Comparing Simulated versus Hardware Training Data for Learning-Based Quantum Error Mitigation

Seyed Mohamad Ali Tousi (IEEE Student Member)
Vision-Guided and Intelligent Robotics Lab (ViGIR)
University of Missouri
Columbia, US
STousi@missouri.edu

G. N. DeSouza (IEEE Senior Member)
Vision-Guided and Intelligent Robotics Lab (ViGIR)
University of Missouri
Columbia, US
DeSouzaG@missouri.edu

Abstract*—Learning-based Quantum Error Mitigation (QEM) has emerged as a promising approach for reducing noise-induced bias on near-term Noisy Intermediate-Scale Quantum (NISQ) hardware. While simulated noise models enable scalable training of learning-based mitigators, real hardware data more accurately reflect device-specific error mechanisms, but are expensive and sparse. Understanding how these two data regimes interact is therefore critical for the practical deployment of learning-based QEM. In this paper, we present a systematic and empirical analysis on the nature of the data (simulated versus hardware noise) for learning-based QEM. Using a recently developed attention-based graph transformer for small- and large-scale quantum error mitigation, we compare various training regimes, starting from exclusively simulated noise and progressively integrating more measurements from real quantum hardware. Our study spans two distinct circuit families: 1) randomly generated circuits with relatively homogeneous structure; and 2) targeted quantum circuits derived from a specific problem, in this case the one-dimensional Burgers' equation. Across five-fold cross validation, we analyze how hardware data affects generalization performance as measured by mean absolute error (MAE). Our results demonstrate that the impact of hardware data strongly depends on circuit structure and diversity. For random circuits, simulated data already guarantees lower absolute error, while hardware data yields consistent, but moderate improvements. In contrast, for targeted circuits, like the structured Burgers' circuits, simulated noise can be less representative of the actual (testing) quantum error, and incorporating hardware data produces substantially larger relative reductions in MAE – despite the intrinsically higher MAE. These findings indicate the value of the nature of the training data in learning-based QEM, which is governed not only by the data's ability to approximate reality – as expected – but also by how difficult it may be to capture circuit-induced error mechanisms in quantum simulations. In general, this work highlights the urgency of quantum circuit coverage in designing scalable and cost-effective learning-based QEM pipelines for devices in this new NISQ-era.

I. Introduction

Quantum computing has progressed rapidly over the past decade, driven by advances in superconducting qubits, trapped ions, photonics, and quantum control hardware. These developments have enabled the execution of increasingly complex quantum circuits on real devices, marking the era of *noisy intermediate-scale quantum* (NISQ) computing [1]. Despite notable experimental milestones and demonstrations of quantum advantage [2], [3], practical quantum computation remains fundamentally constrained by noise, decoherence, limited qubit counts, shallow circuit depths, and hardware-specific imperfections. Addressing these limitations without the overhead of full fault-tolerant error correction remains a central challenge for near-term quantum devices.

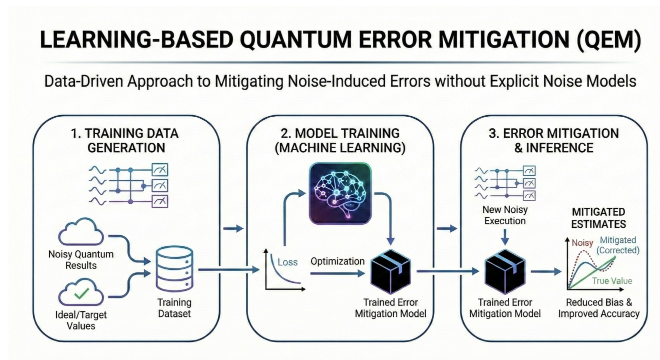


Fig. 1: **Schematic of Learning-Based Quantum Error Mitigation.** The workflow proceeds in three stages: (1) *Training Data Generation*, where a labeled dataset is constructed from pairs of noisy experimental results and calculated ideal values; (2) *Model Training*, where a classical machine learning algorithm optimizes a mapping to minimize the prediction error; and (3) *Inference*, utilizing the trained model to transform new noisy executions into mitigated estimates. This data-driven approach reduces noise-induced bias without requiring explicit characterization of the underlying physical noise channel.

In parallel, machine learning—and neural networks in particular—has emerged as a powerful framework for modeling complex, high-dimensional, and noisy systems. This convergence has led to a growing body of work in *neural networks for quantum computing*, where classical learning models are used to mitigate hardware limitations, improve circuit execution, and extract reliable information from imperfect quantum measurements [4], [5]. Among these applications, learning-based quantum error mitigation (QEM) has attracted particular interest due to its low overhead and compatibility with NISQ-era hardware.

Neural networks are well suited to QEM because many sources of quantum noise are complex, non-Markovian, and strongly hardware-dependent, making them difficult to capture using analytical or first-principles models. Learning-based approaches can instead infer implicit noise-to-correction mappings directly from data, potentially adapting to device drift and capturing effects that are poorly represented in simplified noise models [6]–[9]. However, a critical and largely unresolved question concerns the role of training data: specifically, how models trained on simulated noise compare to those trained on real hardware data, and under what conditions access to hardware measurements meaningfully improves generalization.

Simulated data is inexpensive and can be generated at scale, but its fidelity is limited by the accuracy of the assumed noise models. Real hardware data, by contrast, reflects true device behavior but is expensive, noisy, and often sparsely distributed across circuit families. Understanding how these two data regimes interact in learning-based QEM is essential for designing practical and scalable mitigation pipelines, yet systematic empirical analyses remain limited.

This paper directly addresses this gap by presenting a focused experimental study on incorporating real hardware data into learning-based quantum error mitigation. Using a recently developed attention-based graph transformer model named QAGT-MLP [8], we compare training regimes that rely exclusively on simulated data with regimes that incrementally incorporate hardware measurements. Our analysis spans two distinct circuit families—random circuits and structured quantum simulation circuits derived from one-dimensional Burgers' equations—allowing us to examine how circuit structure and diversity influence the value of hardware data. Through controlled experiments and cross-validation, we analyze performance trends, generalization behavior, and the diminishing returns of hardware data under realistic execution budgets.

The remainder of the paper is organized as follows. Section II introduces background on NISQ noise and learning-based QEM. Section III describes the model architecture, data generation, and training protocol. It also presents empirical results comparing simulated and real hardware training regimes across circuit families. Section IV discusses the implications of these findings for data-efficient learning-based mitigation, followed by concluding remarks in Section VII.

II. Background and Related Work

A. Quantum Error Mitigation vs. Quantum Error Correction

Quantum Error Mitigation (QEM) refers to *hybrid quantum–classical* techniques that reduce noise-induced bias in estimated observables obtained from imperfect quantum hardware, without employing full fault-tolerant encoding [10], [11]. Given an ideal observable expectation

$$\mu \triangleq \mathrm{Tr}(O\rho), \tag{1}$$

and a noisy implementation producing $\tilde{\mu} \approx \mathrm{Tr}(O\tilde{\rho})$, QEM constructs a mitigated estimator $\hat{\mu}_{\mathrm{mit}}$ from additional quantum data (e.g., noise-scaled executions, symmetry checks, or training circuits) such that $\mathbb{E}[\hat{\mu}_{\mathrm{mit}}]$ is closer to μ than $\tilde{\mu}$ [8], [11]–[13]. Figure 1 shows a schematic of a learning-based QEM pipeline from data generation to error mitigation.

QEM differs fundamentally from *Quantum Error Correction* (QEC), which relies on redundancy, syndrome extraction, and fault-tolerant protocols to actively detect and correct errors at the physical layer [14], [15]. While QEM avoids qubit overhead, it typically incurs increased sampling or runtime cost and often relies on stronger assumptions about noise structure or circuit families [10]. Recent theoretical results show that, for broad classes of QEM protocols, achieving small residual error can require rapidly growing sampling overhead with circuit depth or system size [16]–[18].

This paper focuses on *learning-based* QEM approaches, rather than classical analytical methods such as zero-noise extrapolation (ZNE) [11], [12], [19], probabilistic error cancellation (PEC) [7], [12], or calibration-matrix-based measurement mitigation [6].

B. Learning-Based QEM: Formulation

Learning-based QEM replaces hand-crafted mitigation rules with a learned mapping from noisy observations to mitigated estimates. In a generic supervised setting, features x extracted from noisy executions (e.g., expectation values, outcome statistics) and optional circuit or device descriptors $\phi(\mathcal{C})$ are used to train a model

$$f_\theta : \ (x, \phi(\mathcal{C})) \mapsto \hat{\mu}_{\mathrm{mit}}, \tag{2}$$

by minimizing a task-dependent loss against ideal, proxy, or self-supervised targets [13], [20], [21]. A key challenge is that mitigation performance depends not only on model architecture, but critically on how supervision is obtained and how well training data match deployment conditions.

C. Neural and ML Approaches to QEM

a) Supervised denoising and regression.: Early neural QEM methods use ideal labels obtained from classical simulation of small or tractable circuits. Kim *et al.* demonstrate neural correction of noisy measurements but note sensitivity to noise mismatch between training and deployment [22]. Bennewitz *et al.* propose neural error mitigation (NEM) for VQE, reporting improvements in both simulations and hardware experiments [21]. Such approaches are effective when training circuits are well matched to deployment tasks, but face scalability limitations due to label cost.

b) Near-Clifford surrogates and hardware-in-the-loop training.: To reduce reliance on ideal labels, several methods exploit classically simulable surrogates. Learning-based QEM [20] and Clifford data regression (CDR) [23], [24] train models on near-Clifford circuits with exact classical labels and transfer learned corrections to target circuits. These methods scale better but can fail under circuit distribution shift if training circuits are insufficiently representative [10].

c) Hardware-only and noise-scaling supervision.: Some approaches avoid classical simulation by leveraging consistency across executions at different effective noise levels. Zhukov and Pogosov train neural networks on noise-amplified hardware data to learn denoising mappings without ideal labels [25]. These methods reduce classical overhead but rely on stable noise-scaling procedures and may struggle with coherent or correlated errors.

d) Mimicking and amortizing classical QEM.: Another line of work trains models to mimic the outputs of existing QEM pipelines, thereby amortizing mitigation cost. ML-QEM [13] and attention-based extensions such as QAGT-MLP [8] demonstrate that learned models can approximate ZNE or related methods while substantially reducing runtime overhead after training.

e) Noise-agnostic and data-augmentation approaches.: To address label scarcity, DAEM [26] proposes noise-agnostic training via quantum data augmentation using noisy-only supervision. While promising, such methods raise questions about identifiability and robustness under strong drift or non-stationary noise [10].

III. Methodology and Results

The effectiveness of learning-based QEM depends strongly on the training data regime. **Simulation-trained models** offer scalability but risk mismatch with real hardware due to simplified noise models [10], [22]. **Hardware-trained models** capture device-specific effects but are costly, noisy, and subject to drift. **Hybrid and transfer-based approaches** combine simulated and hardware data or adapt models across devices and circuit families, offering a practical compromise [13], [20], [26]. Understanding how these data regimes affect generalization is central to the analysis presented in the remainder of this paper.

In this section, we study the impact of real hardware data on learning-based quantum error mitigation by comparing training regimes that rely purely on simulated noise with those that progressively incorporate measurements from real quantum devices. All experiments are conducted using an attention-based graph transformer model named QAGT-MLP [8] and evaluated using mean absolute error (MAE) in held-out test circuits with five-fold cross validation.

A. Experimental Setup

We consider two distinct circuit families with different structural characteristics. The first consists of randomly generated quantum circuits, designed to span a relatively narrow and controlled distribution of depths and gate counts. The second family comprises a targeted application, leading to a specific set of quantum circuits for Computational Fluid Dynamics (CFD). These circuits were derived from the quantum simulation of the one-dimensional viscous Burgers' equation, which exhibit significantly broader variability in circuit depth, gate composition, and entanglement structure.

For both circuit families, we adopt an incremental training protocol. Models are first trained exclusively on simulated data generated using calibrated noise models. Real hardware data (collected using IBM Quantum Computer `ibm_torino` and Qiskit framework [27]) is then gradually introduced into the training set, while keeping the test set fixed. This protocol isolates the effect of hardware data on generalization performance under otherwise identical training conditions.

TABLE I: Graph statistics summary for the RandomData and BurgerData datasets used in the learning-based QEM analysis. Values are reported as mean ± standard deviation, with ranges shown where applicable.

Statistic	**RandomData**	**BurgerData**
Total graphs	477	174
Nodes	69.06 ± 90.62	300.88 ± 138.87
Edges	114.21 ± 157.15	365.03 ± 171.66
Entangling density	1.137 ± 0.415	1.208 ± 0.013
Connectivity pattern	Partial (100%)	Partial (100%)
Observable type	Partial / Sparse	Sparse only

Comparison Metric	**Value**
Dataset size ratio (Burger / Random)	0.36
Avg. node count difference	+231.81
Avg. edge count difference	+250.83
Entangling density difference	+0.072

B. Results on Random Circuits

Figure 2 summarizes the effect of progressively adding hardware data for both random and CFD circuit families.

For random circuits, introducing real hardware data consistently reduces the test MAE. The absolute error scale in this regime is relatively small, with MAE values on the order of 10^{-2}. While the improvement is steady and accompanied by reduced variance across folds, the overall percentage reduction is moderate. This suggests that, for circuits drawn from a relatively homogeneous distribution, simulated noise models already capture much of the relevant error structure, and hardware data primarily provides fine-grained corrections that align the learned model more closely with true device behavior.

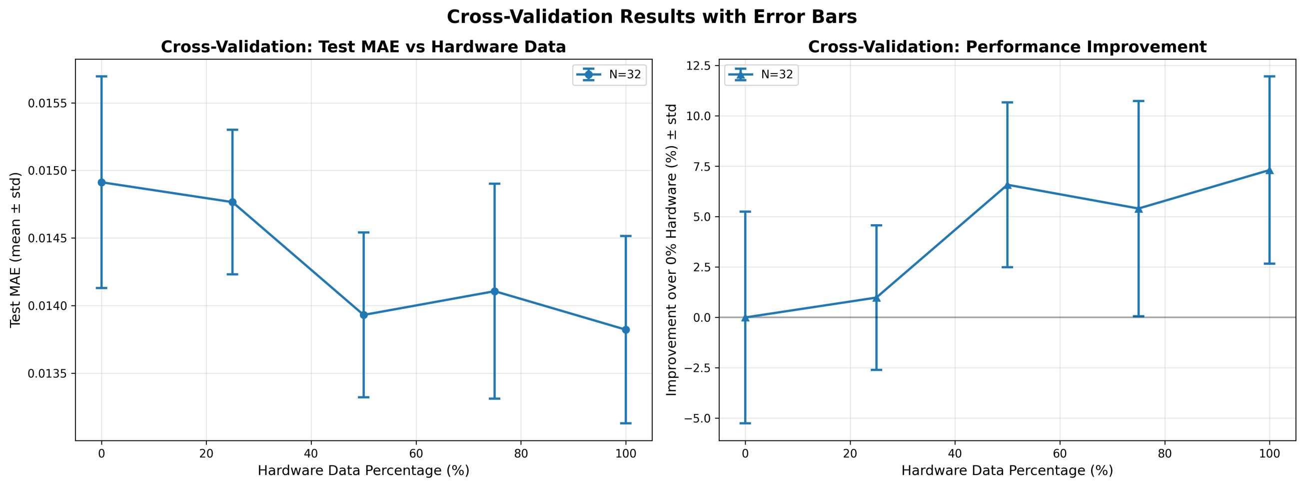
C. Results on Burgers' Equation Circuits

For the Burgers' equation circuits, adding hardware data also leads to a reduction in MAE, but with a notably different profile. The absolute error scale in this setting is higher, with MAE values on the order of 10^{-1}. However, the *relative* improvement—measured as a percentage reduction from the simulation-only baseline—is larger than that observed for random circuits.

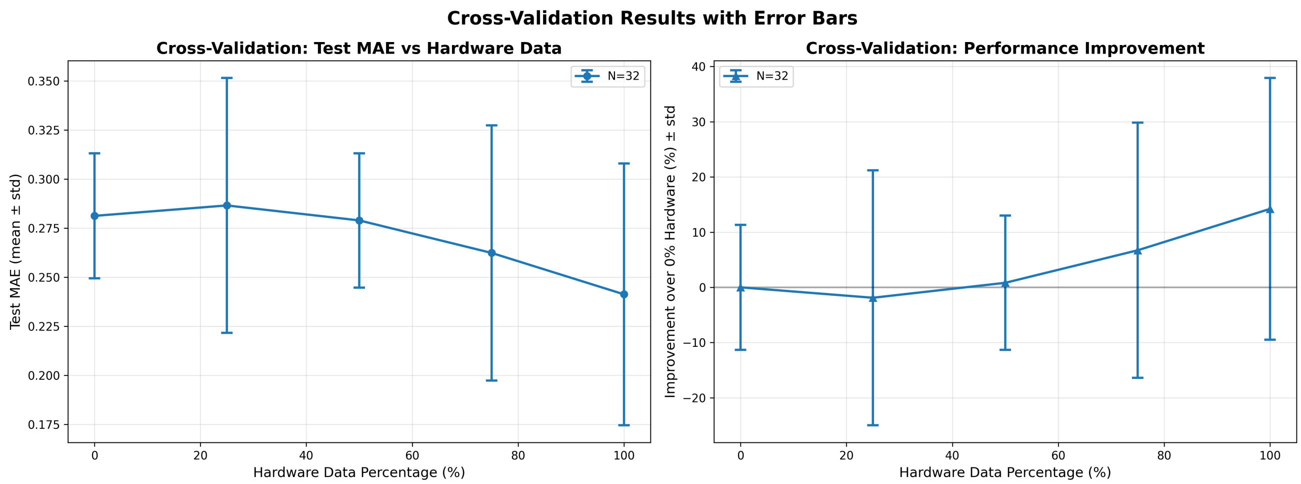
This behavior reflects the increased structural complexity of the quantum CFD circuits. As shown in Table I and Figure 3, these circuits span a much broader distribution of depths, gate counts, and entanglement patterns. In this regime, simulated noise models are less representative of the true error mechanisms, and the introduction of hardware data provides more substantial corrective information, resulting in a larger proportional reduction in error even though the absolute MAE remains higher.

D. Interpretation of the Results

Taken together, these results highlight an important distinction between absolute performance and relative gains from hardware data. For random circuits, simulated training already yields strong performance, leaving limited room for improvement; hardware data therefore produces smaller percentage gains despite lowering an already small MAE. In contrast,



(a) Random circuits



(b) Burgers' equation circuits

Fig. 2: Mean absolute error (MAE) as a function of the fraction of real hardware data included during training for the QAGT-MLP model. (a) Random circuits show low absolute MAE and consistent improvement as hardware data is introduced. (b) Burgers' equation circuits exhibit higher absolute MAE but a larger relative reduction when hardware data is incorporated.

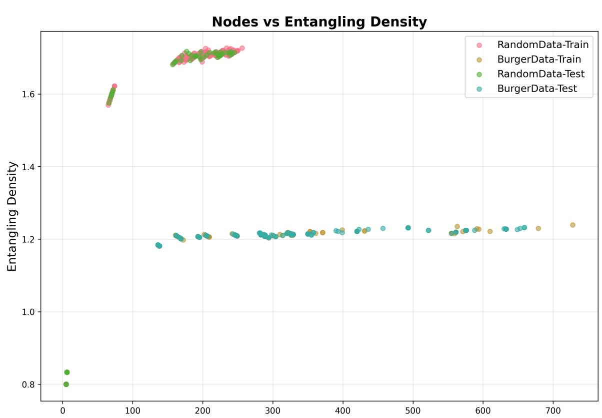


Fig. 3: Relationship between circuit size and entangling density across datasets, illustrating the broader structural diversity of the Burgers' equation circuits.

for the Burgers' equation circuits, the mismatch between simulated noise and real device behavior is more pronounced, and hardware data yields larger relative improvements by correcting errors that are difficult to capture in simulation.

Executing circuits on real quantum hardware remains expensive in terms of shot budgets and access time. For this reason, we deliberately avoided duplicating or oversampling random circuits to artificially match the depth and gate-count distribution of the Burgers' equation circuits. Such duplication would distort the data regime and obscure the practical trade-offs faced in realistic deployment scenarios, where hardware data is inherently sparse and unevenly distributed across circuit families.

Overall, these findings indicate that the value of hardware data in learning-based QEM depends not only on its quantity, but critically on the degree to which simulated noise fails to capture the dominant error mechanisms of the target circuit class. In complex, structurally diverse regimes, hardware data can yield disproportionately large relative gains, even when absolute errors remain higher. This insight has direct implications for designing cost-effective training and data collection strategies for scalable, learning-based quantum error mitigation.

IV. Discussion

Learning-based QEM is fundamentally sensitive to the circuit distribution.: The experiments demonstrated that learning-based QEM should be viewed as a problem of generalization of the distribution of the circuits, rather than the training data. For random circuits, which occupy a relatively narrow and homogeneous region of circuit space, simulated training already achieved low absolute error. Incorporating hardware data further reduced MAE and variance, but the relative improvement was moderate due to the limited gap between simulated and real noise behavior in this regime.

Hardware data yielded larger relative gains for targeted circuits, such as Burgers' equation circuits.: For the Burgers' equation circuits, the absolute MAE remained higher than in the random-circuit setting, reflecting the increased difficulty of the task as a function of the wide range of the quantum circuits, even though the training data was representative of the testing data. In that case, introducing real hardware data produced a *larger relative reduction* in error compared to the simulation-only baseline. This indicates that simulated noise models were less representative for these structured quantum CFD circuits, and that hardware data provided more informative and corrective signals. The broader distributions of circuit depth, gate count, and entanglement patterns led to error mechanisms that were difficult to capture accurately in simulation, making hardware measurements particularly valuable in spite of their sparsity.

Circuit coverage, not realism alone, determined the utility of hardware data.: While hardware data was inherently more realistic, its effectiveness depended on how well it samples the structural diversity of the target quantum circuit family. In the random-circuit regime, coverage was already adequate, limiting the marginal benefit of hardware data. In contrast, for the Burgers' circuits, even limited hardware data was able to substantially reduce error because it captured device-specific effects that are poorly modeled in simulation. These results suggest that the benefit of hardware data is governed less by its absolute quantity and more by its ability to cover circuit-induced variability.

Implications for learning-based QEM.: Because executing circuits on real quantum hardware is expensive in terms of shot budgets and access time, we deliberately avoided duplicating or oversampling random circuits to artificially match the depth and gate-count distribution of the Burgers' circuits. Such duplication would have produced an unrealistic data regime and obscured the practical trade-offs faced in deployment. Instead, our findings emphasized that effective learning-based QEM requires *strategic* use of hardware data, particularly for circuit families where simulated noise fails to capture dominant error mechanisms.

V. Future Works

This research can potentially lead to future extensions in multiple directions:

1) **Coverage-aware hardware data acquisition strategies**: A key insight from this work is that quantum circuit coverage drives the value of hardware data more than data quality alone. Future work should investigate systematic, coverage-aware strategies for selecting which circuits to execute on hardware, minimizing cost while maximizing the diversity of circuit-induced error mechanisms captured in the training set.
2) **Expanding the circuit families**: While the two circuit families used in this work span a range of circuit characteristics, including a broader variety of physics-based quantum circuits (e.g., variational algorithms, Hamiltonian simulation) would strengthen the correlation between circuit application domains and their respective noise behaviors on simulation versus hardware runs, enabling more targeted mitigation strategies.
3) **Different learning-based QEM models**: Experimenting with other QEM models with different feature extraction and embedding capabilities can provide valuable insight into how model expressiveness interacts with circuit diversity and training data regime, particularly for structured circuits where simulation fidelity is limited.
4) **Scalable mitigation pipelines for NISQ devices**: Translating the findings of this study into practical, scalable QEM pipelines for NISQ-era devices requires addressing the trade-off between hardware execution cost and mitigation accuracy. Future work could explore active learning or transfer learning frameworks that efficiently incorporate hardware data as circuit complexity and device scale grow.

VI. Acknowledgment

The computation for this work was performed on the University of Missouri's Quantum Innovation Center, in partnership with IBM Quantum and facilitated by Research Support Services at the University of Missouri, Columbia MO. DOI: 10.32469/10355/107781

VII. Conclusion

This paper presented an empirical analysis of incorporating real hardware data into learning-based quantum error mitigation, focusing on how different natures of training data affect generalization performance. Using a SOTA learning-based QEM model, we systematically compared training on simulated noise alone with training regimes that progressively integrate measurements from real quantum hardware across distinct circuit families.

Our results demonstrated that the impact of hardware data depends strongly on the quantum circuit structure and diversity, more than on the quality of the training data. For random circuits, with similar structures, simulated training quickly achieved low absolute error, while the error for more diverse targeted quantum CFD circuits was intrinsically larger for simulated noise. Also, while both cases benefited from hardware data, the introduction of the same hardware data in problems with diverse quantum circuits yielded substantially larger relative reductions in error. These findings indicated that the value of hardware data is governed not only by the ability of the training data to approximate the testing data, but by how effectively the training/testing data captures circuit-induced error mechanisms that are difficult to model in simulation.

Overall, this study highlights quantum circuit coverage as a central factor in effective learning-based QEM and underscores the need for coverage-aware, cost-efficient strategies when incorporating hardware data. The insights presented here inform the design of scalable learning-based mitigation pipelines for NISQ-era devices, particularly for complex and structured quantum applications where simulation fidelity is limited.

REFERENCES

[1] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[2] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[3] H.-S. Zhong *et al.*, "Quantum computational advantage using photons," *Science*, vol. 370, no. 6523, pp. 1460–1463, 2020.

[4] G. Carleo *et al.*, "Machine learning and the physical sciences," *Reviews of Modern Physics*, vol. 91, no. 4, p. 045002, 2019.

[5] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain," *Reports on Progress in Physics*, vol. 81, no. 7, p. 074001, 2018.

[6] S. Bravyi, S. Sheldon, A. Kandala, D. C. Mckay, and J. M. Gambetta, "Mitigating measurement errors in multiqubit experiments," *Physical Review A*, vol. 103, no. 4, p. 042605, 2021.

[7] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, "Hybrid quantum–classical algorithms and quantum error mitigation," *Journal of the Physical Society of Japan*, vol. 90, no. 3, p. 032001, 2021.

[8] S. M. A. Tousi and G. DeSouza, "Qagt-mlp: An attention-based graph transformer for small and large-scale quantum error mitigation," *arXiv preprint arXiv:2511.03119*, 2025.

[9] S. M. A. Tousi, A. Bazgir, Y. Zhang, and G. DeSouza, "Quantum error mitigation with attention graph transformers for burgers equation solvers on nisq hardware," *arXiv preprint arXiv:2512.23817*, 2025.

[10] Z. Cai *et al.*, "Quantum error mitigation," *Reviews of Modern Physics*, vol. 95, no. 4, p. 045005, 2023.

[11] S. Endo, S. C. Benjamin, and Y. Li, "Practical quantum error mitigation for near-future applications," *Physical Review X*, vol. 8, no. 3, p. 031027, 2018.

[12] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," *Physical Review Letters*, vol. 119, no. 18, p. 180509, 2017.

[13] H. Liao, D. S. Wang, I. Sitdikov, C. Salcedo, A. Seif, and Z. K. Minev, "Machine learning for practical quantum error mitigation," *Nature Machine Intelligence*, vol. 6, no. 12, pp. 1478–1486, 2024.

[14] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge, UK: Cambridge University Press, 2010.

[15] B. M. Terhal, "Quantum error correction for quantum memories," *Reviews of Modern Physics*, vol. 87, no. 2, pp. 307–346, 2015.

[16] R. Takagi, S. Endo, S. Minagawa, and M. Gu, "Fundamental limits of quantum error mitigation," *npj Quantum Information*, vol. 8, no. 1, p. 114, 2022.

[17] Y. Quek, D. Stilck França, S. Khatri, J. J. Meyer, and J. Eisert, "Exponentially tighter bounds on limitations of quantum error mitigation," *Nature Physics*, vol. 20, pp. 1648–1658, 2024.

[18] K. Tsubouchi, T. Sagawa, and N. Yoshioka, "Universal cost bound of quantum error mitigation based on quantum estimation theory," *Physical Review Letters*, vol. 131, no. 21, p. 210601, 2023.

[19] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization," *Physical Review X*, vol. 7, no. 2, p. 021050, 2017.

[20] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, "Learning-based quantum error mitigation," *PRX Quantum*, vol. 2, no. 4, p. 040330, 2021.

[21] E. R. Bennewitz, F. Hopfmueller, B. Kulchytskyy, J. Carrasquilla, and P. Ronagh, "Neural error mitigation of near-term quantum simulations," *Nature Machine Intelligence*, vol. 4, pp. 618–624, 2022.

[22] C. Kim, K. D. Park, and J.-K. Rhee, "Quantum error mitigation with artificial neural network," *IEEE Access*, vol. 8, pp. 188853–188860, 2020.

[23] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, "Error mitigation with Clifford quantum-circuit data," *Quantum*, vol. 5, p. 592, 2021.

[24] A. Lowe, M. H. Gordon, P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, "Unified approach to data-driven quantum error mitigation," *Physical Review Research*, vol. 3, no. 3, p. 033098, 2021.

[25] A. Zhukov and W. Pogosov, "Quantum error reduction with deep neural network applied at the post-processing stage," *Quantum Information Processing*, vol. 21, no. 3, p. 93, 2022.

[26] L. Placidi, I. Williams, E. Rinaldi, D. Mills, C. Crstoiu, V. Eccles, and R. Duncan, "Deep learning approaches to quantum error mitigation." arXiv:2601.14226, 2026.

[27] A. Javadi-Abhari, M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, *et al.*, "Quantum computing with qiskit," *arXiv preprint arXiv:2405.08810*, 2024.