PROCEEDINGS OF SCIENCE

Group Invariance in Quantum Fidelity Kernels for Vector Boson Scattering at the LHC

Väinö Mehtola,[a,b,c,*] **Santeri Laurila**[b,c] **and Michele Grossi**[c]

[a] *VTT Technical Research Centre,*
[b] *Helsinki Institute of Physics*
[c] *CERN*

E-mail: vaino.mehtola@vtt.fi

In this work, we study permutation-invariance and bandwidth-tuning in quantum fidelity kernels for classifying vector boson scattering (VBS) against QCD multijet background using proton-proton collisions simulated at $\sqrt{s} = 13.6$ TeV, with detector simulation corresponding to Run-2 CMS detector. We employ partially and fully symmetrized hardware-efficient kernel variants that embed S_n symmetry of the final-state jets. In noiseless simulation, some symmetry-enhanced kernels outperform a non-invariant HEA kernel, a classical Gaussian kernel baseline, and a product state quantum kernel in terms of the area under the curve (AUC) of the receiver operator characteristic curve. The bandwidth γ crucially affects the classification performance, and its optimal value is noticeably affected by the introduction of finite-shot estimation (10 000 shots) leading to the vanishing of the AUC gains over the classical baseline. Finally, accompanying runs on VTT's Q50 quantum computer exhibit increased noise challenges for the symmetrized entanglement pattern.

The European Physical Society Conference on High Energy Physics (EPS-HEP2025)
7-11 July 2025
Marseille, France

*Speaker

PoS(EPS-HEP2025)634

1. Introduction

Quantum machine learning (QML) has shown advantages in specific learning settings [1, 2, 3] and is being explored in high-energy physics (HEP) [4], where data often exhibit known symmetries. In this work, we study fidelity-based quantum kernels that exploit final-state permutation invariance for event classification, using the embedding quantum-kernel framework in which inputs are mapped to quantum states and similarities are computed from overlaps between states [5]. Using a hardware-efficient ansatz [6] as a baseline, we construct increasingly symmetrized variants via Pauli twirling [7] and tune the kernel bandwidth γ, a data-scaling hyperparameter shown to be important for quantum kernels [8]. These quantum models are compared against a classical Gaussian radial basis function (RBF) kernel and a reference quantum angle-embedding (Pauli-X) kernel without entanglement, following Ref. [9].

The application targets discrimination between two classes of proton-proton collision events—signal events involving vector boson scattering (VBS) and background events from quantum chromodynamics (QCD) multijet production—simulated at the CMS detector of the LHC. In the VBS events, the final state exhibits a natural pairwise permutation symmetry structure among the two jet pairs. We evaluate the models under ideal (fully noise-free) simulation, finite-shot (shot-noisy) simulation, and on VTT's Q50 device to reflect current implementation conditions. Performance is measured by the area under the receiver operating characteristic curve (ROC AUC), with additional analysis of kernel-value concentration as a function of qubit count.

2. Symmetrized quantum fidelity kernels

Parametrized quantum circuits (PQCs) implement data–dependent unitaries

$$U(\mathbf{x}) = \prod_j \exp\left[-i\, f_j(\mathbf{x};\boldsymbol{\theta})\, G_j\right] \tag{1}$$

where G_j are Hermitian generator matrices (typically tensor products of Pauli operators) and $f_j(\mathbf{x};\boldsymbol{\theta})$ real–valued functions of the inputs and trainable parameters. A PQC defines a quantum feature map via $\rho(\mathbf{x}) = U(\mathbf{x})\,|0\rangle\langle 0|\,U^\dagger(\mathbf{x})$. The associated fidelity kernel

$$\kappa_{\mathrm{Q}}(\mathbf{x},\mathbf{x}') = \mathrm{Tr}[\rho(\mathbf{x})\,\rho(\mathbf{x}')] \tag{2}$$

can be estimated on hardware with overlap tests and finite shots [10]. As the number of qubits grows, a key concern is exponential concentration, where kernel values collapse toward an input-independent mean, driving up the number of shots required for useful precision [11]. Hyperparameters that rescale inputs (e.g. a bandwidth γ via $\rho(\mathbf{x}) \mapsto \rho(\gamma\mathbf{x})$) can strongly influence both concentration behaviour and alignment with the task structure [8].

HEP data often exhibit symmetries that provide useful inductive biases, making it natural to design models that are fully or partially invariant under them. Given a symmetry group S acting on inputs, invariance is encoded as $f(s\cdot\mathbf{x}) = f(\mathbf{x})$ for all $s \in S$. For PQCs, finite-group symmetries can be embedded by constraining generators and observables using Pauli twirling. Relative to an initial embedding $U_{\mathrm{init}}(\mathbf{x})$ and its induced unitaries $\{U_s\}$ of a finite group S, the twirl

$$\mathcal{T}[X] = \frac{1}{|S|}\sum_{s\in S} U_s X U_s^\dagger \tag{3}$$

PoS(EPS-HEP2025)634

yields a generator set $\{G : [G, U_s] = 0 \ \forall s\}$ suitable for building symmetry-respecting circuits and measurements [7]. For quantum fidelity kernels, however, achieving full invariance in both arguments is nontrivial and, at least in naive constructions, requires an initial embedding that effectively encodes invariant aggregate feature combinations.

3. Experimental Setup and Method

Permutation symmetry is incorporated into a hardware-efficient feature map by Pauli twirling over the finite permutation group acting on a pair of features from two jets pairwise, yielding partially or fully symmetrized feature maps. To apply twirling, we first fix an initial feature-per-qubit embedding

$$U_{\text{init}}(\mathbf{x}) = \bigotimes_{i=1}^{n_{\text{qubits}}} R_x(x_i). \tag{4}$$

For this choice, the induced unitary representation of a permutation between two features is simply a SWAP gate, as illustrated in Fig. 1.

Applying the twirl in Eq. 3 to the generator pool replaces the original generators with symmetry-respecting ones and yields the symmetrized circuits shown in Fig. 2 for a four-qubit example with two symmetry-respecting and two non-symmetry-respecting qubits. In the remainder, we compare: (a) a non-invariant hardware-efficient ansatz (HEA); (b) HEA with symmetrized single-qubit rotations (HEA-R); (c) HEA with symmetrized entanglers between permutable and non-permutable qubits (HEA-E); (d) HEA with both symmetrized rotations and entanglers (HEA-ER); and, as reference models: (e) an angle-embedding kernel without entanglement; and (f) a classical RBF kernel.

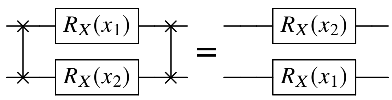

Figure 1: Induced unitary representation of a feature permutation for the initial embedding in Eq. 4: swapping two input features is implemented by SWAP gates. This is used only in constructing the invariant generator set via Eq. 3.

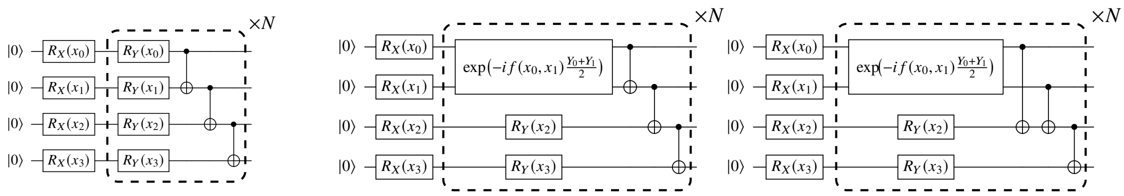

Figure 2: Circuit illustration of (left) the hardware-efficient ansatz (HEA), (middle) a symmetrized-rotation (HEA-R), and (right) an entanglement-symmetrized variant (HEA-E) with one symmetrized pair of qubits and two regular ones, all for $n_{\text{qubits}} = 4$.

The dataset is preselected to enhance the fully hadronic VBS topology while suppressing generic QCD multijet background. Events must pass at least one of the standard CMS trigger paths requiring either a high-p_T jet or large jet p_T sum in the event to target high jet-activity final states. A minimum of two small-R jets ($n_{\text{jet}} \geq 2$) and two large-R "fat" jets ($n_{\text{FatJet}} \geq 2$) is required. Fat jets must satisfy $p_T > 200\text{GeV}$ and $|\eta| < 2.4$, and VBS candidates are chosen as the pair of fat jets obeying these cuts and, for the second candidate, $\Delta\phi > 2.7$ with respect to the first; events are retained only if exactly two fat jets pass these criteria. From these jets we compute m_{VV}, the individual fat-jet masses, and their pseudorapidity separation. Forward tag jets are selected from the remaining small-R jets with $p_T > 30\text{GeV}$ and $|\eta| < 4.7$, required to be non-overlapping with the fat jets; the tag-jet pair must satisfy $m_{jj} > 500\text{GeV}$ and $\Delta\eta > 2.4$, with the leading tag jet constrained to $p_T > 50\text{GeV}$. These requirements select the characteristic two-fat-jet plus two-tag-jet topology of fully hadronic VBS and define the sample used for the training and testing splits.

For each kernel, we randomly select 2000 events from a dataset of $O(10^4)$ events to form a 2000×2000 Gram matrix, and train a support vector machine (SVM) with a convex objective and L_2 regularization. Performance is reported as the area under the receiver operating characteristic curve (ROC AUC) on a held-out test set of 2000 events. A scalar bandwidth γ rescales inputs in the quantum feature maps (analogous to the RBF case) and is scanned together with the SVM regularization parameter. Experiments cover noiseless simulation, finite-shot simulation (10 000 shots), and runs on VTT's Q50 device (10 000 shots), using feature sets of size 4–14 consisting of symmetric and non-symmetric features, with their total number matched to the qubit count.

4. Results

Figure 3 highlights the importance of optimizing the bandwidth γ to obtain competitive mean AUCs. A naive choice such as $\gamma = 1$ leads to substantially degraded performance, and this degradation becomes more severe as n_{qubits} increases (not shown), since the sharp AUC drop at large γ shifts to lower values of γ with growing system size. Apart from the early rise of the RX kernel around $\gamma \approx 10^{-4}$, no strong model–dependent differences appear in the overall γ–dependence.

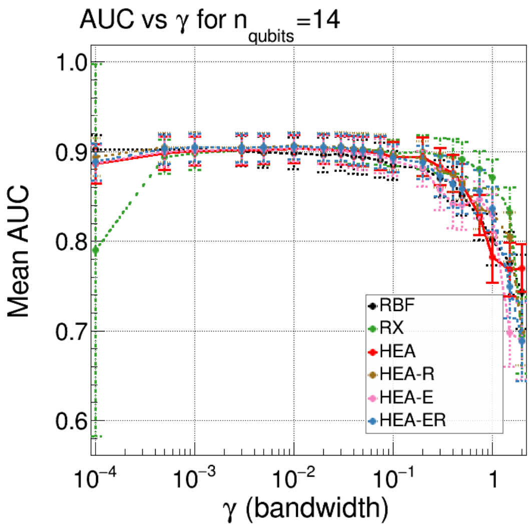

Figure 3: Mean ROC AUC versus kernel bandwidth γ for $n_{\text{qubits}} = 14$ in the noiseless setting. The same qualitative behaviour is observed for smaller qubit counts (not shown).

In fully noiseless conditions, the HEA-R and HEA-ER kernels consistently achieve higher mean AUCs than the baseline HEA, the RX kernel, the classical RBF kernel, and the HEA-E model with only symmetrized entanglement, across all qubit counts considered (cf. Fig. 4). These gains are statistically significant, as illustrated in Table 1, which reports paired t–tests comparing HEA-R against both HEA and RBF at $n_{\text{qubits}} = 10, 12$, and 14. However, once finite sampling is introduced (10 000 shots), the AUC advantage of the symmetrized kernels largely disappears, and the classical RBF kernel attains the highest mean AUC for $n_{\text{qubits}} \geq 8$. This behaviour is reflected in the optimal bandwidths reported in Table 2: in almost all cases, the presence of shot noise shifts the best–performing γ to larger values, effectively compensating for the increased statistical uncertainty.

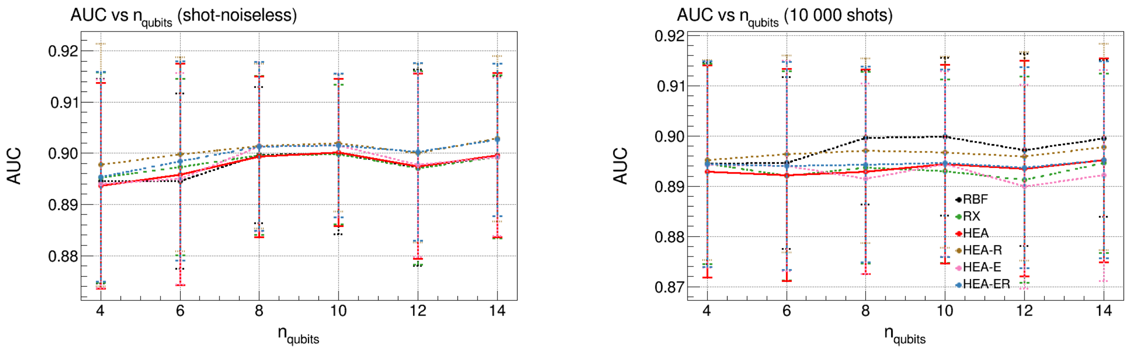

Figure 4: Mean ROC AUC vs. feature/qubit count. Left: ideal (shot–noiseless). Right: ideal simulator with 10k shots. Error bars show seed variation.

n_{qubits}	Comparison	t	p	Sig.
10	HEA vs HEA-R	2.4835	0.022	*
	RBF vs HEA-R	2.0019	0.047	*
12	HEA vs HEA-R	2.4835	0.022	*
	RBF vs HEA-R	2.8537	5.2×10^{-4}	***
14	HEA vs HEA-R	5.0506	7.1×10^{-6}	***
	RBF vs HEA-R	1.9867	0.049	*

Table 1: Paired t–tests for mean AUC differences (seedwise pairing). Significance codes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The hardware experiments on VTT's Q50 device [12] further illustrate the practical challenges. For a fixed subset of 30 events at $n_{\text{qubits}} = 10$, the 30×30 Gram matrices measured on Q50 show non-negligible deviations from ideal simulation, even after dynamical decoupling and Bayesian readout-mitigation [13] are applied. As summarized in Table 3, mitigation roughly halves the mean squared error for both HEA and HEA-ER, but the symmetrized entangling pattern (HEA-ER) remains somewhat more sensitive, resulting in larger residual discrepancies compared to the non-invariant HEA. Overall, these results suggest that while symmetry-enhanced kernels can improve performance in idealized settings, their practical advantage on current noisy hardware may be strongly constrained by shot noise and device-level imperfections.

n_{qubits}	Model	γ (noiseless)	γ (10 000 shots)	Rel. diff.
12	HEA	0.030	0.100	2.33
	HEA-E	0.0005	0.100	199.0
	HEA-R	0.010	0.100	9.00
14	HEA	0.030	0.100	2.33
	HEA-E	0.010	0.075	6.50
	HEA-R	0.0005	0.075	149.0

Table 2: Best-performing bandwidth value γ (per mean AUC) for the HEA, HEA-E, and HEA-R models at $n_{\text{qubits}} = 12$ and 14 in the shot-noiseless and shot-noisy (10 000 shots) settings. Shot noise increases the optimal value of γ in all cases.

Model	Condition	MSE	Median SE
HEA	Unmitigated	0.086	0.086
	Mitigated	0.041	0.038
HEA-ER	Unmitigated	0.106	0.101
	Mitigated	0.058	0.048

Table 3: Squared errors between simulated and Q50–measured 30×30 Gram matrices (10 000 shots, dynamical decoupling & Bayesian readout mitigation).

5. Conclusion

We studied symmetry–enhanced and bandwidth-tuned quantum fidelity kernels for fully hadronic VBS vs. QCD multijet classification using simulated proton-proton collisions. Starting from a hardware–efficient ansatz, we constructed progressively symmetrized variants via Pauli twirling over the pairwise jet–permutation symmetry and optimized γ with respect to AUC, benchmarking against a classical Gaussian RBF kernel and a product–state RX kernel.

In ideal, shot–noiseless simulations, the rotation–symmetrized kernels (HEA–R, HEA–ER) yield statistically significant ROC AUC gains over both the non–invariant HEA and the classical baseline, provided that γ is tuned to avoid concentration. At 10 000 shots these gains disappear and the RBF kernel becomes competitive or superior, with the optimal γ shifting to larger values. Q50 hardware runs show noticeable discrepancies between simulated and measured Gram matrices, even with dynamical decoupling and readout mitigation, and the symmetrized entanglement pattern appears more sensitive to device noise. Overall, embedding permutation structure into quantum kernels can improve inductive bias in ideal settings, but near–term hardware noise and sampling constraints may limit any practical advantage. This motivates further study of more noise–resilient symmetry–aware architectures and of the interplay between shot noise and bandwidth tuning.

References

[1] Riccardo Molteni, Casper Gyurik, and Vedran Dunjko. "Exponential quantum advantages in learning quantum observables from classical data". In: *arXiv preprint arXiv:2405.02027v2* (2024). arXiv: 2405.02027v2 [quant-ph].

[2] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. "A rigorous and robust quantum speed-up in supervised machine learning". In: *Nature Physics* 17.9 (Sept. 2021), pp. 1013–1017.

[3] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. "Quantum machine learning". In: *Nature* 549.7671 (Sept. 2017), pp. 195–202.

[4] Alberto Di Meglio et al. "Quantum Computing for High-Energy Physics: State of the Art and Challenges". In: *PRX Quantum* 5 (3 Aug. 2024), p. 037001.

[5] Vojtěch Havlíček et al. "Supervised learning with quantum-enhanced feature spaces". In: *Nature* 567.7747 (Mar. 2019), pp. 209–212.

[6] Abhinav Kandala et al. "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets". In: *Nature* 549.7671 (Sept. 2017), pp. 242–246.

[7] Johannes Jakob Meyer et al. "Exploiting Symmetry in Variational Quantum Machine Learning". In: *PRX Quantum* 4 (1 Mar. 2023), p. 010328.

[8] Ruslan Shaydulin and Stefan M. Wild. "Importance of kernel bandwidth in quantum machine learning". In: *Phys. Rev. A* 106 (4 Oct. 2022), p. 042407.

[9] Joseph Bowles, Shahnawaz Ahmed, and Maria Schuld. *Better than classical? The subtle art of benchmarking quantum machine learning models*. 2024. arXiv: 2403.07059 [quant-ph].

[10] Maria Schuld and Nathan Killoran. "Quantum Machine Learning in Feature Hilbert Spaces". In: *Phys. Rev. Lett.* 122 (4 Feb. 2019), p. 040504.

[11] Supanut Thanasilp, Samson Wang, M. Cerezo, and Zoë Holmes. "Exponential concentration in quantum kernel methods". In: *Nature Communications* 15.1 (June 2024), p. 5200.

[12] VTT Technical Research Centre of Finland. *VTT Quantum Computer Documentation*. Accessed: 30 November 2025. 2025. URL: https://qx.vtt.fi/docs/.

[13] F. Cosco, F. Plastina, and N. Lo Gullo. *Bayesian mitigation of measurement errors in multi-qubit experiments*. 2024. arXiv: 2408.00869 [quant-ph].

PoS(EPS-HEP2025)634