

Common Applications:

Face Verification: Siamese networks are commonly used in face verification systems to determine whether two images belong to the same person.

Signature Verification: They can also verify if two signatures are from the same individual.

Image Similarity: Siamese networks can be used to classify whether two images are similar based on learned features.

Can Geometric Quantum Machine Learning Lead to Advantage in Barcode Classification?

Chukwudubem Umeano, Stefano Scali, and Oleksandr Kyriienko

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

(Dated: September 4, 2024)

A Siamese architecture is a type of neural network design that consists of two or more identical subnetworks (with shared weights) that work in parallel. These subnetworks are used to process two inputs simultaneously and compare them, typically to determine similarity or relationship between the inputs.

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 largely outperform their classical counterparts. We explain this difference in performance by analyzing correlated distributions used for composing our dataset. We relate the similarity testing with problems that showcase a proven maximal separation between the **BQP complexity class** and the **polynomial hierarchy**. While the ability to achieve advantage largely depends on how data are loaded, we discuss how similar problems can benefit from quantum machine learning.

The BQP (Bounded-Error Quantum Polynomial Time) complexity class is a concept from computational complexity theory that describes the kinds of problems that quantum computers can solve efficiently with a high probability of success.

Global correlation is essential for tasks like pattern recognition because it allows the model to detect relationships between data pairs on a larger scale, rather than just comparing local features.

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 [22–28], and physics-informed solving of differential equations [29–34], showing potential in areas where classical methods struggle. A distinct set of tools has been developed for embedding data into quantum computers [35–37], including various feature maps designed to exploit the high-dimensional space of quantum states for enhanced data representation and processing efficiency [38–44]. The use-cases of QML include image recognition [45–48], data analysis in high-energy physics [49–53], financial risk assessment [21, 54, 55], fraud detection [56–59], natural language processing [60–62], graph-based learning [63–66], power flow analysis [67], enhanced molecular simulations in chemistry [68], among many others.

Several open questions are yet to be answered before QML delivers its full promise. **First, we need to decide how to parameterize quantum models in the most advantageous way.** Do we get an advantage from training variational circuits, or is quantum embedding enough [69]? Or shall we use the structure of quantum kernels [70–72]? **Second, we still need to understand where a proven advantage in quantum machine learning can come from [73].** Does it mainly concern improving runtime and circuit depth [6, 74, 75]? Does it come from sampling (generative modelling) [26, 76–78], or do we gain from better **generalization of quantum models** [19, 79–82]?

One idea fuelling efficient quantum models leans on us-

ing symmetries of considered datasets and problems [83–85]. This approach, inspired by geometric deep learning [86], was dubbed as geometric quantum machine learning (GQML) [87–90]. 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 [91, 92].

To answer the second question several approaches were suggested. One can relate QML with probably approximately correct (PAC) learning [93], proving a superpolynomial advantage for density and Boolean function learning [94, 95], and for specific concept classes [71, 96]. 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 provable advantage based on **BQP problems** [97, 98] and **kernel-based quantum models** with **Grover's advantage** [99]. Recently, we have shown that geometric quantum machine learning can be used to learn **Simon's algorithm**

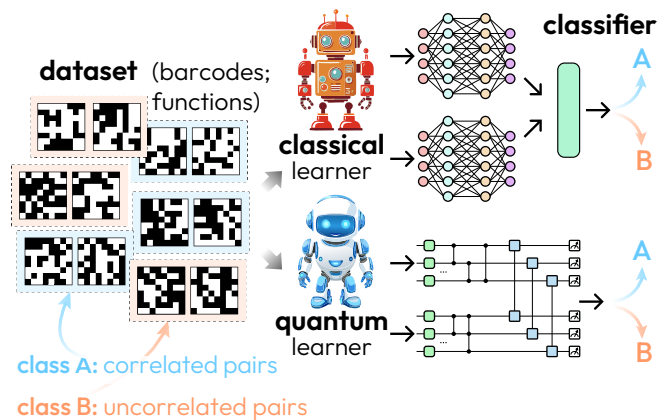


FIG. 1. Sketch of the machine learning workflows for distinguishing images (barcodes) based on their correlations. Pairs of barcodes are processed by a classical learner via Siamese deep neural networks. A quantum learner processes data embedded into quantum states via parametrized equivariant measurements,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

in BQP^A [100], suggesting that both QML architectures and optimal circuits can be learnt to represent known solutions with proven advantages.

In this paper, we tell a story of GQML being applied to the similarity testing problem. We show that for data sampled from specific distributions, certain classes of GQML models can provably outperform classical learners in terms of generalization, **discovering an optimal decision boundary with just few samples**. We set up the problem as distinguishing pairs of pixelated images—barcodes—each represented by a **binary string**. The task is to understand 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 a *generalization* advantage in learning. We explain the results based on established problems for superpolynomial separation between BQP^A and PH.

The problem.—We consider a dataset \mathcal{D} consisting of samples labelled 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 filled and empty 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** among samples coming in pairs.

We suggest that the data can be efficiently encoded into quantum states using the **phase states** $|\phi_x\rangle := \left(\sum_{j=1}^N (-1)^{x_1^{[j]}} |j\rangle\right) \otimes \left(\sum_{j=1}^N (-1)^{x_2^{[j]}} |j\rangle\right) / N =: |\phi_{x_1}\rangle \otimes |\phi_{x_2}\rangle$, where $x_\alpha^{[j]}$ ($\alpha = 1, 2$) 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 S_ϕ that corresponds to a remapping of bitstrings according to rules given by particular Boolean circuits, such that x_k and $\sigma(x_k)$ for $\sigma \in S_\phi$ 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.

The 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

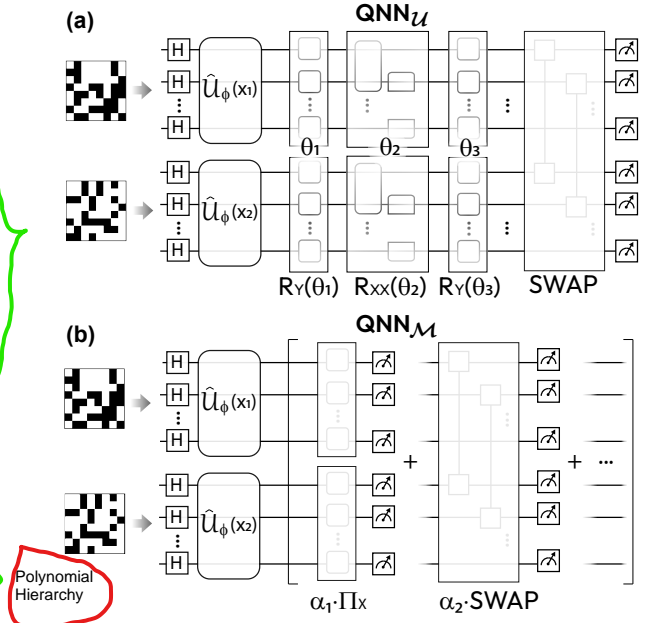


FIG. 2. Architectures for geometric quantum machine learning used in the study. (a) Quantum neural network approach based on equivariant unitary circuit (QNN_U) parametrized by a vector θ that processes a pair of barcodes encoded into quantum states. (b) QNN_M based on equivariant measurements that are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

in various studies [83–85, 87, 101]. We wish to determine whether each pair of barcodes x_m drawn from the dataset \mathcal{D} is correlated or not. 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

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

with $|\psi_0\rangle = H^{\otimes 2n} |0\rangle^{2n}$ being a uniform state and $\hat{U}_\phi(x)$ being the feature map that creates a phase state as $|\phi_x\rangle =: \hat{U}_\phi(x) |\psi_0\rangle$. 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 $V_\sigma : \mathcal{S} \rightarrow \text{Aut}(\mathcal{X}^2)$ [84, 87], 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 $\mathcal{S} = \mathbb{Z}_2 \times S_\phi$. The first symmetry corresponds to the exchange of two barcodes within the pair, $g(V_{\mathbb{Z}_2}[\mathbf{x}_{1,m}, \mathbf{x}_{2,m}]) = g(\mathbf{x}_{2,m}, \mathbf{x}_{1,m}) = g(\mathbf{x}_{1,m}, \mathbf{x}_{2,m})$, 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: $\hat{U}_{\mathbb{Z}_2} = \prod_{i=1}^n \text{SWAP}_{i,i+n}$ 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

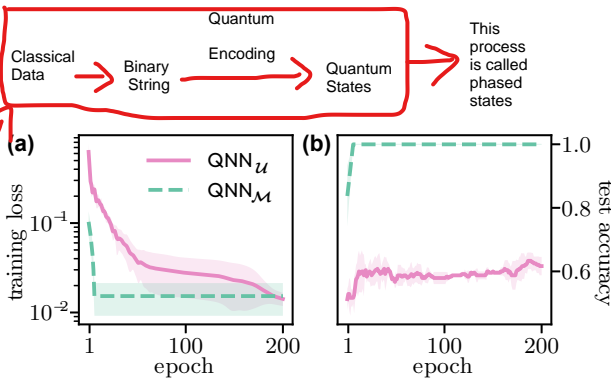


FIG. 3. Comparison of the training performance (a) and generalization (b) between the two GQML architectures, QNN_U and QNN_M . Both models were trained and tested on pairs of barcodes with 16 pixels each, corresponding to an 8-qubit quantum state for each input datum. Solid/dashed lines indicate the mean values across the 10 trials, while the shaded region indicates the standard deviation.

ingredients [84]: 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 [102] or **hypergraph states** [103], 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 and \hat{U}_ϕ . 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 data set, $\mathcal{L}_{\text{MSE}} = \frac{1}{M} \sum_{m=1}^M (ah_\theta(x_m) + b - y_m)^2$, 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.* [104]. 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 $\hat{U}(x) = e^{i\hat{H}(x)}$. 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 [104].

Let us detail the method, adapting it specifically to our task of determining whether pairs of barcodes are correlated or not. 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, $\varphi(x_m) = [\langle \hat{O}_1 \rangle_m, \langle \hat{O}_2 \rangle_m, \dots, \langle \hat{O}_K \rangle_m]$, where $\langle \hat{O}_i \rangle_m = \langle x_m | \hat{O}_i | x_m \rangle$. 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 **LASSO (least absolute shrinkage and selection operator) method**, a machine learning model that finds the optimal weights to minimize the **loss function** as [105]

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

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 non-correlated 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 non-local 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 $K = 10$ 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 10 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 trialled the two quantum models on pairs of 4×4 barcodes, averaging results for each model over 10 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

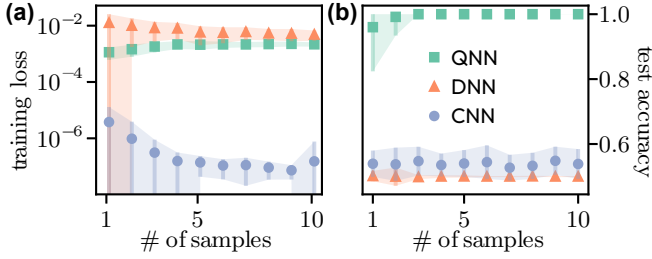


FIG. 4. **Generalization** of quantum and classical learners for the barcode similarity problem shown as a function of the number of samples from each class used in training. We show final training loss (a) and test accuracy (b) of quantum and classical networks, trained and tested on pairs of 1024-pixel 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).

Comparison Function: After the inputs pass through the subnetworks, their outputs are compared using a distance metric such as Euclidean distance or cosine similarity. This determines how similar or different the two inputs are.

be overfitting the training data, indicating that the features encoding the correlations are difficult to extract using the parameterized 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 $\text{QNN}_{\mathcal{U}}$ 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.

The results.—We analyze the described GQML workflows and compare them to classical similarity testing with Siamese neural networks [106, 107]. 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 $\text{QNN}_{\mathcal{M}}$ against those for classical networks, **classifying pairs** of 32×32 barcodes. We consider two types of networks with Siamese structures, namely a deep neural network (DNN) and a convolutional neural network (CNN). Siamese neural networks consist of two identical sub-networks that share the same weights and architecture. Let $f_{\theta}(\mathbf{x})$ represent the function learned by the sub-network 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 $\text{QNN}_{\mathcal{M}}$). 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 semi-logarithmic 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** (binary cross-entropy), added dropouts to the classical neural networks, used **cosine similarity** as a metric, and varied the batch size. None of these settings significantly improved results. Meanwhile, the quantum model again generalizes excellently at this larger system size; training on just 3 samples from each class ($M = 6$) is enough to classify all the pairs of 1024-pixel barcodes loaded into $2n = 20$ -qubit register.

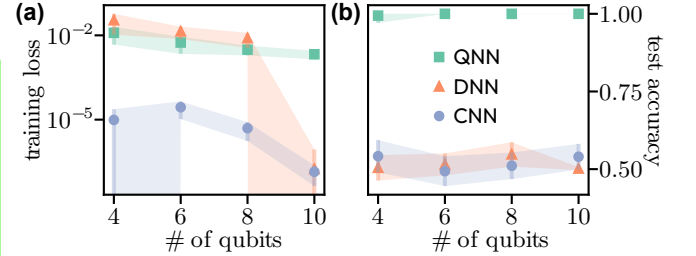


FIG. 5. **Generalization** of learners for the barcode similarity problem shown as a function of system size (number of qubits n for each image, $2n$ in total). Both, training loss (a) and test accuracy (b) remain steady across all system sizes, indicating that the quantum–classical separation is ubiquitous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

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.

The explanation.—Previously we have shown that the problem of barcode similarity testing can be solved with excel-

lent **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 [108], motivated by Aaronson [109], 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. [108] 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 [110] 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 $\hat{U}_\phi(x)$]. This $O(1)$ query complexity for distinguishing distributions showcases advantage over $\tilde{\Omega}(\sqrt{N})$ for the classical query complexity [110, 111]. The essence of the quantum algorithm is the estimation of the overlap between quantum representations of \mathbf{x}_1 and \mathbf{x}_2 as

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

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 $\hat{O}(\alpha^*) = \hat{H}^{\otimes 2n} \cdot \left(\prod_{i=1}^n \text{SWAP}_{i,i+n} \right)$ applied on $|\phi_{\mathbf{x}_1}\rangle \otimes |\phi_{\mathbf{x}_2}\rangle$. Such a Hermitian operator is readily available in 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 QNN_U 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 [98]. Jäger and Krams 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.**

Further discussion.—So, answering the question in the title, can we indeed get an advantage in learning global correlations for pixelated images or barcodes? The answer is still no, at this point. The reason lies in the classical nature of the data that we are dealing with. For pixels that are stored in a memory we can apply the same steps as in the quantum algorithm and obtain an efficient quantum-inspired solution as the Hadamard transform is easy to apply on a classical machine. The drawback of non-oracular data loading is that the process is much easier to dequantize as long as we look inside a black box [112]. It is also known that the 2-fold forrelation is classically simulable [111]. However, there is a scenario where a similar approach can give a genuine advantage. When generalizing Eq. (3) to 4-fold forrelation, the problem becomes classically intractable [111]. The data loading becomes more complicated, as consecutive applications of Hadamards and multi-CZ gates moves the embedding outside of the REW state subspace. At the same time, it hints to us that there are patterns (i.e. more complicated barcodes) and distributions that can be loaded into quantum states, and classified by building a decision boundary based on generalized forrelation.

Conclusion.—In this paper, we described a geometric quantum machine learning (GQML) approach that is able to learn 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 suggest that quantum machine learning surpasses classical deep learning, particularly in scenarios involving correlated distributions and limited training data. This conclusion was supported by the established maximal separations between the BQP complexity class and the polynomial hierarchy for classifying samples from forrelated and uniform distributions. Our work suggests the area where quantum machine learning can excel and gain advantage over classical machine learning.

Acknowledgement.—We acknowledge the funding from UK EPSRC award under the Agreement EP/Y005090/1.

The paper introduces an adaptive measurement-based approach (QNNM), where the quantum neural network is designed to respect the symmetries of the problem.

- [1] Y. LeCun, Y. Bengio, and G. Hinton, *Nature* **521**, 436 (2015).
- [2] A. Moskvichev, V. V. Odouard, and M. Mitchell, *The conceptarc benchmark: Evaluating understanding and generalization in the arc domain* (2023), arXiv:2305.07141 [cs.LG].

- [3] C. Bievers, *Nature* **619**, 686 (2023).
- [4] G. BakIr, T. Hofmann, B. Schölkopf, A. J. Smola, B. Taskar, and S. V. N. Vishwanathan, 'A Tutorial on Energy-Based Learning' in 'Predicting Structured Data' (The MIT Press,

- 2007).
- [5] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, *Quantum Science and Technology* **4**, 043001 (2019).
 - [6] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, *Nature* **549**, 195 (2017).
 - [7] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, *Nature* **567**, 209 (2019).
 - [8] J. Liu, M. Liu, J.-P. Liu, Z. Ye, Y. Wang, Y. Alexeev, J. Eisert, and L. Jiang, *Nature Communications* **15**, 434 (2024).
 - [9] P. W. Shor, *SIAM Journal on Computing* **26**, 1484 (1997), <https://doi.org/10.1137/S0097539795293172>.
 - [10] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information* (Cambridge university press, 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* (2023), [arXiv:2310.03011](https://arxiv.org/abs/2310.03011) [quant-ph].
 - [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, *Nature Reviews Physics* **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, *Physics Reports* **986**, 1–128 (2022).
 - [14] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, *Physical Review A* **98**, 032309 (2018).
 - [15] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, *Quantum* **4**, 226 (2020).
 - [16] H.-Y. Huang, M. Broughton, M. Mohseni, R. Babbush, S. Boixo, H. Neven, and J. R. McClean, *Nature Communications* **12**, 2631 (2021).
 - [17] N. A. Nghiem, S. Y.-C. Chen, and T.-C. Wei, *Phys. Rev. Res.* **3**, 033056 (2021).
 - [18] W. Li and D.-L. Deng, *Science China Physics, Mechanics & Astronomy* **65**, 220301 (2021).
 - [19] I. Cong, S. Choi, and M. D. Lukin, *Nature Physics* **15**, 1273 (2019).
 - [20] S. Monaco, O. Kiss, A. Mandarino, S. Vallecorsa, and M. Grossi, *Phys. Rev. B* **107**, L081105 (2023).
 - [21] D. Herman, C. Googin, X. Liu, Y. Sun, A. Galda, I. Safro, M. Pistoia, and Y. Alexeev, *Nature Reviews Physics* **5**, 450 (2023).
 - [22] J.-G. Liu and L. Wang, *Phys. Rev. A* **98**, 062324 (2018).
 - [23] C. Zoufal, A. Lucchi, and S. Woerner, *npj Quantum Information* **5**, 103 (2019).
 - [24] B. Coyle, D. Mills, V. Danos, and E. Kashefi, *npj Quantum Information* **6**, 60 (2020).
 - [25] A. E. Paine, V. E. Elfving, and O. Kyriienko, *Advanced Quantum Technologies* **n/a**, 2300065 (2023).
 - [26] O. Kyriienko, A. E. Paine, and V. E. Elfving, *arXiv preprint arXiv:2202.08253* (2022).
 - [27] S. Kasture, O. Kyriienko, and V. E. Elfving, *Phys. Rev. A* **108**, 042406 (2023).
 - [28] M. Hibat-Allah, M. Mauri, J. Carrasquilla, and A. Perdomo-Ortiz, *Communications Physics* **7**, 68 (2024).
 - [29] O. Kyriienko, A. E. Paine, and V. E. Elfving, *Physical Review A* **103**, 052416 (2021).
 - [30] M. Lubasch, J. Joo, P. Moinier, M. Kiffner, and D. Jaksch, *Physical Review A* **101**, 010301 (2020).
 - [31] S. Markidis, *Frontiers in Applied Mathematics and Statistics* **8**, 10.3389/fams.2022.1036711 (2022).
 - [32] N. Heim, A. Ghosh, O. Kyriienko, and V. E. Elfving, *Quantum model-discovery* (2021), [arXiv:2111.06376](https://arxiv.org/abs/2111.06376) [quant-ph].
 - [33] A. E. Paine, V. E. Elfving, and O. Kyriienko, *Physical Review A* **107**, 032428 (2023).
 - [34] 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* (2023), [arXiv:2308.01827](https://arxiv.org/abs/2308.01827) [quant-ph].
 - [35] M. Schuld and N. Killoran, *Phys. Rev. Lett.* **122**, 040504 (2019).
 - [36] A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner, *Nature Computational Science* **1**, 403 (2021).
 - [37] T. Goto, Q. H. Tran, and K. Nakajima, *Phys. Rev. Lett.* **127**, 090506 (2021).
 - [38] S. Altares-López, A. Ribeiro, and J. J. García-Ripoll, *Quantum Science and Technology* **6**, 045015 (2021).
 - [39] C. A. Williams, A. E. Paine, H.-Y. Wu, V. E. Elfving, and O. Kyriienko, *Quantum chebyshev transform: Mapping, embedding, learning and sampling distributions* (2023), [arXiv:2306.17026](https://arxiv.org/abs/2306.17026) [quant-ph].
 - [40] 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, *Phys. Rev. A* **107**, 042615 (2023).
 - [41] B. Jaderberg, A. A. Gentile, Y. A. Berrada, E. Shishenina, and V. E. Elfving, *Let quantum neural networks choose their own frequencies* (2023), [arXiv:2309.03279](https://arxiv.org/abs/2309.03279) [quant-ph].
 - [42] C. Umeano and O. Kyriienko, *Ground state-based quantum feature maps* (2024), [arXiv:2404.07174](https://arxiv.org/abs/2404.07174) [quant-ph].
 - [43] H.-Y. Wu, V. E. Elfving, and O. Kyriienko, *Multidimensional quantum generative modeling by quantum hartley transform* (2024), [arXiv:2406.03856](https://arxiv.org/abs/2406.03856) [quant-ph].
 - [44] S. Thabet, M. Djellabi, I. Sokolov, S. Kasture, L.-P. Henry, and L. Henriët, *Quantum positional encodings for graph neural networks* (2024), [arXiv:2406.06547](https://arxiv.org/abs/2406.06547) [quant-ph].
 - [45] S. Y. Chang, B. L. Saux, S. Vallecorsa, and M. Grossi, in *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium* (2022) pp. 4907–4910.
 - [46] N. Schetakis, D. Aghamalyan, P. Griffin, and M. Boguslavsky, *Scientific Reports* **12**, 11927 (2022).
 - [47] A. Senokosov, A. Sedykh, A. Sagingalieva, B. Kyriacou, and A. Melnikov, *Machine Learning: Science and Technology* **5**, 015040 (2024).
 - [48] K. Shen, B. Jobst, E. Shishenina, and F. Pollmann, *Classification of the fashion-mnist dataset on a quantum computer* (2024), [arXiv:2403.02405](https://arxiv.org/abs/2403.02405) [quant-ph].
 - [49] Belis, Vasilis, González-Castillo, Samuel, Reissel, Christina, Vallecorsa, Sofia, Combarro, Elías F., Dissertori, Günther, and Reiter, Florentin, *EPJ Web Conf.* **251**, 03070 (2021).
 - [50] S. Y.-C. Chen, T.-C. Wei, C. Zhang, H. Yu, and S. Yoo, *Phys. Rev. Res.* **4**, 013231 (2022).
 - [51] K. A. Woźniak, V. Belis, 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* (2023), [arXiv:2301.10780](https://arxiv.org/abs/2301.10780) [quant-ph].
 - [52] V. Belis, P. Odagiu, M. Grossi, F. Reiter, G. Dissertori, and S. Vallecorsa, *Machine Learning: Science and Technology* **5**, 035010 (2024).
 - [53] 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, Z. Holmes, S. Kühn, D. Lacroix, R. Lewis, D. Lucchesi, M. L. Martínez, F. Meloni, A. Mezzacapo, S. Montangero, L. Nagano,

- V. R. Pascuzzi, V. Radescu, E. R. Ortega, A. Roggero, J. Schuhmacher, J. Seixas, P. Silvi, P. Spentzouris, F. Tacchino, K. Temme, K. Terashi, J. Tura, C. Tüysüz, S. Vallecorsa, U.-J. Wiese, S. Yoo, and J. Zhang, *PRX Quantum* **5**, 037001 (2024).
- [54] D. J. Egger, C. Gambella, J. Marecek, S. McFaddin, M. Mevissen, R. Raymond, A. Simonetto, S. Woerner, and E. Yndurain, *IEEE Transactions on Quantum Engineering* **1**, 1 (2020).
- [55] 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, H. Isselane, and D. M'tamon, *Phys. Rev. Res.* **5**, 043117 (2023).
- [56] O. Kyriienko and E. B. Magnusson, Unsupervised quantum machine learning for fraud detection (2022), [arXiv:2208.01203 \[quant-ph\]](#).
- [57] M. Grossi, N. Ibrahim, V. Radescu, R. Loredó, K. Voigt, C. von Altrock, and A. Rudnik, *IEEE Transactions on Quantum Engineering* **3**, 1 (2022).
- [58] 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* (2024), [arXiv:2407.12618 \[quant-ph\]](#).
- [59] N. Innan, A. Sawaika, A. Dhor, S. Dutta, S. Thota, H. Gokal, N. Patel, M. A.-Z. Khan, I. Theodonis, and M. Bennai, *Quantum Machine Intelligence* **6**, 7 (2024).
- [60] B. Coecke, G. de Felice, K. Meichanetzidis, and A. Toumi, *Foundations for near-term quantum natural language processing* (2020), [arXiv:2012.03755 \[quant-ph\]](#).
- [61] D. Widdows, A. Alexander, D. Zhu, C. Zimmerman, and A. Majumder, *Annals of Mathematics and Artificial Intelligence* **10.1007/s10472-024-09940-y** (2024).
- [62] D. Widdows, W. Aboumradi, D. Kim, S. Ray, and J. Mei, *Quantum natural language processing* (2024), [arXiv:2403.19758 \[quant-ph\]](#).
- [63] G. Verdon, T. McCourt, E. Luzhnica, V. Singh, S. Leichenauer, and J. Hidary, *Quantum graph neural networks* (2019), [arXiv:1909.12264 \[quant-ph\]](#).
- [64] L.-P. Henry, S. Thabet, C. Dalyac, and L. Henriët, *Phys. Rev. A* **104**, 032416 (2021).
- [65] 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* (2024), [arXiv:2403.11931 \[quant-ph\]](#).
- [66] S. Yu, C. Peng, Y. Wang, A. Shehzad, F. Xia, and E. R. Hancock, *Quantum graph learning: Frontiers and outlook* (2023), [arXiv:2302.00892 \[cs.LG\]](#).
- [67] Z. Kaseb, M. Möller, G. T. Balducci, P. Palensky, and P. P. Vergara, *Electric Power Systems Research* **235**, 110677 (2024).
- [68] O. Kiss, F. Tacchino, S. Vallecorsa, and I. Tavernelli, *Machine Learning: Science and Technology* **3**, 035004 (2022).
- [69] M. Cerezo, M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, E. R. Anschuetz, and Z. Holmes, Does provable absence of barren plateaus imply classical simulability? or, why we need to rethink variational quantum computing (2023), [arXiv:2312.09121 \[quant-ph\]](#).
- [70] M. Schuld, Supervised quantum machine learning models are kernel methods (2021), [arXiv:2101.11020 \[quant-ph\]](#).
- [71] C. Gyurik, R. Molteni, and V. Dunjko, *Limitations of measure-first protocols in quantum machine learning* (2023), [arXiv:2311.12618 \[quant-ph\]](#).
- [72] S. Jerbi, C. Gyurik, S. C. Marshall, R. Molteni, and V. Dunjko, *Nature Communications* **15**, 5676 (2024).
- [73] M. Schuld and N. Killoran, *PRX Quantum* **3**, 030101 (2022).
- [74] N. Wiebe, D. Braun, and S. Lloyd, *Phys. Rev. Lett.* **109**, 050505 (2012).
- [75] S. Lloyd, M. Mohseni, and P. Rebentrost, *Quantum algorithms for supervised and unsupervised machine learning* (2013), [arXiv:1307.0411 \[quant-ph\]](#).
- [76] M. B. Mansky, J. Nüßlein, D. Bucher, D. Schuman, S. Zielinski, and C. Linnhoff-Popien, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01 (2023) pp. 485–495.
- [77] D. Hangleiter and J. Eisert, *Rev. Mod. Phys.* **95**, 035001 (2023).
- [78] K. Gili, M. Hibat-Allah, M. Mauri, C. Ballance, and A. Perdomo-Ortiz, *Quantum Science and Technology* **8**, 035021 (2023).
- [79] 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, *Science* **376**, 1182 (2022).
- [80] M. C. Caro, H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles, *Nature Communications* **13**, 4919 (2022).
- [81] C. Umeano, A. E. Paine, V. E. Elfving, and O. Kyriienko, What can we learn from quantum convolutional neural networks? (2023), [arXiv:2308.16664 \[quant-ph\]](#).
- [82] E. Gil-Fuster, J. Eisert, and C. Bravo-Prieto, Understanding quantum machine learning also requires rethinking generalization (2023), [arXiv:2306.13461 \[quant-ph\]](#).
- [83] H. Zheng, Z. Li, J. Liu, S. Strelchuk, and R. Kondor, *PRX Quantum* **4**, 020327 (2023).
- [84] J. J. Meyer, M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert, *PRX Quantum* **4**, 010328 (2023).
- [85] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, *PRX Quantum* **3**, 030341 (2022).
- [86] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, Geometric deep learning: Grids, groups, graphs, geodesics, and gauges (2021), [arXiv:2104.13478 \[cs.LG\]](#).
- [87] Q. T. Nguyen, L. Schatzki, P. Braccia, M. Ragone, P. J. Coles, F. Sauvage, M. Larocca, and M. Cerezo, *PRX Quantum* **5**, 020328 (2024).
- [88] 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* (2023), [arXiv:2210.07980 \[quant-ph\]](#).
- [89] C. Tüysüz, S. Y. Chang, M. Demidik, K. Jansen, S. Vallecorsa, and M. Grossi, *PRX Quantum* **5**, 030314 (2024).
- [90] I. N. M. Le, O. Kiss, J. Schuhmacher, I. Tavernelli, and F. Tacchino, *Symmetry-invariant quantum machine learning force fields* (2023), [arXiv:2311.11362 \[quant-ph\]](#).
- [91] E. R. Anschuetz, A. Bauer, B. T. Kiani, and S. Lloyd, *Quantum* **7**, 1189 (2023).
- [92] T. A. Bespalova and O. Kyriienko, Quantum simulation and ground state preparation for the honeycomb kitaev model (2021), [arXiv:2109.13883 \[quant-ph\]](#).
- [93] L. G. Valiant, *Commun. ACM* **27**, 1134–1142 (1984).
- [94] R. Sweke, J.-P. Seifert, D. Hangleiter, and J. Eisert, *Quantum* **5**, 417 (2021).
- [95] N. Pirnay, R. Sweke, J. Eisert, and J.-P. Seifert, *Phys. Rev. A* **107**, 042416 (2023).
- [96] C. Gyurik and V. Dunjko, *Exponential separations between classical and quantum learners* (2023), [arXiv:2306.16028 \[quant-ph\]](#).
- [97] Y. Liu, S. Arunachalam, and K. Temme, *Nature Physics* **17**, 1013 (2021).

- [98] J. Jäger and R. V. Krems, *Nature Communications* **14**, 576 (2023).
- [99] T. Muser, E. Zapusek, V. Belis, and F. Reiter, *Provable advantages of kernel-based quantum learners and quantum preprocessing based on grover’s algorithm* (2023), [arXiv:2309.14406 \[quant-ph\]](#).
- [100] C. Umeano, V. E. Elfving, and O. Kyriienko, *Geometric quantum machine learning of bqp^a protocols and latent graph classifiers* (2024), [arXiv:2402.03871 \[quant-ph\]](#).
- [101] J. Bowles, V. J. Wright, M. Farkas, N. Killoran, and M. Schuld, Contextuality and inductive bias in quantum machine learning (2023), [arXiv:2302.01365 \[quant-ph\]](#).
- [102] R. Qu and Y. Bao, *arXiv preprint arXiv:1207.2518* (2012).
- [103] M. Rossi, M. Huber, D. Bruß, and C. Macchiavello, *New Journal of Physics* **15**, 113022 (2013).
- [104] R. Molteni, C. Gyurik, and V. Dunjko, *arXiv preprint arXiv:2405.02027* (2024).
- [105] J. Friedman, T. Hastie, and R. Tibshirani, *Journal of statistical software* **33**, 1 (2010).
- [106] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, in *Advances in Neural Information Processing Systems*, Vol. 6, edited by J. Cowan, G. Tesauro, and J. Alspector (Morgan-Kaufmann, 1993).
- [107] G. R. Koch, R. Zemel, and R. Salakhutdinov, in *Computer Science* (2015).
- [108] R. Raz and A. Tal, in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019 (Association for Computing Machinery, New York, NY, USA, 2019) p. 13–23.
- [109] S. Aaronson, in *Proceedings of the Forty-Second ACM Symposium on Theory of Computing*, STOC ’10 (Association for Computing Machinery, New York, NY, USA, 2010) p. 141–150.
- [110] S. Aaronson and A. Ambainis, in *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC ’15 (Association for Computing Machinery, New York, NY, USA, 2015) p. 307–316.
- [111] S. Bravyi, D. Gosset, D. Grier, and L. Schaeffer, *Classical algorithms for forrelation* (2021), [arXiv:2102.06963 \[quant-ph\]](#).
- [112] E. M. Stoudenmire and X. Waintal, Grover’s algorithm offers no quantum advantage (2023), [arXiv:2303.11317 \[quant-ph\]](#).