



Contribution ID: 212

Type: poster

Permutation and Lorentz invariance in quantum kernels

Permutation and Lorentz Invariance in Quantum Kernels

Authors:

Michele Grossi (CERN)

Santeri Laurila (CERN & Helsinki Institute of Physics HIP)

Massimiliano Incudini (Intel)

Väinö Mehtola (HIP & VTT Technical Research Centre)

Date: December 2024

Introduction

Symmetries are central to the standard model and many other frameworks in physics. When addressing a learning task where the data is expected to originate from a process well-described by such physical frameworks, it is crucial to incorporate these symmetries into the inductive bias of the machine learning (ML) model. This leads to a theoretically provable improvement in the performance of the ML model in terms of sample complexity [1] and generalization error [2]. Empirical evidence is provided by the remarkable success of convolutional neural networks in image processing [3] and equivariant graph neural networks in network analysis under certain assumptions [4].

Quantum machine learning (QML) is a toolbox capable of solving learning tasks relevant to fundamental physics provably beyond classical means [5]. To fully exploit this potential, it is crucial to develop techniques that integrate symmetries directly into QML algorithms. Meyer et al. (2023) [6] has focused on variational algorithms for near-term quantum devices, introducing a circuit-first strategy to enforce symmetries by designing an equivariant set of gates. A modification of this method has been successfully employed in the classification of high-energy particle decays while respecting $SO(1, 3)$ -Lorentz invariance using quantum neural networks [7]. Forming your circuit out of an equivariant gate set increases the potential of inductive bias of the model and reduces the risk of exponential concentration by restricting the potential function space of the model.

In this work, we aim to expand the range of near-term quantum machine learning techniques that enforce symmetries and are applicable to tasks in high-energy physics. Specifically, we focus on developing *invariant quantum kernels*. Quantum kernels are similarity measures computed with the assistance of a quantum computer, with the *embedding quantum kernel* [8] being one of the most widely studied formulations. In this approach, classical data is encoded into the density matrix of a quantum system through a parameterized unitary or a quantum channel. The similarity measure between pairs of classical data is then given by the trace of the product of their corresponding density matrices. The mathematical framework of kernel methods enables some theoretical analysis, such as the task-model alignment [9] and the exponential concentration of kernel values [10], which cannot be applied to quantum neural networks due to the non-convexity of the training landscape of the latter¹.

Bringing permutation invariance of certain features of a datapoint to a kernel brings about subtleties one must consider regarding the kernel type, embedding strategy, (un)intentional feature engineering, and the observable. First, permutation invariance in global fidelity kernels can only be achieved with embeddings which

effectively define new, combined features that are combinations of the permutable features and invariant under permutation, such as $(x_i + x_j)/2$ or $x_i x_j$. In other words, the features that are permutation invariant do not affect the output similarity individually, but together via a permutation invariant pre-processing function $g(x_i, x_j)$ that effectively defines new aggregate features. To tackle this issue and provide functional dependence directly on the raw features, we propose a variant of the linear projected quantum kernel that utilizes a partial measurement.

¹ Some analyses on QNN share similarities with the ones on quantum kernels, e.g., the ones focusing on exponential concentration, or again by approximating the model as a quantum neural tangent kernel.

Contributions

We recall the definition of the embedding quantum kernel.

Definition (Embedding Quantum Kernel)

Given a feature embedding $\mathbf{x} \in \mathcal{X} \mapsto \rho(\mathbf{x})$, which maps classical data into a quantum density operator, the embedding quantum kernel κ is a mapping:

$$\kappa : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}, \kappa(\mathbf{x}, \mathbf{x}') = \text{Tr}[\rho(\mathbf{x})\rho(\mathbf{x}')].$$

An invariant generator set is always defined with respect to an initial data embedding unitary $U(\mathbf{x})$ and the relevant symmetry group $\mathcal{S}$ for whose elements one needs to find induced unitary representations. These induced representations are defined as follows.

Definition (Induced Unitary Representations of Symmetry Transformations) Let $\mathcal{S}$ be a symmetry group under whose action the data domain $\mathcal{X}$ remains invariant for the given learning task, i.e., $f(\mathbf{x}) = f(V_s[\mathbf{x}])$ where f is the learned function, $s \in \mathcal{S}$ and $V_s[\mathbf{x}]$ represents the action of s on $\mathbf{x} \in \mathcal{X}$. Given then an embedding unitary $U(\mathbf{x})$ embedding the datum $\mathbf{x} \in \mathcal{X}$, the set of unitaries $\{U_s | s \in \mathcal{S}\}$ induced by the symmetry group $\mathcal{S}$ implement the symmetry transformations for the embedding in:

$$U(V_s[\mathbf{x}]) = U_s U(\mathbf{x}) U_s^\dagger$$

With this, we can formally define the invariant generator set as follows.

Definition (Invariant Generator Set) Let $\mathcal{S}$ be a symmetry group under whose action the data domain $\mathcal{X}$ remains invariant for the given learning task, i.e., $f(\mathbf{x}) = f(V_s[\mathbf{x}])$, and the set of unitaries induced by the symmetry transformations $\{U_s | s \in \mathcal{S}\}$. An invariant generator set consists precisely of those unitaries W for which it holds that, for all $s \in \mathcal{S}$:

$$[W, U_s] = 0$$

The set of generators for which this condition holds can be found using the Twirling formula:

$$T_U[X] = \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} U_s X U_s^\dagger.$$

We consider two different kinds of permutation invariances: datapoint-wise and feature-wise.

Definition (Feature-Wise Permutation) Given a data vector $\mathbf{x} = [x_1, \dots, x_i, \dots, x_j, \dots, x_n]$, a feature-wise permutation of $\mathbf{x}$ is defined as $\pi_{i,j}(\mathbf{x}) = [x_1, \dots, x_j, \dots, x_i, \dots, x_n]$.

Definition (Datapoint-Wise Permutation) Given a function $g : \mathcal{X} \times \mathcal{X} \mapsto \mathbb{R}$ evaluated as $g(\mathbf{x}, \mathbf{x}')$, a datapoint-wise permutation yields $g(\mathbf{x}', \mathbf{x})$.

Definition (Feature-Wise Invariant Embedding Quantum Kernel) Let κ be an embedding quantum kernel on the data domain $\mathcal{X}$. Let $\mathcal{S}$ be a symmetry group under whose action $\mathcal{X}$ remains invariant for the given learning task. Then, a feature-wise permutation invariant embedding quantum kernel has the property:

$$\kappa(\mathbf{x}, \mathbf{x}') = \kappa(\pi(\mathbf{x}), \pi'(\mathbf{x}'))$$

for all permutation transformations $\pi, \pi' \in \mathcal{S}$ of features in $\mathbf{x}, \mathbf{x}'$.

Figure here, see pdf

(a): Fidelity test. **(b):** Quantum circuit implementing the function γ in Equation 7.

We focus here on specific kinds of symmetry, the permutation and Lorentz invariance, and a specific kind of initial embedding unitary, the angle embedding. Specifically, we rely on a quantum circuit

obtained as a variant of the *fidelity* or *overlap test* (cf. Figure 1a). Such a circuit is shown in Figure 1b and implements the following transformation:

$$\begin{aligned}\gamma(x, x') &= \text{Tr} [U^\dagger(x')U(x)\rho_0U^\dagger(x)U(x')(\mathbb{I} \otimes |0\rangle\langle 0|)] \quad (7) \\ &= \text{Tr} [U(x)\rho_0U^\dagger(x)U(x')(\mathbb{I} \otimes |0\rangle\langle 0|)U^\dagger(x')] \\ &= \text{Tr} [\rho_x \bar{\rho}_{x'}]\end{aligned}$$

Here, $\rho_0 = |0\rangle\langle 0|$ is the initial state of the computation, ρ_x is obtained by applying the invariant feature embedding U (representing angle embedding followed by U_{inv}) over the datapoint x to ρ_0 , and $\bar{\rho}_{x'}$ is obtained by applying the invariant feature embedding U (representing angle embedding followed by U_{inv}) over the datapoint x' over an initial state $\mathbb{I} \otimes |0\rangle\langle 0|$ (not ρ_0). Notably, this function is not datapoint-wise permutation invariant, and it is not a kernel, as we are practically encoding the two data points using different embeddings.

A kernel can be defined via the mapping:

$$\kappa(\mathbf{x}, \mathbf{x}') = \frac{\gamma(\mathbf{x}, \mathbf{x}') + \gamma(\mathbf{x}', \mathbf{x})}{2} \quad (10)$$

The kernel in Equation (10) is a proper Mercer kernel, symmetric in its argument by construction and whose positive semi-definiteness holds by:

$$\gamma(x, x) = \text{Tr}[U^\dagger(x)U(x)\rho_0U^\dagger(x)U(x)(\mathbb{I} \otimes |0\rangle\langle 0|)] = \text{Tr}[\rho_0(\mathbb{I} \otimes |0\rangle\langle 0|)] = 1 \geq 0$$

The projection onto the subset of qubits associated with non-symmetric features is a key ingredient in obtaining the feature-wise permutation invariance. Lorentz invariance is obtained with Weyl's theorem [11].

Application to Vector Boson Scattering

An eventual goal of this study is to demonstrate the applicability of this symmetry-aware embedding quantum kernel technique in a realistic LHC data analysis use case. In particular, it will be applied to classify proton-proton collision events recorded by the CMS experiment in order to identify rare vector boson scattering (VBS) events, reconstructed in an all-hadronic final state, from an overwhelming background of multijet events induced by quantum chromodynamics.

Besides its importance for understanding the electroweak symmetry breaking, the VBS process is chosen as a use case because its experimental signature exhibits well-defined characteristic symmetries, making it suitable to demonstrate this symmetry-aware approach: (1) permutation invariance among the two massive vector bosons reconstructed as central large-radius jets, (2) permutation invariance among the two small-radius forward jets in the opposite ends of the detector (a characteristic feature of the VBS process that distinguishes it from other diboson production processes), and (3) the Lorentz invariance of the energy-momentum four-vectors of these four jets.

Dataset Monte Carlo simulation of signal (VBS) and background (QCD) processes, including full simulation of detector response, is performed to produce samples used in training and testing the algorithm. The eventual goal is to apply the trained model in the analysis of collision data recorded by CMS. The features consist of permutation invariant and variant ones.

Experimental Setup We first create a hardware-efficient invariant quantum kernel of form Eq. (7) and compare it with a kernel that uses the same circuit but with the global all-zero measurement $\bigotimes_{i=1}^n |0\rangle\langle 0|$ breaking the feature-wise invariance. This does not change anything within the original circuit operations, only extends the observable to the rest of the qubits and therefore provides a natural comparison between an invariant and a non-invariant model. We perform hyperparameter optimization with respect to regularization λ and the bandwidth ω . The quantum kernels are ideally simulated with the eventual aim of being inspected for exponential concentration, geometric difference, and target alignment in the number of qubits.

Analysis We aim to evaluate our approach in terms of the area-under-curve (AUC) of the receiving operator characteristic (ROC), true positive rate working points, and concentration of the off-diagonal Gram matrix values. Relative differences between AUCs of the permutation invariant and variant models are displayed in Table 1 with respect to different dataset sizes.

Table 1: Preliminary Results | **Dataset size** | ** $|D_{\text{train}}| = 500$ ** | ** $|D_{\text{train}}| = 750$ ** | ** $|D_{\text{train}}| = 1000$ ** |

8. For each column, $|\mathcal{D}_{\text{train}}| = |\mathcal{D}_{\text{test}}|$ with balance of signal and background samples in $\mathcal{D}_{\text{train}} \cup \mathcal{D}_{\text{test}}$.

References

1. **Tahmasebi, B., and Jegelka, S., 2023** - *The exact sample complexity gain from invariances for kernel regression*. Advances in Neural Information Processing Systems, 36.
2. **Sokolic, J., et al., 2017** - *Generalization error of invariant classifiers*. In *Artificial Intelligence and Statistics* (pp. 1094–1103). PMLR.
3. **Krizhevsky, A., et al., 2017** - *ImageNet classification with deep convolutional neural networks*. Communications of the ACM, 60(6), 84–90. AcM New York, NY, USA.
4. **Satorras, V. G., et al., 2021** - *E(n) equivariant graph neural networks*. In *International Conference on Machine Learning* (pp. 9323–9332). PMLR.
5. **Molteni, R., et al., 2024** - *Exponential quantum advantages in learning quantum observables from classical data*. arXiv preprint arXiv:2405.02027.
6. **Meyer, J. J., et al., 2023** - *Exploiting symmetry in variational quantum machine learning*. PRX Quantum, 4(1), 010328. APS.
7. **Li, Z., et al., 2024** - *Enforcing exact permutation and rotational symmetries in the application of quantum neural networks on point cloud datasets*. Physical Review Research, 6(4), 043028. APS.
8. **Gil-Fuster, E., et al., 2024** - *On the expressivity of embedding quantum kernels*. Machine Learning: Science and Technology, 5(2), 025003. IOP Publishing.
9. **Canatar, A., et al., 2021** - *Spectral bias and task-model alignment explain generalization in kernel regression and infinitely wide neural networks*. Nature Communications, 12(1), 2914. Nature Publishing Group UK London.
10. **Thanasilp, S., et al., 2024** - *Exponential concentration in quantum kernel methods*. Nature Communications, 15(1), 5200. Nature Publishing Group UK London.
11. **Weyl, H., 1946** - *The classical groups: their invariants and representations*. Princeton University Press.

Email Address of submitter

vaino.matinpoika.mehtola@cern.ch

Short summary

Symmetries play a fundamental role in high-energy physics, and integrating these symmetries into quantum machine learning (QML) models can enhance their performance. This work extends recent approaches in QML to incorporate symmetries such as permutation and Lorentz invariance into quantum kernels. Our method introduces feature-wise permutation invariance, a symmetry particularly relevant for high-energy physics applications, such as vector boson scattering (VBS) classification. We study the potential benefits of symmetry-aware quantum kernels on Monte Carlo simulated data for VBS identification against the quantum chromodynamics background, with preliminary results showing promise for a preliminary model comparison, reflected in higher area-under-curve (AUC) values for the permutation-invariant model.

Authors: INCUDINI, Massimiliano (Intel); MEHTOLA, Vaino (Helsinki Institute of Physics (FI))

Co-authors: Dr GROSSI, Michele (CERN); LAURILA, Santeri (CERN & Helsinki Institute of Physics (FI))

Presenter: MEHTOLA, Vaino (Helsinki Institute of Physics (FI))

Session Classification: Networking cocktail and Poster Session