



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/243093/>

Version: Published Version

Article:

Umeano, C., Scali, S. and Kyriienko, O. (2026) Geometric quantum machine learning for barcode similarity classification. Physical Review A, 113 (5). 052425. ISSN: 2469-9926

<https://doi.org/10.1103/sylc-gclc>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Geometric quantum machine learning for barcode similarity classification

Chukwudubem Umeano¹, Stefano Scali¹, and Oleksandr Kyriienko^{1,2}

¹*Department of Physics and Astronomy, University of Exeter, Stocker Road, Exeter EX4 4QL, United Kingdom*

²*School of Mathematical and Physical Sciences, University of Sheffield, Sheffield S3 7RH, United Kingdom*

 (Received 6 February 2025; revised 10 March 2026; accepted 20 March 2026; published 12 May 2026)

We consider the problem of distinguishing two vectors (visualized as images or barcodes) and learning if they are related to one another. For this, we develop a geometric quantum machine learning (GQML) approach with embedded symmetries that allows for the classification of similar and dissimilar pairs based on global correlations, and enables generalization from just a few samples. Unlike GQML algorithms developed to date, we propose to focus on symmetry-aware measurement adaptation that outperforms unitary parametrizations. We compare GQML for similarity testing against classical deep neural networks and convolutional neural networks with Siamese architectures. We show that quantum networks demonstrate strong empirical performance gains over their classical counterparts in generalization. We explain this difference in performance by analyzing correlated distributions used for composing our dataset. We relate the similarity testing to problems that showcase a proven maximal separation between the bounded-error quantum polynomial time complexity class and the polynomial hierarchy. While the ability to achieve a formal advantage depends on how data are loaded, we discuss how similar problems can benefit from quantum machine learning. Finally, we present a 40-qubit hardware implementation of our quantum model, using a superconducting processor (IBM Kyiv), showing remarkable scalability and resilience to noise.

DOI: [10.1103/syhc-gclc](https://doi.org/10.1103/syhc-gclc)

I. INTRODUCTION

Machine learning (ML) became a crucial tool for information processing [1], bringing us closer to advanced artificial intelligence [2,3]. Its further progress largely depends on the development of qualitatively new hardware for computing and algorithmic breakthroughs [4]. From this perspective, quantum machine learning (QML) [5]—an area of ML devoted to developing quantum computing algorithms to solve learning-based problems—has attracted much attention in recent years [6–8]. Generally motivated by the capability of quantum computers of getting a superpolynomial improvement in scaling as compared to classical algorithms [9–11], QML has evolved both in terms of applications to be considered, and approaches to be used. Notably, the adoption of variational quantum algorithms (VQAs) [12,13] led to the variety of quantum circuit optimization-based protocols, applied to tasks such as classification [7,14–18], phase recognition [19–21], generative modeling [22–29], and physics-informed solving of differential equations [30–35], showing potential in areas where classical methods struggle. A distinct set of tools has been developed for embedding data into quantum computers [36–38], including various feature maps designed to exploit the high-dimensional space of quantum states for enhanced data representation and processing efficiency [39–45]. The

use-cases of QML include image recognition [46–49], data analysis in high-energy physics [50–54], financial risk assessment [21,55,56], fraud detection [57–60], natural language processing [61–63], graph-based learning [64–67], power flow analysis [68], enhanced molecular simulations in chemistry [69], among many others.

Several open questions are yet to be answered before QML delivers its full promise. First, we need to decide how to parametrize quantum models in the most advantageous way. Do we get an advantage from training variational circuits, or is quantum embedding enough [70,71]? Or shall we use the structure of quantum kernels [72–74]? Second, we still need to understand where a proven advantage in quantum machine learning can come from [75]. Does it mainly concern improving runtime and circuit depth [6,76,77]? Does it come from sampling (generative modeling) [26,78–81], including proven formal separations for learning quantum data distributions [82], or do we gain from better generalization of quantum models [19,83–86]?

One idea fueling efficient quantum models leans on using symmetries of considered datasets and problems [87–89]. This approach, inspired by geometric deep learning [90], was dubbed as geometric quantum machine learning (GQML) [91–94]. GQML excels in problems with apparent symmetries and offers a guide for optimal ansatz parametrizations. But we need to keep in mind that symmetries also help classical ML making problems more tractable [95,96].

To answer the second question several approaches were suggested. One can relate QML with probably approximately correct (PAC) learning [97], proving a superpolynomial advantage for density and Boolean function learning [98,99],

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

and for specific concept classes [73,100]. Another approach to proving quantum-classical separation relies on connecting QML models to some of the known quantum algorithms that have exponential separation. This includes developing quantum kernels with a provable advantage based on bounded-error quantum polynomial time (BQP) problems [101,102] and kernel-based quantum models with Grover's advantage [103]. Recently, we have shown that geometric quantum machine learning can be used to learn Simon's algorithm in BQP^A [104], suggesting that both QML architectures and optimal circuits can be learnt to represent known solutions with proven advantages.

In this paper, we apply GQML to the similarity testing problem. We show that for data sampled from specific distributions, certain classes of GQML models exhibit strong empirical generalization benefits over classical learners, discovering an optimal decision boundary with just a few samples. We set up the problem as distinguishing pairs of pixelated images—barcodes—each represented by a binary string. The task is to determine which pairs are related to each other (i.e., coming from a correlated distribution), or are dissimilar. We develop several architectures for an equivariant quantum neural network (QNN) to address the task. We show that, while the unitarily parametrized QNN does not train well on the supplied dataset, the adaptive measurement-based QNN trains and provides strong empirical generalization benefits. Embracing its remarkable scalability and resilience to noise, we are able to run a 40-qubit realization of the model on IBM Kyiv. Finally, we explain the results based on established problems for superpolynomial separation between BQP^A and polynomial hierarchy (PH).

II. THE PROBLEM

We consider a dataset $\mathcal{D}$ consisting of samples labeled as correlated ($y_m = 0$, class A) and uncorrelated ($y_m = 1$, class B). The dataset $\mathcal{D} = \{x_m = (\mathbf{x}_{1,m}, \mathbf{x}_{2,m}), y_m\}_{m=1}^M$ of M samples includes information about x_m being pairs of binary strings $\mathbf{x}_{1,m} \in \mathcal{X}$ and $\mathbf{x}_{2,m} \in \mathcal{X}$, where $\mathcal{X} = \{0, 1\}^N$. The first part of the pair (e.g., $\mathbf{x}_{1,1} = 010001\dots$) and the second part ($\mathbf{x}_{2,1} = 101100\dots$) need to be loaded and processed in parallel. Classically, these samples can be represented as two-dimensional barcodes with solid and open squares. The task is to establish a machine learning workflow for the similarity testing between the pairs (Fig. 1). We are given the information that samples come from latent distributions as $x_{|y=0} \sim P_A(x)$ and $x_{|y=1} \sim P_B(x)$, where $x \in \mathcal{X} \times \mathcal{X} =: \mathcal{X}^2$. Therefore, the underlying goal is to distinguish the probability distributions $P_A(x)$ and $P_B(x)$ in terms of their properties related to *global* correlations between samples coming in pairs.

We suggest that the data can be efficiently encoded into quantum states using the phase states $|\phi_x\rangle := (\sum_{j=1}^N (-1)^{x_1^{[j]}} |j\rangle) \otimes (\sum_{j=1}^N (-1)^{x_2^{[j]}} |j\rangle) / N =: |\phi_{x_1}\rangle \otimes |\phi_{x_2}\rangle$, where $\mathbf{x}_\alpha^{[j]}$ ($\alpha = 1, 2$) denote the j th digit of the binary converted into a $0/\pi$ phase in front of the corresponding computational basis state [10]. In this case, N -digit strings are compressed into $n = \lceil \log_2 N \rceil$ qubit registers, and we need $2n$ qubits for loading a sample x_m . We are also given an associated symmetry $\mathcal{S}_\Phi$ that corresponds to a remapping of bitstrings according to rules given by particular Boolean

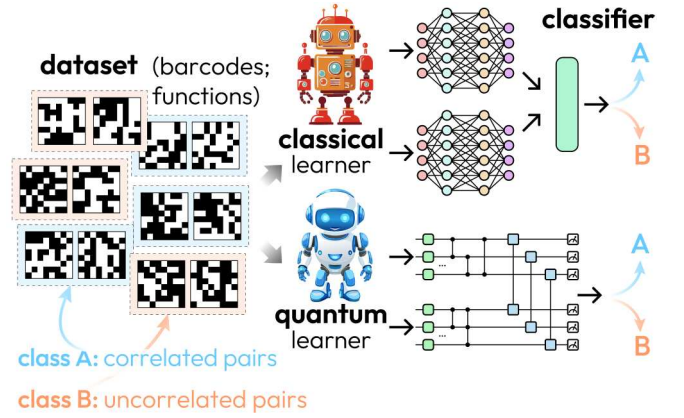


FIG. 1. Sketch of the machine learning workflows for distinguishing images (barcodes) based on their correlations. Pairs of barcodes are processed by a classical learner via Siamese deep neural networks. A quantum learner processes data embedded into quantum states via parametrized equivariant measurements, classifying barcodes as correlated or uncorrelated.

circuits, such that x_k and $\sigma(x_k)$ for $\sigma \in \mathcal{S}_\Phi$ share the same label y_k . This has the representation of $\sigma : x_k \mapsto \sigma(x_k) \cong \bar{x}_k$, where $\bar{x}_k = (\bar{x}_{1,k}, \bar{x}_{2,k})$ is the complement of x_k , crucially containing the same global correlations. At this point, we are ready to propose machine learning approaches that can address the problem.

III. MODELS AND ARCHITECTURES

Next, we detail distinct machine learning approaches and neural network architecture that can be used to solve the problem of similarity testing. Specifically, we describe several QNNs and loss functions originating from geometric quantum machine learning. Then, we describe the classical workflows to set a baseline for benchmarking.

For the GQML, we consider two different options. The first is the standard formalism for GQML that has been detailed in various studies [87–89,91,105]. We wish to determine whether or not each pair of barcodes x_m drawn from the dataset $\mathcal{D}$ is correlated. In other words, we have a target function $g : \mathcal{X} \times \mathcal{X} \rightarrow \mathcal{Y}$ mapping samples to the binary domain such that $g(x_m) = y_m$, and we wish to train a model to approximate g across all inputs x_m . The quantum model is defined as

$$h_\theta(x_m) = \langle \psi_0 | \hat{\mathcal{U}}_\phi^\dagger(x_m) \hat{W}^\dagger(\theta) \hat{O} \hat{W}(\theta) \hat{\mathcal{U}}_\phi(x_m) | \psi_0 \rangle, \quad (1)$$

with $|\psi_0\rangle = H^{\otimes 2n} |0\rangle^{2n}$ being a uniform state and $\hat{\mathcal{U}}_\phi(x)$ being the feature map that creates a phase state as $|\phi_x\rangle := \hat{\mathcal{U}}_\phi(x) |\psi_0\rangle$. Next, the state is processed with an adjustable ansatz $\hat{W}(\theta)$ followed by the measurement of a Hermitian observable $\hat{O}$.

To move from QML to GQML, we choose the components of our model in a way that incorporates symmetries and ensures label invariance. Given a symmetry group $\mathcal{S}$ with a corresponding representation $V_\sigma : \mathcal{S} \rightarrow \text{Aut}(\mathcal{X}^2)$ [88,91], we say that the function g is invariant under the action of $\mathcal{S}$ if $g(V_\sigma[x_m]) = g(x_m) \forall x_m \in \mathcal{X}^2, \sigma \in \mathcal{S}$. In our case, we have the symmetry group $\mathcal{S} = \mathbb{Z}_2 \times \mathcal{S}_\Phi$. The first symmetry corresponds to the exchange of two barcodes within the

pair, $g(V_{Z_2}[\mathbf{x}_{1,m}, \mathbf{x}_{2,m}]) = g(\mathbf{x}_{2,m}, \mathbf{x}_{1,m}) = g(\mathbf{x}_{1,m}, \mathbf{x}_{2,m})$, while the second corresponds to the bitstring remapping symmetry applied on each barcode, as described in the previous section: $g(V_{S_\Phi}[\mathbf{x}_{1,m}, \mathbf{x}_{2,m}]) = g(\bar{\mathbf{x}}_{1,m}, \bar{\mathbf{x}}_{2,m}) = g(\mathbf{x}_{1,m}, \mathbf{x}_{2,m})$.

To encode these symmetries on the Hilbert space, we use the unitary representations of our symmetries: $\hat{U}_{Z_2} = \prod_{i=1}^n \text{SWAP}_{i,i+n}$ for the exchange symmetry, and $\hat{U}_\Phi = Y^{\otimes n} \otimes Y^{\otimes n}$ for the bitstring remapping symmetry. Guided by these, we can build an invariant quantum model using the following ingredients [88]: (1) *invariant* initial state, $\hat{U}_\sigma|\psi_0\rangle = |\psi_0\rangle$; (2) *equivariant* embedding, $\hat{U}(V_\sigma[x_m]) = \hat{U}_\sigma \hat{U}(x_m) \hat{U}_\sigma^\dagger$; (3) *equivariant* ansatz, $[\hat{W}(\theta), \hat{U}_\sigma] = 0$; and (4) *invariant* measured observable satisfying $\hat{U}_\sigma \hat{O} \hat{U}_\sigma^\dagger = \hat{O}$.

The equivariant embedding of the pairs of barcodes is done by preparing the aforementioned phase states $|\phi_{x_m}\rangle$. These are also known as real equally weighted (REW) states [106] or hypergraph states [107], and the corresponding unitary encoding circuit $\hat{U}_\phi(x_m) = \hat{U}(\mathbf{x}_{1,m}) \otimes \hat{U}(\mathbf{x}_{2,m})$ consists of a series of (multi)controlled-Z operations. Combined with the initial state $|\psi_0\rangle = |+\rangle^{\otimes 2n}$, it is clear to see that our encoding is equivariant with respect to both symmetries. As for the equivariant ansatz, we select generators that commute with $\hat{U}_{Z_2}$ and $\hat{U}_\Phi$. These include local operators $\{\sum_i \hat{Y}_i, \sum_i \hat{X}_i \hat{X}_{i+1}, \sum_i \hat{Y}_i \hat{Y}_{i+1}, \sum_i \hat{Z}_i \hat{Z}_{i+1}\}$ as well as global operators $\{\hat{X}^{\otimes 2n}, \hat{Z}^{\otimes 2n}, \prod_{i=1}^n \text{SWAP}_{i,i+n}, \hat{H}^{\otimes 2n}, \dots\}$. We can extend this pool by noting that if operators $\hat{A}$ and $\hat{B}$ both commute with $\hat{C}$, then so does the product $\hat{A}\hat{B}$. This allows us to insert operators such as $\{\prod_{i=1}^n \text{SWAP}_{i,i+n} \cdot \hat{X}^{\otimes 2n}, \hat{H}^{\otimes 2n} \cdot \hat{Z}^{\otimes 2n}, \prod_{i=1}^n \text{SWAP}_{i,i+n} \cdot \hat{H}^{\otimes 2n}, \dots\}$ into the pool. We select from this same pool of operators when choosing the invariant measurement observable.

We train the quantum model by minimizing the mean squared error (MSE) between the model predictions and the data labels over the training dataset, $\mathcal{L}_{\text{MSE}} = \frac{1}{M} \sum_{m=1}^M [a h_\theta(x_m) + b - y_m]^2$, where M is the number of training samples and a and b are trainable scaling and bias parameters. In this variational GQML approach [denoted as QNN_U as in Fig. 2(a)], we optimize the parameters within the variational ansatz and measure a single observable.

Next, we present a second option for GQML, involving the measurement of multiple observables with adjustable weights QNN_M , Fig. 2(b)]. We take inspiration from the ideas presented by Molteni *et al.* [108]. There, the authors prove quantum advantages for the task of learning quantum observables of the form $\hat{O}(\alpha) = \sum_{k=1}^K \alpha_k \hat{P}_k$, where α_k are real, unknown coefficients and $\hat{P}_k$ are Pauli strings. Specifically, the task is to learn a model $h(x)$ which approximates some unknown concept $f^\alpha(x) = \text{tr}[\rho_H(x) \hat{O}(\alpha)]$. Here, $\rho_H(x) = \hat{U}(x)|0\rangle\langle 0| \hat{U}^\dagger(x)$, with $\hat{U}(x) = e^{iH(x)}$. This problem is provably hard for classical algorithms under certain conditions, but the proposed quantum protocol can learn the concept with a training set of polynomial size [108].

Let us detail the method, adapting it specifically to our task of determining whether or not pairs of barcodes are correlated. We first encode the barcodes into quantum states in the same way as the standard GQML approach, $|\phi_{x_m}\rangle = \hat{U}_\phi(x_m)|\psi_0\rangle$. We then build a feature vector $\varphi(x_m)$ for each data point x_m by measuring the expectation values of a set of observables $\hat{O}$, directly on the quantum states, $\varphi(x_m) =$

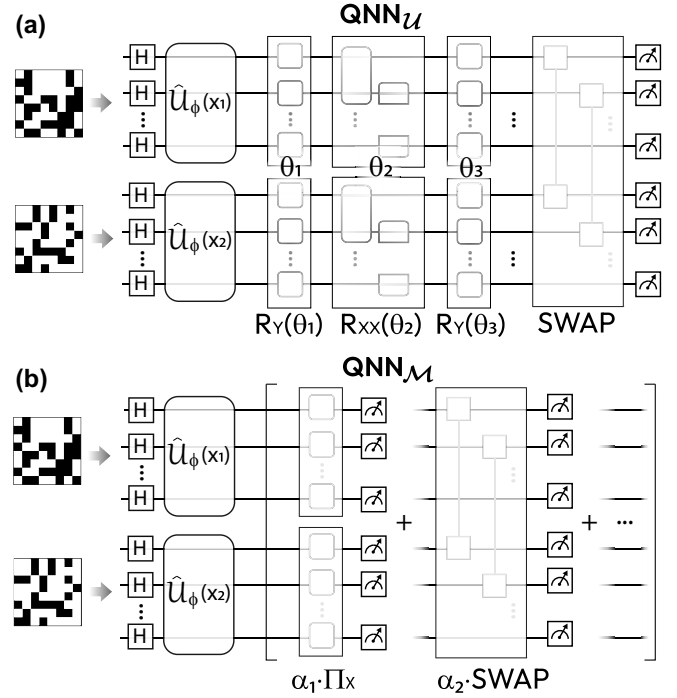


FIG. 2. Architectures for geometric quantum machine learning used in the study. (a) Quantum neural network approach based on equivariant unitary circuit (QNN_U) parametrized by a vector θ that processes a pair of barcodes encoded into quantum states. (b) QNN_M based on equivariant measurements that are weighted adaptively to form a model for correlated sample processing.

$[\langle \hat{O}_1 \rangle_m, \langle \hat{O}_2 \rangle_m, \dots, \langle \hat{O}_K \rangle_m]$, where $\langle \hat{O}_i \rangle_m = \langle \phi_{x_m} | \hat{O}_i | \phi_{x_m} \rangle$. We select this set of observables from the aforementioned pool of equivariant operators, such that our model maintains symmetry invariance. We then define the model as $h(x_m) = \alpha \cdot \varphi(x_m)$, where α is a vector of real-valued weights. We make use of the least absolute shrinkage and selection operator (LASSO) method, a machine learning model that finds the optimal weights to minimize the loss function as [109]

$$\text{LASSO} : \min_{\alpha} \frac{1}{2M} \sum_{m=1}^M [\alpha \cdot \varphi(x_m) - y_m]^2 + \lambda \|\alpha\|_{l_1}, \quad (2)$$

where λ is a regularization constant which constrains the size of the weights. In this case, our labels $y_m = f^\alpha(x_m)$ represent samples from our concept class. By optimizing the loss function, we can determine the combination of observables that lead to a clear separation between correlated and noncorrelated samples $x_m = (\mathbf{x}_{1,m}, \mathbf{x}_{2,m})$.

We proceed to compare the performance of the two different quantum architectures, QNN_U and QNN_M . For the variational GQML, we selected the generators $\{\sum_i \hat{Y}_i, \sum_i \hat{X}_i \hat{X}_{i+1}, \sum_i \hat{Y}_i \hat{Y}_{i+1}, \prod_{i=1}^n \text{SWAP}_{i,i+n}\}$ to build our model [Fig. 2(a)], balancing layers of trainable local operations with the measurement of a nonlocal SWAP operator in an attempt to capture global correlations between the input pair. The variational parameters were optimized using the Adam optimizer with a learning rate of 0.1, and training was performed over 200 epochs. As for the adaptive measurement protocol, we selected a set of $K = 10$

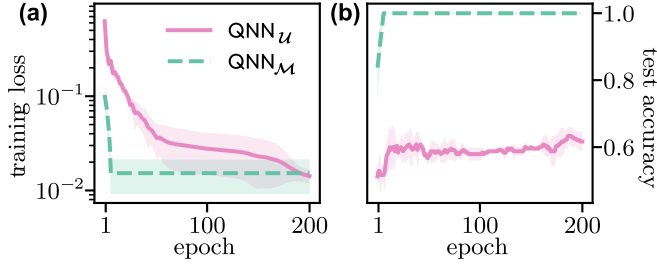


FIG. 3. Comparison of the training performance (a) and generalization (b) between the two GQML architectures, QNN_U and QNN_M. Both models were trained and tested on pairs of barcodes with 16 pixels each, corresponding to an 8-qubit quantum state for each input datum. Solid/dashed lines indicate the mean values across the ten trials, while the shaded region indicates the standard deviation.

measurement observables from our equivariant pool of operators, including most of the operators already listed, and the weights of the LASSO model were optimized via coordinate descent. For both architectures, we used a training dataset of ten samples from each class ($M = 20$), and tested on 40 samples from each class, thus studying the models' ability to generalize to unseen data from relatively few samples. We trialed the two quantum models on pairs of 4×4 barcodes, averaging results for each model over ten trials (for each trial, variational parameters/weights were randomly initialized and a different training set was selected). The results are shown in Fig. 3. We can see that both models successfully train; the training loss of QNN_U steadily decreases over the training period, while the QNN_M model converges extremely quickly. However, we observe a stark difference in generalization; while the adaptive measurement model achieves virtually perfect test accuracy across all runs, the variational circuit model struggles to reach test accuracies over 60%. QNN_U may be overfitting the training data, indicating that the features encoding the correlations are difficult to extract using the parametrized quantum circuit framework. However, measuring a linear combination of specific observables proves to be much more successful in distinguishing correlated and uncorrelated samples. The struggles of QNN_U are even more pertinent as we increase the system size, with an increasing number of ansatz layers required just to train the model, with similarly poor generalization. Based on these trials, we proceed with the adaptive measurement approach as the GQML protocol to test on larger images.

IV. RESULTS

A. *In silico*: classical versus quantum model

We analyze the described GQML workflows and compare them to classical similarity testing with Siamese neural networks [110,111]. To emphasize the resilience of the separation between quantum and classical learners, we present generalization over the number of training samples and system size. In Fig. 4, we show the training loss and test accuracy of the QNN_M against those for classical networks, classifying pairs of 32×32 barcodes. We consider two types of networks with

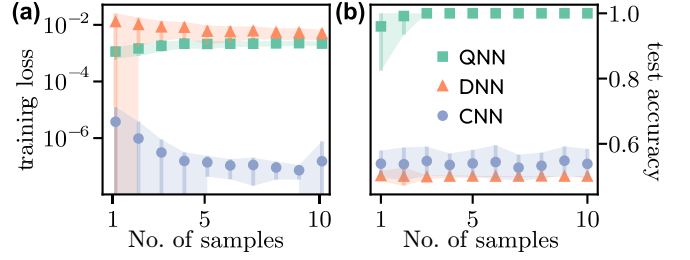


FIG. 4. Generalization of quantum and classical learners for the barcode similarity problem shown as a function of the number of samples from each class used in training. We show final training loss (a) and test accuracy (b) of quantum and classical networks, trained and tested on pairs of 1024-pixel barcodes (20-qubit system). Classical benchmarks correspond to deep (DNN) and convolutional (CNN) Siamese neural networks, which learn the similarity via a metric distance evaluated in feature space (see details in the main text).

Siamese structures, namely a deep neural network (DNN) and a convolutional neural network (CNN). Siamese neural networks consist of two identical subnetworks that share the same weights and architecture. Let $f_{\theta}(\mathbf{x})$ represent the function learned by the subnetwork parametrized by weights θ . Given a dataset pair $x_m = (\mathbf{x}_{1,m}, \mathbf{x}_{2,m})$, the network computes the outputs $\mathbf{h}_{1,m} = f_{\theta}(\mathbf{x}_{1,m})$ and $\mathbf{h}_{2,m} = f_{\theta}(\mathbf{x}_{2,m})$. The idea is to map similar inputs close to each other in feature space while mapping dissimilar inputs far apart. To achieve this, we compute the commonly used Euclidean distance metric between the two outputs, $d(\mathbf{h}_{1,m}, \mathbf{h}_{2,m}) = \|\mathbf{h}_{1,m} - \mathbf{h}_{2,m}\|^2$. The learning objective is to minimize the loss function $\mathcal{L}_{\text{MSE}} = (1/M) \sum_{m=1}^M (\tilde{y}_m - y_m)^2$ where $\tilde{y}_m$ is the predicted label for the pair $(\mathbf{x}_{1,m}, \mathbf{x}_{2,m})$. Note that the mean square error is chosen for better comparison with the quantum models. We used the following training and test parameters for the simulations with quantum and classical networks. We average over 50 randomizations of the training dataset for each training data set size, training over (up to) 300 epochs for both models (noting again that far fewer epochs were required for the convergence of QNN_M). For each run, we randomize the M samples x_m selected for the training dataset, with $\frac{M}{2} = 1, 2, \dots, 10$ selected from each category (correlated and uncorrelated samples). Testing is done on 40 unseen samples from each class. The error bar represents the standard deviation over the training randomizations. Note that the “tearing” effect of the lower error bars in Fig. 4(a) is due to the semilogarithmic scale.

Both quantum and classical ML protocols showcase good training performances across all datasets; all three models reach 100% training accuracy at all dataset sizes M , with MSE for CNN converging to particularly low values. The real differences between the architectures are seen in their generalization, where both classical models struggle to perform on unseen data, achieving test accuracies only slightly better than random guessing. This suggests that the classical models are prone to overfitting the training data, failing to discover the true features underlying the correlations between the barcodes. Several adjustments were attempted to improve the classical performance; we tested different loss functions (e.g., MSE and binary cross-entropy), added dropouts, used

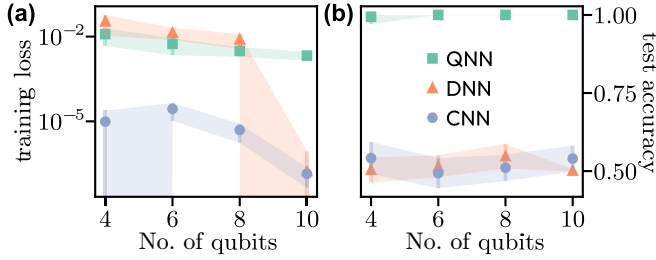


FIG. 5. Generalization of learners for the barcode similarity problem shown as a function of system size (number of qubits n for each image, $2n$ in total). Both training loss (a) and test accuracy (b) remain steady across all system sizes, indicating that the quantum-classical separation is ubiquitous for smaller and larger pairs of barcodes. The plots show simulations run with the same parameters as described in Fig. 4, with five training samples taken from each class ($M = 10$).

cosine similarity as a metric, and varied batch sizes. None of these attempts significantly improved results. Meanwhile, the quantum model again generalizes excellently at this larger system size; training on just three samples from each class ($M = 6$) is enough to classify all the pairs of 1024-pixel barcodes loaded into a $2n = 20$ -qubit register.

Next, we compare the ability of different models to detect image correlation at different system sizes fixing the number of training samples. The results are shown in Fig. 5. We plot the average training loss and test accuracy as a function of the number of qubits n used to encode each barcode within the quantum model; an N -pixel barcode requires $n = \log_2 N$ qubits. Both quantum and classical models show no significant performance variation across the range of image dimensions, showcasing that quantum-classical separation persists at different system sizes. While we fix the number of training samples to 5 from each class ($M = 10$), all other parameters and the running procedure are the same as the ones of Fig. 4.

B. Hardware: 40-qubit quantum model

To evaluate the scalability and noise resilience of the proposed QNN model, we implemented a 40-qubit protocol ($2n = 40$) on IBM Kyiv using the Eagle r3 architecture [112].

Following the structure in Fig. 2(b), we started by generating the instruction sets $\hat{U}(\mathbf{x}_{1,m})$ and $\hat{U}(\mathbf{x}_{2,m})$ to prepare 50 random 1048576-pixel barcode pairs and 50 correlated pairs. These instructions defined the unitary encoding circuits $\hat{U}_\phi(x_m) = \hat{U}(\mathbf{x}_{1,m}) \otimes \hat{U}(\mathbf{x}_{2,m})$, which we used to create the states $|\phi_{x_m}\rangle = \hat{U}_\phi(x_m)|\psi_0\rangle$. We calculated the expectation values on these states of seven observables: $\sum_i \hat{Y}_i$, $\sum_i \hat{X}_i \hat{X}_{i+1}$, $\sum_i \hat{Y}_i \hat{Y}_{i+1}$, $\sum_i \hat{Z}_i \hat{Z}_{i+1}$, $\prod_{i=1}^n \text{SWAP}_{i,i+n}$, $\hat{H}^{\otimes 2n}$, and $\prod_{i=1}^n \text{SWAP}_{i,i+n} \cdot \hat{H}^{\otimes 2n}$. We evaluated these expectation values by sampling return probabilities from the state $|0\rangle^{2n}$, then constructed the feature vectors $\varphi(x_m)$ for the LASSO model as described in Eq. (2).

Figure 6 shows the results. In Fig. 6(a), we plot the expectation values for 50 random and 50 correlated pairs. Figure 6(b) displays the training loss and test accuracy of the LASSO model across 50 random data splits. Figure 6(c)

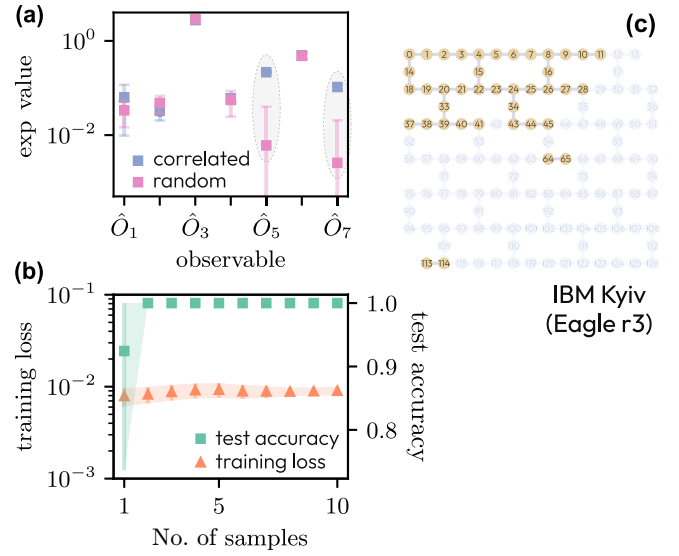


FIG. 6. Implementation of a $2n = 40$ model on IBM Kyiv with Eagle r3 architecture. We run 50 randomizations of random and correlated barcode pairs. In (a), we show the sampled expectation values and relative standard deviations of the randomizations of the following observables, $\hat{O}_1 = \sum_i \hat{Y}_i$, $\hat{O}_2 = \sum_i \hat{X}_i \hat{X}_{i+1}$, $\hat{O}_3 = \sum_i \hat{Y}_i \hat{Y}_{i+1}$, $\hat{O}_4 = \sum_i \hat{Z}_i \hat{Z}_{i+1}$, $\hat{O}_5 = \prod_{i=1}^n \text{SWAP}_{i,i+n}$, $\hat{O}_6 = \hat{H}^{\otimes 2n}$, $\hat{O}_7 = \prod_{i=1}^n \text{SWAP}_{i,i+n} \cdot \hat{H}^{\otimes 2n}$. Both observables including the SWAP operations, $\hat{O}_5$ and $\hat{O}_7$, show a clear separation (highlighted in gray) between random and correlated pairs. In (b), these values are then used in the LASSO model at different sample sizes. We show training loss and test accuracy of the model averaged over 50 sample randomizations. The model remains solid against hardware noise and showed perfect accuracy already training over two samples. In (c), we show an example of 40-qubit configuration used during the sampling runs on IBM Kyiv. The disconnected qubits are likely chosen during the optimization passes (level 3) to compensate for qubit and connection errors.

highlights a 40-qubit configuration used during the sampling routine on IBM Kyiv. Expectation values for random and correlated pairs are comparable (within their standard deviations) for the local and global operators $\sum_i \hat{Y}_i$, $\sum_i \hat{X}_i \hat{X}_{i+1}$, $\sum_i \hat{Y}_i \hat{Y}_{i+1}$, $\sum_i \hat{Z}_i \hat{Z}_{i+1}$, and $\hat{H}^{\otimes 2n}$. However, $\prod_{i=1}^n \text{SWAP}_{i,i+n}$ and $\prod_{i=1}^n \text{SWAP}_{i,i+n} \cdot \hat{H}^{\otimes 2n}$ show significant differences, highlighted in gray in Fig. 6(a). This separation is crucial for training the model, as shown in Fig. 6(b). For datasets comprising a single correlated pair and a single random pair, the model does not achieve 100% average test accuracy. Instead, accuracy strongly depends on the specific training instance selected; smaller observable separations result in reduced model accuracy. However, already with two training samples the test accuracy of a quantum model reaches 100%, largely overcoming the random guess limit characteristic to the classical analysis.

To perform the sampling, we sampled circuits coded in Qiskit and shipped to the quantum processing unit (QPU) IBM Kyiv via IBM Cloud. We evaluated the expectation values running 2048 shots per pair. We optimized the circuits for the Eagle r3 QPU architecture using the preset pass manager at optimization level 3 and grouped them into batches. Figure 6(c) shows an example of the 40-qubit configuration

used. After running the batches through the sampler, we collected the $\mathbf{0}_{2n}$ string counts and processed the expectation values through the LASSO model.

V. THE EXPLANATION

Previously we have shown that the problem of barcode similarity testing can be solved with excellent generalization. This is ensured by the considered dataset, symmetries, and trained quantum models. Let us unravel why GQML did so well when learning the similarity (global correlations), and show that the task can be related to problems leading to a proven superpolynomial speedup. To form the dataset we employed the strategy presented by Raz and Tal [113], motivated by Aaronson [114], and developed to show a separation between BQP and PH. Specifically, we take a normal multivariate probability distribution $\tilde{P}_A(x) := \mathcal{N}(0, \varepsilon \cdot \Sigma)$ over $\tilde{x} \in \mathbb{R}^N \times \mathbb{R}^N$ with zero mean, variance of ε , and covariances defined as a 2×2 -block matrix $\Sigma = [\mathbb{1}^{\otimes n}, \hat{H}^{\otimes n}; \hat{H}^{\otimes n}, \mathbb{1}^{\otimes n}]$, for $n = \lceil \log_2 N \rceil$. Drawing samples from this distribution $\tilde{X} \sim \tilde{P}_A(\tilde{x})$ we get a list of correlated vectors. This can be seen as drawing the N -dimensional first part of the sample $\tilde{\mathbf{x}}_1$ and the second part being related via the Hadamard transform as $\tilde{\mathbf{x}}_2 \approx \hat{H}^{\otimes n} \tilde{\mathbf{x}}_1$. Following Ref. [113] we perform truncation of vector entries and get the samples $x|_{y=0}$. Next, sampling the distribution $\tilde{P}_B(x) := \mathcal{N}(0, \varepsilon \cdot [\mathbb{1}^{\otimes n}, \mathbf{0}; \mathbf{0}, \mathbb{1}^{\otimes n}])$, then performing the same truncation gives us uncorrelated samples from effectively independent distributions, $x|_{y=1}$. Raz and Tal then showed that following the quantum algorithm for the forrelation problem by Aaronson and Ambainis [115] one can solve this Promise-BQP problem of distinguishing forrelated and uniform distributions, calling a single query to load the data as a quantum state [similar to our feature map $\hat{\mathcal{U}}_\phi(x)$]. This $O(1)$ query complexity for distinguishing distributions showcases advantage over $\tilde{\Omega}(\sqrt{N})$ for the classical query complexity [115, 116]. The essence of the quantum algorithm is the estimation of the overlap between quantum representations of $\mathbf{x}_1$ and $\mathbf{x}_2$ as

$$F = |\langle \phi_{\mathbf{x}_1} | \hat{H}^{\otimes n} | \phi_{\mathbf{x}_2} \rangle|^2 = |\langle 0 | \hat{H}^{\otimes n} \hat{\mathcal{U}}_\phi^\dagger(\mathbf{x}_1) \hat{H}^{\otimes n} \hat{\mathcal{U}}_\phi(\mathbf{x}_2) \hat{H}^{\otimes n} | 0 \rangle|^2, \quad (3)$$

which takes high values for pairs from $P_A(x)$ as we are effectively undoing the Hadamard transform on $\mathbf{x}_2$. Conversely, for uncorrelated samples $\hat{H}^{\otimes n}$ simply reshuffles the amplitudes, and F remains exponentially small in N . This represents a feature that can be used by the GQML classifier. To pick up this feature of forrelated distributions during similarity testing QNNs require learning a specific observable for constructing a decision boundary, corresponding to $\hat{O}(\alpha^*) = \hat{H}^{\otimes 2n} \cdot (\prod_{i=1}^n \text{SWAP}_{i, i+n})$ applied on $|\phi_{\mathbf{x}_1}\rangle \otimes |\phi_{\mathbf{x}_2}\rangle$. Such a Hermitian operator is readily available in the QNN_M model, and can be learnt by LASSO optimization. At the same time, we note that there is no equivariant unitary transformation and observable such that $\hat{W}^\dagger(\theta) \hat{O} \hat{W}(\theta)$ represents the required product of SWAPs and Hadamards. This explains why the QNN_M model does not generalize well for the task.

We stress that forrelation also underpins the first example of provably efficient quantum machine learning solvers of BQP problems in the kernel-based operation mode [102]. Jäger and Krems have shown that by using a feature map specific to k -forrelation one can produce a classifier that outperforms any classical analog. In our work, we show that GQML achieves a generalization advantage without tailored feature maps, and aided by simple symmetries can distinguish distributions and patterns of global correlations.

VI. CONCLUSION

To summarize, we described a geometric quantum machine learning (GQML) approach that enables learning an optimal algorithm for similarity testing between two vectors (visualized as images or barcodes). Our approach embedded symmetries to enable generalization from just a few samples. Unlike existing GQML algorithms, we focused on symmetry-aware measurement adaptation, which we demonstrated to outperform unitary adaptation of observables. We compared the performance of our GQML approach to that of classical deep neural networks employing a Siamese architecture. Our findings show that the quantum model achieves strong empirical performance gains over classical deep learning, particularly in scenarios involving correlated distributions and limited training data. We explained these empirical generalization benefits through the connection to the established maximal separations between the BQP complexity class and the PH for classifying samples from forrelated and uniform distributions. Our work identifies a setting where quantum machine learning can offer empirical benefits over classical approaches. In this direction, we demonstrated the suitability of the quantum model based on equivariant measurements for hardware implementation by successfully running a 40-qubit protocol on IBM Eagle architecture, which represents the largest hardware run of a forrelation test to date.

Importantly, the demonstrated advantage for using GQML depends on the considered data access. For the setting with quantum data access, the GQML approach consistently outperformed classical Siamese architectures. At the same time, for classical data access and no black-box assumptions, we note that the process becomes more prone to dequantization [117]. For pixels that are stored in memory, one can apply the same steps as in the quantum algorithm and obtain an efficient quantum-inspired solution, since the Hadamard transform is easy to apply on a classical machine. An interesting scenario corresponds to data loading schemes, where certain families of REW stated are prepared by construction [118], but the circuit structure is unknown and families of circuit distributions need to be classified. Discovering problems that benefit from quantum similarity detection in the quantum data setting is a motivation for future works.

ACKNOWLEDGMENTS

We acknowledge the funding from UK EPSRC awards under the Agreements No. EP/Y005090/1 and No. EP/Z53318X/1 (QCi3 Hub). We also acknowledge the

use of IBM Quantum services for this work supported via the Quantum Program. The views expressed are those of the authors, and do not reflect the official policy or position of IBM or the IBM Quantum team.

DATA AVAILABILITY

The data that support the findings of this article are openly available [119], embargo periods may apply.

- [1] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature (London)* **521**, 436 (2015).
- [2] A. Moskvichev, V. V. Odouard, and M. Mitchell, The conceptarc benchmark: Evaluating understanding and generalization in the ARC domain, *arXiv:2305.07141*.
- [3] C. Biever, ChatGPT broke the turing test - the race is on for new ways to assess AI, *Nature (London)* **619**, 686 (2023).
- [4] Y. LeCun, S. Chopra, R. Hadsell, M. A. Ranzato, and F. J. Huang, Energy based models, in *Predicting Structured Data*, edited by G. BakIr, T. Hofmann, B. Schölkopf, A. J. Smola, B. Taskar, and S. V. N. Vishwanathan (The MIT Press, Cambridge, MA, 2007).
- [5] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, Parameterized quantum circuits as machine learning models, *Quantum Sci. Technol.* **4**, 043001 (2019).
- [6] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Quantum machine learning, *Nature (London)* **549**, 195 (2017).
- [7] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, *Nature (London)* **567**, 209 (2019).
- [8] J. Liu, M. Liu, J.-P. Liu, Z. Ye, Y. Wang, Y. Alexeev, J. Eisert, and L. Jiang, Towards provably efficient quantum algorithms for large-scale machine-learning models, *Nat. Commun.* **15**, 434 (2024).
- [9] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.* **26**, 1484 (1997).
- [10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).
- [11] A. M. Dalzell, S. McArdle, M. Berta, P. Bienias, C.-F. Chen, A. Gilyén, C. T. Hann, M. J. Kastoryano, E. T. Khabiboulline, A. Kubica, G. Salton, S. Wang, and F. G. S. L. Brandão, *Quantum Algorithms: A Survey of Applications and End-To-End Complexities* (Cambridge University Press, Cambridge, UK, 2023).
- [12] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational quantum algorithms, *Nat. Rev. Phys.* **3**, 625 (2021).
- [13] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, The variational quantum eigensolver: A review of methods and best practices, *Phys. Rep.* **986**, 1 (2022).
- [14] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning, *Phys. Rev. A* **98**, 032309 (2018).
- [15] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, Data re-uploading for a universal quantum classifier, *Quantum* **4**, 226 (2020).
- [16] H.-Y. Huang, M. Broughton, M. Mohseni, R. Babbush, S. Boixo, H. Neven, and J. R. McClean, Power of data in quantum machine learning, *Nat. Commun.* **12**, 2631 (2021).
- [17] N. A. Nghiem, S. Y.-C. Chen, and T.-C. Wei, Unified framework for quantum classification, *Phys. Rev. Res.* **3**, 033056 (2021).
- [18] W. Li and D.-L. Deng, Recent advances for quantum classifiers, *Sci. China Phys., Mech. Astron.* **65**, 220301 (2021).
- [19] I. Cong, S. Choi, and M. D. Lukin, Quantum convolutional neural networks, *Nat. Phys.* **15**, 1273 (2019).
- [20] S. Monaco, O. Kiss, A. Mandarino, S. Vallecorsa, and M. Grossi, Quantum phase detection generalization from marginal quantum neural network models, *Phys. Rev. B* **107**, L081105 (2023).
- [21] D. Herman, C. Googin, X. Liu, Y. Sun, A. Galda, I. Safro, M. Pistoia, and Y. Alexeev, Quantum computing for finance, *Nat. Rev. Phys.* **5**, 450 (2023).
- [22] J.-G. Liu and L. Wang, Differentiable learning of quantum circuit born machines, *Phys. Rev. A* **98**, 062324 (2018).
- [23] C. Zoufal, A. Lucchi, and S. Woerner, Quantum generative adversarial networks for learning and loading random distributions, *npj Quantum Inf.* **5**, 103 (2019).
- [24] B. Coyle, D. Mills, V. Danos, and E. Kashefi, The born supremacy: Quantum advantage and training of an Ising born machine, *npj Quantum Inf.* **6**, 60 (2020).
- [25] A. E. Paine, V. E. Elfving, and O. Kyriienko, Quantum quantile mechanics: Solving stochastic differential equations for generating time-series, *Adv. Quantum Technol.* **6**, 2300065 (2023).
- [26] O. Kyriienko, A. E. Paine, and V. E. Elfving, Protocols for trainable and differentiable quantum generative modeling, *Phys. Rev. Res.* **6**, 033291 (2024).
- [27] S. Kasture, O. Kyriienko, and V. E. Elfving, Protocols for classically training quantum generative models on probability distributions, *Phys. Rev. A* **108**, 042406 (2023).
- [28] M. Hibat-Allah, M. Mauri, J. Carrasquilla, and A. Perdomo-Ortiz, A framework for demonstrating practical quantum advantage: Comparing quantum against classical generative models, *Commun. Phys.* **7**, 68 (2024).
- [29] W. Hou, M. Li, and Y.-Z. You, Sequential learning on a tensor network born machine with trainable token embedding, *Mach. Learn.: Sci. Technol.* **6**, 025071 (2025).
- [30] O. Kyriienko, A. E. Paine, and V. E. Elfving, Solving nonlinear differential equations with differentiable quantum circuits, *Phys. Rev. A* **103**, 052416 (2021).
- [31] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, Variational quantum algorithms for nonlinear problems, *Phys. Rev. A* **101**, 010301(R) (2020).
- [32] S. Markidis, On physics-informed neural networks for quantum computers, *Front. Appl. Math. Stat.* **8**, 1036711 (2022).
- [33] N. Heim, A. Ghosh, O. Kyriienko, and V. E. Elfving, Quantum model-discovery, *arXiv:2111.06376*.

- [34] A. E. Paine, V. E. Elfving, and O. Kyriienko, Quantum kernel methods for solving regression problems and differential equations, *Phys. Rev. A* **107**, 032428 (2023).
- [35] A. E. Paine, V. E. Elfving, and O. Kyriienko, Physics-informed quantum machine learning: Solving nonlinear differential equations in latent spaces without costly grid evaluations, *arXiv:2308.01827*.
- [36] M. Schuld and N. Killoran, Quantum machine learning in feature Hilbert spaces, *Phys. Rev. Lett.* **122**, 040504 (2019).
- [37] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, The power of quantum neural networks, *Nat. Comput. Sci.* **1**, 403 (2021).
- [38] T. Goto, Q. H. Tran, and K. Nakajima, Universal approximation property of quantum machine learning models in quantum-enhanced feature spaces, *Phys. Rev. Lett.* **127**, 090506 (2021).
- [39] S. Altares-López, A. Ribeiro, and J. J. García-Ripoll, Automatic design of quantum feature maps, *Quantum Sci. Technol.* **6**, 045015 (2021).
- [40] C. A. Williams, A. E. Paine, H.-Y. Wu, V. E. Elfving, and O. Kyriienko, Quantum chebyshev transform: Mapping, embedding, learning and sampling distributions, *arXiv:2306.17026*.
- [41] B. Albrecht, C. Dalyac, L. Leclerc, L. Ortiz-Gutiérrez, S. Thabet, M. D’Arcangelo, J. R. K. Cline, V. E. Elfving, L. Lassablière, H. Silvério, B. Ximenez, L.-P. Henry, A. Signoles, and L. Henriët, Quantum feature maps for graph machine learning on a neutral atom quantum processor, *Phys. Rev. A* **107**, 042615 (2023).
- [42] B. Jaderberg, A. A. Gentile, Y. Achari Berrada, E. Shishenina, and V. E. Elfving, Let quantum neural networks choose their own frequencies *Phys. Rev. A* **109**, 042421 (2024).
- [43] C. Umeano and O. Kyriienko, Ground state-based quantum feature maps, *arXiv:2404.07174*.
- [44] H.-Y. Wu, V. E. Elfving, and O. Kyriienko, Multidimensional quantum generative modeling by quantum Hartley transform *Adv. Quantum Technol.* **8**, 2400337 (2025).
- [45] S. Thabet, M. Djellabi, I. Sokolov, S. Kasture, L.-P. Henry, and L. Henriët, Quantum positional encodings for graph neural networks, *arXiv:2406.06547*.
- [46] S. Y. Chang, B. L. Saux, S. Vallecorsa, and M. Grossi, Quantum convolutional circuits for earth observation image classification, in *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium* (IEEE, New York, 2022), pp. 4907–4910.
- [47] N. Schetakakis, D. Aghamalyan, P. Griffin, and M. Boguslavsky, Review of some existing QML frameworks and novel hybrid classical–quantum neural networks realising binary classification for the noisy datasets, *Sci. Rep.* **12**, 11927 (2022).
- [48] A. Senokosov, A. Sedykh, A. Sagingalieva, B. Kyriacou, and A. Melnikov, Quantum machine learning for image classification, *Mach. Learn.: Sci. Technol.* **5**, 015040 (2024).
- [49] K. Shen, B. Jobst, E. Shishenina, and F. Pollmann, Classification of the fashion-MNIST dataset on a quantum computer, *arXiv:2403.02405*.
- [50] V. Belis, S. González-Castillo, C. Reissel, S. Vallecorsa, E. F. Combarro, G. Dissertori, and F. Reiter, Higgs analysis with quantum classifiers, *EPJ Web Conf.* **251**, 03070 (2021).
- [51] S. Y.-C. Chen, T.-C. Wei, C. Zhang, H. Yu, and S. Yoo, Quantum convolutional neural networks for high energy physics data analysis, *Phys. Rev. Res.* **4**, 013231 (2022).
- [52] V. Belis, K. Anna Woźniak, E. Puljak, P. Barkoutsos, G. Dissertori, M. Grossi, M. Pierini, F. Reiter, I. Tavernelli, and S. Vallecorsa, Quantum anomaly detection in the latent space of proton collision events at the LHC, *Commun. Phys.* **7**, 334 (2023).
- [53] V. Belis, P. Odagiu, M. Grossi, F. Reiter, G. Dissertori, and S. Vallecorsa, Guided quantum compression for high dimensional data classification, *Mach. Learn.: Sci. Technol.* **5**, 035010 (2024).
- [54] A. Di Meglio, K. Jansen, I. Tavernelli, C. Alexandrou, S. Arunachalam, C. W. Bauer, K. Borras, S. Carrazza, A. Crippa, V. Croft, R. de Putter, A. Delgado, V. Dunjko, D. J. Egger, E. Fernández-Combarro, E. Fuchs, L. Funcke, D. González-Cuadra, M. Grossi, J. C. Halimeh, *et al.*, Quantum computing for high-energy physics: State of the art and challenges, *PRX Quantum* **5**, 037001 (2024).
- [55] D. J. Egger, C. Gambella, J. Marecek, S. McFaddin, M. Mevissen, R. Raymond, A. Simonetto, S. Woerner, and E. Yndurain, Quantum computing for finance: State-of-the-art and future prospects, *IEEE Trans. Quantum Eng.* **1**, 3101724 (2020).
- [56] L. Leclerc, L. Ortiz-Gutiérrez, S. Grijalva, B. Albrecht, J. R. K. Cline, V. E. Elfving, A. Signoles, L. Henriët, G. Del Bimbo, U. A. Sheikh, M. Shah, L. Andrea, F. Ishtiaq, A. Duarte, S. Mugel, I. Cáceres, M. Kurek, R. Orús, A. Seddik, O. Hammami, *et al.*, Financial risk management on a neutral atom quantum processor, *Phys. Rev. Res.* **5**, 043117 (2023).
- [57] O. Kyriienko and E. B. Magnusson, Unsupervised quantum machine learning for fraud detection, *arXiv:2208.01203*.
- [58] M. Grossi, N. Ibrahim, V. Radescu, R. Loredó, K. Voigt, C. von Altrock, and A. Rudnik, Mixed quantum–classical method for fraud detection with quantum feature selection, *IEEE Trans. Quantum Eng.* **3**, 3102812 (2022).
- [59] M. Doosti, P. Wallden, C. B. Hamill, R. Hankache, O. T. Brown, and C. Heunen, A brief review of quantum machine learning for financial services, *arXiv:2407.12618*.
- [60] N. Innan, A. Sawaika, A. Dhor, S. Dutta, S. Thota, H. Gokal, N. Patel, M. A.-Z. Khan, I. Theodonis, and M. Bennai, Financial fraud detection using quantum graph neural networks, *Quantum Mach. Intell.* **6**, 7 (2024).
- [61] B. Coecke, G. de Felice, K. Meichanetzidis, and A. Toumi, Foundations for near-term quantum natural language processing, *arXiv:2012.03755*.
- [62] D. Widdows, A. Alexander, D. Zhu, C. Zimmerman, and A. Majumder, Near-term advances in quantum natural language processing, *Ann. Math. Artif. Intell.* **92**, 1249 (2024).
- [63] D. Widdows, W. Aboumradi, D. Kim, S. Ray, and J. Mei, Quantum natural language processing *Künstl. Intell.* **38**, 293 (2024).
- [64] G. Verdon, T. McCourt, E. Luzhnica, V. Singh, S. Leichenauer, and J. Hidary, Quantum graph neural networks, *arXiv:1909.12264*.
- [65] L.-P. Henry, S. Thabet, C. Dalyac, and L. Henriët, Quantum evolution Kernel: Machine learning on graphs with programmable arrays of qubits, *Phys. Rev. A* **104**, 032416 (2021).
- [66] C. Dalyac, L. Leclerc, L. Vignoli, M. Djellabi, W. da Silva Coelho, B. Ximenez, A. Dareau, D. Dreon, V. E. Elfving, A. Signoles, L.-P. Henry, and L. Henriët, Graph algorithms with neutral atom quantum processors, *arXiv:2403.11931*.

- [67] S. Yu, C. Peng, Y. Wang, A. Shehzad, F. Xia, and E. R. Hancock, Quantum graph learning: Frontiers and outlook, [arXiv:2302.00892](#).
- [68] Z. Kaseb, M. Möller, G. T. Balducci, P. Palensky, and P. P. Vergara, Quantum neural networks for power flow analysis, *Electr. Power Syst. Res.* **235**, 110677 (2024).
- [69] O. Kiss, F. Tacchino, S. Vallecorsa, and I. Tavernelli, Quantum neural networks force fields generation, *Mach. Learn.: Sci. Technol.* **3**, 035004 (2022).
- [70] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, *et al.*, Does provable absence of barren plateaus imply classical simulability? *Nat. Commun.* **16**, 7907 (2025).
- [71] O. Kyriienko, C. Umeano, and Z. Holmes, Advantage for discrete variational quantum algorithms in circuit recompilation, [arXiv:2510.01154](#).
- [72] M. Schuld, Supervised quantum machine learning models are kernel methods, [arXiv:2101.11020](#).
- [73] C. Gyurik, R. Molteni, and V. Dunjko, Limitations of measure-first protocols in quantum machine learning, [arXiv:2311.12618](#).
- [74] S. Jerbi, C. Gyurik, S. C. Marshall, R. Molteni, and V. Dunjko, Shadows of quantum machine learning, *Nat. Commun.* **15**, 5676 (2024).
- [75] M. Schuld and N. Killoran, Is quantum advantage the right goal for quantum machine learning? *PRX Quantum* **3**, 030101 (2022).
- [76] N. Wiebe, D. Braun, and S. Lloyd, Quantum algorithm for data fitting, *Phys. Rev. Lett.* **109**, 050505 (2012).
- [77] S. Lloyd, M. Mohseni, and P. Rebentrost, Quantum algorithms for supervised and unsupervised machine learning, [arXiv:1307.0411](#).
- [78] M. B. Mansky, J. Nüßlein, D. Bucher, D. Schuman, S. Zielinski, and C. Linnhoff-Popien, Sampling problems on a quantum computer, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01 (IEEE, New York, 2023), pp. 485–495.
- [79] D. Hangleiter and J. Eisert, Computational advantage of quantum random sampling, *Rev. Mod. Phys.* **95**, 035001 (2023).
- [80] K. Gili, M. Hibat-Allah, M. Mauri, C. Ballance, and A. Perdomo-Ortiz, Do quantum circuit born machines generalize? *Quantum Sci. Technol.* **8**, 035021 (2023).
- [81] Z.-Y. Han, J. Wang, H. Fan, L. Wang, and P. Zhang, Unsupervised generative modeling using matrix product states, *Phys. Rev. X* **8**, 031012 (2018).
- [82] H.-Y. Huang, M. Broughton, N. Eassa, H. Neven, R. Babbush, and J. R. McClean, Generative quantum advantage for classical and quantum problems, [arXiv:2509.09033](#).
- [83] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean, Quantum advantage in learning from experiments, *Science* **376**, 1182 (2022).
- [84] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles, Generalization in quantum machine learning from few training data, *Nat. Commun.* **13**, 4919 (2022).
- [85] C. Umeano, A. E. Paine, V. E. Elfving, and O. Kyriienko, What can we learn from quantum convolutional neural networks? *Adv. Quantum Technol.* **8**, 2400325 (2023).
- [86] E. Gil-Fuster, J. Eisert, and C. Bravo-Prieto, Understanding quantum machine learning also requires rethinking generalization *Nat. Commun.* **15**, 2277 (2023).
- [87] H. Zheng, Z. Li, J. Liu, S. Strelchuk, and R. Kondor, Speeding up learning quantum states through group equivariant convolutional quantum ansätze, *PRX Quantum* **4**, 020327 (2023).
- [88] J. J. Meyer, M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert, Exploiting symmetry in variational quantum machine learning, *PRX Quantum* **4**, 010328 (2023).
- [89] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, Group-invariant quantum machine learning, *PRX Quantum* **3**, 030341 (2022).
- [90] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, Geometric deep learning: Grids, groups, graphs, geodesics, and gauges, [arXiv:2104.13478](#).
- [91] Q. T. Nguyen, L. Schatzki, P. Braccia, M. Ragone, P. J. Coles, F. Sauvage, M. Larocca, and M. Cerezo, Theory for equivariant quantum neural networks, *PRX Quantum* **5**, 020328 (2024).
- [92] M. Ragone, P. Braccia, Q. T. Nguyen, L. Schatzki, P. J. Coles, F. Sauvage, M. Larocca, and M. Cerezo, Representation theory for geometric quantum machine learning, [arXiv:2210.07980](#).
- [93] C. Tüysüz, S. Y. Chang, M. Demidik, K. Jansen, S. Vallecorsa, and M. Grossi, Symmetry breaking in geometric quantum machine learning in the presence of noise, *PRX Quantum* **5**, 030314 (2024).
- [94] I. Nha Minh Le, O. Kiss, J. Schuhmacher, I. Tavernelli, and F. Tacchino, Symmetry-invariant quantum machine learning force fields, *New J. Phys.* **27**, 023015 (2025).
- [95] E. R. Anschuetz, A. Bauer, B. T. Kiani, and S. Lloyd, Efficient classical algorithms for simulating symmetric quantum systems, *Quantum* **7**, 1189 (2023).
- [96] T. A. Bespalova and O. Kyriienko, Quantum simulation and ground state preparation for the honeycomb Kitaev model, [arXiv:2109.13883](#).
- [97] L. G. Valiant, A theory of the learnable, *Commun. ACM* **27**, 1134 (1984).
- [98] R. Sweke, J.-P. Seifert, D. Hangleiter, and J. Eisert, On the quantum versus classical learnability of discrete distributions, *Quantum* **5**, 417 (2021).
- [99] N. Pirnay, R. Sweke, J. Eisert, and J.-P. Seifert, Superpolynomial quantum-classical separation for density modeling, *Phys. Rev. A* **107**, 042416 (2023).
- [100] C. Gyurik and V. Dunjko, Exponential separations between classical and quantum learners, [arXiv:2306.16028](#).
- [101] Y. Liu, S. Arunachalam, and K. Temme, A rigorous and robust quantum speed-up in supervised machine learning, *Nat. Phys.* **17**, 1013 (2021).
- [102] J. Jäger and R. V. Krems, Universal expressiveness of variational quantum classifiers and quantum kernels for support vector machines, *Nat. Commun.* **14**, 576 (2023).
- [103] T. Muser, E. Zapusek, V. Belis, and F. Reiter, Provable advantages of kernel-based quantum learners and quantum preprocessing based on Grover’s algorithm, [arXiv:2309.14406](#).
- [104] C. Umeano, V. E. Elfving, and O. Kyriienko, Geometric quantum machine learning of BQP^A protocols and latent graph classifiers, [arXiv:2402.03871](#).

- [105] J. Bowles, V. J. Wright, M. Farkas, N. Killoran, and M. Schuld, Contextuality and inductive bias in quantum machine learning, [arXiv:2302.01365](#).
- [106] R. Qu and Y. Bao, Quantum entanglement in n-qubit real equally weighted states, [arXiv:1207.2518](#).
- [107] M. Rossi, M. Huber, D. Bruß, and C. Macchiavello, Quantum hypergraph states, *New J. Phys.* **15**, 113022 (2013).
- [108] R. Molteni, C. Gyurik, and V. Dunjko, Exponential quantum advantages in learning quantum observables from classical data, *npj Quantum Inf.* **12**, 19 (2026).
- [109] J. Friedman, T. Hastie, and R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, *J. Stat. Softw.* **33**, 1 (2010).
- [110] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, Signature verification using a “siamese” time delay neural network, in *Advances in Neural Information Processing Systems*, edited by J. Cowan, G. Tesauro, and J. Alspecter, Vol. 6 (Morgan-Kaufmann, San Francisco, 1993).
- [111] G. R. Koch, Siamese neural networks for one-shot image recognition, in *Proceedings of the 32nd International Conference on Machine Learning, Lille, France* (JMLR, 2015), Vol. 37.
- [112] IBM quantum resources, <https://cloud.ibm.com/quantum/resources/systems>.
- [113] R. Raz and A. Tal, Oracle separation of bqp and ph, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019 (Association for Computing Machinery, New York, 2019), pp. 13–23.
- [114] S. Aaronson, BQP and the polynomial hierarchy, in *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC '10 (Association for Computing Machinery, New York, 2010), pp. 141–150.
- [115] S. Aaronson and A. Ambainis, Forrelation: A problem that optimally separates quantum from classical computing, in *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '15 (Association for Computing Machinery, New York, 2015), pp. 307–316.
- [116] S. Bravyi, D. Gosset, D. Grier, and L. Schaeffer, Classical algorithms for forrelation, [arXiv:2102.06963](#).
- [117] E. M. Stoudenmire and X. Waintal, Grover’s algorithm offers no quantum advantage, *Phys. Rev. X* **14**, 041029 (2024).
- [118] C. Umeano, S. Scali, and O. Kyriienko, Quantum community detection via deterministic elimination, *Phys. Rev. A* **112**, 052422 (2025).
- [119] S. Scali and C. Umeano, Qu-dos/dataset-barcode [Dataset], Zenodo (2024), <https://doi.org/10.5281/zenodo.13691943>.