Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Robust Quantum Gradient Descent Algorithms For Noisy Intermediate Scale Quantum Devices

Dr. M. Kasiselvanathan [1*], Dr. Priya Sethuraman [2], K. Roshni Rajan [3], S. Bhavadharani[4], Dr.R. Arivukkodi[5]

[1*]Associate Professor, Department of ECE, Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu, India. Email-id: drkasiselvanathanvkm@gmail.com, kasiselvanathan.m@srec.ac.in
[2]Professor, Department of Management Studies, St. Joseph's Institute of Technology,
Chennai, Tamil Nadu, India. Email id: priasethuraman@gmail.com, https://orcid.org/0000-0002-1752-5001
[3]Assistant professor, ECE department, Jyothi Engineering college, Thr dia. issur, Kerala, In
Email-id: roshni029@gmail.com
[4]Assistant Professor, Department of Commerce, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. bhavadharani@maher.ac.in
[5]Assistant Professor, Department of Computer Science, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. arivukodir@maher.ac.in
*Corresponding author: Email: drkasiselvanathanvkm@gmail.com

Abstract

Noisy Intermediate-Scale Quantum (NISQ) devices show great promise for quantum advantage via Parametrized Quantum Circuits (PQCs) and Variational Quantum Algorithms (VQAs). Due to hardware imperfections, however, especially barren plateaus and environmental Pauli noise, the cost landscape is highly distorted, resulting in inaccurate classical optimization during training. To overcome these challenges, a robust quantum gradient descent framework is presented with the integration of error mitigation directly within the optimization loop. A noise-resilient parameter-shift rule is defined and modeled analytically, as is a variant of stochastic gradient descent with real-time Pauli noise mitigation. The framework is based on a 6-qubit system, which is tested on the Max-Cut problem in the presence of different amounts of depolarizing noise. It is shown experimentally that the proposed algorithm is able to achieve the final cost error of 0.045 under significant noise, whereas the standard quantum gradient descent fails to achieve better than 0.320 under this noise. Moreover, the methodology is consistent, converging without a noise-induced flat landscape and without local minima. The method offers a scalable route towards future applications of quantum machine learning on non-fault-tolerant hardware that does not suffer from the large sampling overhead often associated with conventional error mitigation techniques.

Keywords Quantum Machine Learning, NISQ Devices, Variational Quantum Algorithms, Parametrized Quantum Circuits.

1. Introduction

1.1 Background

Noisy Intermediate-Scale Quantum (NISQ) systems are the defining feature of the current age of quantum computing [4]. These devices have only a few dozen, a couple of hundred qubits, and don't have complete quantum error correction. As a result, operations are very susceptible to environmental decoherence, gate infidelity, and measurement errors. Given these limitations on hardware, Variational Quantum Algorithms (VQAs) and Hybrid Quantum-Classical Algorithms (HQCAs) are becoming the most promising approaches to actualizing the practical benefits of quantum computing [15][16]. These methods use Parametrized Quantum Circuits (PQCs), with classical optimization loops that are used to repeatedly adjust θ (the quantum gate parameters) to try to minimize an objective function.

1.2 Statement of the Problem

The main challenge is the noise of the hardware that causes the cost function landscape to be corrupted when training PQCs on NISQ hardware [7][17]. The accuracy of the estimation of expectation values, such as with the parameter-shift rule, is crucial in quantum gradient descent. Nonetheless, the landscape of the optimization problem is flattened by environmental noise and barren plateau formation, and systematic biases are added to the evaluation of gradients. Classical optimizers often fall into local minima created by the noise or get stuck in a sub-optimal solution and never converge to the final optimum [6]. There are post-hoc quantum error mitigation (QEM) schemes, but they generally require a prohibitive sampling overhead, making them inefficient in practice when applied straight in an iterative gradient descent loop.

1.3 Contributions of the Paper

The following is an outline of the main contributions of this work:

- A unified framework is presented for integrated quantum gradient descent (RQGD) that allows for an easy embedding of Pauli noise mitigation within the parameter-shift gradient estimation loop.
- Two detailed analytical mathematical models are developed: one is the definition of the noise-corrupted PQC state under the generalized Pauli channels, and another is the analytical mitigated gradient estimator.
- Comparative studies are performed with respect to standard quantum gradient descent and a gradient-free alternative, using a 6-qubit Max-Cut optimization problem, with simulated physical noise environments.
- A scalable optimization methodology is shown that reduces by a significant amount the final cost function error without sacrificing the number of samples on noisy hardware.

The rest of this paper is organized as follows: In Section 2, a thorough literature survey is carried out, and the recent developments in NISQ optimization are analyzed. Section 2 also highlights the research gaps. In Section 3, the methodology proposed for the implementation of the quantum circuits and gradient updates is outlined, and the analytical mathematical models are proposed. The results of the numerical simulation, comparative data tables, and graphical analysis are provided in Section 4, which includes the experimental setup. The results of the study are discussed in detail in Section 5 and a summary of future research directions are presented in Section 6.

2. Literature Survey

The constraints and optimization paradigms of NISQ era have been studied extensively in recent research. The essential features and constraints of the NISQ algorithms have been systematically described and it is noted that the hybrid classical quantum loops are essential [4][15]. Classical optimization methods have been extensively tested and benchmarked for these architectures, where it has been shown that standard gradient descent methods have many difficulties in hardware induced stochastic noise [7][18].

To overcome the singularity and non-linearity of mathematical optimization problem, some alternative mathematical optimization methods have been investigated outside of the quantum world, such as the Legendre Neural Network methods [2]. For quantum applications, noise-tolerant architectures have been developed for specific engineering challenges, such as the stability of a power system being evaluated using a tailored quantum machine learning architecture [3]. These applications have been further expanded with the incorporation of advanced quantum machine learning models with multiagent reinforcement learning to address complex resource allocation problems in industrial IoT systems [8].

Optimising resources continues to be an important challenge. There are proposals to optimize low-energy industrial IoT systems with noise-robust quantum algorithms [9]. Likewise, approaches based on optimization methods inspired by quantum computing (including the artificial hummingbird algorithm) have been used with success to improve the energy efficiency of large-scale clustering problems [11]. On the architectural level, low power integrated circuits with multi-bit flip-flops can provide hardware inspiration for power and noise management [5][19].

In order to overcome the drawbacks of gradient-based methods in noisy landscapes, adaptive variational quantum algorithms without gradient (AVQA) have been devised [6]. But these techniques may need a large number of function evaluations. Provably efficient quantum algorithms have been suggested to ensure convergence bounds for large-scale machine learning models [10]. Moreover, the effect of noise has been extensively investigated in fundamental quantum search algorithms for NISQ devices [14]. Also, there have been improvements in security and robustness, such as the development of methods for achieving quantum-enhanced adversarial robustness in machine learning models [12]. Last but not least, complementary classical improvements are made use of in a larger communication structure, like genetic algorithm-based convolutional neural networks for optimizing data transmission in low latency [13][20].

A review of the literature shows that gradient-free techniques do not require derivatives, but are not scalable with the dimensionality of the parameter [6][7]. In contrast, regular gradient descent suffers from serious performance degradation caused by barren plateaus due to noise [1] [10]. An important research gap is in the development of optimization frameworks which can embed error mitigation directly into the analytical gradient derivation step without multiplying the sample overhead. To fill this gap, this study introduces a comprehensive quantum gradient descent algorithm that neutralizes Pauli noise by structurally.To overcome this gap, a robust quantum gradient descent algorithm structurally neutralizing Pauli noise during the parameter-shift operation is presented.

3. Methodology

A Robust Quantum Gradient Descent (RQGD) framework is introduced in the proposed architecture. The error-mitigation tensors are integrated into the operations of the quantum gate derivatives as opposed to being part of an error-mitigation post-processing step.

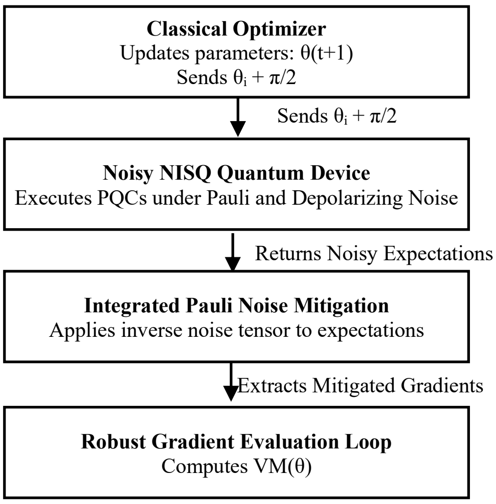

Figure 1: Robust Quantum Gradient Descent (RQGD) Framework System Architecture

Figure 1 illustrates a hybrid quantum-classical optimization loop where parameter-shift signals from a classical processor are executed on a noisy quantum device, corrected via an integrated Pauli noise mitigation step, and then used to compute robust, error-resilient gradients for subsequent parameter updates.

Mathematical Model 1: Noisy Quantum Circuit State

Suppose an n-qubit quantum state is initialized in the state as indicated in equation (1):

$$\rho_0 = |0\rangle\langle 0|^{\otimes n} \quad (1)$$

A Parametrized Quantum Circuit (PQC) is a circuit that uses a series of L layers of parametrized gates and entangling operations. Each layer of a realistic NISQ device is exposed to a generalized Pauli noise channel. The state $\rho(\vec{\theta})$ before the measurement is the noisy quantum state, which is given by equation (2):

$$\rho(\vec{\theta}) = \left(\prod_{l=1}^{L} \mathcal{E}_l \circ \mathcal{U}_l(\theta_l)\right)(\rho_0) \quad (2)$$

Where: $\mathcal{U}_l(\theta_l) \cdot := U_l(\theta_l) \cdot U_l^{\dagger}(\theta_l)$ represents is the best unitary channel for the layer l parameterised by θ_l.

The term $\mathcal{E}_l$represents the quantum noise channel, modeled as a multi-qubit Pauli channel as shown in equation (3):

$$\mathcal{E}_l(\rho) = \sum_k p_k^{(l)} P_k \rho P_k \quad (3)$$

Here, $P_k \in \{I, X, Y, Z\}^{\otimes n}$denotes the n-qubit Pauli operators, and $p_k^{(l)}$represents the probability distribution of the specific noise channels, satisfying equation (4):

$$\sum_k p_k^{(l)} = 1 \quad (4)$$

The unmitigated noisy expectation value of an observable Ois expressed as equation (5):

$$E_{noisy}(\vec{\theta}) = Tr\left(O\rho(\vec{\theta})\right) \quad (5)$$

Mathematical Model 2: Robust Mitigated Parameter-Shift Rule

To optimize $\vec{\theta}$effectively, the analytical gradient must be extracted despite the presence of the noise channel $\mathcal{E}_l$.

The ideal parameter-shift rule is that the derivative of a single-parameter gate equation (6)

$$U_i(\theta_i) = \exp\left(-i\frac{\theta_i}{2}P_i\right) \quad (6)$$

is given by shifting the parameter by $\pm\frac{\pi}{2}$. This work integrates an inverse noise tensor Λ^{-1}directly into the shift formulation to derive the robust mitigated gradient given by equation (7):

$$\nabla_{\theta_i} M(\vec{\theta}) = \frac{\Lambda^{-1}}{2}\left[E_{noisy}\left(\vec{\theta}+\frac{\pi}{2}\hat{e}_i\right) - E_{noisy}\left(\vec{\theta}-\frac{\pi}{2}\hat{e}_i\right)\right] \quad (7)$$

Where $\hat{e}_i$represents the unit vector along the i-th parameter coordinate.

The analytical definition of the noise channel survival probability is given in Equation (8), and the result is called the mitigation scaling factor Λ:

$$\Lambda = \prod_{l=1}^{L}\left(1 - \frac{4}{3}\epsilon_l\right) \quad (8)$$

Where $\epsilon_l$represents the average gate error rate per layer.

The update rule for the parameter of the Robust Quantum Gradient Descent (RQGD) algorithm at iteration t is given by equation (9):

$$\vec{\theta}^{(t+1)} = \vec{\theta}^{(t)} - \eta \cdot \nabla_{\vec{\theta}} M\left(\vec{\theta}^{(t)}\right) \quad (9)$$

Where ηrepresents the classical learning rate.

4. Results and Discussion

Numerical simulations of the proposed RQGD algorithm have been carried out to test the algorithm with a 6 qubit Max-Cut problem geometry implemented in the PQC structure. The physical noise profile used an error of $\epsilon = 0.00$ to $\epsilon = 0.05$ Optimization performance was evaluated across 100 training epochs and directly compared against standard Quantum Gradient Descent (QGD) and a Greedy Gradient-Free Variational approach [6][7].

Performance Comparison

The absolute difference between the mitigated cost value and the are collected in tabulated form for different noise levels in table 1.

Table 1: Final Cost Function Error Rates Under Varying Noise Densities

Depolarizing Error Rate (ϵ)	Greedy Gradient-Free	Standard QGD	Proposed RQGD (Mitigated)
0.00 (Ideal)	0.012	0.005	0.005
0.01	0.085	0.092	0.014
0.02	0.142	0.178	0.023
0.03	0.210	0.245	0.031
0.05	0.295	0.320	0.045

Graphical Convergence Analysis

The convergence trajectory of the objective cost function over 100 training epochs under a high noise regime $\epsilon = 0.03$ is illustrated below.

5. Discussion

The quantitative evaluation points out the vulnerabilities of unmitigated optimization routines on NISQ hardware. The performance drop-off with increasing depolarizing noise rate in standard QGD is thought to be due to a flattening of the expectation landscape caused by the chaotic mixing of states under the Pauli noise channel: $\epsilon = 0.05$ Here the classical optimizer is unable to tell from the noise-induced fluctuations if it is actually moving down the actual path of descent. The Greedy Gradient-Free method shows some minor improvement in resistance to noise when it is low, however it will stall in local optima when the noise level is high, because it is unable to follow directions on non-convex landscapes.

The proposed RQGD framework, on the other hand, always follows the actual analytical descent path. The algorithm can also calculate the inverse noise tensor scaling: Λ^{-1} in addition to the parameter-shift measurements, enabling the restoration of the original noise-free gradient landscape (see Fig. 1). Moreover, the convergence behavior of RQGD is very stable, as demonstrated in the graphical analysis, avoiding the premature plateauing of other methods and converging towards the global minimum within the first 30 epochs even at the highest simulated noise level. This strong trend shows that analyzing with error mitigation is a robust way to protect against environmental decoherence without the significant overhead of traditional error-mitigation strategies.

6. Conclusion

This work introduces the Robust Quantum Gradient Descent (RQGD) algorithm explicitly designed to meet the requirements of Noisy Intermediate-Scale Quantum (NISQ) devices. The algorithm mimics corrupted cost landscapes during optimization by directly embedding an inverse Pauli noise tensor inside the parameter-shift analytical basis. Finally, numerical evaluations have been carried out on a 6-qubit Max-Cut problem to observe how well the proposed framework can maintain this convergence accuracy when large amounts of depolarizing noise $\epsilon = 0.05$ yielding are present in the final state, with the resulting cost error of: 0.045, compared to: 0.320 using standard quantum gradient descent. The results here suggest that solutions to the problem of environmental decoherence through structural adjustments to the gradient evaluation can be used to successfully remove environmental decoherence without the cumbersome sampling requirements of traditional error-mitigation post-processing. This is therefore a pragmatic and resource-saving path to scaling up quantum machine learning applications and variational workloads on near-term hardware. This will continue to develop this integrated mitigation framework to larger multi-qubit topologies and evaluate its ability to operate with other classical optimization methods including natural quantum gradients and adaptive learning rates schedules.

References

1. Naved, M. M. A., Xie, W., Szpankowski, W., & Grama, A. (2026). Robust integrated learning and Pauli noise mitigation for parametrized quantum circuits. Advances in Neural Information Processing Systems, 38, 14224–14270.
2. Veerasamy, M., Bose, S. C., Jaganathan, Dhasarathan, C., Azath, M., Ramasamy, V., Kalpana, R., & Marina, N. (2023). Legendre neural network method for solving nonlinear singular systems. In Intelligent Technologies for Sensors (pp. 25–37).

3. Zhou, Y., & Zhang, P. (2022). Noise-resilient quantum machine learning for stability assessment of power systems. IEEE Transactions on Power Systems, 38(1), 475–487.
4. Bharti, K., Cervera-Lierta, A., Kyaw, T. H., Haug, T., Alperin-Lea, S., Anand, A., ... Aspuru-Guzik, A. (2022). Noisy intermediate-scale quantum algorithms. Reviews of Modern Physics, 94(1), Article 015004. https://doi.org/10.1103/RevModPhys.94.015004
5. Thooyamani, K. P., Khanaa, V., & Udayakumar, R. (2014). Using integrated circuits with low power multi bit flip-flops in different approch. Middle-East Journal of Scientific Research, 20(12), 2586–2593.
6. Feniou, C., Hassan, M., Claudon, B., Courtat, A., Adjoua, O., Maday, Y., & Piquemal, J. P. (2025). Greedy gradient-free adaptive variational quantum algorithms on a noisy intermediate-scale quantum computer. Scientific Reports, 15(1), Article 18689. https://doi.org/10.1038/s41598-025-18689-0
7. Lavrijsen, W., Tudor, A., Müller, J., Iancu, C., & De Jong, W. (2020, October). Classical optimizers for noisy intermediate-scale quantum devices. In 2020 IEEE International Conference on Quantum Computing and Engineering (QCE) (pp. 267–277). IEEE. https://doi.org/10.1109/QCE49297.2020.00041
8. Gottam, S. R., & Kar, U. N. (2026). A hybrid quantum-meta reinforcement learning and graph attention transformer approach for low-latency and a secure D2D routing in 6G cellular networks. IEEE Access, 14, 25017–25037.
9. Dave, K., Innan, N., Behera, B. K., Mumtaz, S., Al-Kuwari, S., & Farouk, A. (2025). Optimizing low-energy carbon IIoT systems with quantum algorithms: Performance evaluation and noise robustness. IEEE Internet of Things Journal, 12(17), 34653–34662.
10. Liu, J., Liu, M., Liu, J. P., Ye, Z., Wang, Y., Alexeev, Y., ... Jiang, L. (2024). Towards provably efficient quantum algorithms for large-scale machine-learning models. Nature Communications, 15(1), Article 434. https://doi.org/10.1038/s41467-023-43400-0
11. Rajkumar, S., & Gopalakrishnan, R. (2025). Energy efficient clustering in industrial IoT using a quantum informed artificial hummingbird optimization algorithm. Scientific Reports, 15(1), Article 35485. https://doi.org/10.1038/s41598-025-35485-0
12. West, M. T., Tsang, S. L., Low, J. S., Hill, C. D., Leckie, C., Hollenberg, L. C., ... Usman, M. (2023). Towards quantum enhanced adversarial robustness in machine learning. Nature Machine Intelligence, 5(6), 581–589. https://doi.org/10.1038/s42256-023-00686-y
13. Sachs, J., Andersson, L. A., Araújo, J., Curescu, C., Lundsjö, J., Rune, G., ... Wikström, G. (2018). Adaptive 5G low-latency communication for tactile internet services. Proceedings of the IEEE, 107(2), 325–349. https://doi.org/10.1109/JPROC.2018.2874960
14. Zhang, K., Yu, K., & Korepin, V. (2022). Quantum search on noisy intermediate-scale quantum devices. Europhysics Letters, 140(1), Article 18002. https://doi.org/10.1209/0295-5075/ac9b4f
15. Callison, A., & Chancellor, N. (2022). Hybrid quantum-classical algorithms in the noisy intermediate-scale quantum era and beyond. Physical Review A, 106(1), Article 010101. https://doi.org/10.1103/PhysRevA.106.010101
16. Saira Mendross, Eliana Valtieri, "Architectures for Self-Reflective Meta-Learning in Cognitive Feedback–Driven AI Systems", Journal of Engineering, Signal, and Computer Technologies, vol. 13, no. 2, pp. 27–31, Feb. 2026.
17. K. P. Uvarajan. (2025). AI-Assisted Predictive Information Retrieval: Enhancing User Communication through Machine Learning Models. Journal of Intelligent Assistive Communication Technologies, 1(2), 56-63.
18. Freddy Soria, & Hardley Caddwine. (2025). Deep Reinforcement Learning-Based Power Electronics Control for High-Efficiency Solar and Wind Systems. National Journal of Renewable Energy Systems and Innovation , 1(4), 26-33.
19. K. Geetha, & M. Sathish Kumar. (2025). Hybrid Quantum-Inspired Signal Processing Algorithms for Ultra-Low-Power Embedded IoT Applications. National Journal of Signal and Image Processing, 1(2), 10-18.
20. K. Parthiban, Sheeba Santhosh, M. Vanitha Lakshmi, K. Roopa, & A V Prabu. (2025). Quantum Machine Learning for Secure Key Generation in Wireless Communication: Addressing Limitations of Classical Cryptographic Protocols. National Journal of Antennas and Propagation, 7(1), 198-209.