on Trotterized circuits

<u>Circuits:</u>
First-order Trotterized 1D 4Q TFIM.

<u>Train:</u>
14 Trotter steps
300 circuits per step, each different J/h

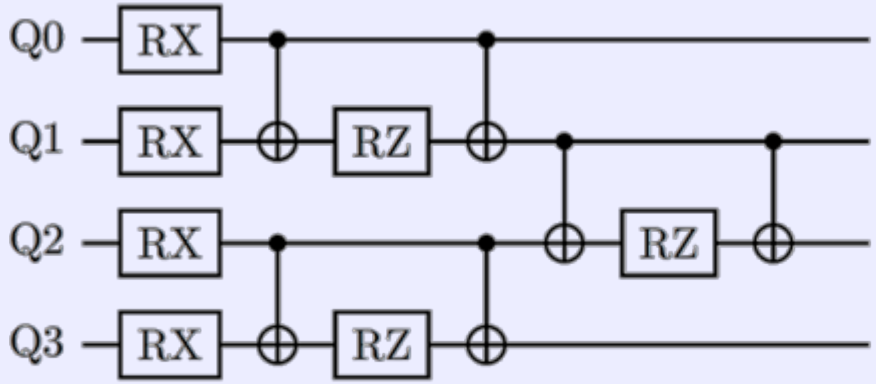
<u>Test:</u>
28 Trotter steps
300 circuits per step, each different J/h

<u>Measure:</u>
Random Pauli basis, weight-1 observables

Transverse-field Ising model (TFIM) in 1D

$$\hat{H} = -J\sum_j \hat{Z}_j\hat{Z}_{j+1} + h\sum_j \hat{X}_j = -J\hat{H}_{ZZ} + h\hat{H}_X$$

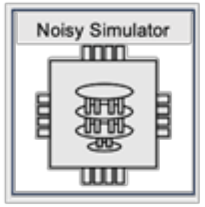
J: coupling strength h: transverse-field strength



Circuit for a single step in first-order Trotter time expansion

Performance on Trotterized circuits

Testing results in both interpolation and extrapolation regimes.

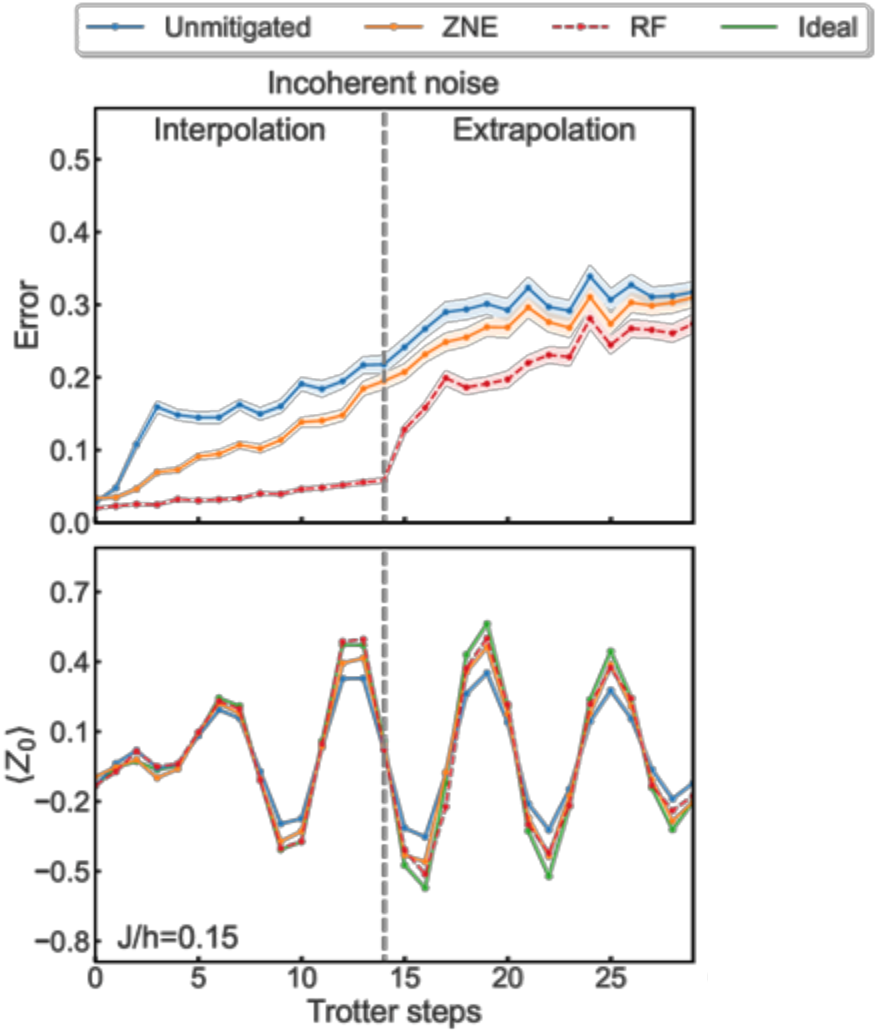


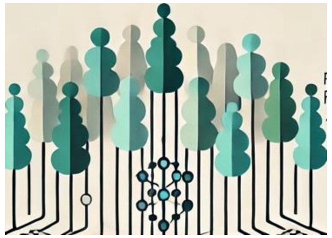
<u>Circuits:</u>
First-order Trotterized
1D 4Q TFIM.

<u>Train:</u>
14 Trotter steps
300 circuits per step,
each different J/h

<u>Test:</u>
28 Trotter steps
300 circuits per step,
each different J/h

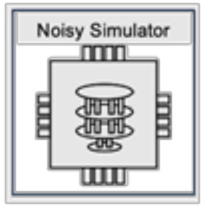
<u>Measure:</u>
Random Pauli basis,
weight-1 observables

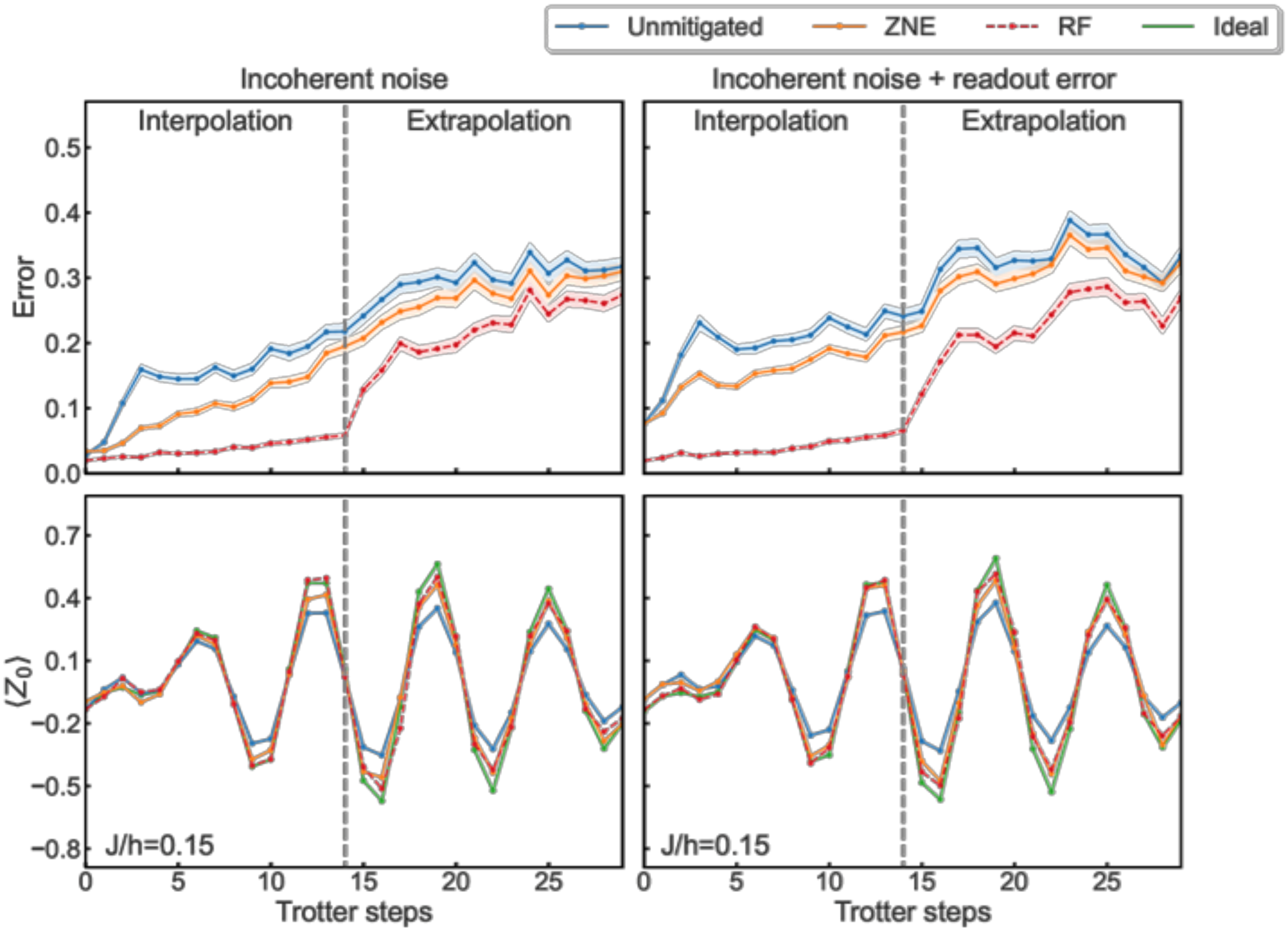




Performance on Trotterized circuits

Testing results in both interpolation and extrapolation regimes.





Circuits:
First-order Trotterized 1D 4Q TFIM.

Train:
14 Trotter steps
300 circuits per step, each different J/h

Test:
28 Trotter steps
300 circuits per step, each different J/h

Measure:
Random Pauli basis, weight-1 observables

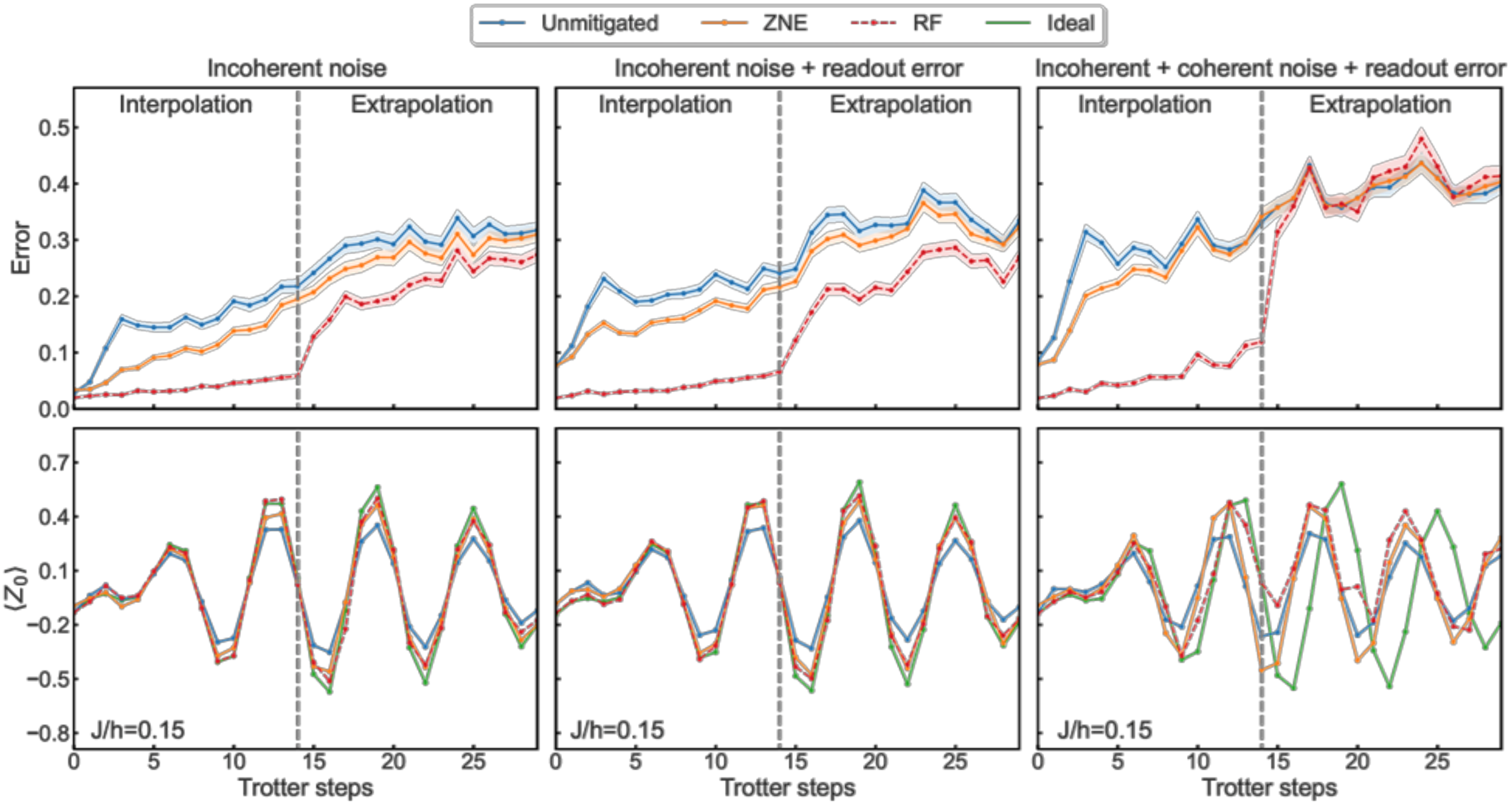
Performance on Trotterized circuits

Circuits:
First-order Trotterized 1D 4Q TFIM.

Train:
14 Trotter steps
300 circuits per step, each different J/h

Test:
28 Trotter steps
300 circuits per step, each different J/h

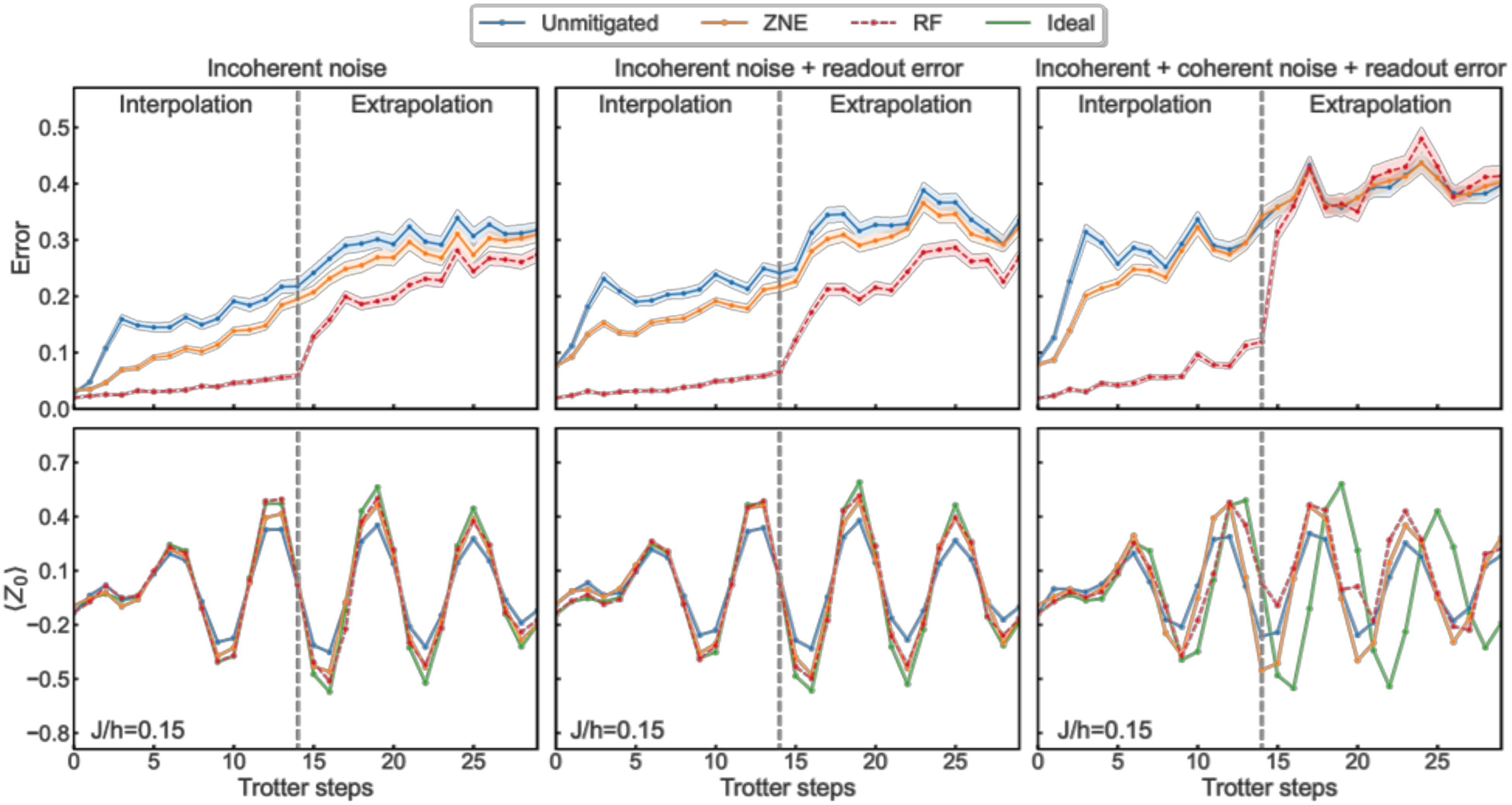
Measure:
Random Pauli basis, weight-1 observables



Performance on Trotterized circuits

ML-QEM with random forest outperforms digital ZNE

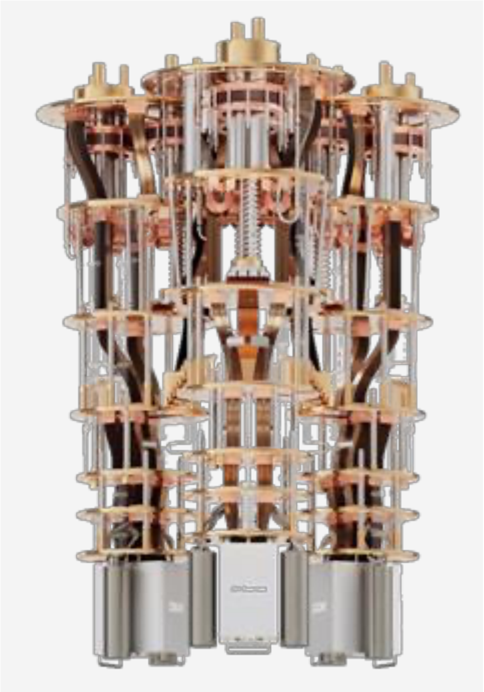
It wins even in extrapolation when there is no coherent error



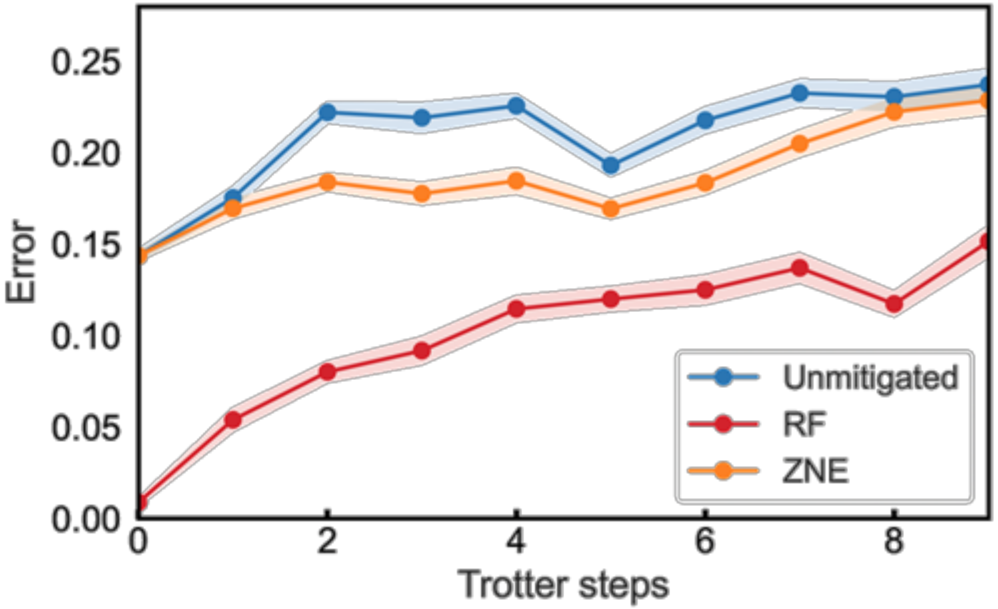
Nature Machine Intelligence v. 6, p. 1478 (2024)

Performance on Trotterized circuits on hardware

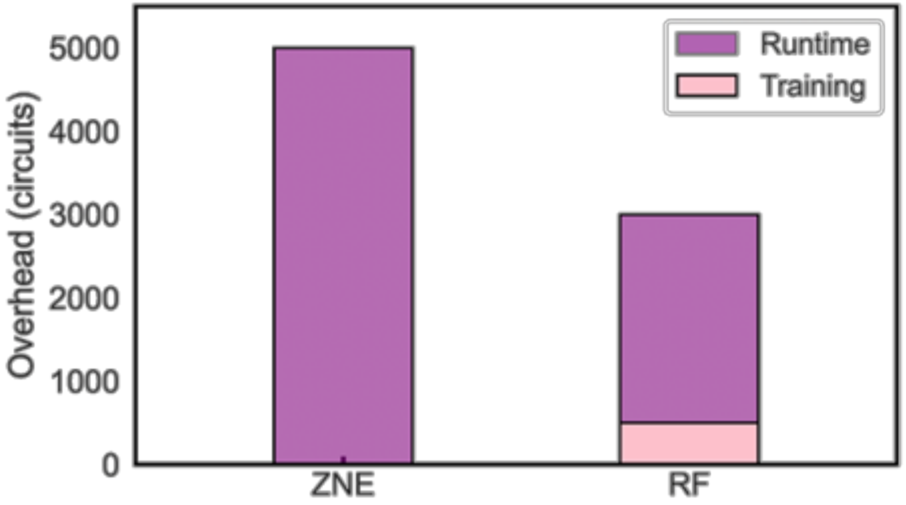
Results from
ibm_algiers



Circuits: 500 training, 2500 testing circuits
Hardness: We compute ideal expectation values.
Execution: Device (Subject to incoherent + coherent noise + readout error)



Compared to MVP ZNE with only 2 noise factors, RF has 50% lower runtime overhead, and 40% lower overall overhead



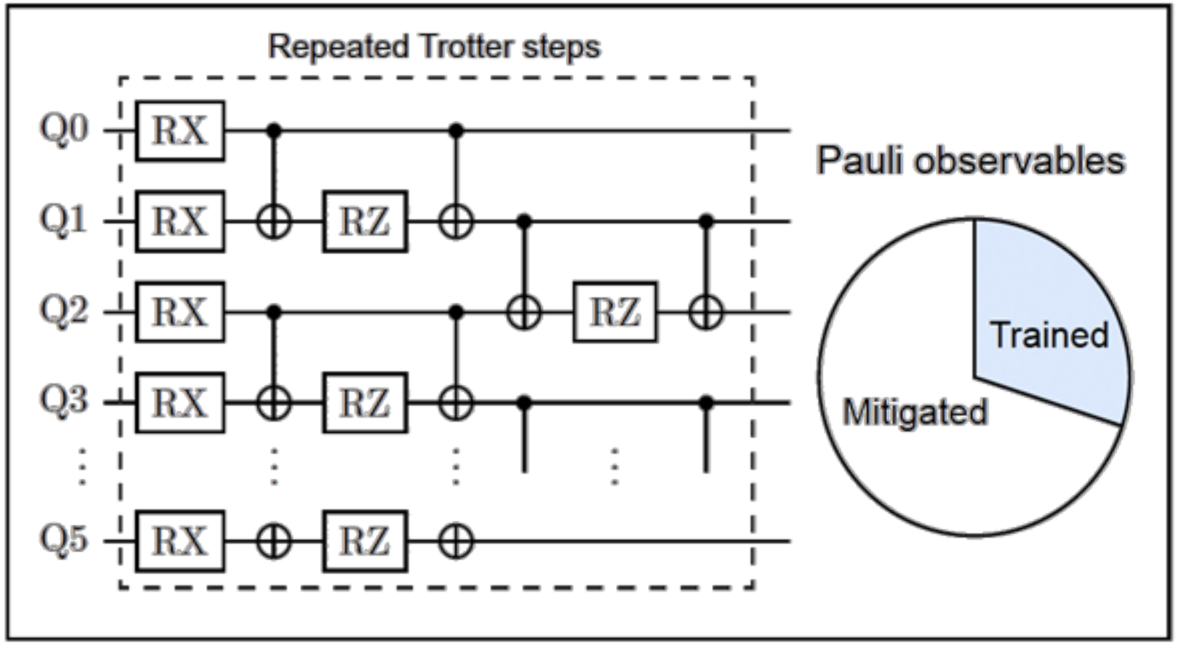
0.7 QPU hours to generate all training data*

ML-QEM with random forest outperforms digital ZNE on hardware

In addition to higher runtime efficiency, we observe higher overall efficiency

*Estimate based on 10k shots and a conservative sampling rate of 2 kHz:
Y. Kim, A. Eddins, S. Anand et al., Nature 618, 500–505 (2023)

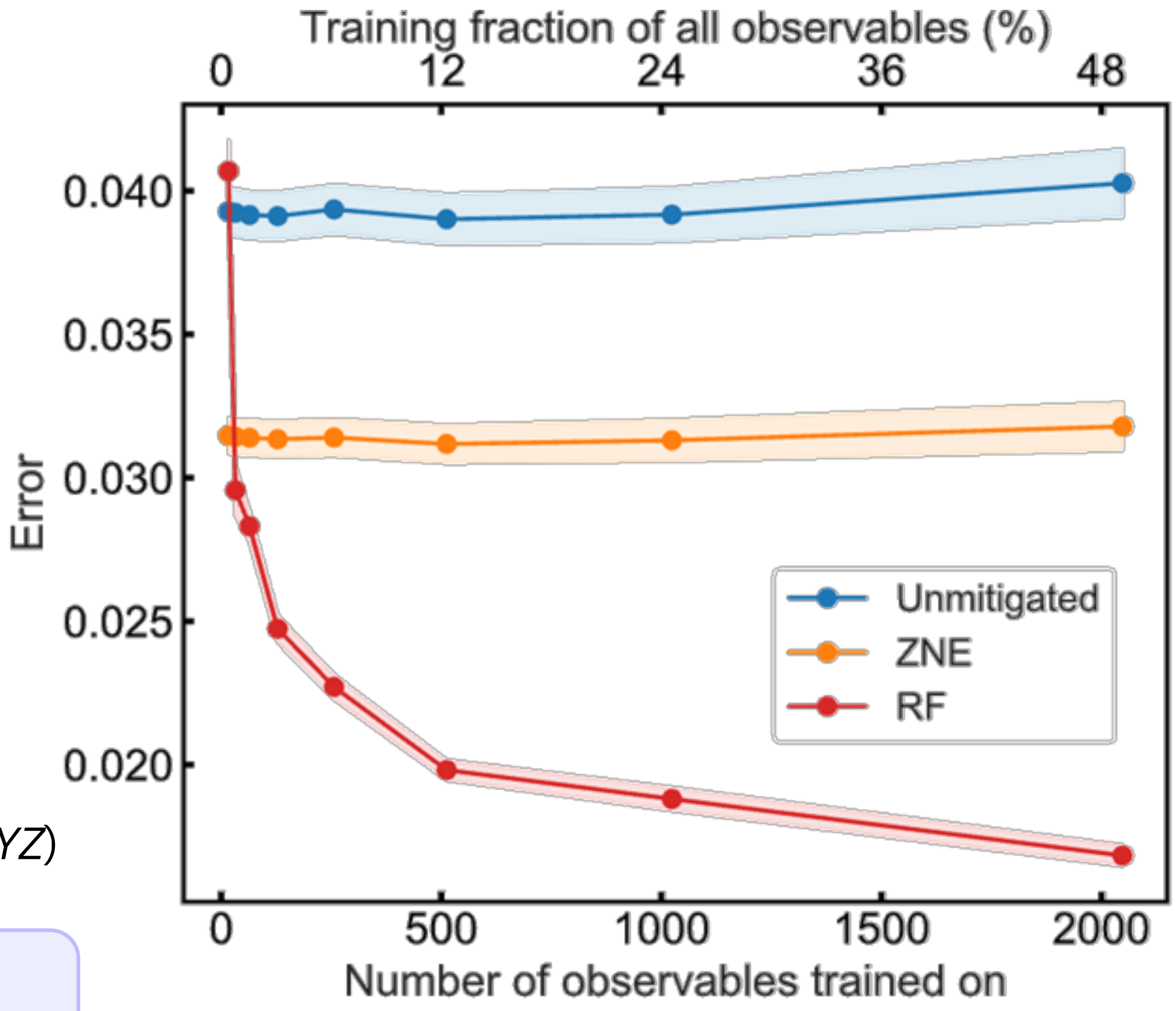
Mitigating observables not seen in training



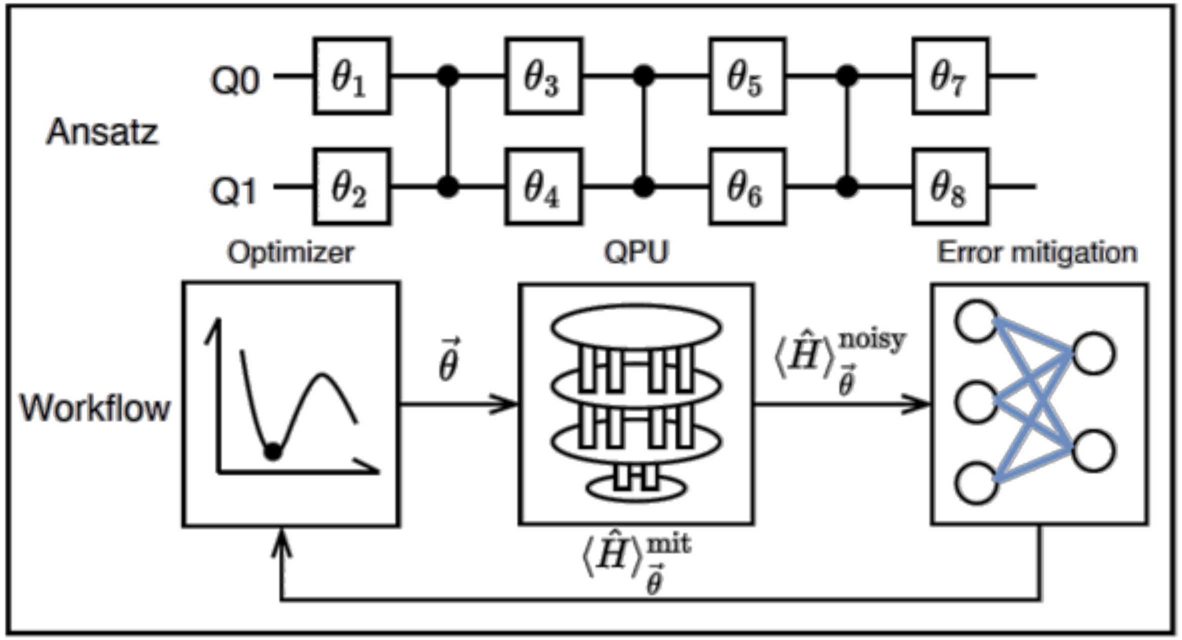
Train: on a portion of all Pauli observables (e.g., *XXIZYX*) measured on the 6Q 1D TFIM with 5 Trotter steps

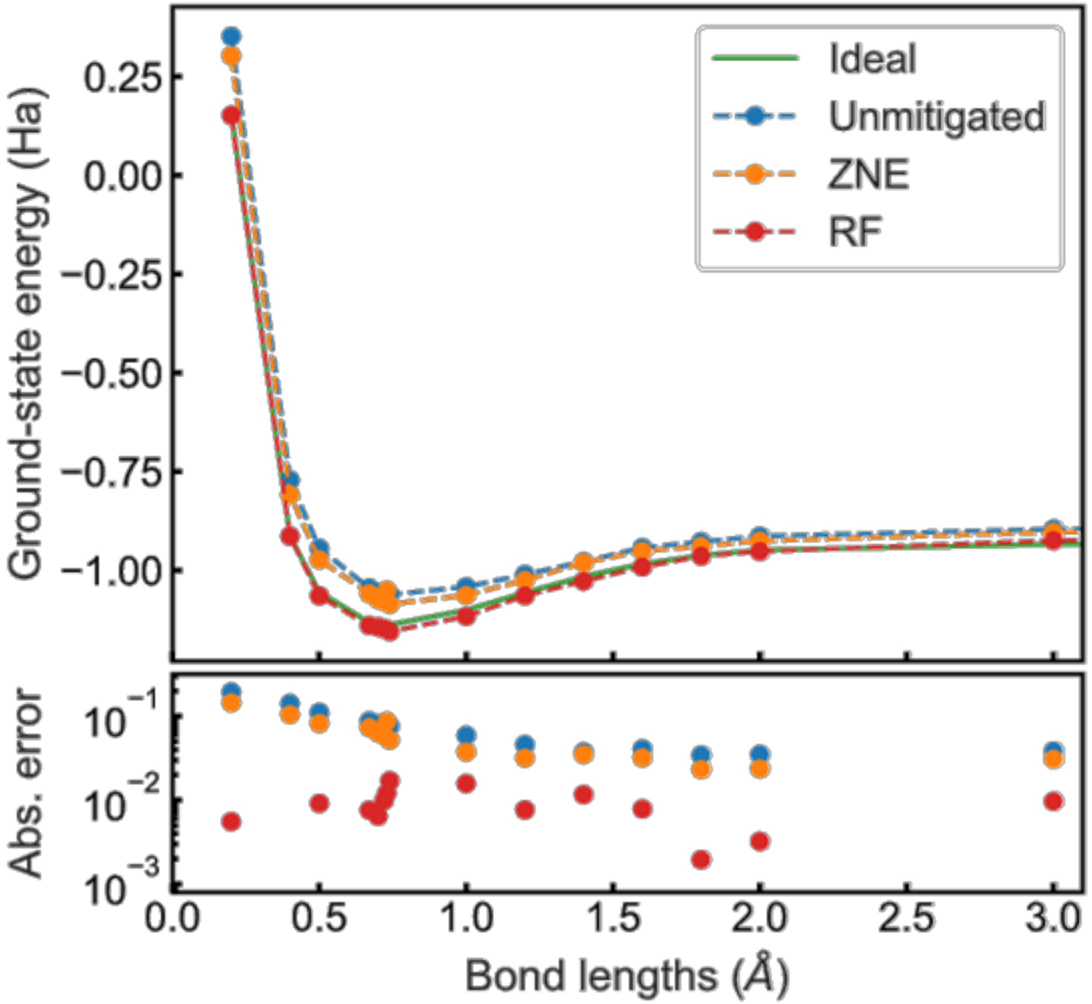
Test: on the remaining of the Pauli observables (e.g., *YIIXYZ*)

ML-QEM only needs to train on a very small fraction to outperform digital ZNE



Enhancing variational algorithms (e.g., VQE)





Train: on 2000 2Q two-local ansatz with randomly sampled parameters measured in either *XX* or *ZZ*

One trained model for all bond lengths

Run VQE to find the ground-state energy of H_2 molecule whose Hamiltonian is composed of *XX*, *ZI*, *IZ*, *ZZ* terms

ML-QEM outperforms digital ZNE

Scaling & Adaptability

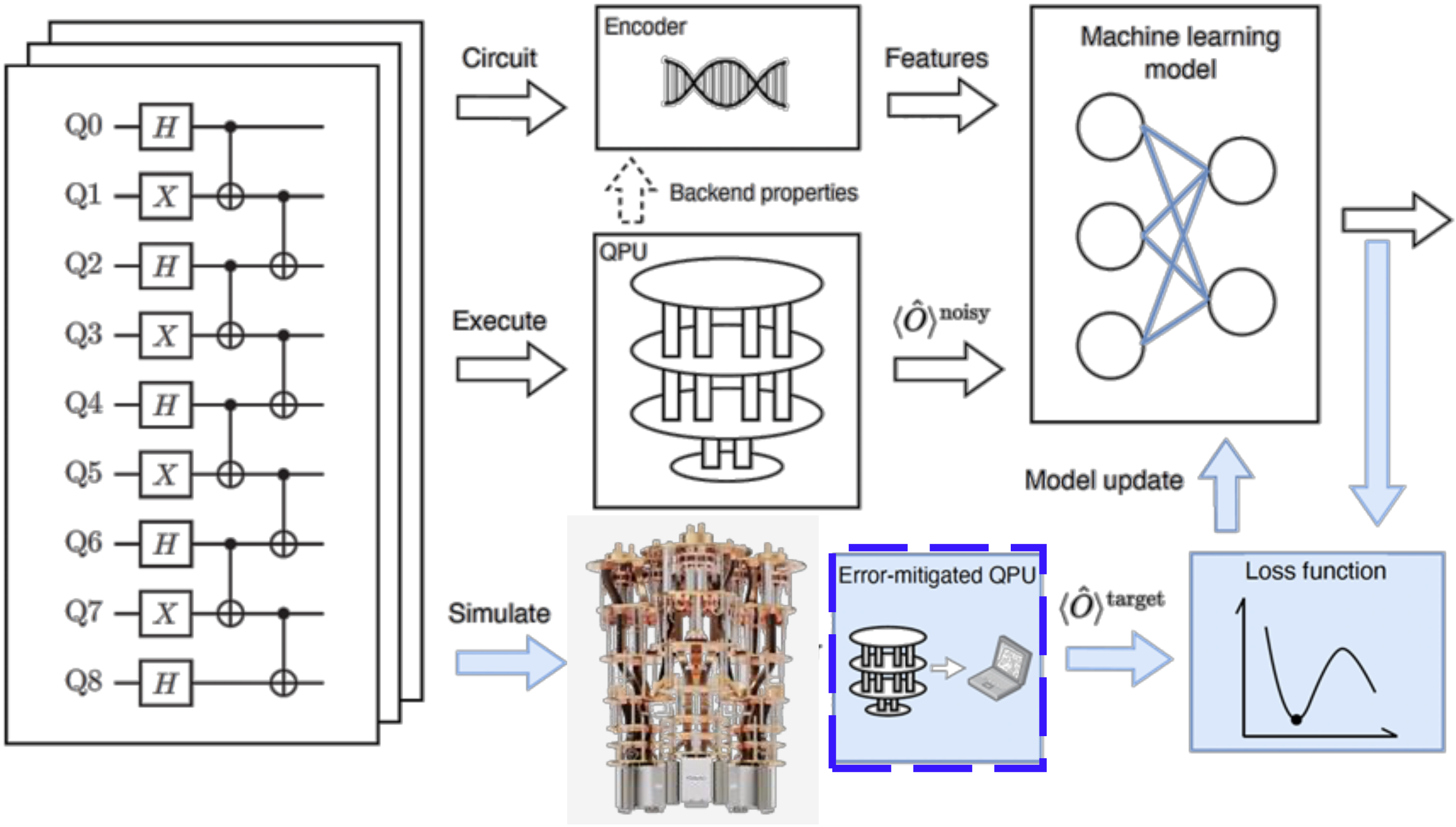
Scalability through mimicry

Setting:

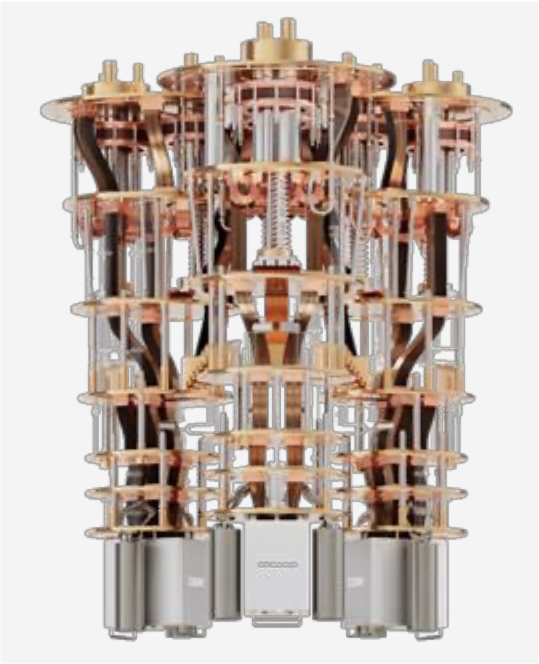
100Q+ circuits

No classical sims available

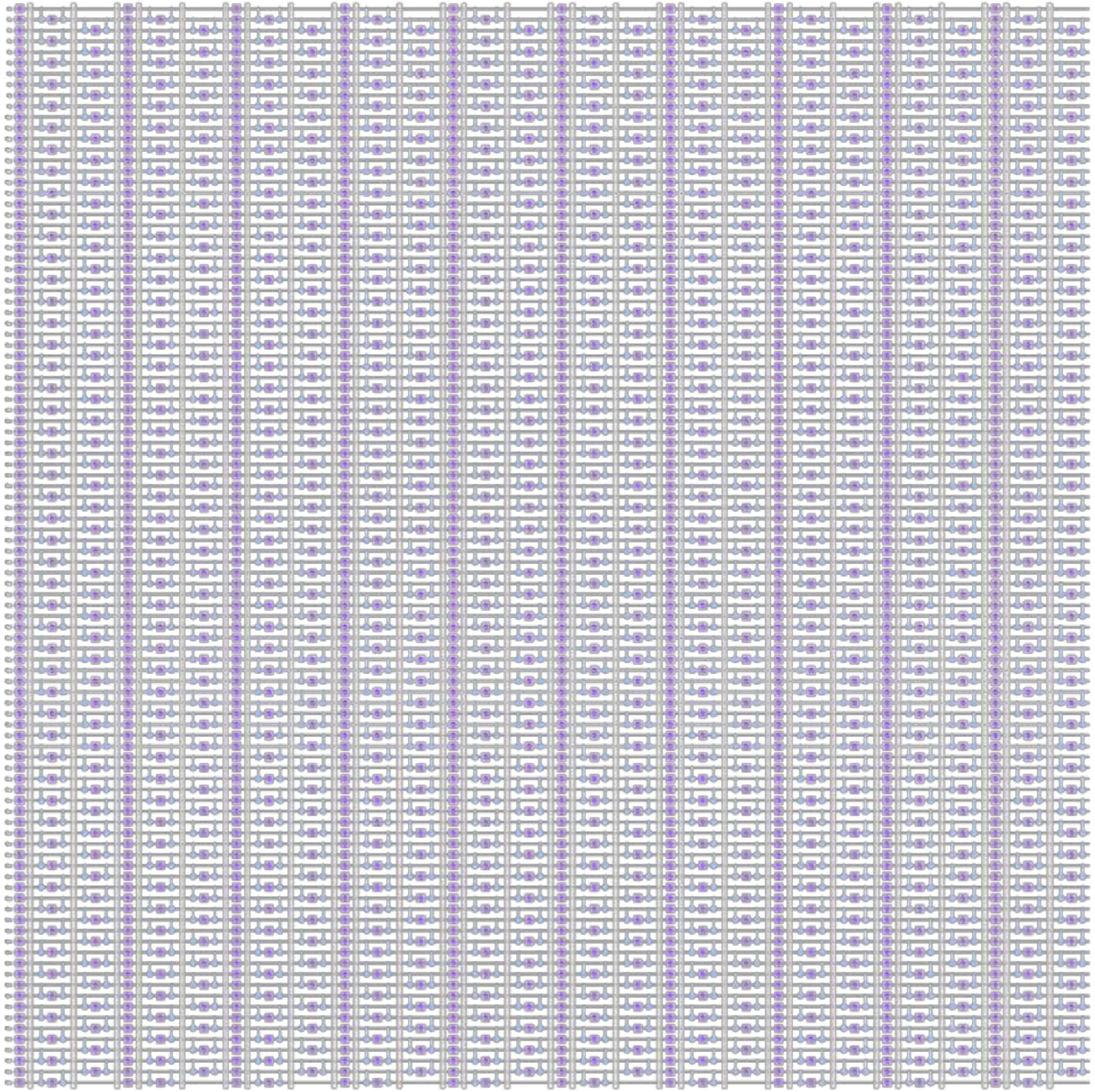
Aim:
Produce results matching gold-standard conventional method (e.g., ZNE, PEC, PEA) but with higher runtime efficiency.



Example: Mimicry on a 100-qubit experiment



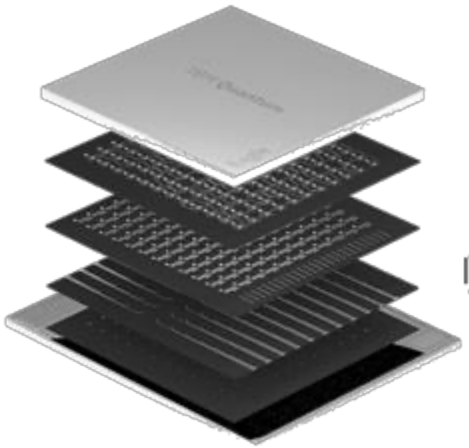
ibm_brisbane



<u>Circuits:</u> 100-qubit Trotterized 1D transverse-field Ising model (TFIM) experiment up to 10 Trotter steps (40 CNOT depth, ~2000 CNOTs) with twirling.

<u>Measure:</u> <Z> expectation values

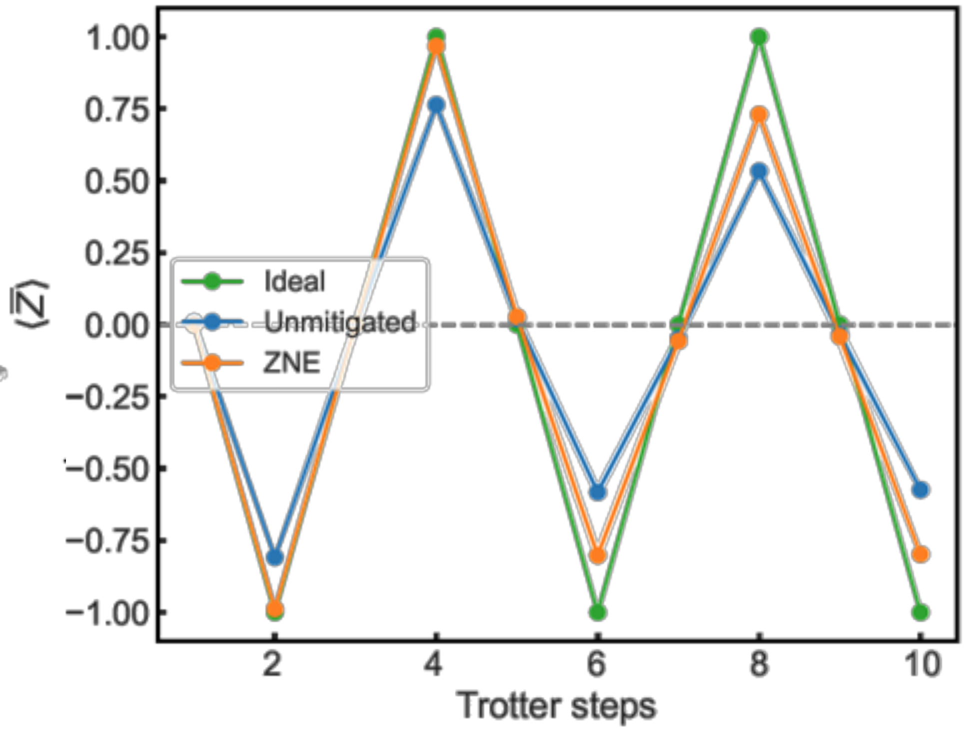
Mimicry on a 100-qubit experiment



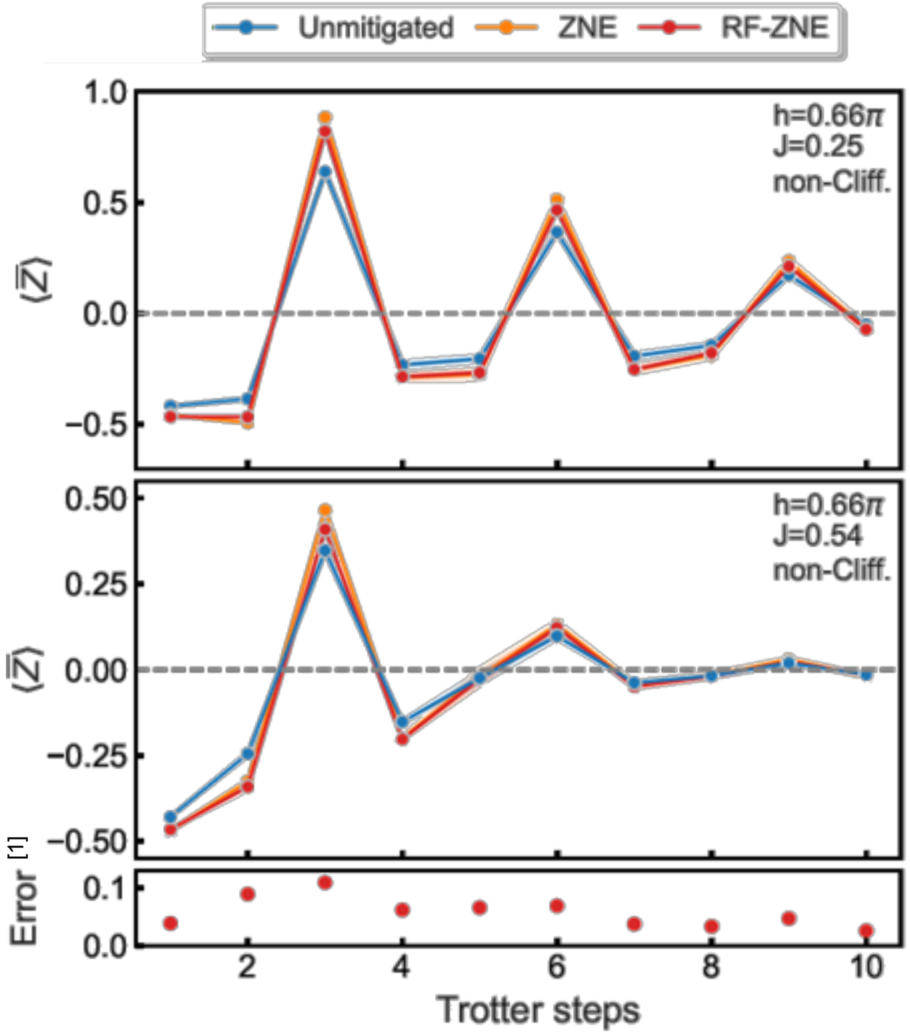
ibm_brisbane
(with twirling)

100 training, 400 testing circuits in total

Verification on Clifford
circuits (h=0.5π, J=0)

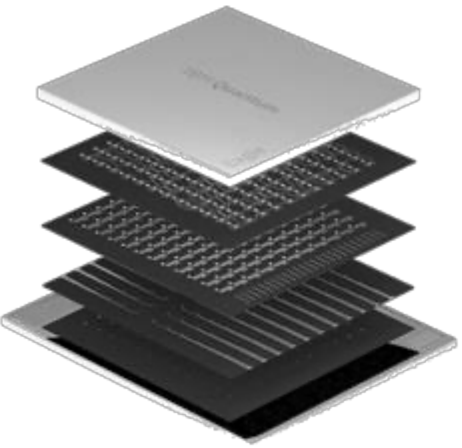


Mimicry on non-Clifford
circuits with different J/h



[1] Errors are between ZNE and RF-ZNE, also averaged over 40 different J/h at each step

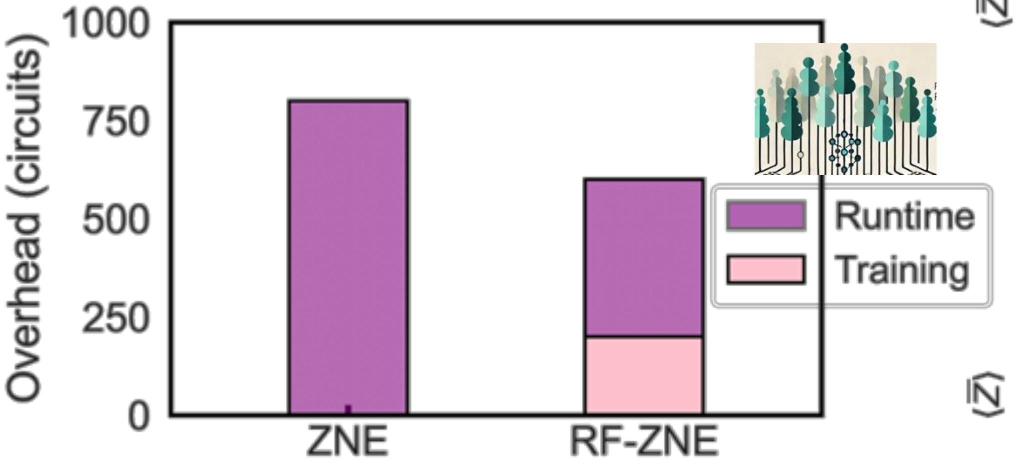
Mimicry on a 100-qubit experiment



ibm_brisbane
(with twirling)

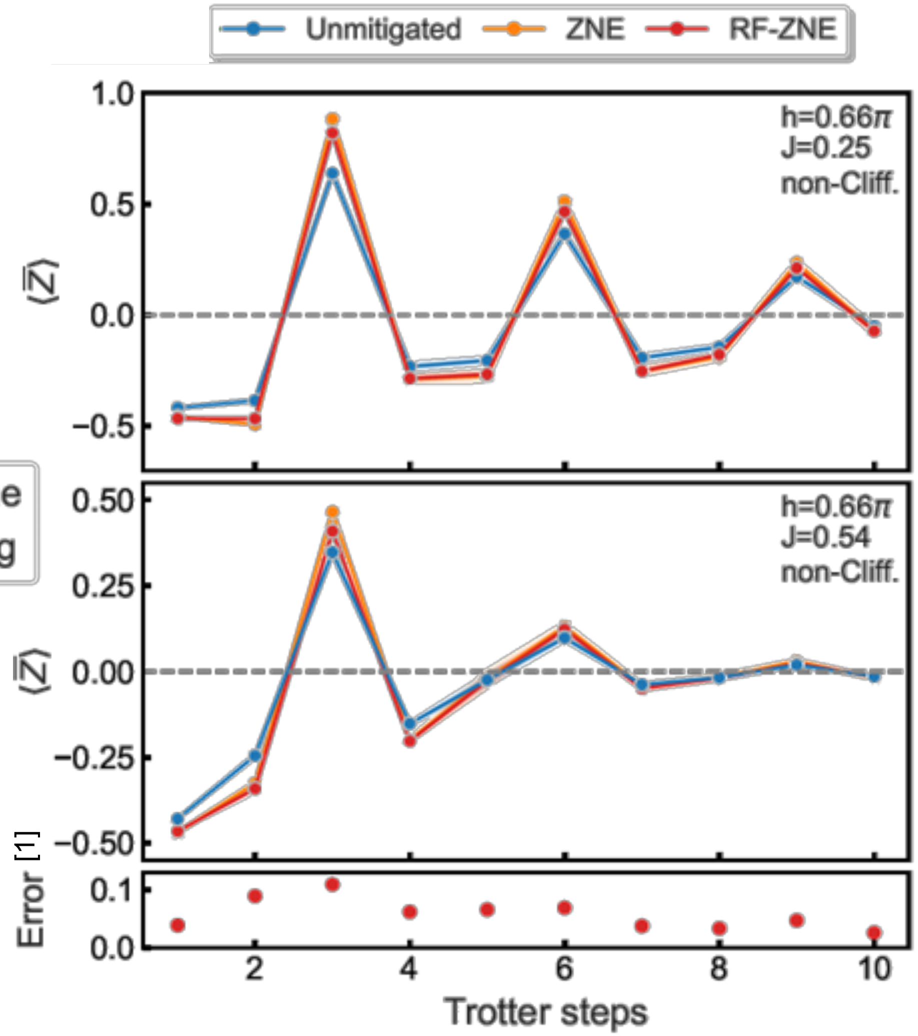
100 training, 400 testing circuits in total

Compared to ZNE with 2 noise factors,
Runtime overhead of RF-ZNE is 50% lower
Overall overhead of RF-ZNE is 25% lower



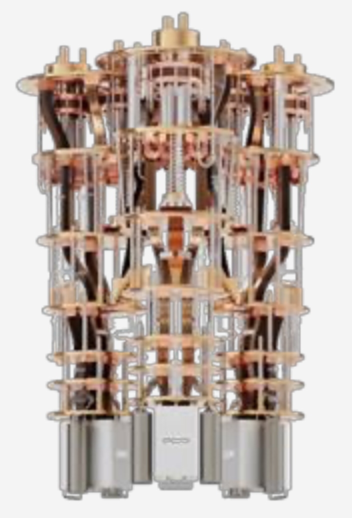
0.14 QPU hours to generate all training data
(excluding randomized compiling) [2]

Mimicry on non-Clifford
circuits with different J/h



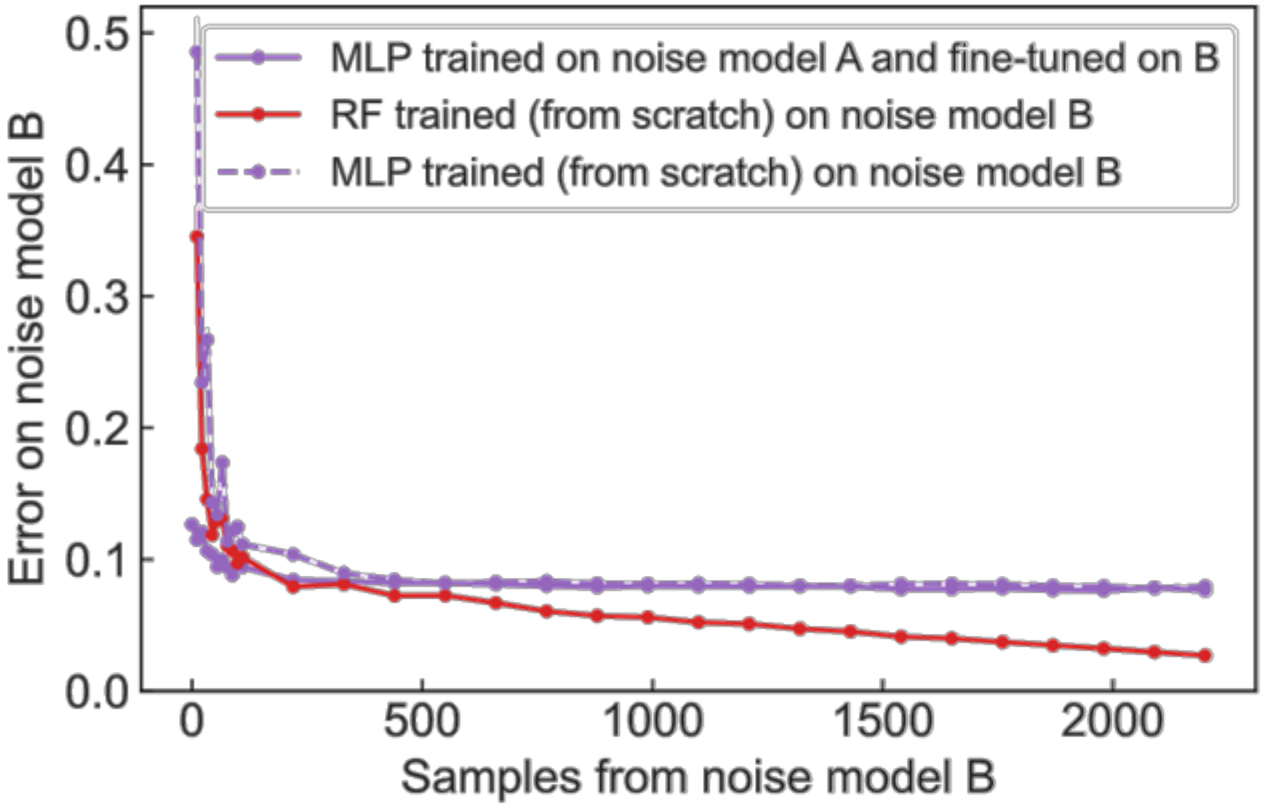
[2] Estimate based on 10k shots and a conservative sampling rate of 2 kHz: Y. Kim, A. Eddins, S. Anand et al., Nature 618, 500–505 (2023)
[1] Errors are between ZNE and RF-ZNE, also averaged over 40 different J/h at each step

Adapt to device noise drift



MLP can be fine-tuned on a different noise model after initial training, results in faster convergence than training new MLP from scratch

RF trained from scratch converges very quickly, demonstrates good efficiency in training an updated RF model

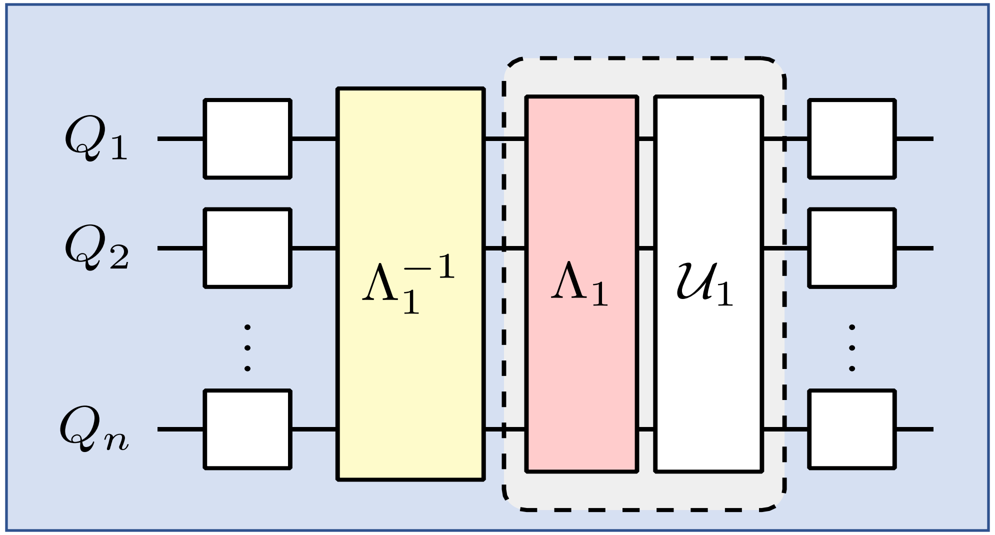


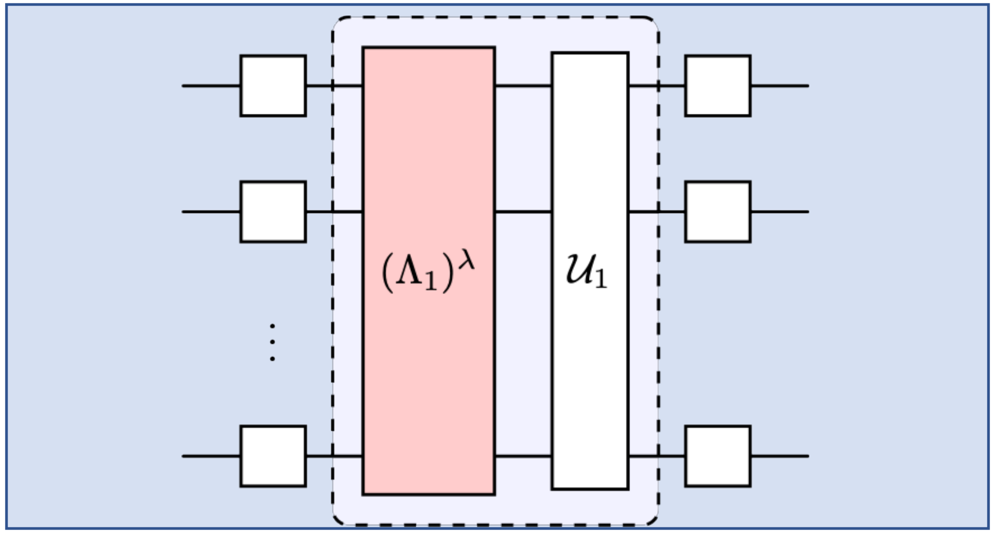
MLP and RF can efficiently adapt to time-varying noise models

Nature Machine Intelligence v. 6, p. 1478 (2024)

Why could ML-QEM be working?

Physics based methods





ML-QEM

Skip learning quantum noise

Complex general object; We care about the relationship of classical objects. Instead, show similar circuits.

Learn classical output-only relationships

Learn non-linear relationship b/w <O_noisy> and <O_target> + circuit features directly. Generically hard, but here simple.

Convergence of global noise

Experiment: Empirically, noise effect is to damp same-weight Paulis by very similar amount in most circuits

Theory: log(d) convergence for random circuits.*

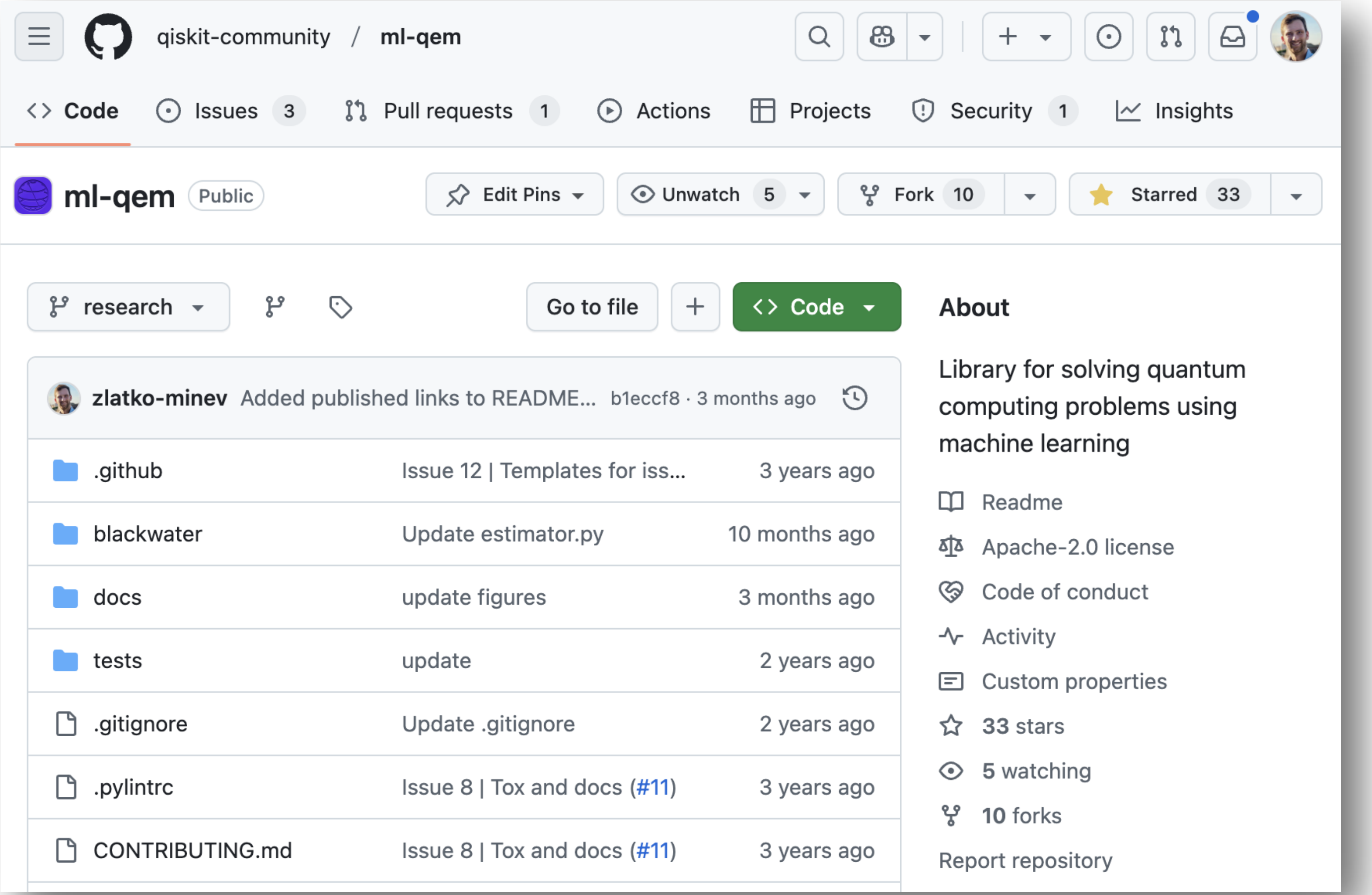
Tailor the noise

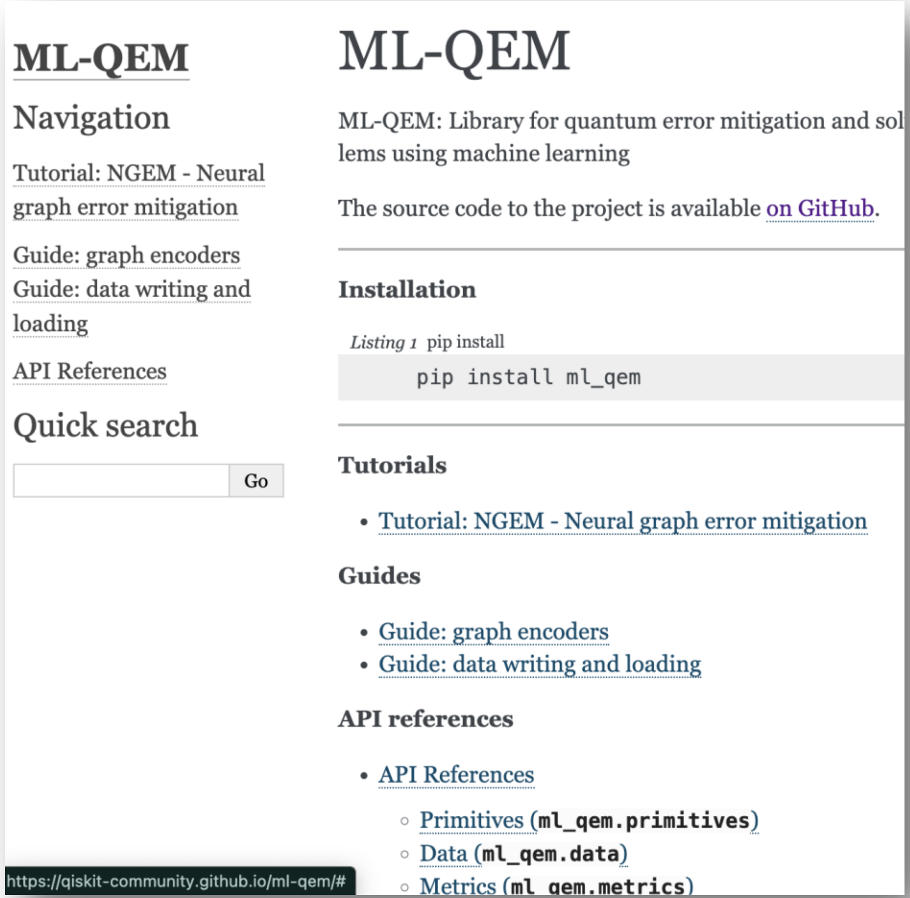
We can tailor & simplify the noise (twirl, DD, etc.)

* Dalzell, et al. Random quantum circuits transform local noise into global white noise, Commun. Math. Phys. (2024).

Try it yourself: Code

Code: [github.com/qiskit-community/ml-qem] **Docs:** [qiskit-community.github.io/ml-qem]





Conclusion: ML-QEM can significantly *accelerate* (reduce cost of) quantum error mitigation without sacrificing *accuracy* even at at utility *scale*.

General framework
(can test other ML and QEM models)

Applicable at utility scale (100Q+)

Best model: RF 

Adaptable to noise fluctuation

Cloud deployable

Limit set by conventional methods

Lots more to test and try!

Haoran Liao, Derek Wang, Iskandar Sitdikov, Ciro Salcedo, Alireza Seif, **Zlatko Minev**
Nature Machine Intelligence v. 6, p. 1478 (2024); IBM Quantum Team

AI compiling:
arXiv:2405.13196, arXiv:2407.21225S

Slides? zlatko-minev.com/blog @zlatko_minev

Machine learning for practical quantum error mitigation

Haoran Liao, Derek Wang, Iskandar Sitdikov, Ciro Salcedo, Alireza Seif, **Zlatko Minev**

Nature Machine Intelligence v. 6, p. 1478 (2024)

Acknowledgements: Brian Quanz, Patrick Rall, Oles Shtanko, Roland De Putter, Kunal Sharma, Barbara Jones, Sona Najafi, Thaddeus Pelligrini, Grace Harper, Vinay Tripathi, Antonio Mezzacapo, Christa Zoufal, Travis Scholten, Bryce Fuller, Swarnadeep Majumder, Sarah Sheldon, Youngseok Kim, Pedro Rivero, Will Bradbury, Nate Gruver, Minh Tran, Kristan Temme, and the broader IBM Quantum team

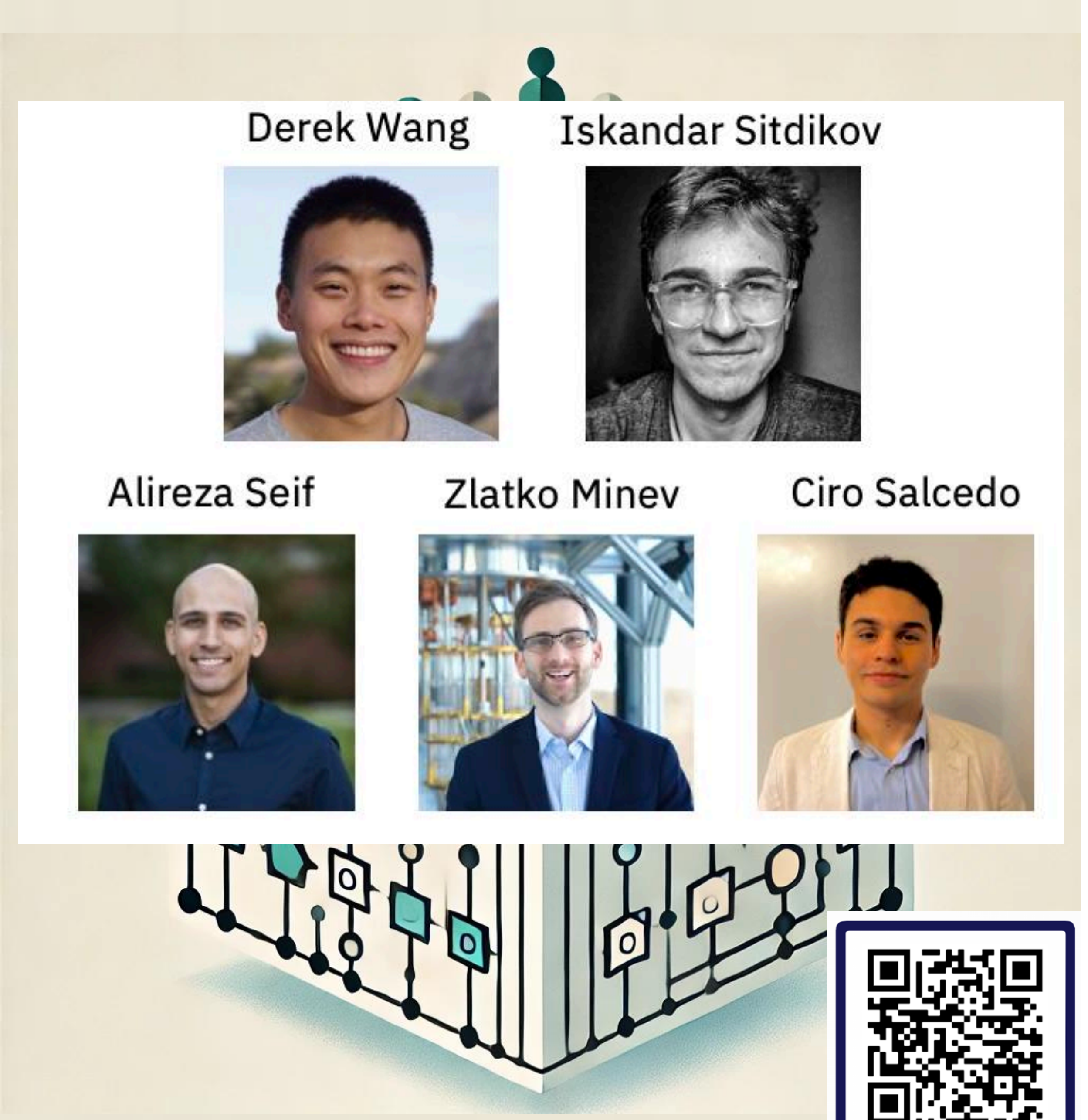


Illustration generated using AI (DALL·E)

Slides

IBM

Slides?
zlatko-minev.com/blog @zlatko_minev

EXTRA

Some future directions

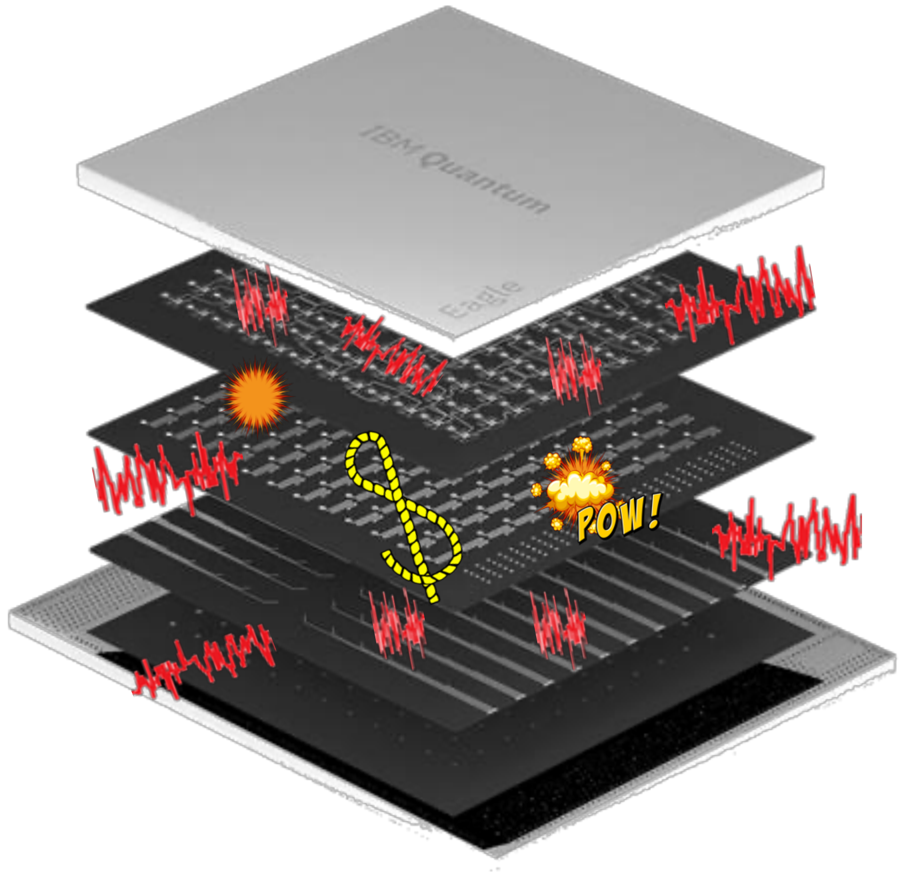
1. When trained with different noise parameters and their corresponding noisy expectation values, can we set all noise parameters encoded in GNN to the "good" values (i.e., set encoded gate error to 0, coherence times to infinity) to predict the ideal expectation values without ever training on ideal expectation values (and thus scalable)?

2. Can we instead regress the ideal and noisy expectation values of small-scale circuits onto the noise parameters to extract some noise parameters (e.g., avg. 2q gate error) of the backend, without running benchmarking experiments with additional quantum resources?

3. In ML-QEM, the training set can be optimized in terms of both size and type of the training circuits subject to design principles

4. ML-QEM can be more rigorously benchmarked against leading methods, such as PEC, PEA, and pulse-stretching ZNE

Error mitigation and error correction



Error mitigation: working with what you have

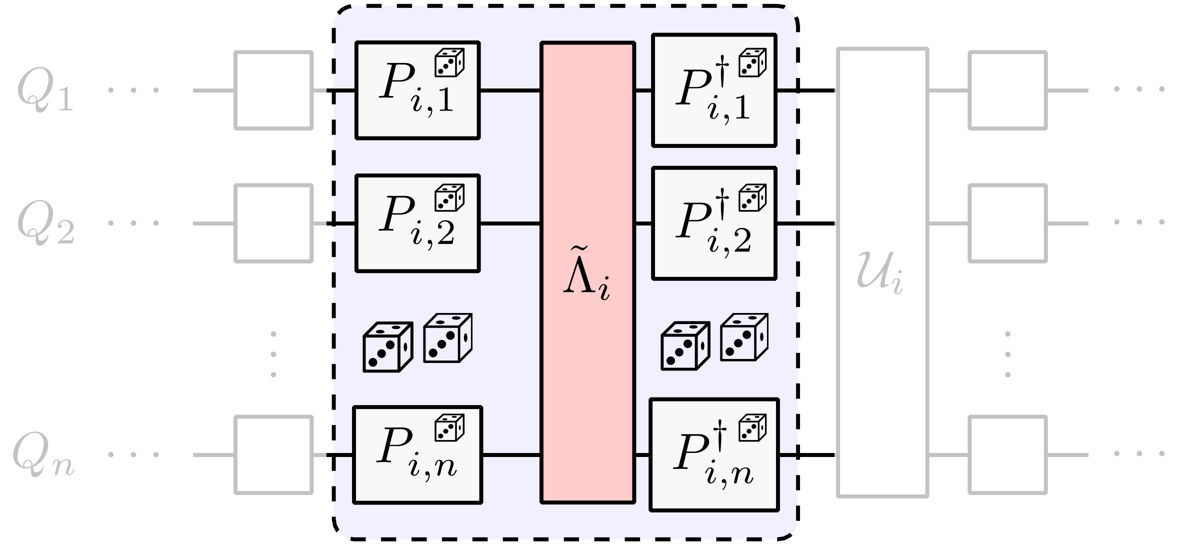
- **benefit** suppress errors on classical results (expectation values)
- **q-cost** no extra qubits or hardware resources needed
- **c-cost** trades classical resources (post-processing) for lower error
- **limitation** bad asymptotic scaling: high number of samples & circuits

Error correction: protecting quantum information

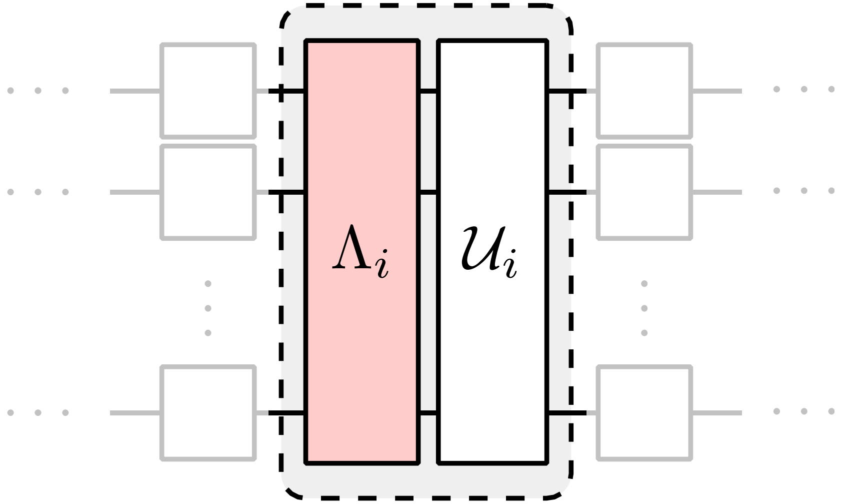
- **benefit** suppress & correct errors to arbitrarily small level
- **q-cost** very large qubit and hardware overhead
- **c-cost** decoding and encoding can be classically costly
- **challenge** requires fault-tolerant operations and readout

Simplify the noise: twirl

twirl reduces to noise $4^n \times 4^n$ matrix to diagonal one with 4^n entries in Pauli basis



average over twirl instances*



(Stochastic) Pauli channel

Twirling references

1. C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, W. K. Wootters, et al., Phys. Rev. Lett. 76, 722 (1996).
2. E. Knill, arXiv:0404104 (2004).
3. O. Kern, G. Alber, D. L. Shepelyansky, EPJ D 32, 153 (2005).
4. M. R. Geller, Z. Zhou, Physical Review A 88, 012314 (2013).
5. J. J. Wallman, J. Emerson, Physical Review A 94, 052325 (2016)
6. Hashim *et al*., Phys. Rev. X 11, 041039 (2021)
7. Tutorial: zlatko-minev.com/blog/twirling (2022)
8. ...

Tutorial

SCAN ME