

TWIRLATOR: A Pipeline for Analyzing Subgroup Symmetry Effects in Quantum Machine Learning Ansatzes

Valter Uotila
valter.uotila@aalto.fi
Aalto University &
University of Helsinki
Espoo, Finland

Väinö Mehtola
vaino.mehtola@vtt.fi
VTT
Espoo, Finland

Ilmo Salmenperä
ilmo.salmenpera@helsinki.fi
University of Helsinki
Helsinki, Finland

Bo Zhao
bo.zhao@aalto.fi
Aalto University
Espoo, Finland

Abstract

Symmetry is a strong inductive bias in geometric deep learning and its quantum counterpart, and has attracted increasing attention for improving the trainability of QML models. Yet incorporating symmetries into quantum machine learning (QML) ansatzes is not free: symmetrization often adds gates and constrains the circuits. To understand these effects, we present **TWIRLATOR**, which is an automated pipeline that symmetrizes parameterized QML ansatzes and quantifies the trade-offs as the amount of symmetry increases. TWIRLATOR models partial symmetries by the size of a subgroup of the symmetric group, enabling analysis between the “no symmetry” and “full symmetry” extremes. Across 19 common ansatz patterns, TWIRLATOR symmetrizes circuits with respect to any subgroup of S_n and measures (1) generator drift, (2) circuit overhead (depth and size), and (3) expressibility and entangling capability. The experimental evaluation focuses on subgroups of S_4 and S_5 . TWIRLATOR reveals that larger subgroups typically increase circuit overhead, reduce expressibility, and often increase entangling capability. The pipeline and results provide practical guidance for selecting ansatz patterns and symmetry levels that balance hardware cost and model performance in symmetry-aware QML applications.

Keywords

geometric quantum machine learning, symmetries, subgroup twirling, expressibility, entangling capability

ACM Reference Format:

Valter Uotila, Väinö Mehtola, Ilmo Salmenperä, and Bo Zhao. 2026. TWIRLATOR: A Pipeline for Analyzing Subgroup Symmetry Effects in Quantum Machine Learning Ansatzes. In *17th International Workshop on Quantum Software Engineering (Q-SE '26), April 12–18, 2026, Rio de Janeiro, Brazil*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3786150.3788613>

1 Introduction

Many foundational classical machine learning models benefit from additional structure imposed by the learning tasks. Convolutional models exploit translation equivariance in vision [12, 13], and graph neural networks enforce permutation symmetry for relational data [10, 23]. Beyond conceptual elegance, symmetry constraints can produce concrete gains in accuracy and data efficiency. For instance, group-equivariant CNNs reduce test error on rotated

MNIST from 5.03% (standard CNN) to 2.28% (rotation-equivariant P4CNN) under matched settings [5]. In scientific machine learning, equivariant architectures can also reduce data requirements by orders of magnitude: NequIP [27] reports competitive accuracy with 1,000× fewer training samples in a molecular modeling benchmark [3]. These successes underpin geometric deep learning [4] and have recently carried over to *geometric* and *equivariant* quantum machine learning (QML) [18–21, 24], where symmetry-aware models can possibly improve trainability, generalization and avoid certain problems such as barren plateaus [8, 14, 29].

From a *quantum software engineering* perspective, “using symmetry” is a pipeline decision that should be identified, engineered, validated, and benchmarked. Incorporating symmetries into QML models affects the QML software stack end-to-end: after identifying symmetries in the learning task, they influence circuit construction, compilation/transpilation, the trainability and learnability of QML models, and resource estimations under hardware constraints. Concretely, stronger symmetry can turn a circuit that is deployable under a depth/size budget into one that is not. The current tooling provides limited reusable automation for applying symmetry transformations consistently across different ansatzes and learning tasks and for producing metric-driven analysis that makes these trade-offs explicit to practitioners.

Although many learning tasks have natural underlying symmetries, especially in physics [15], there is no quantum software that would enable us to symmetrize quantum machine learning ansatzes in an automated manner. For example, Qiskit, Pennylane, and Cirq provide limited tools for symmetrizing QML models. As noted in [21], the practical realization of these models requires a solid understanding of group representation theory.

Symmetry is often treated as a binary choice (none vs. full symmetry), while *partial* symmetry—which may provide a better cost-performance balance—is rarely explored systematically [21]. To quantify the impact of increasing symmetry, we approach the symmetrization in terms of subgroups. While most of the previous research has not explicitly formulated symmetrization in terms of subgroups, we present this straightforward generalization. Subgroup-based symmetrization can be considered to produce partially equivariant models, which have been more extensively studied in classical geometric and equivariant machine learning [1, 6, 22]. Nevertheless, this study shows that employing subgroup-based symmetries rather than the whole group produces more expressive and shorter ansatzes.

TWIRLATOR structures the evaluation around three research questions: **(RQ1)** How does subgroup symmetrization alter ansatz generators as symmetry increases? **(RQ2)** How does circuit overhead



This work is licensed under a Creative Commons Attribution 4.0 International License.
Q-SE '26, Rio de Janeiro, Brazil
© 2026 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-2383-4/2026/04
<https://doi.org/10.1145/3786150.3788613>

(depth and gate count) scale with subgroup size across common ansatzes? **(RQ3)** How do expressibility and entangling capability change as symmetry increases?

This paper makes following contributions:

(1) A reusable symmetry-aware QML pipeline. Based on angle encoding, induced unitary representations of symmetry groups, and the Pauli twirling formula, we implement an automated pipeline that symmetrizes some of the most common quantum machine learning ansatz patterns with respect to the subgroups. TWIRLATOR automatically generalizes to any common ansatzes.

(2) Subgroup-based symmetry as a tunable knob in the QML model construction. We propose that symmetry can be understood as a tunable knob in QML models, since one need not use the entire group but can incorporate only subgroups, as this work demonstrates. This reframes symmetry as a continuum rather than a binary choice, thereby supporting the systematic exploration of symmetry effects.

(3) A benchmark-driven characterization of symmetry. We concretely study the effects of the gate symmetrization process by computing a difference between original and symmetric generators, which we call *generator drift*. Additionally, we obtain circuit-related metrics, such as depth, and the expressibility and entangling capabilities of the symmetrized ansatzes for varying degrees of symmetry. We have open sourced TWIRLATOR¹.

2 Background

TWIRLATOR treats symmetry as an explicit engineering knob in QML ansatz design. This section introduces the background needed to operationalize that knob: (i) how we model task symmetries with subgroups of a symmetric group, (ii) how TWIRLATOR enforces a target subgroup via the twirling formula, and (iii) the expressibility and entangling capability metrics used in our evaluation.

2.1 Symmetries in learning tasks

Symmetries can appear in (quantum) machine learning in multiple ways [4, 15, 18]. Symmetries can be discrete or continuous, and they are often divided into equivariant or invariant symmetries [15]. Symmetries can appear only in the source data, or we can identify that the mapping, i.e., the learning problem, respects certain symmetries due to the problem’s nature. In this work, we focus on symmetries that appear in the learning problem.

In this work, we assume an input space \mathcal{X} and an output space \mathcal{Y} for a given machine learning problem [25]. Given a set of samples $\mathcal{S} = \{(x_i, y_i) \in \mathcal{X} \times \mathcal{Y}\}_{i=1}^N$, the goal is to learn a function $f: \mathcal{X} \rightarrow \mathcal{Y}$ so that it approximates the unknown distribution $D \subset \mathcal{X} \times \mathcal{Y}$. Depending on the learning task, it is possible to identify that the distribution $D \subset \mathcal{X} \times \mathcal{Y}$ respects certain symmetries, which can be encoded in the model.

In this work, symmetries are modeled using symmetric groups S_n and their subgroups $S'_k \subset S_n$. The connection between the group theory and the vector spaces (data) is given by representation theory [21]. We can define a representation of a group S_n on a vector space V as a mapping $\varphi: S_n \times V \rightarrow V$, which satisfies that $\varphi(s): V \rightarrow V$ is linear for every $s \in S_n$, $\varphi(e, v) = v$ for the neutral

¹<https://github.com/valterUo/twirlator>

element $e \in S_n$ and $\varphi(s_1, \varphi(s_2, v)) = \varphi(s_1 s_2, v)$ for every $s_1, s_2 \in S_n$ and $v \in V$. Using this representation, we obtain $\varphi(S'_k)$, which is the corresponding subgroup of linear mappings $V \rightarrow V$.

Next, we assume that $\mathcal{X} \subset V$. Then, the mapping f is invariant under the symmetric subgroup S'_k if $f(\varphi(s, x)) = f(x)$ for all $x \in \mathcal{X}$ and for all $s \in S'_k$. With this definition, the level of symmetry in the mapping can be quantified by considering the size of the subgroup $S'_k \subset S_n$, for which the invariance holds. If the subgroup is a trivial one-element group consisting of only the neutral element, the data does not admit any symmetries. Neutral element is always mapped to the identity matrix I and this does not change the vector, so that $f(I[x]) = f(x)$ is trivially satisfied. On the other hand, if the subgroup satisfies $S'_k = S_n$, the data is maximally symmetric with respect to the fixed symmetry group S_n .

In this work, we consider that the representation $\varphi: S_n \times V \rightarrow V$ is also realized by unitary matrices such that $\varphi(S_n) = \{U_s \mid s \in S_n\}$. Let U_{init} be the initial data encoding layer for the quantum machine learning model as a unitary operator. We consider special *induced unitary representations of symmetry groups* which are required to satisfy the following condition

$$U_{\text{init}}(\varphi(s, x)) = U_s U_{\text{init}}(x) U_s^\dagger, \quad (1)$$

for all $x \in \mathcal{X}$. This condition necessarily encodes the fact that the induced representation $\{U_s \mid s \in S_n\}$ in a certain sense commutes with the data encoding layer $U_{\text{init}}(x)$ with respect to the group S_n .

Now the symmetry in the learning problem is modeled with unitaries $\{U_s \mid s \in S_n\}$ which also satisfy Equation 1. The equation establishes the connection between data encoding and symmetries. It is also important to note that Equation 1 is defined at the operator level, not the state level.

2.2 Gate symmetrization

This work implements gate symmetrization, which is performed with the Pauli twirling formula [18]. For this subsection, we assume that for a fixed data encoding U_{init} and a subgroup S'_k of the symmetry group S_n , we have computed the induced unitary representations as defined in Equation 1.

Since the construction involves parameterized quantum circuits, the symmetrization relies on generators. All of the gates in the ansatzes can be expressed in terms of fixed generators as $R_G(\theta) = e^{-i\theta G}$, where G is the generator for the parametrized gate $R_G(\theta)$. Based on Proposition 1 (Commuting generators) [18], it suffices to apply the symmetrization process only to the generators. Hence, assume a fixed gate set expressed in terms of generators. Then, we define the Pauli twirling formula [7, 18] over the subgroup S'_k :

$$\mathcal{T}[G] = \frac{1}{|S'_k|} \sum_{s \in S'_k} U_s G U_s^\dagger. \quad (2)$$

Applying the twirling formula to the generators \mathcal{G} of the original gate set, we obtain the equivariant generator set $\mathcal{T}[\mathcal{G}]$. These generators are used to construct the corresponding symmetric parameterized circuit.

2.3 Expressibility

Expressibility of a parametrized quantum circuit refers to the circuit's ability to express pure states so that they cover the Hilbert space [26]. In practice, the expressibility is computed based on fidelities which are estimated so that we sample pairs of states $|\varphi_\theta\rangle$ and $|\varphi_\phi\rangle$ for different randomly initialized parametrizations θ and ϕ . Then, the corresponding fidelities $F = |\langle\varphi_\theta|\varphi_\phi\rangle|^2$ are considered as random variables. Let $P(F, \theta)$ be the estimated probability distribution of fidelities for parameters θ , which results from sampling states from the parametrized circuit. Let $P_{\text{Haar}}(F)$ be the Haar distribution [16], which can be computed with the closed form formula $P_{\text{Haar}}(F) = (N-1)(1-F)^{N-2}$. The distributions $P(F, \theta)$ and $P_{\text{Haar}}(F)$ are compared with Kullback-Leibler (KL) divergence, i.e., relative entropy [11]

$$D_{\text{KL}}(P(F, \theta) \| P_{\text{Haar}}(F)) = \sum_x P(F, \theta)[x] \log \frac{P(F, \theta)[x]}{P_{\text{Haar}}[x]}.$$

The expressibility of a parametrized quantum circuit is defined as its KL-divergence.

2.4 Entangling capability

The entangling capability of a parameterized quantum circuit quantifies how close, on average, the state it produces is to the maximally entangled state. Following the description in [26], we have computed Meyer-Wallach (MW) entanglement measure [17]. Let $\Theta = \{\theta_i \mid 1 \leq i \leq n\}$ be a collection of parameter vectors for a parameterized quantum circuit. Previous work [2] defines Meyer-Wallach entanglement measure as

$$Q = \frac{2}{|\Theta|} \sum_{\theta_i \in \Theta} \left(1 - \frac{1}{n} \sum_{k=1}^n \text{Tr}(\rho_k^2(\theta_i)) \right),$$

where n is the number of qubits and ρ_k is the reduced density matrix of qubit k . The value $\text{Tr}(\rho_k^2)$ is the purity of the qubit k . If qubit k is entangled with others, then the state ρ_k becomes mixed and $\text{Tr}(\rho_k^2) < 1$. Hence, this formula can be used to estimate the total entangling capability of the parameterized circuit. According to this measure, if a parameterized circuit produces only separable (product) states, its entangling capability is 0. In contrast, a circuit that creates highly entangled states attains a value close to 1. Previous research [26] identified that the entangling capability of Haar random states is around 0.82.

3 Methodology

Using the symmetry knob and metrics defined in §2, we now describe TWIRLATOR's evaluation methodology. The methodology consists of three steps: (i) selecting a representative collection of ansatzes [26] and enumerating subgroups of S_4 , (ii) generating subgroup-symmetrized circuits via the twirling formula, and (iii) comparing original and symmetrized circuits using generator drift, circuit overhead, expressibility, and entangling capability metrics.

3.1 Selected ansatzes and subgroups

TWIRLATOR studies various symmetry-related properties of 19 common ansatz patterns from [26]. The ansatz implementations used in this work originate from [9], which provides them as PennyLane

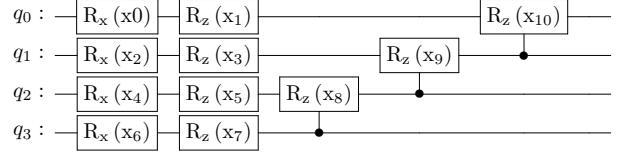


Fig. 1: Ansatz 3 from [26] with four qubits

circuits. An example ansatz with id 3 is presented in Fig. 1. Making the results partially comparable to [26], we focus on four-qubit versions of these ansatzes. Because of selecting four qubits, we construct the symmetric group S_4 and consider its possible subgroups, which have orders as 1 (trivial one-element subgroup), 2, 3, 4, 6, 8, 12, and 24 (complete S_4 has $4! = 24$ elements). We have considered all of the subgroups of S_4 .

3.2 Subgroup-based symmetrization

Since most of the gates in the ansatzes are parametrized, the system relies on Proposition 1 about commuting generators [18], which states that it suffices to symmetrize the generators instead of parametrized gates as unitaries. To compute the generators for the gates in the ansatzes, we use PennyLane's `qml.generator` function. Then, the system applies the previously described induced unitary representations as permutation matrices to the generators. Next, it uses the Pauli twirling formula in Equation 2 to obtain the equivariant set of symmetrized generators. Finally, the same set of parameters is assigned for each symmetrized generator, and the resulting set of generators is synthesized into parametrized quantum circuits using Qiskit. When this process is applied to the example circuit in Fig. 1 using a fixed four-element subgroup of the symmetric group S_4 , we obtain a circuit in Fig. 2.

3.3 Evaluation metrics

Next, we introduce the evaluation metrics used to assess performance differences between the original ansatzes and their symmetrized counterparts. Let A_i be the ansatz from [26] for $1 \leq i \leq 19$. Let $\mathcal{G}(A_i)$ be the set of generators for the gates in the ansatz. Let $\mathcal{T}[\mathcal{G}(A_i)]$ be the equivariant gateset for the gates defined by generators in $\mathcal{G}(A_i)$, when they are symmetrized with respect to a fixed subgroup. In the following, the norms are Frobenius norms. Then, we consider the three classes of metrics as follows.

Generator drift. To quantify the difference between the original ansatzes and their symmetrized versions, we compute the average difference of the norm

$$\|G - G_{\text{twirl}}\|,$$

for $G_{\text{twirl}} \in \mathcal{T}[\mathcal{G}(A_i)]$ and $G \in \mathcal{G}(A_i)$, where G_{twirl} is the symmetrized version of G . In other words, this means that the average norm for the projection onto the symmetric subspace is defined as

$$\frac{1}{|\mathcal{T}[\mathcal{G}(A_i)]|} \sum_{G_{\text{twirl}} \in \mathcal{T}[\mathcal{G}(A_i)]} \frac{1}{|\mathcal{G}(A_i)|} \sum_{G \in \mathcal{G}(A_i)} \|G - G_{\text{twirl}}\|. \quad (3)$$

Note that the previous value is interestingly depth-invariant. The value is computed for each gate's generator and then averaged over all gates in the ansatz. Since increasing the number of layers in the

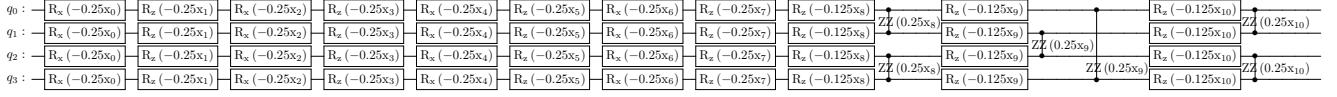


Fig. 2: Symmetrized circuit corresponding the circuit in Fig. 1 using a four element subgroup of symmetric group S_4

ansatz structures repeats the same gates multiple times, the value does not depend on the ansatz depth. Hence, in the results section, we present the results without referencing the depth.

Circuit-related metrics. When the gates in ansatz A_i are replaced with the corresponding symmetrized gates based on generators in $\mathcal{T}_{\mathcal{U}}[\mathcal{G}(A_i)]$, we synthesize the circuit using the same gates as the original ansatzes have. Synthesis is performed using the Qiskit transpile function with the highest optimization level 3. The results slightly depend on the synthesis method, and Qiskit does not perform all possible optimizations because the synthesized circuits contain parameters. For example, one can identify that some R_z gates could still be combined in the circuit in Fig. 2, reducing the gate count and depth slightly. Nevertheless, these metrics provide a concrete quantification of the overhead introduced by symmetrization as a function of the system’s degree of symmetry. Circuit-related metrics also enable us to compare those ansatzes that are expensive to symmetrize in terms of gate overhead. We observed that the overall trends for the total number of gates, two-qubit gates, and circuit depth are consistent. Therefore, we present the results using absolute gate counts in terms of the subgroup size.

Expressibility and entangling capability. For the same symmetrized circuits as in the previous point, we compute the expressibility and entangling capability as defined in [26]. The values are calculated with open-source QLeet software [2], which we updated². We use the hyperparameters from [26] so that we sample 10 000 states for which we compute the independent fidelities (i.e., the same state does not appear twice in the fidelity computations). Although QLeet allows continuous expressibility comparison, we follow the description in [26] and approximate the distributions $P(F, \theta)$ and $P_{\text{Haar}}(F)$ with 75 bins. With this setup, we obtained results that were closely similar to those of the previous research [26] for the original ansatzes without symmetrization. We also used 10 000 states to compute the entangling capability. This sample size was found statistically robust [26].

4 Implementation

TWIRLATOR implements subgroup-based ansatz symmetrization for parameterized circuits under discrete permutation symmetries. It takes as input a circuit and a subgroup, and outputs a symmetrized circuit along with artifacts such as induced representations and symmetrized generators.

Precomputed subgroups. Using the precomputed subgroups for S_4 , we construct the unitary representations that satisfy the induced unitary representation definition in Equation 1. For a fixed symmetric group S_n and its subgroup $S'_k \subset S_n$, the elements in the group are permutations. Since we focus on angle encoding, which is expressed in terms of a diagonal unitary U_{init} , this means that the induced representations U_s are simply permutation matrices,

²<https://github.com/valterUo/qleet-light>

which are easy to construct. If we introduce entanglement into the data-encoding unitary U_{init} and apply, for example, amplitude encoding, the resulting representations are more complex to build. Considering induced unitary representations for advanced data encoding methods is part of future research.

The system also contains other precomputed symmetric group structures for extended experiments. Using Sage Math, we have sampled at most 30 random subgroups of size k for each $1 \leq k \leq |S_n|$ for $3 \leq n \leq 9$. Some experiments also rely on the corresponding five-qubit circuits and the group S_5 . Generally, the implementation handles any parametrized circuit with respect to any subgroup of a symmetric group, provided the model uses angle encoding.

Induced unitary representations. Based on the permutations in the subgroup, the corresponding permutation matrix, i.e., the induced unitary representation, is constructed as follows. Let n be the number of qubits in the system and σ be a permutation on $\{0, 1, \dots, n-1\}$ such that $\sigma(i)$ gives us the target position of qubit i . This means we want to encode an action that moves the bit originally at position i to position $\sigma(i)$. We define σ^{-1} as the inverse permutation satisfying $\sigma^{-1}(\sigma(i)) = i$. The corresponding qubit permutation unitary $U_\sigma \in \mathbb{C}^{2^n \times 2^n}$ operates on computational basis states as

$$U_\sigma |b_0 b_1 \dots b_{n-1}\rangle = |b_{\sigma^{-1}(0)} b_{\sigma^{-1}(1)} \dots b_{\sigma^{-1}(n-1)}\rangle.$$

The bit originally at position i moves to position $\sigma(i)$.

5 Evaluation

Computationally, the most expensive experiments are those for expressibility and entangling capability, since they require simulating circuits. For those, we utilized 760 CPUs divided into 95 tasks. Each task was allocated 8 CPUs with 4GB of memory, and the running time ranged from 30 minutes to 2 hours, depending on the depth of the circuits. The software used for quantum computing simulations was Qiskit, and Pennylane was used for ansatz construction and generator computation.

5.1 RQ1: How much does symmetrization change the generators?

Four qubits (S_4). The results for the operator difference norm, i.e., generator drift, which we defined in Equation 3, are presented in Fig. 3a for four-qubit ansatzes. When the subgroup size is 1, meaning that the subgroup consists of only the group’s neutral element, we correctly obtain 0 as a difference since $G = G_{\text{twirl}}$ holds always. The norm increases with the size of the symmetry group, as a larger number of symmetries in the problem necessitates more extensive modifications to the generator G . The average generator drift across ansatzes increases from 0 (subgroup size 1) to 1.72 (subgroup size 24). Some interesting ansatzes are 2, 9, 11 and 15, whose generator drifts are 2.11, 2.26, 2.01, and 2.42 respectively

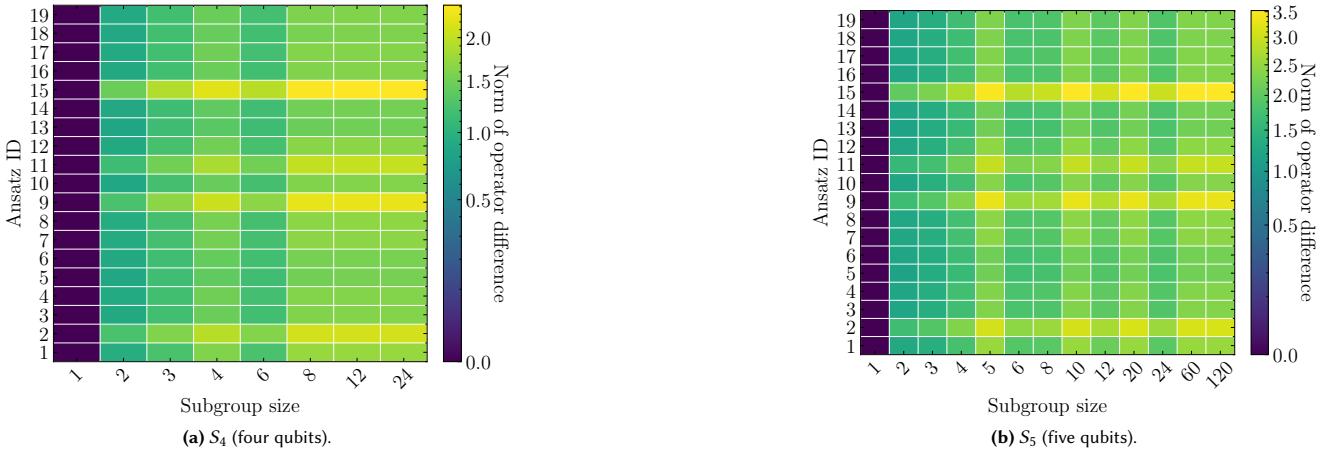


Fig. 3: Generator drift under subgroup-based symmetrization. We report the operator difference norm (Equation 3) between original gate generators and their subgroup-twirled counterparts as symmetry strength increases (subgroup size on the x-axis). Higher values indicate larger generator changes induced by symmetrization.

when the subgroup size is 24, and these are visible as yellow stripes. The commonality between these ansatzes is that they all contain non-parametrized gates, such as CNOTs, CZs, and Hadamard gates, and we suspect that their symmetrization slightly increases the difference. We can also see that the norm for subgroups of size 6 is slightly lower than for subgroups of sizes 4 and 8, which makes this case stand out.

Five qubits (S_5). A similar trend is observed for larger instances. Corresponding results for the 5-qubit ansatzes computed over the subgroups of S_5 are presented in Fig. 3b. In these cases, the average generator drift across ansatzes increases from 0 to 2.51. Ansatzes 2, 9, 11, and 15 again perform differently. We also observe that subgroups of sizes 5, 10, 20, 60, and 120 show higher norms. Looking at the subgroups more closely, one can identify that many subgroups of size 5 are contained in subgroups of size 10, many of size 10 are contained in subgroups of size 20, and so on. Subgroups that contain smaller subgroups consequently inherit their properties. This would explain the stripes in Fig. 3a and Fig. 3b.

5.2 RQ2: How much circuit overhead does symmetrization introduce?

We start by comparing the sizes of the original circuit to the symmetrized circuits. The circuit size is the total number of instructions in the circuit. The sizes of depth-1 ansatzes for four qubits are presented in Fig. 4a, and for five qubits in Fig. 4b.

In Fig. 4a, most ansatzes grow moderately as symmetry increases, remaining below 200 operations even at $|S'_k|=24$. However, a small number of circuits dominate the cost tail. The most expensive circuit to implement is the circuit with id 6. While symmetrization creates substantial overhead for a couple of circuits, in many cases, the circuits do not become so long that the hardware cannot support them. Interestingly, in all cases, the cost of implementing symmetric circuits for subgroups of sizes 12 and 24 is the same.

We also observe that gate counts do not grow monotonically. This is especially apparent in Fig. 4b, where circuits corresponding to subgroups of size 6, 12, and 24 show lower overhead compared

to the cases around them. We believe that the reason for this is the same as the reason why we observed the lower operator norm for the size 6 subgroup in Fig. 3a. A similar phenomenon can be seen in the five-qubit cases in Fig. 4b. The figure shows peaks at the same subgroups as we observe higher norm in Fig. 3b.

5.3 RQ3: How does symmetry affect expressibility and entangling capability?

Expressibility. To highlight differences in the sizes of the symmetry groups, we again fix the circuit depth to 1 and plot the results as a function of subgroup sizes. Fig. 5 shows the expressibility results for the studied circuits. The results are ordered from the least to the most expressible, according to the expressibility value of the original circuit.

One can observe that the most expressive ansatzes are also the most expensive to implement, as shown in Fig. 4. The values for the original ansatzes are close to those obtained in the original study [26]. We consistently observe that expressibility decreases as symmetry increases. In most cases, the relative decrease in expressibility stays approximately the same.

Furthermore, we see that the expressibility of symmetric ansatzes corresponding to subgroups of sizes 3 and 6 yields similar KL-divergence values. We assume that this phenomenon again follows from the reasons that we discussed around Fig. 3a and Fig. 3b. Subgroups of size 3 are often subgroups for subgroups of size 6, and thus they show similar performance. The KL-divergence values also seem to be the same for symmetric ansatzes corresponding to subgroups of sizes 12 and 24, presumably for the same reason.

In the studied cases, the expressibility results indicate that most of the symmetrized circuits are not in the previously identified favorable expressibility region [26]. Fig. 7a demonstrates that even increasing the depth does not seem to increase reduced expressibility, and this applies to the other circuits in the study. Thus, the

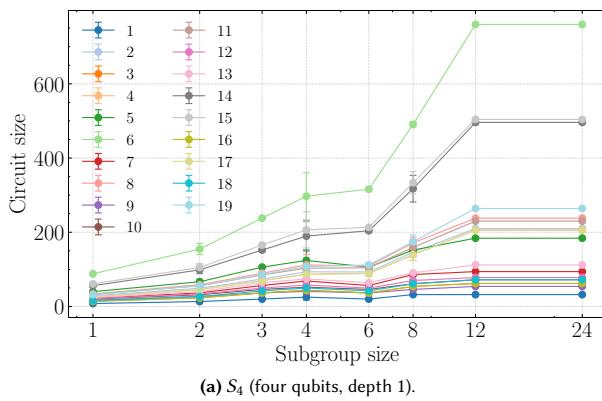
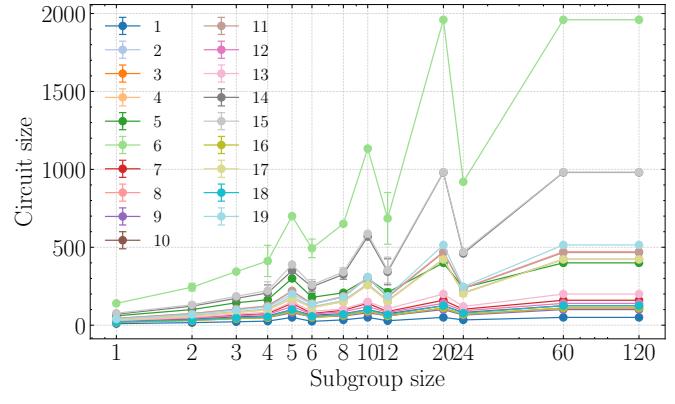
(a) S_4 (four qubits, depth 1).(b) S_5 (five qubits, depth 1).

Fig. 4: Circuit overheads under subgroup-based symmetrization. We report the synthesized circuit size (total operations after Qiskit transpilation, opt-level 3) as symmetry increases (subgroup size on the x-axis). Most ansatzes grow moderately, while a few families incur sharp blow-ups and occasional non-monotonic changes across subgroup orders.

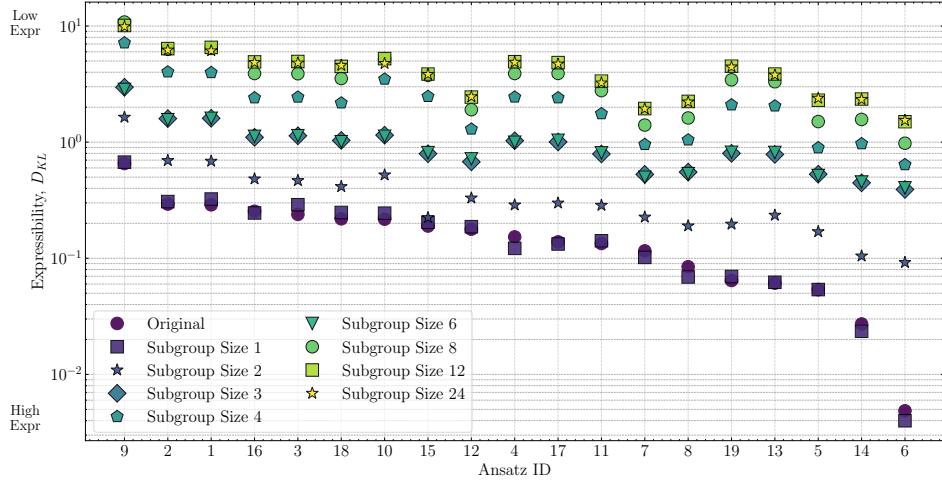


Fig. 5: Expressibility under subgroup-based symmetrization (depth 1): larger subgroups increase D_{KL} (lower expressibility).

apparent solution is to use only sufficiently small or partial symmetries. It is possible that analyzing symmetries in the learning problem becomes part of the QML architecture design, allowing us to find the optimal model design.

Entanglement capability. Fig. 6 presents the results of the entangling capability studies. While it is intuitive that expressibility decreases as symmetry increases, it is less intuitive how entanglement behaves in this context. We did not observe similarly consistent patterns as in the case of expressibility.

The general trend that we observe is that entanglement increases in most cases. This is most likely because the symmetrization introduces parametrized gates that create entanglement. Interesting exceptions are the ansatzes 2, 9, 10, 11, and 15 that have non-parametric two-qubit gates. The results indicate that these non-parametrized entangling gates are transformed into gates that create less entanglement in the circuit (ansatzes 2, 9, 11, 15) or that

maintain approximately the same entanglement capability (ansatz 10). We also note that ansatzes 3 and 16 exhibit very similar characteristics, which is understandable given their structures.

None of the previous metrics examined the effect of increasing the circuit depth. We found that although increasing the circuit depth changes the absolute values, the relative influence of symmetries is consistent across depths for each ansatz. In essence, the patterns observed at depth 1 already reflect the main effects of symmetrization, and greater depths rarely alter this picture. For example, present the expressibility values for ansatz 6 with different depths in Fig. 7a, which shows that the effects of symmetrization affect similarly. In the same way, Fig. 7b shows the entangling capability results for ansatz 18.

6 Discussion

Across 19 ansatz patterns, subgroup-based symmetrization exposes a consistent symmetry–overhead–performance trade-off across ansatzes. As subgroup size increases, (i) generator drift grows, (ii)

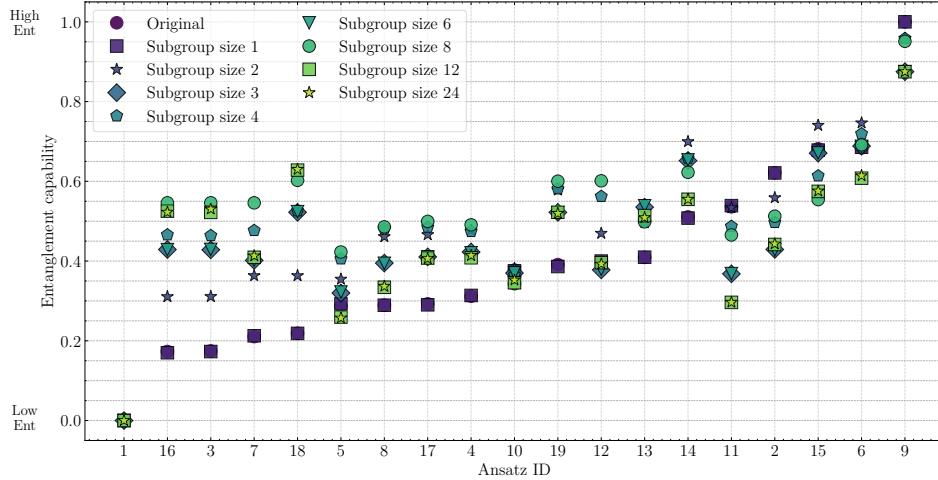
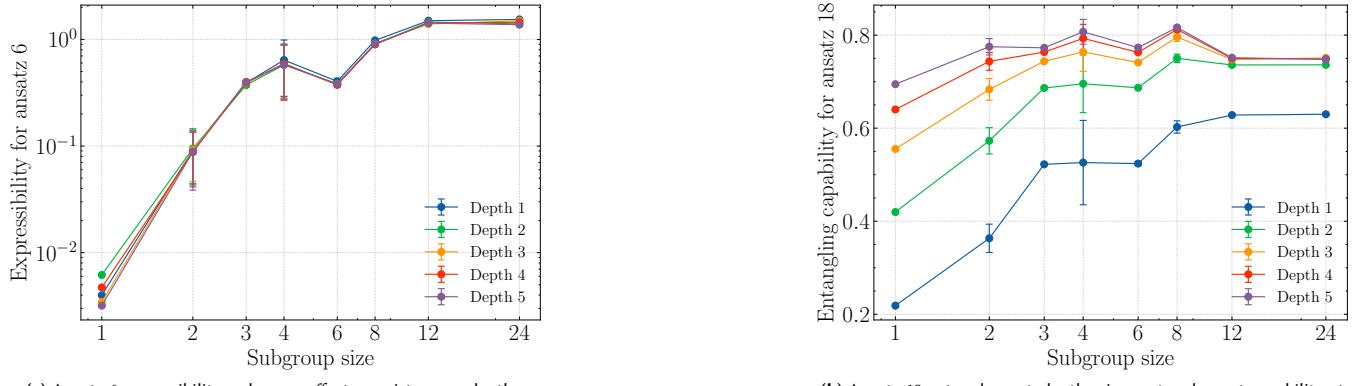


Fig. 6: Entanglement capability of the ansatzes when depth is 1



(a) Ansatz 6 expressibility: subgroup effects persist across depths.

(b) Ansatz 18 entanglement: depth raises entanglement capability at weak symmetry and saturates under strong symmetry.

Fig. 7: Depth sensitivity under subgroup-based symmetrization: depth shifts absolute values, while relative subgroup trends remain stable.

synthesized circuits become larger, and (iii) expressibility decreases, while entangling capability often increases. Without symmetrization (subgroup size 1), our evaluation reproduced the reference baselines reported in [26], confirming that TWIRLATOR preserves the original benchmark.

Why expressibility drops under stronger symmetry. Large subgroups cause the largest decrease in expressibility because they enforce invariance across many qubits simultaneously. Learning problems that exhibit all these symmetries are often relatively constrained and structured. We still need to determine the exact practical consequences, because the low expressibility of symmetric models should still perform well on data that share the same symmetries. Thus, whether low or high expressibility and entangling capability values are desirable depends on the specific use case and the underlying hardware constraints.

Implications for quantum software engineering. This work and the implementation open possibilities for tooling and automation for creating and analyzing symmetry effects in QML. The

pipeline demonstrates that properties such as expressibility, entangling capability, and symmetry-induced overhead can be automatically estimated. TWIRLATOR and similar tools could be used in many ways, such as ranking ansatz choices, deciding whether to adjust the degree of symmetry in ansatzes, stopping training early, or switching between ansatzes. When implementing such tools, one should keep in mind that expressibility and entangling capability are proxy metrics, and the models' true performance will depend on many factors.

Scope and limitations. This was not a study of real-world use cases, and thus, we did not examine how symmetries affect trainability. This will be part of future work. Another promising direction is to investigate the effect of symmetries on measurement operations, which have been identified as less studied than other parts of quantum machine learning models [28].

We excluded studies on more complex data encoding methods, such as amplitude encoding. Computing an induced unitary representation for amplitude encoding is more challenging. It might be easier to compute a specific relaxation for induced unitary representations. Fixing the state $|00 \dots 0\rangle$, we obtain $U_{\text{init}}(\varphi(s, x))|00 \dots 0\rangle =$

$U_s U_{\text{init}}(x) U_s^\dagger |00 \dots 0\rangle$. We call these state-level induced unitary representations. Then, we compute U_s only so that the previous state-level equation holds. This creates an induced unitary representation that does not necessarily satisfy the operator-level Equation 1.

7 Conclusion

This paper presented TWIRLATOR, an automated symmetry-aware QML pipeline that operationalizes subgroup-based symmetrization as a practical knob between no symmetry and full-group symmetry. Across 19 common ansatz patterns, TWIRLATOR quantifies how increasing symmetry reshapes generator structure, circuit overhead, expressibility, and entangling capability, revealing a consistent trade-off: larger subgroups create higher overhead, reduce expressibility, and often increase entanglement. TWIRLATOR and the results help practitioners reason about deployability and the expected performance of ansatzes early in the design of QML models. Our current scope focuses on permutation symmetries and angle encoding; future work includes extending TWIRLATOR to richer symmetry groups and representations, supporting more complex encodings and measurements, and validating the trade-offs on end-to-end tasks and real-world datasets.

Acknowledgments

This work is funded by Business Finland (grant number 169/31/2024), Research Council of Finland (grant number 362729) and the *Finnish Quantum Flagship Exploratory Project* to PI Bo Zhao.

References

- [1] Matthew Ashman, Cristiana Diaconu, Adrian Weller, Wessel Bruinsma, and Richard E. Turner. 2025. Approximately equivariant neural processes. In *Proceedings of the 38th International Conference on Neural Information Processing Systems* (Vancouver, BC, Canada) (NIPS '24). Curran Associates Inc., Red Hook, NY, USA, Article 3078, 36 pages.
- [2] Utkarsh Azad and Animesh Sinha. 2023. qLEET: visualizing loss landscapes, expressibility, entangling power and training trajectories for parameterized quantum circuits. *Quantum Information Processing* 22, 6 (6 2023), 256. doi:10.1007/s11128-023-03998-z
- [3] Simon Batzner, Albert Musaelian, Lixin Sun, Mario Geiger, Jonathan P. Mailoa, Mordechai Kornbluth, Nicola Molinari, Tess E. Smidt, and Boris Kozinsky. 2022. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials. *Nature Communications* 13 (2022), 2453. doi:10.1038/s41467-022-29939-5
- [4] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. 2021. Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges. doi:10.48550/arXiv.2104.13478 arXiv:2104.13478 [cs].
- [5] Taco Cohen and Max Welling. 2016. Group Equivariant Convolutional Networks. In *Proceedings of The 33rd International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 48)*. Maria Florina Balcan and Kilian Q. Weinberger (Eds.). PMLR, New York, New York, USA, 2990–2999. https://proceedings.mlr.press/v48/cohen16.html
- [6] Marc Finzi, Greg Benton, and Andrew Gordon Wilson. 2021. Residual pathway priors for soft equivariance constraints. In *Proceedings of the 35th International Conference on Neural Information Processing Systems (NIPS '21)*. Curran Associates Inc., Red Hook, NY, USA, Article 2299, 13 pages.
- [7] Jonas Helsen. 2019. *Quantum information in the real world*. Ph. D. Dissertation. TU Delft. doi:10.4233/uuid:312b719d-32bc-4219-82bb-86febc2abcc
- [8] Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles. 2022. Connecting Ansatz Expressibility to Gradient Magnitudes and Barren Plateaus. *PRX Quantum* 3, 1 (Jan. 2022), 010313. doi:10.1103/PRXQuantum.3.010313
- [9] Linus Jern, Valter Uotila, Cong Yu, and Bo Zhao. 2025. Agent-Q: Fine-Tuning Large Language Models for Quantum Circuit Generation and Optimization. In *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01. 1621–1632. doi:10.1109/QCE5121.2025.00179
- [10] Thomas N. Kipf and Max Welling. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017, Conference Track Proceedings*. OpenReview.net. doi:10.48550/arXiv.1609.02907 arXiv:1609.02907 [cs].
- [11] S. Kullback and R. A. Leibler. 1951. On Information and Sufficiency. *The Annals of Mathematical Statistics* 22, 1 (March 1951), 79–86. doi:10.1214/aoms/1177729694
- [12] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. 1989. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* 1, 4 (Dec. 1989), 541–551. doi:10.1162/neco.1989.1.4.541
- [13] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86, 11 (Nov. 1998), 2278–2324. doi:10.1109/5.726791
- [14] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. 2018. Barren plateaus in quantum neural network training landscapes. *Nature Communications* 9, 1 (Nov. 2018), 4812. doi:10.1038/s41467-018-07090-4
- [15] Väinö Mehtola. 2025. *Group Invariance in Quantum Kernels for Vector Boson Scattering Identification at the LHC*. Master's thesis. Aalto University. https://aaltodoc.aalto.fi/handle/123456789/136648
- [16] Antonio Anna Mele. 2024. Introduction to Haar Measure Tools in Quantum Information: A Beginner's Tutorial. *Quantum* 8 (May 2024), 1340. doi:10.22331/q-2024-05-08-1340
- [17] David A. Meyer and Nolan R. Wallach. 2002. Global entanglement in multiparticle systems. *J. Math. Phys.* 43, 9 (9 2002), 4273–4278. doi:10.1063/1.1497700
- [18] Johannes Jakob Meyer, Marian Mularski, Elies Gil-Fuster, Antonio Anna Mele, Francesco Arzani, Alissa Wilms, and Jens Eisert. 2023. Exploiting symmetry in variational quantum machine learning. *PRX Quantum* 4, 1 (March 2023), 010328. doi:10.1103/PRXQuantum.4.010328 arXiv:2205.06217 [quant-ph].
- [19] Quynh T. Nguyen, Louis Schatzki, Paolo Braccia, Michael Ragone, Patrick J. Coles, Frédéric Sauvage, Martin Larocca, and M. Cerezo. 2024. Theory for Equivariant Quantum Neural Networks. *PRX Quantum* 5, 2 (May 2024), 020328. doi:10.1103/PRXQuantum.5.020328
- [20] Isabel Nha Minh Le, Oriel Kiss, Julian Schuhmacher, Ivano Tavernelli, and Francesco Tacchino. 2025. Symmetry-invariant quantum machine learning for fields. *New Journal of Physics* 27, 2 (Feb. 2025), 023015. doi:10.1088/1367-2630/adad0c
- [21] Michael Ragone, Paolo Braccia, Quynh T. Nguyen, Louis Schatzki, Patrick J. Coles, Frédéric Sauvage, Martin Larocca, and M. Cerezo. 2023. Representation Theory for Geometric Quantum Machine Learning. doi:10.48550/arXiv.2210.07980 arXiv:2210.07980 [quant-ph].
- [22] David W. Romero and Suhas Lohit. 2022. Learning partial equivariances from data (NIPS '22). Curran Associates Inc., Red Hook, NY, USA, Article 2642, 13 pages.
- [23] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. 2009. The Graph Neural Network Model. *IEEE Transactions on Neural Networks* 20, 1 (Jan. 2009), 61–80. doi:10.1109/TNN.2008.2005605
- [24] Louis Schatzki, Martin Larocca, Quynh T. Nguyen, Frédéric Sauvage, and M. Cerezo. 2024. Theoretical guarantees for permutation-equivariant quantum neural networks. *npj Quantum Information* 10, 1 (Jan. 2024), 12. doi:10.1038/s41534-024-00804-1
- [25] Shai Shalev-Shwartz and Shai Ben-David. 2014. *Understanding machine learning: from theory to algorithms*. Cambridge university press, New York.
- [26] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. 2019. Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms. *Advanced Quantum Technologies* 2, 12 (2019), 1900070. arXiv:https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/qute.201900070 doi:10.1002/qute.201900070
- [27] Chuin Wei Tan, Marc L. Descoteaux, Mit Kotak, Gabriel de Miranda Nascimento, Séan R. Kavanagh, Laura Zichi, Menghang Wang, Aadit Saluja, Yizhong R. Hu, Tess Smidt, Anders Johansson, William C. Witt, Boris Kozinsky, and Albert Musaelian. 2025. High-performance training and inference for deep equivariant interatomic potentials. arXiv:2504.16068 [physics.comp-ph] https://arxiv.org/abs/2504.16068
- [28] Valter Uotila, Ilmo Salmenperä, Leo Becker, Arianne Meijer van de Griend, Aakash Ravindra Shinde, and Jukka K. Nurminen. 2025. Perspectives on Utilization of Measurements in Quantum Algorithms. In *2025 IEEE International Conference on Quantum Software (QSW)*. IEEE Computer Society, Los Alamitos, CA, USA, 1–12. doi:10.1109/QSW67625.2025.00016
- [29] Maxwell T. West, Jamie Heredge, Martin Sevior, and Muhammad Usman. 2024. Provably Trainable Rotationally Equivariant Quantum Machine Learning. *PRX Quantum* 5, 3 (July 2024), 030320. doi:10.1103/PRXQuantum.5.030320