



# Article A Modified Depolarization Approach for Efficient Quantum Machine Learning

Bikram Khanal 🕩 and Pablo Rivas\*🕩

School of Engineering and Computer Science, Department of Computer Science, Baylor University, Waco, TX,USA, 76798; bikram\_khanal1@baylor.edu

\* Correspondence: pablo\_rivas@baylor.edu; Tel.: +1-(254)-710-3385

**Abstract:** Quantum Computing in the Noisy Intermediate-Scale Quantum (NISQ) era has shown promising applications in machine learning, optimization, and cryptography. Despite these progresses, challenges persist due to system noise, errors, and decoherence. These system noises complicate the simulation of quantum systems. The depolarization channel is a standard tool for simulating a quantum system's noise. However, modeling such noise for practical applications is computationally expensive when we have limited hardware resources, as is the case in the NISQ era. This work proposes a modified representation for a single-qubit depolarization channel. Our modified channel uses two Kraus operators based only on *X* and *Z* Pauli matrices. Our approach reduces the computational complexity from six to four matrix multiplications per channel execution. Experiments on a Quantum Machine Learning (QML) model on the Iris dataset across various circuit depths and depolarization rates validate that our approach maintains the model's accuracy while improving efficiency. This simplified noise model enables more scalable simulations of quantum circuits under depolarization, advancing capabilities in the NISQ era.

Keywords: NISQ; depolarization channel; quantum machine learning; circuit depth optimization

MSC: 68Q12, 68Q09, 81P68

# 1. Introduction

Quantum Computing has seen significant progress in recent years, with the development of quantum algorithms for a variety of applications, including machine learning [1–6], optimization [7–11], and cryptography [12–15]. However, the development of quantum algorithms is still in its infancy. Many of the developed algorithms are not yet ready for practical use [16,17]. Due to the susceptibility of NISQ device operations to inherent errors and decoherence [18,19], simulating quantum systems remains a major challenge in developing quantum algorithms [17].

In the NISQ era, system noise is not merely a challenge to be addressed but a fundamental tool that shapes the field of QML research. Interestingly, many works have chosen to regard noise not as a challenge but as an opportunity to advance their research. Studies have shown that, unlike classical algorithms, quantum learning of n-bit parity functions is highly resilient to depolarization noise [20]. This early work demonstrated the potential for quantum algorithms to maintain a learning advantage even in noisy conditions. While traditionally viewed as a detrimental factor to quantum computation, depolarization noise under certain conditions can enhance the robustness and functionality of quantum learning algorithms against adversarial attacks [21–25]. This counterintuitive finding highlights the potential of noise to endow quantum models with robustness against malicious attempts that aim to manipulate the model's outputs.

However, harnessing the power of noise as a training tool requires careful consideration. For example, the effectiveness of adversarial training techniques hinges on the



**Citation:** Khanal, B.; Rivas, P. A Modified Depolarization Approach for Efficient Quantum Machine Learning. *Mathematics* **2024**, *1*, 0. https://doi.org/

Academic Editor:

Received: 10 April 2024 Revised: 26 April 2024 Accepted: Published:



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). assumption that the test attack and the training attack employ the same methods to generate adversarial examples. In real-world scenarios where attackers may employ diverse and unknown strategies, this advantage is not guaranteed [26,27]. Therefore, deriving robust guarantees against worst-case scenarios remains crucial for building truly secure and resilient quantum learning algorithms.

The challenges posed by noise extend beyond algorithm design, impacting the very foundations of QML. The inherent noise in the NISQ machines also presents significant challenges to the learning capabilities of Quantum Neural Networks and QML models [18,28]. System noise can significantly diminish the quantum kernel advantage [29], raising concerns about the viability of quantum kernel methods [30,31]. Additionally, calculating numerical gradients on noisy qubits presents a delicate balancing act: reducing the step size to improve accuracy can obscure subtle differences in the cost function for nearby parameter values [32]. Further research into controlled noise simulations, such as the depolarization channel, is necessary to comprehend better and mitigate these complexities.

In the worst-case scenario, we can use the depolarization channel to simulate the quantum system's noise [30]. However, executing depolarizing noise in a controlled manner on quantum hardware presents critical challenges and intriguing opportunities for advancing our understanding and mitigation of this noise model. One of the primary challenges in executing depolarizing noise lies in its inherently probabilistic nature. Depolarization introduces errors with a certain probability, often modeled by the Kraus operators, onto the quantum state [33]. Implementing such probabilistic errors precisely on hardware requires sophisticated control techniques and careful calibration procedures. Inaccurate noise injection can lead to deviations from the expected noise model. This deviation can compromise the validity of subsequent experiments and analyses. To fully realize the potential of QML, it is crucial to develop effective error correction techniques [34]. Techniques like surface codes [35] and stabilizer codes [36] offer promising avenues for error correction and safeguard the integrity of quantum computations. Combining noise-estimate circuits with other error correction techniques can accurately estimate and correct noise in circuits even with extensive CNOT gates [37]. We can improve the accuracy of error mitigation using multi-exponential error mitigation techniques [38]. On the one hand, we may be able to model depolarizing noise accurately, but on the other hand, the computational cost can be prohibitive in the NISQ era [39].

The conventional depolarization noise channel utilizes three Pauli matrices, (X, Y, and Z), to capture isotropic noise processes affecting quantum states [33]. Simulating the standard noisy channel requires three Kraus operators for each Pauli matrix, resulting in six matrix multiplications. However, as discussed by [17], this mathematical formalism introduces significant computational overhead, particularly in near-term systems with limited hardware resources.

To address these challenges, we propose a modified approach for single-qubit depolarization utilizing only two Kraus operators and the *X* and *Z* Pauli matrices. This reduced model simplifies the mathematical representation and decreases the computational complexity from six to four matrix multiplications for each noise channel execution. Our approach provides a more efficient means of simulating depolarization in resource-constrained quantum hardware, an essential capability in the NISQ era where computational resources are scarce. By developing simplified yet representative noise models, our work aims to enable more efficient and scalable approaches to simulating and correcting depolarization noise in deep quantum circuits.

## 1.1. Contribution

We summarize the contribution of our work as follows:

- 1. Depolarization Channel Representation: We propose a modified representation of the depolarization channel for single-qubit quantum states.
- 2. Kraus Operators Configuration: The proposed method contains only two Kraus operators.

- 3. Pauli Matrices Utilization: Unlike the standard approach that uses three Pauli matrices, our channel only uses two, *X* and *Z*, Pauli matrices.
- 4. Computational Efficiency: The proposed representation reduces the computational complexity from six to four matrix multiplications for each channel execution.
- 5. Theoretical Verification: We rigorously prove the validity of our proposed Kraus operators and the modified channel.
- 6. Experimental Validation: We empirically tested the proposed channel representation using a QML model on the Iris dataset. We evaluated the model performance across various circuit depths and depolarization rates.

## 1.2. Organization

The rest of the paper is organized as follows. Section 2 provides background on the standard depolarizing channel, and Section 3 derives the proposed modified channel. Section 4 experimentally analyzes the modified channel on the QML task on the Iris dataset. Section 5 discusses the implications of our method, and Section 6 concludes with a summary of our contributions and an outlook on future research directions.

## 2. Preliminaries

This section provides the preliminaries of quantum computing, quantum machine learning, and depolarizing channel representation. Chapter 3 and 8 from [33] is an excellent reference for quantum computing and quantum noise. Readers can refer to [40] for quantum machine learning and quantum neural networks. In the following, we provide a brief overview of the quantum state, quantum gates, depolarizing channel, and quantum machine learning.

# 2.1. Quantum State and Quantum Gates

A two-dimensional complex vector represents the single qubit quantum state. The quantum vector space for a single qubit in a computation basis space is spanned by the basis vectors  $|0\rangle$  and  $|1\rangle$ , which are represented as:

$$|0\rangle = \begin{bmatrix} 1\\ 0 \end{bmatrix}$$
 and  $|1\rangle = \begin{bmatrix} 0\\ 1 \end{bmatrix}$ . (1)

The general single qubit quantum state can be represented as:

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha |0\rangle + \beta |1\rangle,$$
 (2)

where  $\alpha$  and  $\beta$  are complex numbers, and  $|\alpha|^2 + |\beta|^2 = 1$ . Similarly, the quantum gates are the unitary operators that act on the quantum state to perform quantum operations. The Pauli matrices, *X*, *Y*, and *Z*, are the fundamental quantum gates in quantum computing. These single-qubit gates are at the heart of quantum computing and are used to perform quantum operations on the quantum state. Quantum gates and the quantum state form the building blocks of quantum circuits. This circuit is measured at the end to obtain the final output. Figure 1 shows an arbitrary single-qubit quantum circuit.

$$|0\rangle - H - \cdots - X$$

**Figure 1.** An arbitrary single qubit quantum circuit starting at  $|0\rangle$ , applying a Hadamard gate, followed by a sequence of unspecified single quantum gates, then a Pauli-X gate, and finally measurement.

The measurement operator O, also called observable, is used to measure the quantum state at the end of the quantum circuit. A Hermitian matrix represents the measurement operator. The measurement outcome is probabilistic and real, and the probability of

obtaining a particular outcome is given by the Born rule [33]. Given a quantum state  $|\psi\rangle$  and a measurement operator  $\mathcal{O}$ , the expectation value of  $\mathcal{O}$  is given by :

$$\langle \mathcal{O} \rangle = \langle \psi | \mathcal{O} | \psi \rangle. \tag{3}$$

In addition to the pure states described above, quantum systems can also be in mixed states. A mixed state is a statistical ensemble of pure states, and a density matrix represents it. The density matrix  $\rho$  for a quantum state  $|\psi\rangle$  is given by:

$$= \left|\psi\right\rangle \left\langle\psi\right|. \tag{4}$$

Similar to (3), the expectation value of the measurement operator O for a mixed state  $\rho$  is given by:

ρ

<

$$\mathcal{O}\rangle = \mathrm{Tr}\{\rho\mathcal{O}\},\tag{5}$$

where Tr denotes the trace operation. The density matrix representation is essential for modeling quantum noise and errors in quantum systems. In practice, quantum systems are often subjected to noise and errors that can affect the quantum state. The depolarizing channel is a standard tool for modeling such noise in quantum systems.

#### 2.2. Depolarization Channel Representation

The depolarization channel is a quantum channel that characterizes the loss of purity in quantum information due to specific types of noise. This channel randomly applies one of the Pauli operations or leaves the quantum state unchanged, each with certain probabilities. As a result, the quantum state becomes more mixed. The depolarization channel is important in quantum computing for modeling errors and understanding how quantum information degrades in real-world quantum systems. The standard expression of the depolarizing channel  $\mathcal{E}$  for a single qubit is given by:

$$\rho' = \mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z)$$
(6)

where *p* is the depolarization rate; and *X*, *Y*, and *Z* are the Pauli matrices, for the *X*, *Y*, and *Z* quantum gates, respectively.

Equation 6 implies that with probability (1 - p), the qubit state remains unchanged, and with probability p, it is subjected to equal mixtures of bit-flip, phase-flip, and both bit and phase-flip errors.

We can represent (6) using Kraus operators as:

$$\rho' = K_0 \rho K_0^{\dagger} + K_1 \rho K_1^{\dagger} + K_2 \rho K_2^{\dagger} + K_3 \rho K_3^{\dagger} = \sum_i K_i \rho K_i^{\dagger}, \tag{7}$$

where

$$K_0 = \sqrt{1-p}\mathbb{I}, \quad K_1 = \sqrt{\frac{p}{3}}X, \quad K_2 = \sqrt{\frac{p}{3}}Y, \quad K_3 = \sqrt{\frac{p}{3}}Z,$$

and  $\mathbb{I}$  is an identity matrix. As mentioned in the previous section, modeling such noise for practical applications is computationally expensive when we have limited hardware resources. We propose a modified representation of the depolarizing channel to address this challenge and improve the efficiency of simulating depolarization noise. We aim to reduce the computational complexity of the depolarization channel while maintaining its accuracy and effectiveness in modeling quantum noise. This improvement will enable more efficient and scalable simulations of quantum circuits under depolarization, advancing capabilities in the NISQ era. In Section 3, we derive an alternative representation of this channel.

#### 2.3. Quantum Machine Learning

Quantum Machine Learning is an emerging field that combines quantum computing and machine learning to develop quantum algorithms for solving computationally hard problems. It leverages the principles of quantum mechanics to enhance the performance of machine learning models, offering the potential to solve computationally intractable problems more efficiently than classical methods [40,41]. A parameterized quantum circuit model also called a variational circuit, can be trained as a QML model. The variational quantum circuit model is a QML model represented by a quantum circuit [40,42]. The model *f* is expressed mathematically as:

$$f_{\theta}(x) = \langle \psi(x,\theta) | \mathcal{O} | \psi(x,\theta) \rangle \tag{8}$$

Here,  $|\psi(x,\theta)\rangle$  represents a quantum state, which is generated by the parameterized quantum circuit  $U(x,\theta)$ . This circuit includes input vector x and trainable parameters  $\theta$  that can be optimized during training [43]. For a quantum system with n qubits, the circuit  $U(x,\theta)$  begins with a set of quantum gates starting from an initial state, typically  $|0\rangle^{\otimes n}$  [41]. The following equation can define the connection between the PQC and the unitary operation:

$$|\psi(x,\theta)\rangle = U(x,\theta) |0\rangle^{\otimes n}$$
(9)

During the model training process, the parameters  $\theta$  are optimized to minimize the cost function  $C(\theta)$  using a classical optimization algorithm. The goal of the training process is to find the optimal parameters  $\theta^*$  that minimize the cost function  $C(\theta)$  and maximize the model's performance. The noise introduced in the quantum system can affect the model's performance. In this work, we will analyze the effect of depolarizing noise on the QML model's performance. In Section 4, we will experimentally evaluate the performance of the QML model on the Iris dataset under the proposed modified depolarizing channel.

#### 3. Derivation

In this section, we derive an alternative representation of the depolarizing channel characterized by reduced matrix multiplication operations using only the *X* and *Z* Pauli matrices. We will also define the Kraus operators for the modified depolarizing channel and prove their validity. To establish the validity of the modified channel, we prove that the modified channel is equivalent to the standard depolarizing channel, the density matrix is preserved, and the Kraus operators are valid.

#### 3.1. Alternative Expression of the Depolarizing Channel

We define the alternative representation of the depolarizing channel as:

$$\rho'_m = (1 - \frac{2p}{3})\rho + \frac{2p}{3}Z((\rho X)^T X)Z.$$
(10)

In this representation, the state is partly retained with a coefficient of  $(1 - \frac{2p}{3})$  and partly subjected to a specific combination of Pauli *X* and *Z* operations with a coefficient of  $\frac{2p}{3}$ . This alternative expression is validated below to produce the same results as (6).

#### **Theorem 1.** Equations 6 and 10 are equivalent.

Proof. Consider the Pauli Matrices:

$$\mathbf{X} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \mathbf{Y} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \mathbf{Z} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \ \mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Consider an arbitrary single-qubit density matrix  $\rho = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ . Substituting  $\rho$  in (6) and (10) and with trivial algebraic work, we get:

$$\rho' = \rho'_m = \begin{bmatrix} -2a\frac{p}{3} + 2ad\frac{p}{3} & -4b\frac{p}{3} + b\\ -4c\frac{p}{3} + c & -2d\frac{p}{3} + d \end{bmatrix}.$$
(11)

Next, we will define Kraus operators and prove their validity.

**Theorem 2.** The following Kraus operators are valid operators for (10).

$$K_0 = \sqrt{1 - \frac{2p}{3}} \mathbb{I}, \quad K_1 = i \sqrt{\frac{2p}{3}} Z X.$$

**Proof.** From (10), one can immediately see that the corresponding Kraus operator corresponding to the term  $(1 - \frac{2p}{3})\rho$  is:  $\sqrt{1 - \frac{2p}{3}}\mathbb{I}$ . Now let us consider the second terminology of (10), i.e.,  $\frac{2p}{3}Z((\rho X)^T X)Z$ , which enables us to re-write without loss of generality the following:

$$\frac{2p}{3}Z((\rho X)^T X)Z = \frac{2p}{3}ZX\rho XZ \quad (\because \rho = \rho^T \text{ and } AX^T = XA), \text{ for any } A \in \mathbb{R}^{2\times 2}.$$

Next, we want a Kraus operator  $K_1$  s.t.  $K_1\rho K_1^{\dagger} = ZX\rho XZ$ . Thus, intuitively,

$$K_1 \propto ZX$$
,  $K_1 = xZX$ , (*x* is a scalar).

Following the Kraus operator completeness constraint, we can write:

$$K_0 K_0^{\dagger} + K_1 K_1^{\dagger} = \mathbb{I},$$

or

$$K_1 K_1^{\dagger} = \frac{2p}{3} \mathbb{I}. \tag{12}$$

To satisfy (12),  $K_1$  must have a magnitude of  $\sqrt{\frac{2p}{3}}$ . Therefore,

$$K_1 = \sqrt{\frac{2p}{3}} Z X.$$

We added "i" to correct a sign discrepancy while validating the operators, resulting in:

$$K_1 = i\sqrt{\frac{2p}{3}}ZX.$$

Next, we show that the derived Kraus operators are valid. We know that any set of Kraus operators satisfies the completeness property. That is,

$$\sum_{i} K_i^{\dagger} K_i = K_0^{\dagger} K_0 + K_1^{\dagger} K_1 = \mathbb{I}.$$

Solving each of the Kraus operators squared individually, we can get,

$$K_0^{\dagger}K_0 = (1 - \frac{2p}{3})\mathbb{I}, \quad K_1^{\dagger}K_1 = \frac{2p}{3}\mathbb{I},$$
$$K_0^{\dagger}K_0 + K_1^{\dagger}K_1 = (1 - \frac{2p}{3})\mathbb{I} + \frac{2p}{3}\mathbb{I} = \mathbb{I}.$$

This proves that the Kraus operators proposed in Theorem 2 are valid Kraus operators.  $\Box$ 

It follows from (7) that we can re-define (10) as:

$$\rho'_m = K_0 \rho K_0^{\dagger} + K_1 \rho K_1^{\dagger}. \tag{13}$$

In general, for a single qubit representation (6), (7), (10) and (13) yield the same result. The above derivations show that the modified depolarization channel expression is equivalent to the standard equation. We further proved that the proposed Kraus operators for (10) are valid Kraus operators. The next step is to show that the matrix given by the modified channel is a valid density matrix. We must prove that (10) is Hermitian, positive, semi-definite, and has a unit trace.

**Theorem 3.** *The matrix given by* (10) *is a valid density matrix.* 

**Proof.** First, we show that  $\rho'_m$  is Hermitian.

A matrix is Hermitian if it equals its conjugate transpose,  $A = A^{\dagger}$ .

To show  $\rho'_m$  is Hermitian, we calculate its conjugate transpose using its definition given by (13):

$$\begin{aligned} (\rho'_m)^{\dagger} &= (K_0 \rho K_0^{\dagger} + K_1 \rho K_1^{\dagger})^{\dagger} \\ &= (K_0 \rho K_0^{\dagger})^{\dagger} + (K_1 \rho K_1^{\dagger})^{\dagger} \\ &= (K_0^{\dagger})^{\dagger} \rho^{\dagger} K_0^{\dagger} + (K_1^{\dagger})^{\dagger} \rho^{\dagger} K_1^{\dagger} \quad (\because (AB)^{\dagger} = B^{\dagger} A^{\dagger}) \\ &= K_0 \rho K_0^{\dagger} + K_1 \rho K_1^{\dagger} \\ &= \rho'_m. \end{aligned}$$

Since  $K_0$  and  $K_1$  are constructed from unitary matrices and complex numbers, their conjugate transposes are their adjoints. Hence  $\rho'_m$  is Hermitian.

Second, we show that  $\rho'_m$  is Positive Semi-Definite.

Given the Kraus operators  $K_0$  and  $K_1$ , and density matrix  $\rho'_m$  we want to show that for any vector v, the expectation value  $v^{\dagger}\rho'_m v$  is non-negative.

Starting with the expression for  $v^{\dagger}\rho'_{m}v$ :

$$v^{\dagger}\rho'_{m}v = v^{\dagger}(K_{0}\rho K_{0}^{\dagger} + K_{1}\rho K_{1}^{\dagger})v.$$

This expands to:

$$v^{\dagger}\rho'_{m}v = v^{\dagger}K_{0}\rho K_{0}^{\dagger}v + v^{\dagger}K_{1}\rho K_{1}^{\dagger}v.$$

We can express each term as:

$$v^{\mathsf{T}}K_0\rho K_0^{\mathsf{T}}v = (K_0^{\mathsf{T}}v)^{\mathsf{T}}\rho(K_0^{\mathsf{T}}v),$$

$$v^{\mathsf{T}}K_1\rho K_1^{\mathsf{T}}v = (K_1^{\mathsf{T}}v)^{\mathsf{T}}\rho(K_1^{\mathsf{T}}v),$$

where  $w_0 = K_0^{\dagger} v$  and  $w_1 = K_1^{\dagger} v$  are vectors in the Hilbert space.

Since  $\rho$  is a positive semi-definite density matrix, we have for any vector w,  $w^{\dagger}\rho w \ge 0$ . Applying this to  $w_0$  and  $w_1$ :

$$(w_0)^{\dagger} \rho w_0 \ge 0,$$
  
 $(w_1)^{\dagger} \rho w_1 \ge 0.$ 

Therefore, the sum is also non-negative:

$$v^{\dagger}\rho'_{m}v = (w_{0})^{\dagger}\rho w_{0} + (w_{1})^{\dagger}\rho w_{1} \ge 0,$$

which establishes that  $\rho'_m$  is positive semi-definite.

Third, we show that  $\rho'_m$  has Unit Trace ,i.e.,  $\text{Tr}(\rho'_m) = 1$ . The trace of  $\rho'_m$  is given by:

$$\operatorname{Tr}(\rho_m') = \operatorname{Tr}(K_0 \rho K_0^{\dagger}) + \operatorname{Tr}(K_1 \rho K_1^{\dagger}).$$

Using the cyclic property of the trace, we can rewrite this as:

$$\operatorname{Tr}(\rho'_m) = \operatorname{Tr}(K_0^{\dagger} K_0 \rho) + \operatorname{Tr}(K_1^{\dagger} K_1 \rho).$$

Computing  $K_0^{\dagger}K_0$  and  $K_1^{\dagger}K_1$ , we get:

$$K_0^{\dagger} K_0 = (1 - \frac{2p}{3})\mathbb{I},$$
  
 $K_1^{\dagger} K_1 = -\frac{2p}{3}\mathbb{I},$ 

Thus, the trace of  $\rho'_m$  simplifies to:

$$\operatorname{Tr}(\rho'_m) = \operatorname{Tr}((1 - \frac{2p}{3})\mathbb{I}\rho) - \operatorname{Tr}(\frac{2p}{3}\mathbb{I}\rho).$$

As  $\rho$  is a density matrix with unit trace,  $\text{Tr}(\rho) = 1$ , this leads to:

$$Tr(\rho'_m) = (1 - \frac{2p}{3}) - \frac{2p}{3} = 1 - 2p + 2p = 1.$$

Therefore,  $\rho'_m$  maintains the unit trace property, as required for any density matrix.  $\Box$ 

Next, we derive the expression on how (10) evolves when the depolarization channel is applied *m* times on a quantum state  $\rho$  that leads us to the following lemma.

**Lemma 1.** When a depolarization channel with p depolarizing rate is applied on a single qubit quantum state  $\rho$  up to m times, the resulting quantum state is defined as follows up to first order in p:

$$\rho'_{mm} = (1 - \frac{2mp}{3})\rho + \frac{2mp}{3}Z((\rho X)^T X)Z + \mathcal{O}(p^2).$$
(14)

And, for an observable O, the expectation value is given as:

$$\langle O \rangle_{\rho'_{mm}} = Tr\{O\rho'_{mm}\} = Tr\{O\rho\} - \frac{2mp}{3}Tr\{O\rho\} + \frac{2mp}{3}Tr\{OZX^T\rho^TXZ\}.$$
(15)

In the following section, we verify that the difference between the results obtained from (15) and the measurement of O on standard channel simulations are negligible. We will demonstrate through experimental evidence that (10), (14) and (15) can be effectively used for training machine learning models.

# 4. Experiment

We start this section by showing that the results from (15) are consistent with simulation results of (6) for multiple values of m and p. Later, we empirically show that the Depolarization rate up to threshold w does not affect the performance of the single qubit QML model for the iris dataset. For the scope of this experiment, we are considering the binary classification. We used the first, Setosa, and third, Virginia, flower classes and only the first two features, Sepal length and sepal width. We experimented on Python with Pennylane for quantum circuit simulations and quantum computation. Otherwise mentioned, we used (10) for the depolarization channel and (15) for the depolarization channel applied m times.

#### 4.1. Quantum Circuit Behavior Analysis under Depolarization Channel up to m Times

Equation 15 posits that when a depolarization channel with a rate p is applied m times to a quantum state  $\rho$ , the resultant state  $\rho'_{mm}$  adheres to a predictable transformation that maintains linearity with respect to p in the first order for small value of p. This suggests that despite the iterative application of noise, the overall system's behavior under the

depolarization channel can be approximated linearly. We conducted a series of simulations to substantiate this theory. We computed the expectation values for a single qubit with varying depths—specifically, 3, 8, and 15 gates. The behavior is assessed across different depolarization rates, p = [0.0, 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01], and depolarization channel repetition, m = [0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

Figure 2 presents a scatter plot visualization for the expectation value differences between (14) and the standard depolarization model as a function of both p and m for Pauli-Z observables. A minimal deviation exists between the standard and modified channels' expectation value for low depolarization rates across all gate counts. This alignment implies that the modified equation retains fidelity to the standard model's predictions in the low-noise regime. There is a uniform trend across the plots where, for small values of p, the difference in expectation values is negligible across all values of m. This insignificant difference remains consistent as the number of gates increases, emphasizing the robustness of the modified model. For a three quantum gates system, the difference between standard and modified channel expectation values across all combinations is 0. When we increased the gate count to 8, the difference had a mean of 0.00023 and a standard deviation of 0.00038. Only about 6% of computed differences were above 0.001. However, for a 15 gates system, the difference has a mean of 0.00065 and a standard deviation of 0.00106. About 22% of computed differences were above 0.001. We present this information in Table 1. We observe that the difference between the standard and modified channels' expectation values increases as the number of gates increases. Regardless, the difference remains minimal, with the majority of values below 0.001. This analysis demonstrates that the modified depolarization channel is a reliable and efficient alternative to the standard model for simulating quantum circuits under depolarization noise.



**Figure 2.** Scatter plots present the difference between the standard channel and modified depolarization channel expectation value. Each channel was applied to a quantum circuit with single qubit gates of 3, 8, and 15, respectively. The result for 3 single qubit gates is presented in plot (**a**), while plot (**b**) and (**c**) represent the results for 8 and 15 gates circuits, respectively. The *x*-axis of each plot represents the number of times the noisy channel was applied and is given by *m*, while the *y*-axis gives the varying depolarization rates.

Table 1. Comparison of Differences between Standard and Modified Channels expectation values.

Gate Count	Mean Difference	Standard Deviation	% of Differences above 0.001
3	0	0	0%
8	0.00023	0.00038	6%
15	0.00065	0.00106	22%

Extending on these results, we analyze the training of the QML model on the Iris dataset under the modified depolarization channel. First, we map the classic data into quantum Hilbert space via a feature map.

(a)

## 4.2. Data Encoding

We start by initializing the qubit in a computational basis  $|0\rangle$ . Let  $\phi : x \mapsto \phi(x)$  be a mapping from input space X to a quantum Hilbert Space  $\mathbb{R}$ . Let us define  $\phi(x) = RX(x_1)RY(x_0)$ , where  $x_0$  and  $x_1$  are features of the input vector, and

$$RY(x_0) = \begin{bmatrix} \cos(\frac{x_0}{2}) & -\sin(\frac{x_0}{2}) \\ \sin(\frac{x_0}{2}) & \cos(\frac{x_0}{2}) \end{bmatrix},$$
$$RX(x_1) = \begin{bmatrix} \cos(\frac{x_1}{2}) & -i\sin(\frac{x_1}{2}) \\ -i\sin(\frac{x_1}{2}) & \cos(\frac{x_1}{2}) \end{bmatrix},$$

are rotational single qubit quantum gates. We chose the angle encoding scheme because it linearly separates input data in the Bloch Sphere better than amplitude encoding, as shown in Figure 3.





**Figure 3.** Feature Mapping of the Iris dataset using Amplitude Encoding and Rotational encoding method. The Rotational encoding scheme, a combination of *RX* and *RY*, provides better mapping results for the classification problem. The red color represents Class 1, the blue color represents Class 2, and the green color represents Class 3. (a) Bloch Sphere representation of the quantum states obtained by Amplitude encoding of the features vectors. (b) Bloch Sphere representation of the quantum states obtained by Angle encoding of the features vectors.

Motivated by the result in Figure 4, we choose RX, RY rotational gates for feature encoding. Thus, for an input vector x, the data-encoded state is defined as:

$$|\psi\rangle = \phi(x) |0\rangle = RX(x_1)RY(x_0) |0\rangle.$$
(16)

Our learning circuit consists of the encoding and variational layers with trainable parameters. In the next section, we describe this variational circuit.

## 4.3. Variational Layers

Similar to the encoding scheme, we applied a series of variational gates (RY and RX with parameters) whose parameters  $\theta$  can be optimized during training. For *N* trainable parameters  $\theta$ , we define the operation of Variational layers *U* as:

$$U(\theta) = \prod_{i=1}^{N} \text{Gate}_{i}(\theta_{i}),$$
(17)

where  $Gate_i(\theta_i)$  is either an *RY* or *RX* gate with parameter  $\theta_i$ . Thus, we can define a variational circuit as:

$$|\Psi\rangle = U(x,\theta) = U(\theta) |\psi\rangle.$$
(18)

Let  $\rho = |\Psi\rangle \langle \Psi|$  be a density matrix. The system undergoes evolution through a depolarization channel. This channel, denoted as  $\mathcal{E}(\rho)$ , transforms the state be  $\rho$  of the qubit by mixing it towards a maximally mixed state as the depolarization rate p increases as given by (6) and (10).



**Figure 4.** Various encoding schemes for single qubits using the rotational encoding. The combination of *RZ* and *RX* gates provides the best mapping for binary classification. The red color represents Class 1, the blue color represents Class 2, and the green color represents Class 3.

Let this state  $\rho'$ . Now we measure an observable Pauli $-Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$  matrix to get a quantum machine learning model that can be defined as:

$$f(x,\theta) = \operatorname{Tr}\{Z\rho'\}.$$
(19)

In the following section, we train this QML model on the Iris dataset under a modified depolarization channel and discuss the results.

#### 4.4. Training

We trained the model for various values of m = [1, 3, 5, 10, 15] and p = [0.0, 0.001, 0.005, 0.005]0.01, 0.05, 0.08, 0.1, 0.5]. This wide range of values was selected to understand the model's behavior under varying noise levels and circuit depths through an empirical trial-and-error approach. The initial learnable parameters were chosen randomly, a common practice in machine learning to break symmetry and ensure diverse initial weights. We used the Adam optimizer because it effectively handles sparse gradients on noisy problems [44]. We determined through empirical testing that a step size of 0.1 effectively balanced convergence speed and stability. We used the standard square loss as the loss function and a parameter-shift rule to compute the gradients. The parameter-shift rule provides an analytical method to estimate gradients of quantum circuits, well-suited for training models involving quantum components in the NISQ era when traditional backpropagation is not feasible [45]. We trained for 30 epochs for each combination of p and m based on empirical observations of the model's convergence speed and performance. The script ran for approx. 3842 seconds to train the model on an M1 pro chip with 10 cores and 32 GB memory for the standard depolarization channel and approximately 2492 seconds for the modified depolarization channel. We discuss the gain in computational efficiency in Section 5.

The pictorial representation of the depolarization noise at each layer of the circuit and its influence on the QML model training is presented in Figure 5. The decision boundaries in the figure's right column illustrate the model's behavior across varying noise and circuit depth levels. The blue data points are from the Setosa class, while the red points represent the Virginia class. Since the possible measurement values for Z observable are between -1and 1, we labeled Setosa as -1 and Virginia as 1. The decision boundary is the line that separates the two classes. Similarly, the blue decision boundary represents the model's prediction for the Setosa class, while the red represents the Virginia class. We observe that with an increase in circuit depth, while the expressivity of the model may enhance, there is a concurrent increase in its vulnerability to noise, adversely affecting the quality of the decision boundary. Notably, the model's performance appears to be robust for circuit depths ranging from 1-5. The model can create a decision boundary at a depth of 5, even when the depolarization rate is 0.1 or at a depth of 15 for a 0.05 error rate. To determine the impact of noise on model training, we can analyze the training dynamics of the model as shown in the left column of Figure 5. The graphs display the model's accuracy and loss over 30 epochs for varying noise levels and circuit depths. The results indicate that the model's performance is consistent across all circuit depths for noise levels up to 0.5. This consistent performance of the model at lower noise levels across all circuit depths indicates that QML models can be robust to noise up to a certain threshold.

The model has difficulty making accurate predictions, even with one or three trainable gates when the noise level is 0.5 or higher. This indicates a possible threshold for both circuit depth and noise level that maximizes QML model performance. In our experiments, reaching a plateau at a depth of five suggests that the model's capacity for feature representation may be sufficiently saturated by a depth of five. Beyond this point, we observed a decline in precision, F1 score, and accuracy. These results suggest that while QML models exhibit robustness in lower noise environments and at shallower circuit depths, their performance diminishes with increased circuit complexity and higher noise levels.

In the next section, we discuss the advantages and limitations of a modified depolarization channel and the trade-off between model complexity and noise resilience.



**Figure 5.** Experimental results for decision boundary evolution presented in the right column and training dynamics in the left column for a QML model on the Iris dataset, with varied noise levels (p) and depolarization channel applied up to (m) times. The decision boundaries are plotted for depths of 1, 3, 5, 10, and 15, at noise levels ranging from 0.0 to 0.5. The results across rows are presented in chronological order in circuit depth. Accuracy and loss graphs display the model's performance over 30 epochs, highlighting the impact of noise rate and circuit depth on learning efficacy.

#### 5. Discussions

We argue that (10) and (15) provide the computation advantage for simulating depolarization noise. As long as p is small, similar to (15), we can define the density matrix of a system under depolarization channel up to m times for (6) as follows:

$$\rho'_{m} = (1 - mp)\rho + \frac{mp}{3}(X\rho X + Y\rho Y + Z\rho Z) + \mathcal{O}(p^{2}).$$
<sup>(20)</sup>

The expectation value of an observable *O* can be defined as:

$$\langle O \rangle_{\rho'_m} = \operatorname{Tr}\{O\rho'_m\} = \operatorname{Tr}\{O\rho\} - mp\operatorname{Tr}\{O\rho\} + \frac{mp}{3}[\operatorname{Tr}(OX\rho X) + \operatorname{Tr}(OY\rho Y) + \operatorname{Tr}(OZ\rho Z)].$$
(21)

We studied to analyze the computational efficiencies of two depolarizing channel models applied to quantum states. These models were given by (6) and (10). We focused on the matrix multiplication overhead and the operational requirements, specifically using Pauli gates.

Our findings suggest that the standard depolarizing channel, given by (6) and (20), requires a higher computation. It requires six matrix multiplications for a single state evolution and an additional four for computing expectation values, as shown in (21), totaling ten multiplications. This approach utilizes (X, Y, Z) Pauli gates. The use of all Pauli gates can introduce complexity in gate operations and potential errors in practical quantum computing environments.

On the other hand, the modified depolarizing channel, given by (10) and (14), presents an efficient alternative. It requires only four matrix multiplications for a single state evolution and two additional ones for expectation value computations, according to (15), totaling six multiplications. The modified model also eliminates the need for direct Pauli Y-gate applications, thereby simplifying the operational framework. Let us consider that n is the number of qubits in a system. Since the quantum state space grows exponentially, let  $d = 2^n$  represent the dimension of  $\rho$ . If we assume the computational complexity of matrix multiplication generally scales at  $O(d^3)$ , the standard depolarizing channel requires  $O(10d^3)$  operations, while the modified channel requires  $O(6d^3)$  operations. Although the overall time complexity is the same due to the exponential nature of the complexity, in practice, the modified model would still be faster due to the difference in time complexity coefficients. As the number of qubits increases, the computational overhead for both models will grow exponentially. However, the modified model will always have a lower computational overhead than the standard model due to the fewer matrix multiplications required.

In practical scenarios like cloud-based quantum computing environments, users often face lengthy queues, leading to extended wait times. These queues usually span for hours to execute a single operation. By reducing the matrix multiplications, our approach effectively reduces the computational load by at least 16 multiplication and 8 addition operations. Considering that a typical user might only perform one operation per hour on real quantum hardware due to these queuing constraints, our method could result in substantial time savings. This makes it ideal for quantum algorithms where efficiency and gate operation minimization are essential.

We further argue that we can extend the modified channel to train the QML model. The results on QML model behavior demonstrate the nuanced interplay between circuit depth, noise levels, and the model's performance. The results suggest that there exists a level of quantum circuit complexity where the representational power of the model is optimal. However, as we extend the circuit depth beyond this optimal point, we observe diminishing returns in model performance, highlighting a critical trade-off between the expressiveness of deeper quantum circuits and their susceptibility to noise. In this study, we proposed a modified depolarization channel that reduces the computational overhead and operational complexity compared to the standard depolarization channel. The modified channel requires only two Kraus operators and eliminates the need for direct Pauli Y-gate applications. By reducing the number of matrix multiplications, the modified channel offers a more efficient alternative for simulating depolarization noise.

The significance of the modified channel lies in its ability to reduce computational overhead and operational complexity. It requires only four matrix multiplications for state evolution and two additional ones for expectation value computations. This reduction in computational load can lead to substantial time savings, particularly in practical scenarios such as cloud-based quantum computing environments with lengthy queues.

Furthermore, we demonstrated that we can efficiently extend the modified channel to train QML models. Our experiments on the Iris dataset classification showed that as circuit complexity and noise levels increase, the performance of the QML model diminishes. This highlights the critical trade-off between the expressiveness of deeper quantum circuits and their susceptibility to noise.

It is important to note that while our work applies only to small values of *p*, we believe it lays a solid foundation for future research. Future research will focus on expanding this approach to multi-qubit systems and exploring its effectiveness with higher values of *p*. We believe that our work will inspire further research in this area and contribute to advancing quantum computing and quantum machine learning.

**Author Contributions:** B.K. and P.R. contributed equally to this work. B.K. conceptualized the study, developed the methodology, and performed the experiments. P.R. supervised the study, provided resources, and reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** Part of this work was performed while P.R. was funded by the National Science Foundation under Grant Nos. CNS-2210091, and CHE-1905043.

**Data Availability Statement:** Data sharing does not apply to this article as no datasets were generated or analyzed during the current study.

Acknowledgments: We would like to thank the reviewers for their valuable feedback during the peer review process.

Conflicts of Interest: The authors declare no conflict of interest.

### References

- 1. Mitarai, K.; Negoro, M.; Kitagawa, M.; Fujii, K. Quantum circuit learning. *Phys. Rev. A* 2018, 98, 032309.
- Havlíček, V.; Córcoles, A.D.; Temme, K.; Harrow, A.W.; Kandala, A.; Chow, J.M.; Gambetta, J.M. Supervised learning with quantum-enhanced feature spaces. *Nature* 2019, 567, 209–212.
- 3. Liu, Y.; Arunachalam, S.; Temme, K. A rigorous and robust quantum speed-up in supervised machine learning. *Nat. Phys.* 2021, 17, 1013–1017.
- 4. Sajjan, M.; Sureshbabu, S.H.; Kais, S. Quantum machine-learning for eigenstate filtration in two-dimensional materials. *J. Am. Chem. Soc.* **2021**, *143*, 18426–18445.
- Cai, X.D.; Wu, D.; Su, Z.E.; Chen, M.C.; Wang, X.L.; Li, L.; Liu, N.L.; Lu, C.Y.; Pan, J.W. Entanglement-based machine learning on a quantum computer. *Phys. Rev. Lett.* 2015, 114, 110504.
- Ciliberto, C.; Herbster, M.; Ialongo, A.D.; Pontil, M.; Rocchetto, A.; Severini, S.; Wossnig, L. Quantum machine learning: A classical perspective. *Proc. R. Soc. A Math. Phys. Eng. Sci.* 2018, 474, 20170551.
- 7. Farhi, E.; Goldstone, J.; Gutmann, S. A quantum approximate optimization algorithm. arXiv 2014, arXiv:1411.4028.
- 8. McClean, J.R.; Romero, J.; Babbush, R.; Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **2016**, *18*, 023023.
- 9. Rebentrost, P.; Schuld, M.; Wossnig, L.; Petruccione, F.; Lloyd, S. Quantum gradient descent and Newton's method for constrained polynomial optimization. *New J. Phys.* **2016**, *21*, 073023. https://doi.org/10.1088/1367-2630/ab2a9e.
- Bittel, L.; Kliesch, M. Training Variational Quantum Algorithms Is NP-Hard. *Phys. Rev. Lett.* 2021, 127, 120502. https://doi.org/10.1 103/PhysRevLett.127.120502.
- 11. Rebentrost, P.; Lloyd, S. Quantum computational finance: Quantum algorithm for portfolio optimization. *arXiv* 2018, arXiv:1811.03975.

- 12. Broadbent, A.; Schaffner, C. Quantum cryptography beyond quantum key distribution. *Des. Codes Cryptogr.* 2015, *78*, 351–382. https://doi.org/10.1007/s10623-015-0157-4.
- Padamvathi, V.; Vardhan, B.; Krishna, A.V. Quantum Cryptography and Quantum Key Distribution Protocols: A Survey. In Proceedings of the 2016 IEEE 6th International Conference on Advanced Computing (IACC), Bhimavaram, India, 27–28 February 2016; pp. 556–562. https://doi.org/10.1109/IACC.2016.109.
- Lai, H.; Luo, M.; Pieprzyk, J.; Zhang, J.; Pan, L.; Li, S.; Orgun, M. Fast and simple high-capacity quantum cryptography with error detection. *Sci. Rep.* 2017, 7, 46302. https://doi.org/10.1038/srep46302.
- 15. Pirandola, S.; Andersen, U.; Banchi, L.; Berta, M.; Bunandar, D.; Colbeck, R.; Englund, D.; Gehring, T.; Lupo, C.; Ottaviani, C.; et al. Advances in Quantum Cryptography. *Adv. Opt. Photonics* 2020, *12*, 1012–1236. https://doi.org/10.1364/AOP.361502.
- Harrow, A.; Montanaro, A. Quantum computational supremacy. *Nature* 2017, 549, 203–209. https://doi.org/10.1038/nature23458.
   Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* 2018, *2*, 79.
- Du, Y.; Hsieh, M.; Liu, T.; You, S.; Tao, D. Learnability of quantum neural networks. PRX Quantum 2021, 2, 040337. https://doi.org/10.1 103/PRXQuantum.2.040337.
- Khanal, B.; Rivas, P.; Orduz, J.; Zhakubayev, A. Quantum machine learning: A case study of grover's algorithm. In Proceedings of the 2021 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 15–17 December 2021; IEEE: Piscateville, NJ, USA, 2021; pp. 79–84.
- Cross, A.; Smith, G.; Smolin, J. Quantum learning robust against noise. *Phys. Rev. A* 2014, 92, 012327. https://doi.org/10.1103/ PhysRevA.92.012327.
- Du, Y.; Hsieh, M.; Liu, T.; Tao, D.; Liu, N. Quantum noise protects quantum classifiers against adversaries. *Phys. Rev. Res.* 2021, 3, 023153. https://doi.org/10.1103/PhysRevResearch.3.023153.
- Huang, J.; Tsai, Y.; Yang, C.; Su, C.; Yu, C.M.; Chen, P.Y.; Kuo, S.Y. Certified Robustness of Quantum Classifiers against Adversarial Examples through Quantum Noise. In Proceedings of the ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 4–10 June 2023.
- 23. West, M.T.; Tsang, S.L.; Low, J.S.; Hill, C.D.; Leckie, C.; Hollenberg, L.C.; Erfani, S.M.; Usman, M. Towards quantum enhanced adversarial robustness in machine learning. *Nat. Mach. Intell.* **2023**, *5*, 581–589.
- 24. Lu, S.; Duan, L.M.; Deng, D.L. Quantum adversarial machine learning. Phys. Rev. Res. 2020, 2, 033212.
- 25. Skolik, A.; Mangini, S.; Bäck, T.; Macchiavello, C.; Dunjko, V. Robustness of quantum reinforcement learning under hardware errors. *EPJ Quantum Technol.* **2023**, *10*, 8. https://doi.org/10.1140/epjqt/s40507-023-00166-1.
- 26. Bai, T.; Luo, J.; Zhao, J.; Wen, B.; Wang, Q. Recent advances in adversarial training for adversarial robustness. *arXiv*, **2021**, arXiv:2102.01356.
- Kang, D.; Sun, Y.; Brown, T.; Hendrycks, D.; Steinhardt, J. Transfer of adversarial robustness between perturbation types. *arXiv*, 2019, arXiv:1905.01034.
- 28. Khanal, B.; Rivas, P. Evaluating the Impact of Noise on Variational Quantum Circuits in NISQ Era Devices. In Proceedings of the International Conference on Emergent and Quantum Technologies (ICEQT 2023), Las Vegas, NV, USA, 24–27 July 2023; pp. 1–7.
- 29. Huang, H.Y.; Broughton, M.; Mohseni, M.; Babbush, R.; Boixo, S.; Neven, H.; McClean, J.R. Power of data in quantum machine learning. *Nat. Commun.* **2021**, *12*, 2631.
- 30. Wang, X.; Du, Y.; Luo, Y.; Tao, D. Towards understanding the power of quantum kernels in the NISQ era. Quantum 2021, 5, 531.
- Khanal, B.; Rivas, P.; Orduz, J. Kernels and Quantum Machine Learning. In Proceedings of the International Conference on Emergent and Quantum Technologies (ICEQT 2022), Las Vegas, NV, USA, 25–28 July 2022; pp. 1–15.
- Piskor, T.; Reiner, J.; Zanker, S.; Vogt, N.; Marthaler, M.; Wilhelm, F.K.; Eich, F.G. Using gradient-based algorithms to determine ground-state energies on a quantum computer. *Phys. Rev. A* 2022, 105, 062415. https://doi.org/10.1103/PhysRevA.105.062415.
- Nielsen, M.A.; Chuang, I.L. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.
   Wootton, J.R.; Loss, D. High threshold error correction for the surface code. *Phys. Rev. Lett.* 2012, 109, 160503. https://doi.org/10.110 3/PhysRevLett.109.160503.
- 35. Fowler, A.G.; Mariantoni, M.; Martinis, J.M.; Cleland, A.N. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A* **2012**, *86*, 032324.
- 36. Gottesman, D. Stabilizer Codes and Quantum Error Correction; California Institute of Technology: Pasadena, CA, USA, 1997.
- 37. Urbanek, M.; Nachman, B.; Pascuzzi, V.R.; He, A.; Bauer, C.W.; de Jong, W.A. Mitigating depolarizing noise on quantum computers with noise-estimation circuits. *Phys. Rev. Lett.* **2021**, *127*, 270502.
- Cai, Z. Multi-exponential error extrapolation and combining error mitigation techniques for NISQ applications. *Npj Quantum Inf.* 2020, 7, 80. https://doi.org/10.1038/s41534-021-00404-3.
- Haug, T.; Self, C.; Kim, M. Quantum machine learning of large datasets using randomized measurements. *Mach. Learn. Sci. Technol.* 2023, 4, 015005. https://doi.org/10.1088/2632-2153/acb0b4.
- 40. Schuld, M.; Petruccione, F. Machine Learning with Quantum Computers; Springer: Berlin/Heidelberg, Germany, 2021.
- 41. Biamonte, J.; Wittek, P.; Pancotti, N.; Rebentrost, P.; Wiebe, N.; Lloyd, S. Quantum machine learning. Nature 2017, 549, 195–202.
- 42. Wittek, P. Quantum Machine Learning: What Quantum Computing Means to Data Mining; Academic Press: Cambridge, MA, USA, 2014.
- 43. Benedetti, M.; Lloyd, E.; Sack, S.; Fiorentini, M. Parameterized quantum circuits as machine learning models. *Quantum Sci. Technol.* **2019**, *4*, 043001.

- 44. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. arXiv 2014, arXiv:1412.6980.
- 45. Schuld, M.; Bergholm, V.; Gogolin, C.; Izaac, J.; Killoran, N. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A* **2019**, *99*, 032331.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.