

From Qubits to Couplings: A Hybrid Quantum Machine Learning Framework for LHC Physics

Marwan Ait Haddou*

*LPMC, Faculty of Sciences Ben M'sick,
Hassan II University of Casablanca, Morocco.*

Mohamed Belfkir[†]

Department of Physics, United Arab Emirates University, Al-Ain, UAE

Salah Eddine El Harrauss[‡]

Department of Physics, United Arab Emirates University, Al-Ain, UAE

(Dated: November 20, 2025)

Abstract

In this paper, we propose a new Hybrid Quantum Machine Learning (HyQML) framework to improve the sensitivity of double Higgs boson searches in the $HH \rightarrow b\bar{b}\gamma\gamma$ final state at $\sqrt{s} = 13.6$ TeV. The proposed model combines parameterized quantum circuits with a classical neural network meta-model, enabling event-level features to be embedded in a quantum feature space while maintaining the optimization stability of classical learning. The hybrid model outperforms both a state-of-the-art XGBoost model and a purely quantum implementation by a factor of two, achieving an expected 95% CL upper limit on the non-resonant double Higgs boson production cross-section of $1.9 \times \sigma_{\text{SM}}$ and $2.1 \times \sigma_{\text{SM}}$ under background normalization uncertainties of 10% and 50%, respectively. In addition, expected constraints on the Higgs boson self-coupling κ_λ and quartic vector-boson–Higgs coupling κ_{2V} are found to be improved compared to the classical and purely quantum models.

I. INTRODUCTION

The discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments at the Large Hadron Collider (LHC) [1, 2] marked a cornerstone in the validation of the Standard Model (SM) of particle physics. Over the past decade, an extensive program of precision measurements has established the properties of the Higgs boson including its mass, spin and couplings to fermions and gauge bosons to be consistent with SM predictions within current experimental uncertainties [3, 4]. However, the shape of the Higgs potential and in particular the strength of the trilinear Higgs self-coupling λ_3 remains largely unconstrained. This coupling governs the dynamics of electroweak symmetry breaking and the stability of the Higgs vacuum. Deviations from its SM expectation value, commonly expressed via the modifier $\kappa_\lambda = \lambda_3^{\text{BSM}}/\lambda_3^{\text{SM}}$, would provide direct evidence of an extended scalar sector or new physics beyond the SM (BSM) [5–7].

The most sensitive process for probing κ_λ at the LHC is Higgs boson pair production ($pp \rightarrow HH$), which proceeds mainly through gluon–gluon fusion (ggF) with a subdominant contribution from vector-boson fusion (VBF). Among the possible decay channels of the

* marwan.aithaddou@edu.uca.ac.ma

† Contact author: m_belfkir@uaeu.ac.ae

‡ selharrauss@uaeu.ac.ae

Higgs pair, $HH \rightarrow b\bar{b}\gamma\gamma$ offers an optimal balance between a clean experimental signature, excellent di-photon mass resolution and manageable backgrounds, despite its small branching fraction of approximately 0.26% for $m_H = 125$ GeV [8]. Recent ATLAS and CMS searches have set expected upper limits on the non-resonant double Higgs cross-section at approximately three to five times the SM prediction, leaving significant room for methodological improvements to enhance sensitivity [7, 9].

Machine learning (ML) techniques have become indispensable in collider physics enabling efficient discrimination between signal and background by exploiting complex high-dimensional correlations among observables [10, 11]. Classical ML approaches such as tree-based algorithms, convolutional neural networks or graph neural networks have delivered substantial gains in searches for new physics [12]. Nevertheless, these methods remain bounded by classical computational architectures and feature representations, motivating the exploration of novel paradigms that may provide a representational advantage.

In this context, *Quantum Machine Learning* (QML) has emerged as a promising new direction leveraging quantum computation to process information in exponentially large Hilbert spaces [13]. QML operates on quantum bits (qubits) instead of classical bits allowing the encoding of data into quantum states and the execution of transformations that can naturally capture high-order correlations in complex datasets. For example, the Quantum Support Vector Machine (QSVM) [14, 15] which employs quantum kernel estimation to achieve accurate classifications for certain benchmark datasets. With rapid advances in superconducting and photonic quantum hardware the feasibility of deploying QML algorithms on noisy intermediate-scale quantum (NISQ) devices has become increasingly realistic [13, 16].

Proof-of-principle studies in high-energy physics have already demonstrated the applicability of QML to detector simulation, particle reconstruction, and signal-background separation tasks [17, 18]. Recent implementations of the QSVM-Kernel algorithm on both quantum simulators and superconducting hardware have achieved performance comparable to classical classifiers on small-scale datasets despite being limited by hardware noise, circuit depth and scalability [17]. These constraints highlight the current gap between the theoretical potential of fully quantum algorithms and their practical deployment on near-term devices. To bridge this gap, this study introduces a *Hybrid Quantum Machine Learning* (HyQML) model combining parameterized quantum circuits with classical neural networks that ex-

exploits the representational power of quantum state spaces while retaining the optimization stability and scalability of classical learning frameworks. Such hybrid architecture provides a promising pathway toward realizing quantum advantage in realistic particle physics analyses.

In this work, we demonstrate the potential of such HyQML to improve the sensitivity to the Higgs boson self-coupling in the $HH \rightarrow b\bar{b}\gamma\gamma$ channel at $\sqrt{s} = 13.6$ TeV. Using the same simulated datasets and event selections as in our previous study [12], we develop a hybrid quantum–classical architecture that maps event-level features into a quantum feature space through parameterized quantum circuits. The objective is to assess whether such hybrid models can yield improved signal significance and tighter constraints on κ_λ compared to established classical classifiers such as the XGBoost described in our previous work [12], while simultaneously probing the prospects of quantum-enhanced learning for collider physics applications.

The rest of this paper is organized as follows: Section II provides an overview of Higgs boson pair production at the LHC. The Monte Carlo (MC) simulation setup and generated samples are described in Section III. Section IV details the physics object definitions and event selection criteria used in the analysis. The HyQML framework including model architecture and training procedure is presented in Section V. The resulting classification performance, statistical interpretation and constraints on the Higgs self-coupling are discussed in Section VI, and the main conclusions are summarized in Section VII.

II. HIGGS BOSON PAIR PRODUCTION

At the LHC, the Higgs boson pair production process proceeds mainly through the ggF production mode, mediated by heavy-quark loops, predominantly the top quark. The ggF amplitude receives contributions from two destructively interfering diagrams: the *triangle* diagram, which explicitly depends on the trilinear coupling, and the *box* diagram, which does not. The destructive interference between these two amplitudes makes the total production rate highly sensitive to deviations in the coupling modifier κ_λ . A smaller but complementary contribution arises from VBF production, which also depends on κ_λ and, in more general parameterizations, on the Higgs–vector boson couplings κ_V and κ_{2V} . Figure 1 shows the leading-order Feynman diagrams contributing to Higgs boson pair production in both ggF and VBF channels.

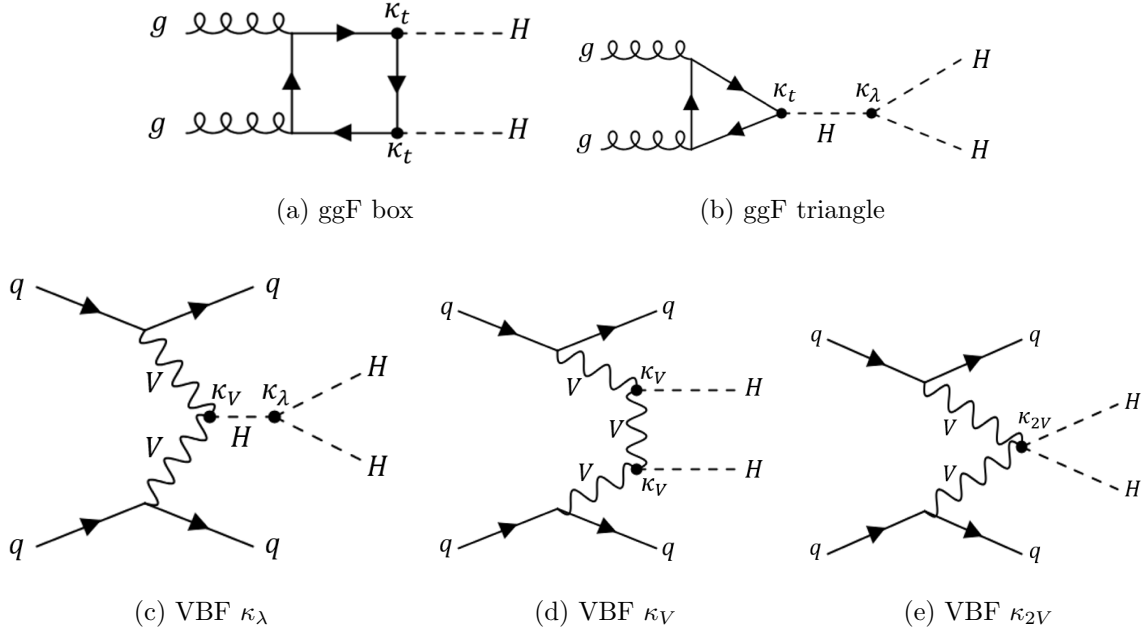


FIG. 1: The leading-order Feynman diagrams for (a–b) gluon–gluon fusion and (c–e) vector boson fusion Higgs boson pair production.

In the SM, the ggF mode dominates with a next-to-next-to-leading order (NNLO) cross-section of $\sigma_{\text{ggF}} = 0.3413$ pb at $\sqrt{s} = 13.6$ TeV and $m_H = 125$ GeV [8], while the VBF contribution is smaller by nearly two orders of magnitude, $\sigma_{\text{VBF}} = 0.001874$ pb [19]. Theoretical uncertainties arise primarily from parton distribution functions (PDFs), α_s , and QCD scale variations, typically at the 2–5% level for the VBF process and up to 20% for ggF [8, 19, 20].

Although small in the SM, the double Higgs boson cross-section can be significantly enhanced in BSM scenarios where new scalar degrees of freedom modify the Higgs potential or introduce resonant production modes. Examples include extended scalar sectors such as two-Higgs-doublet models (2HDM), singlet-extended models (2HDM+S), and supersymmetric frameworks [21–23]. In this work, however, we restrict our study to the non-resonant production regime, where variations in the couplings affect only the overall production rate and event kinematics [24].

The dependence of the ggF and VBF cross-sections on κ_λ can be expressed through quadratic parameterizations [12, 19]:

$$\sigma_{\text{ggF}}(\kappa_\lambda) = 75.76 - 53.29 \kappa_\lambda + 11.61 \kappa_\lambda^2 \text{ [fb]}, \quad (1)$$

$$\sigma_{\text{VBF}}(\kappa_\lambda) = 0.0032 - 0.0029 \kappa_\lambda + 0.00093 \kappa_\lambda^2 \text{ [fb]}. \quad (2)$$

The VBF cross-section is generally parameterized including also κ_{2V} and κ_V [19]. As can be seen in Figure 2, when κ_λ deviates from the SM value, the cross-sections change accordingly demonstrating the sensitivity of these production modes to the Higgs self-coupling and making them important probes for BSM physics.

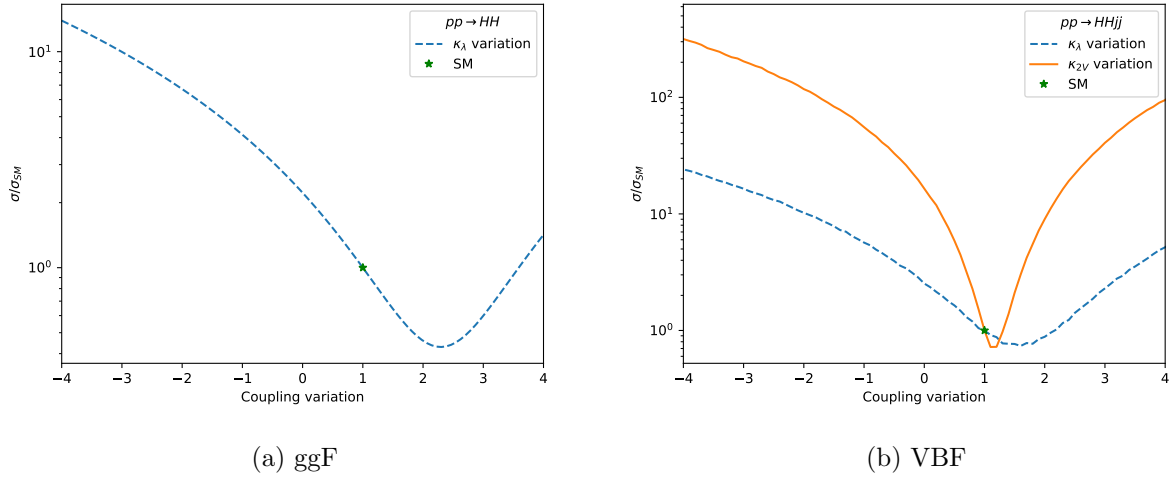


FIG. 2: Variation of (a) ggF and (b) VBF double Higgs boson cross-section as a function of κ_λ and κ_{2V} . If one parameter varies, the remaining parameters are fixed to 1.

III. MONTE CARLO SIMULATION

This study uses simulated proton–proton collision events at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV to explore the potential of HyQML in improving the sensitivity to non-resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$ decay channel. The analysis uses an integrated luminosity of 308 fb^{-1} , corresponding to the partial Run-3 + Run-2 dataset collected by the ATLAS detector between 2018 and 2024. The targeted final state consists of two well-identified photons and two b -quark tagged jets, originating from $H \rightarrow \gamma\gamma$ and $H \rightarrow b\bar{b}$ decays. This channel remains one of the most sensitive probes of the Higgs boson self-coupling among all double Higgs boson decay modes [12].

The simulated events include both signal and dominant SM backgrounds. The signal processes comprise ggF and VBF Higgs boson pair production, generated with the trilinear

Higgs coupling fixed to its SM value. The ggF process is simulated at next-to-leading order (NLO) using **Powheg-Box v2** [25], including finite top-quark mass effects, while VBF events are generated at leading order (LO) using **MadGraph5_aMC@NLO v3.3.0** [26]. Both signal samples are normalized to the most recent theoretical cross-section predictions evaluated at NNLO for ggF, and at N3LO QCD plus NLO electroweak accuracy for VBF at Higgs mass $m_H = 125$ GeV [19, 24].

The background simulation follows the same methodology. Resonant backgrounds arise from single Higgs boson production where $H \rightarrow \gamma\gamma$, including also ggF, VBF, associated production with a Z boson ($Z(\rightarrow b\bar{b}) H(\rightarrow \gamma\gamma)$), and associated production with top quarks ($t\bar{t}H$), are simulated using **Powheg-Box v2**. Non-resonant $\gamma\gamma + \text{jets}$ background originates from QCD-induced di-photon events accompanied by jets, which can mimic the $b\bar{b}\gamma\gamma$ final state without an intermediate Higgs resonance, is simulated with **MadGraph5_aMC@NLO**, with matrix elements including up to two additional partons to model the jet multiplicity spectrum accurately. Table I summarizes the simulated processes with respective number of events and cross-sections.

Process	Event Generator	Number of events	Cross-section [pb]
ggF HH ($\kappa_\lambda = 1$)	Powheg	50k	0.3413 [19]
VBF HH ($\kappa_\lambda = 1$)	MadGraph	100k	0.001874 [19]
ggF Higgs	Powheg	500k	52.17 [8]
VBF Higgs	Powheg	550k	4.075 [8]
$qq \rightarrow ZH$	Powheg	50k	1.834×10^{-3} [8]
$gg \rightarrow ZH$	Powheg	100k	3.087×10^{-4} [8]
$t\bar{t}$ Higgs	Powheg	100k	5.688×10^{-1} [8]
$\gamma\gamma + \text{jets}$	MadGraph	2.5M	48.1

TABLE I: Summary of generated signal and background events.

All generated samples are processed through **Pythia 8.186** [27] for parton showering, hadronization, and underlying-event modeling. The detector response is emulated using the **Delphes** [28] fast simulation framework configured with an ATLAS detector configuration card. The detector card has been tuned to reproduce the latest ATLAS performance calibrations.

IV. OBJECT DEFINITIONS AND EVENT SELECTIONS

A. Object definitions

The event topology for the $HH \rightarrow b\bar{b}\gamma\gamma$ signal consists of two isolated photons and two b -identified jets in the final state. To ensure consistent and accurate object reconstruction, standard Run-3 ATLAS detector definitions are adopted and emulated in the **Delphes** fast simulation framework.

Photons are reconstructed from energy deposits in the electromagnetic calorimeter using a tower-based clustering algorithm. Each photon candidate is required to have a transverse momentum of $p_T > 20$ GeV and lie within the pseudorapidity range $|\eta| < 2.37$, excluding the calorimeter barrel–endcap transition region $1.37 < |\eta| < 1.52$, where the detector response deteriorates due to reduced granularity and inactive material [29]. To avoid double counting of closely spaced clusters and ensure high isolation efficiency, a self-overlap removal criterion is applied: if two reconstructed photon candidates are separated by less than $\Delta R < 0.01$, only the leading (highest- p_T) photon is retained. Photon candidates must also satisfy the *Tight* identification working point (WP) emulated with the Run-3 photon identification efficiency at 13.6 TeV [30].

Jets are reconstructed from calorimeter energy deposits using the anti- k_t algorithm [31] with a radius parameter $R = 0.4$, as implemented in the **FastJet** package [32]. The inputs to the clustering algorithm are calorimeter towers representing the energy flow in the detector. Selected jets must satisfy $p_T > 25$ GeV and a rapidity $|y| < 4.5$, ensuring full containment within the ATLAS calorimeter acceptance. Jets originating from b -quarks are identified using a parameterization of the ATLAS b -tagging WP at the 85% efficiency [33]. This operating point corresponds to mis-tag rates of approximately 0.17 for c -jets and 0.01 for light-flavor jets, allowing realistic modeling of background contamination from mis-identified jets. Only b -tagged jets within the inner-detector tracking region ($|\eta| < 2.5$) are considered in the analysis.

Leptons (electrons and muons) are reconstructed from the particle-flow track collection. Muon candidates are required to have $p_T > 10$ GeV and $|\eta| < 2.7$, while electrons must satisfy $p_T > 10$ GeV and $|\eta| < 2.47$, excluding candidates in the same calorimeter transition region as photons. These criteria ensure that leptons are well-measured and isolated from

hadronic activity, enabling the rejection of events with leptonic signatures from top-quark or electroweak background processes.

Standard object-cleaning criteria are imposed to remove overlaps between reconstructed objects and to ensure isolation of photon and jet candidates. All reconstructed objects are required to be mutually distinct and to pass the default overlap-removal procedure implemented in the `Delphes` configuration. The resulting set of calibrated and isolated photons, jets, b -jets, and leptons forms the input basis for the subsequent event selection and quantum machine-learning classification described later.

B. Event selections

Events are required to pass a di-photon trigger optimized for the selection of final states containing two energetic photons. The trigger demands that the leading and sub-leading photons have transverse energies exceeding 35 GeV and 25 GeV, respectively. The corresponding trigger efficiency is applied to simulated events using the publicly available measurements derived from ATLAS Run-3 data at $\sqrt{s} = 13.6$ TeV [34].

Following the selection strategy adopted in previous ATLAS analyses of the $HH \rightarrow b\bar{b}\gamma\gamma$ channel [9, 35, 36], the two leading photons in each event are required to have an invariant mass within the range $105 < m_{\gamma\gamma} < 160$ GeV. To maintain a uniform efficiency across the diphoton mass spectrum, dynamic transverse momentum thresholds are imposed, requiring $p_T^{\gamma_1} > 0.35 m_{\gamma\gamma}$ for the leading photon and $p_T^{\gamma_2} > 0.25 m_{\gamma\gamma}$ for the sub-leading photon. These conditions prevent biases in the background distribution near the Higgs boson mass peak [37]. The selected photon pair is used to reconstruct the $H \rightarrow \gamma\gamma$ candidate.

The reconstruction of the $H \rightarrow b\bar{b}$ candidate is performed using the two leading b -tagged jets in the event, both satisfying the selection criteria described in Section IV. Events are required to contain at least two b -tagged jets. To suppress background contributions from top-quark processes, particularly $t\bar{t}H$ production, events containing identified electrons or muons are vetoed. In addition, to reduce hadronic contamination from $t\bar{t}H$ and multi-jet processes, the total number of reconstructed central jets in each event is restricted to six or fewer.

Beyond the baseline selection, the topology of the event is further characterized by identifying potential VBF jets. When present, the two highest- p_T jets not associated with the

$H \rightarrow b\bar{b}$ reconstruction are labeled as VBF-tagged jets. These jets typically correspond to forward-scattered quarks in the VBF process and provide valuable information on VBF event topology. However, the presence of VBF-tagged jets is not mandatory, and events lacking such jets are retained in the inclusive selection. A dedicated VBF event category, which could enhance sensitivity to anomalous couplings such as κ_{2V} , is not defined in the present study but will be explored in future analyses.

V. HYBRID-QUANTUM MACHINE LEARNING MODEL

This section describes the HyQML model used to enhance the event categorization of the analysis. The proposed HyQML approach combines a classical neural network encoder with a parameterized quantum circuit (PQC), allowing the system to learn both non-linear and quantum-correlated feature representations. The objective is to improve the separation between signal and background events beyond what is achievable with classical XGBoost model or purely classical machine learning selections.

A. Data pre-processing

Before training the hybrid model, a dedicated data pre-processing pipeline is implemented to standardize the input features and minimize detector-induced asymmetries. In particular, events passing the selection described in Section IV are geometrically rotated in the transverse plane such that the leading photon is aligned with the beam axis. This transformation exploits the cylindrical symmetry of the ATLAS detector in the azimuthal angle ϕ , effectively removing arbitrary rotational variations and allowing the model to focus on the intrinsic event topology [38]. All other reconstructed objects, including the sub-leading photon and the two b -jets, are rotated accordingly to preserve their relative spatial configuration. Empirically, this geometric normalization improves the stability of the training process and enhances the model’s ability to identify kinematic correlations relevant to Higgs boson pair production.

All the input variables listed in Table II are subsequently standardized so that their distributions are centered and scaled to unit variance. This normalization step ensures numerical stability and accelerates convergence during the optimization of both classical and quantum

parameters by bringing all observables to comparable magnitudes [39]. The standardized dataset is then randomly partitioned into two statistically independent subsets: 75% of the events are used for model training, while the remaining 25% are reserved exclusively for performance evaluation and statistical treatment. The model performance is therefore assessed only on unseen events to ensure an unbiased estimation of generalization capability.

Objects	variables
Photons	p_T, η, ϕ
b -jets	p_T, η, ϕ
$\gamma\gamma$ system	p_T, η, ϕ, m
$b\bar{b}$ system	p_T, η, ϕ, m
$b\bar{b}\gamma\gamma$ system	p_T, η, ϕ, m
VBF-jets	$p_T, \eta, \phi, m_{jj}, \Delta\eta_{jj}$

TABLE II: Input variables for the HyQML model, where m_{jj} is the invariant mass of the two VBF jets and $\Delta\eta_{jj}$ is their η separation.

Each event is assigned a weight proportional to its generator-level cross-section, ensuring that the composition of the training sample reflects the expected yield of signal and background processes in data. To further mitigate the strong class imbalance between the simulated signal and backgrounds, an additional weighting factor is applied based on class frequencies, computed using the `compute_sample_weight` function from the `scikit-learn` package [40]. The total event weight is therefore defined as the product of the generator-level cross-section weight and the class weight. This strategy ensures that the hybrid model remains sensitive to both signal and background contributions during training, avoiding bias toward the more abundant background events.

B. HyQML model classifier

The proposed classification model employs a HyQML model designed to integrate classical feature encoding with quantum parameterized circuits in a unified end-to-end training scheme [41, 42]. The model combines two key components:

1. *Meta-Parameter Mapping Network* that generates adaptive quantum circuit parameters conditioned on the event kinematics.
2. *Parameterized Quantum Circuit* that performs quantum feature transformation and outputs expectation values used for event classification.

This hybrid architecture is implemented using `PennyLane` interfaced with `PyTorch` for automatic differentiation [43, 44].

1. *Meta-Parameter Mapping Network*

The meta-parameter mapping network (`MetaParamMapNet`) is a fully connected feed-forward neural network that maps the classical kinematic features of each event to the trainable parameters of the PQC. By conditioning the PQC on event-by-event information, this network enables the quantum model to adapt its state preparation and entanglement structure to the underlying physics, thereby enhancing its expressive power for signal to background discrimination.

For an input vector of dimension d_{in} , the network outputs d_θ rotation angles that parameterize the single-qubit gates in the PQC:

$$d_\theta = N_{\text{qubits}} \times N_{\text{layers}} \times 3,$$

where the factor of 3 corresponds to the independent R_X , R_Y , and R_Z rotations applied to each qubit in every layer.

The network consists of three hidden layers of 64 neurons, each followed by ReLU activation and a normalization layer [45]. This architecture provides sufficient non-linearity to model complex correlations in the input features while maintaining stable gradients during training. The output of the encoder can be expressed as:

$$h = \text{LayerNorm}(\text{ReLU}(W_2 \text{ReLU}(W_1 x + b_1) + b_2)),$$

which is subsequently projected onto the d_θ -dimensional parameter space of the PQC.

Each PQC parameter θ_i in the proposed HyQML model is produced by a dedicated linear head within the `MetaParamMapNet`. After the shared encoder processes the input variables through several nonlinear layers and produces a latent representation h , this vector is passed

to a group of small, independent linear sub-networks, or "heads" each head corresponds to one learnable parameter of the PQC. Specifically, the i -th head applies a simple linear transformation with its own weight vector and bias term to generate a single scalar value $\Delta\theta_i$:

$$\Delta\theta_i = W_i^{(\text{head})}h + b_i^{(\text{head})},$$

and combined with a learnable base parameter vector θ_0 as

$$\theta = \theta_0 + \Delta\theta,$$

which is then squashed into the range $[-\pi, \pi]$ using a sigmoid-based rescaling. This parameterization allows the network to produce a distinct set of quantum rotation angles for every event, effectively learning a data-dependent embedding in the quantum circuit parameter space.

2. *Parameterized Quantum Circuit*

The parameterized quantum circuit in the proposed HyQML framework is implemented as a hardware-efficient ansatz with $N_{\text{qubits}} = 4$ and $N_{\text{layers}} = 4$. A hardware-efficient ansatz means the circuit is designed to be easily realizable on near-term quantum hardware, which uses standard rotation and entangling gates that can be implemented on most quantum processors. This structure strikes a balance between expressiveness and practicality, providing sufficient parameterized gates to capture complex data relationships while maintaining a shallow enough circuit to remain trainable and computationally efficient.

Before quantum computation begins, the input features from the classical dataset must be embedded into a quantum state. In this model, we use amplitude embedding, which encodes the classical feature vector x directly into the amplitude coefficients of a quantum state:

$$|\psi(x)\rangle = \text{AmplitudeEmbedding}(x),$$

which normalizes and loads the input vector into the quantum amplitude space of dimension $2^{N_{\text{qubits}}}$. Each layer of the quantum circuit performs a series of qubit rotations followed by entangling CNOT operations [46]. The rotations R_X , R_Y , and R_Z are parameterized by the trainable angles $\theta_{l,i,1}$, $\theta_{l,i,2}$, and $\theta_{l,i,3}$ for qubit i in layer l . These rotations allow each qubit

to explore the full Bloch sphere, providing a flexible basis for representing complex transformations of the encoded input state. After all qubits in a layer are rotated, entanglement is introduced through a sequence of CNOT gates arranged in a ring topology, where each qubit acts as a control for the next one. This structure ensures that correlations propagate throughout the system, enabling both local and global quantum interactions. This operation is expressed as:

$$U(\theta) = \prod_{i=1}^{N_{\text{qubits}}} \text{CNOT}(i, (i+1) \bmod N_{\text{qubits}}) \cdot \prod_{l=1}^{N_{\text{layers}}} \left[\prod_{i=1}^{N_{\text{qubits}}} R_Z^{(i)}(\theta_{l,i,3}) R_Y^{(i)}(\theta_{l,i,2}) R_X^{(i)}(\theta_{l,i,1}) \right]$$

After that, the expectation values of the Pauli-Z operator are measured on each qubit from the quantum feature vector that represents the output of the parameterized quantum circuit. For the j -th qubit, the expectation value is defined as:

$$z_j = \langle \psi(x, \theta) | Z_j | \psi(x, \theta) \rangle, \quad j = 1, \dots, N_{\text{qubit}}.$$

These measured values capture the quantum correlations and interference effects generated within the circuit, encoding the processed information from the embedded input event kinematics. The resulting vector $\mathbf{z} = [z_1, z_2, \dots, z_{N_{\text{qubits}}}]^\top$ is then passed to a classical linear output layer, which maps the N_{qubits} expectation values to $N_{\text{classes}} = 2$ logits corresponding to the classification outputs. This final step converts the quantum features into classical decision variables, completing the hybrid quantum–classical learning pipeline.

A schematic view of the full hybrid quantum machine learning model is shown in Figure 3.

C. Hybrid Training Procedure

The hybrid training objective defines how the classical and quantum components of the model are combined and optimized together. In this framework, the overall function $f(x)$ represents the hybrid model that takes a classical input and processes it through both the meta-network and the PQC. The quantum circuit then computes the expectation values of observables, which form the quantum feature vector. These values are passed to a classical linear layer that outputs the final prediction. The model demonstrates that the quantum

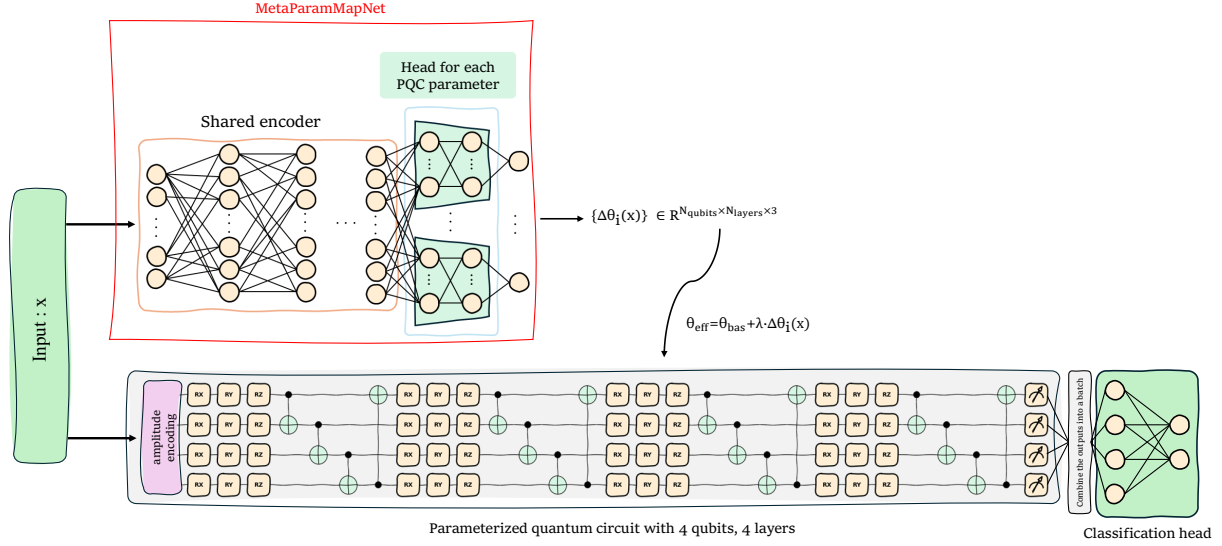


FIG. 3: A schematic view of the hybrid quantum machine learning model.

circuit generates learned features that are linearly combined to produce the classification output. The hybrid model can be simplified as:

$$f(x) = \text{Linear}(\mathbb{E}_{\text{PQC}}(\theta(x))) .$$

where $\theta(x)$ are the parameters predicted by the meta-model. A scalar coefficient λ regulates the contribution of the meta-network to the final circuit parameters [41], allowing gradual control between static and data-dependent parameterization:

$$\theta_{\text{eff}} = \theta_{\text{base}} + \lambda \Delta\theta(x).$$

For different values of the hyperparameter λ , the full model is trained using the standard cross-entropy loss function, which measures the difference between the predicted and true class probabilities, which is defined as:

$$\mathcal{L} = - \sum_i w_i y_i \log \hat{y}_i,$$

where w_i is per-event weight. Training is performed using the Adam optimizer with a learning rate of 10^{-3} for 20 epochs, and gradient clipping on the PQC parameters θ to prevent instability from large updates.

A meta-learning stage precedes the hybrid model training, during which the meta-parameter network is optimized independently to minimize the condition number of the

Fubini–Study metric g_{ij} of the PQC [47]. This metric characterizes the local geometry of the quantum parameter space and provides a diagnostic of barren-plateau behavior [48]. The meta-objective is defined as:

$$\mathcal{L}_{\text{meta}} = \langle \log \kappa(g_{ij}) \rangle ,$$

where $\kappa(g_{ij})$ denotes the condition number of the metric tensor. Minimizing $\mathcal{L}_{\text{meta}}$ improves the expressivity and trainability of the PQC by promoting well-conditioned gradients and mitigating vanishing-gradient regimes [41].

The meta-network is trained using a small batch of randomly sampled events and optimized with the Adam optimizer until convergence. Once this meta-learning phase is completed, the meta-network weights are frozen and remain fixed throughout the subsequent quantum-circuit training stage, ensuring that the PQC is trained on a stable and geometry-optimized parameterization.

D. Hybrid Model Performance

The performance of the trained quantum classifier is evaluated using the Receiver Operating Characteristic (ROC) curve, which quantifies the trade-off between signal efficiency (true positive rate) and background rejection (false positive rate) as the decision threshold on the model output score is varied. The area under the ROC curve (AUC) serves as a global measure of classification performance. Figure 4 shows the ROC curve obtained for the hybrid quantum model for different hyperparameter λ values. It is found that the $\lambda = 0.1$ provides a better discrimination power between the $HH \rightarrow b\bar{b}\gamma\gamma$ signal and the dominant background processes. The case $\lambda = 0$ corresponds to a pure quantum circuit (Pure-QML), which shows very low performance compared to the HyQML model. The rest of the results are shown for $\lambda = 0.1$.

To gain insight into the internal representations learned by the hybrid quantum model, we monitor the evolution of the PQC output during training. At each epoch, the expectation values of the PQC measurements are collected and projected onto a two-dimensional space using Principal Component Analysis (PCA) [49]. Figure 5 shows the PCA projection between the first and the last epoch in the training of the hybrid quantum model. These projections provide an interpretable visualization of the quantum latent space, allowing one

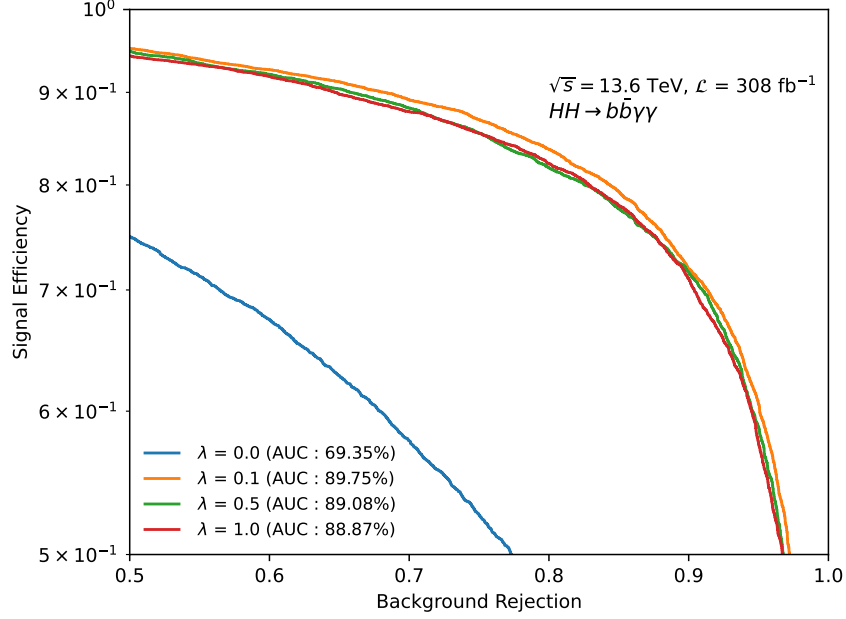


FIG. 4: Weighted ROC curve for the HyQML classifier for different λ values. The ROC curve is computed using event weights on the test dataset.

to monitor the progressive separation between signal and background distributions. At an early stage of training, the signal and background samples occupy largely overlapping regions in the latent space, indicating that the PQC has not yet learned a meaningful transformation of the input kinematic features. At a later training epoch, a clear organization of the latent space emerges: the signal and background distributions begin to separate into distinguishable clusters, and the points align along a more coherent low-dimensional manifold. The latent topology also undergoes a noticeable rotation relative to the early-epoch projection, reflecting the progressive restructuring of the quantum feature space as the model learns. The reduced overlap between the two classes demonstrates that the PQC increasingly captures discriminative patterns relevant for classification. This diagnostic step reveals that the HyQML model learns structured manifolds in the quantum feature space, with a gradual increase in class separation and rotation in the latent topology over successive epochs.

VI. RESULTS

The goal of this analysis is to enhance the sensitivity to non-resonant Higgs boson pair production using a HyQML classifier designed to discriminate between signal and back-

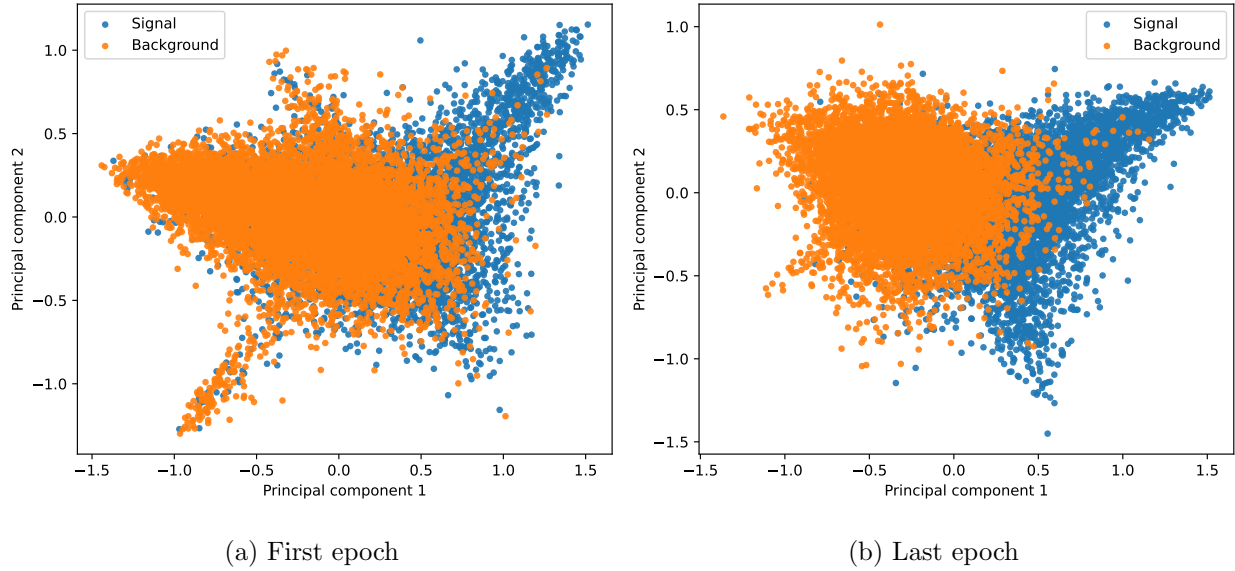


FIG. 5: Two-dimensional projection of the PQC expectation values for the first and the last epochs for $\lambda = 0.1$.

ground events. The model combines quantum feature mapping with classical optimization to exploit both non-linear correlations and quantum-induced entanglement patterns that are otherwise inaccessible to classical architectures. To quantify the statistical sensitivity gain, the output distribution of the HyQML classifier is divided into optimized score regions corresponding to increasing signal purity. In particular, the signal-enriched region is defined for classifier scores above 0.5, where the quantum model exhibits the strongest separation between signal and background. The optimized regions are required to maximize the expected statistical significance computed using the Asimov approximation to the profile likelihood ratio [50] defined as:

$$Z = \sqrt{2 \times ((s + b) \times \ln(1 + s/b) - s)},$$

where s is the total signal yield and b corresponds to the total background yield. In each defined bin a requirement of at least one background event is imposed. The lower-score region, containing predominantly background events, is grouped into a single inclusive bin as it contributes marginally to the overall sensitivity. Figure 6 presents the HyQML output score distribution for both signal and total SM background in the signal region defined before. The vertical dashed line marks the optimized threshold separating the signal-enriched and

inclusive regions. A clear distinction between the two classes is visible, highlighting the HyQML model’s ability to capture non-trivial event correlations in the $b\bar{b}\gamma\gamma$ final state.

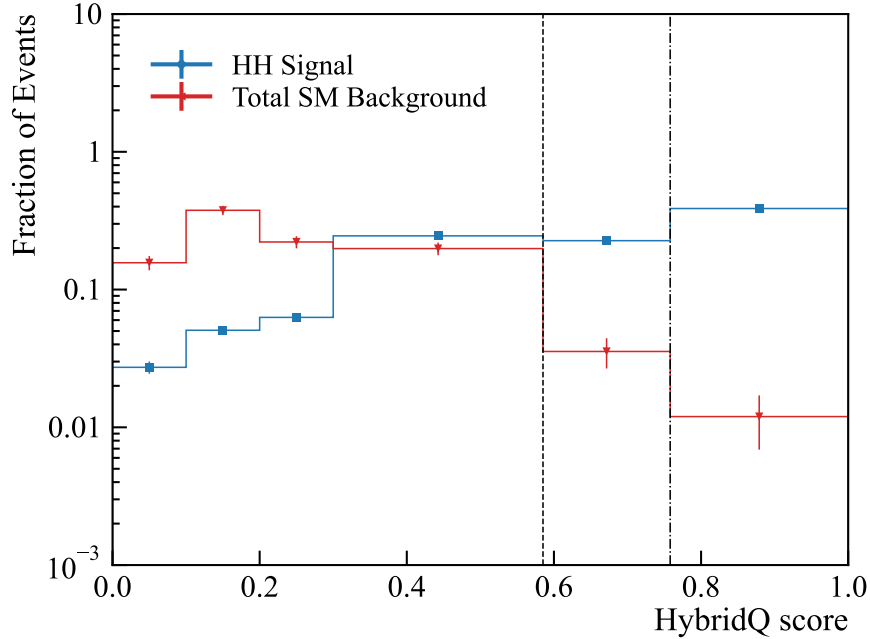


FIG. 6: Distribution of the HyQML score. The vertical dashed line indicates the optimized bins threshold.

Due to the limited size of the simulated continuum $\gamma\gamma$ +jets background, the effective event statistics are low in high-score regions. Consequently, bins with few high-score background events may have large statistical fluctuations. To assess the robustness of the result, the statistical interpretation is therefore performed with a background uncertainty of 10% and 50%. In both cases, the HyQML model maintains stable performance and consistent behavior, confirming that the hybrid quantum approach generalizes well even in low-statistics regimes. The performance of the HyQML is also compared to a pure-QML ($\lambda = 0$) and the XGBoost model presented in the Ref [12] with 10% background uncertainty.

The sensitivity of the analysis is quantified through two key statistical metrics: the expected discovery significance and the 95% confidence level (CL) upper exclusion limit on the Higgs boson pair production cross-section [51]. Both quantities are obtained from a binned likelihood fit to the HyQML classifier output distribution using the optimized binning.

The statistical inference is performed with the `pyhf` package [52], which implements the

profile-likelihood formalism. For a given signal strength parameter $\mu \geq 0$ defined as,

$$\mu = \frac{\sigma(pp \rightarrow HH)}{\sigma_{\text{SM}}(pp \rightarrow HH)},$$

such that $\mu = 1$ corresponds to the SM expectation, and a set of nuisance parameters $\boldsymbol{\theta}$ describing uncertainties corresponding in this analysis to 10% and 50%, the likelihood function is defined as

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{N_{\text{bins}}} \text{Pois}(n_i | \mu s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta})) \times \prod_{k=1}^{N_{\text{nuis}}} \pi_k(\theta_k),$$

where n_i is the observed yield in bin i , while s_i and b_i are the expected signal and background yields, respectively. n_i is defined as the sum of the background yield b_i and the SM ($\kappa_\lambda = 1$) expected signal yield s_i . The constraint terms $\pi_k(\theta_k)$ encode only the background normalization uncertainty of 10% and 50% and is implemented as log-normal distributions.

The profile-likelihood ratio is defined as

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}_\mu)}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})},$$

where $(\hat{\mu}, \hat{\boldsymbol{\theta}})$ are the unconditional maximum-likelihood estimators (MLEs), and $\hat{\boldsymbol{\theta}}_\mu$ are the conditional MLEs for a fixed signal strength μ . The corresponding test statistic for upper limits is

$$q_\mu = \begin{cases} -2 \ln \lambda(\mu), & 0 \leq \hat{\mu} \leq \mu, \\ 0, & \text{otherwise.} \end{cases}$$

The discovery test statistic is obtained by setting $\mu = 0$,

$$q_0 = \begin{cases} -2 \ln \lambda(0), & \hat{\mu} \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Under the asymptotic approximation, the one-sided discovery significance is given by [50]

$$Z = \sqrt{q_0},$$

evaluated on an Asimov dataset generated under the signal-plus-background hypothesis ($\mu = 1$). Similarly, the 95% CL upper limit on the signal strength μ is obtained from the modified frequentist CL_s criterion,

$$CL_s(\mu) = \frac{p(q_\mu | \mu)}{1 - p(q_0 | 0)}, \quad \text{solve for } \mu \text{ such that } CL_s(\mu) = 0.05.$$

The same likelihood fit to the HyQML classifier output is used to extract both the expected discovery significance and the 95% CL upper limit.

Table III summarizes the expected discovery significance obtained with the fit for both 10% and 50% background uncertainties compared to the pure-QML and classical XGBoost, which is found to be consistent with the values reported by the ATLAS experiment [9]. A factor of two improvement is achieved with the HyQML over the pure-QML in the expected discovery significance.

	HyQML (10% sys.)	HyQML (50% sys.)	Pure-QML ($\lambda = 0$)	XGBoost
Significance	1.41	1.12	0.65	1.09

TABLE III: Expected discovery significance for both 10% and 50% background uncertainties compared to the pure-QML and the XGBoost.

Figure 7 shows the expected 95% CL upper limits on the μ_{HH} for the two background normalization uncertainties hypotheses and the pure quantum model. Assuming a total background normalization uncertainty of 10%, the expected limit is approximately $1.9 \times \sigma_{\text{SM}}$. When the uncertainty is increased to 50%, the limit relaxes slightly to about $2.1 \times \sigma_{\text{SM}}$. Despite this increase, the results remain consistent within uncertainties, demonstrating the robustness of the HyQML approach and its ability to maintain stable performance even under limited-statistics conditions. The HyQML demonstrates a factor of 1.75 improvement in the expected 95% CL upper limit compared to the pure quantum model architecture and achieves a 21% improvement in the upper limit compared to a classical XGBoost.

Despite being based on a simplified detector simulation, the quantum-enhanced classifier achieves improvement compared to classical method used in the latest ATLAS results [9]. The ATLAS collaboration reports an observed (expected) upper limits on the Higgs boson pair production cross-section of 3.8 (3.7) times the SM prediction. Even though a direct comparison is not optimal, the proposed HyQML achieves a significant improvement. The performance gain is attributed to the hybrid quantum model's ability to encode correlations in a higher-dimensional feature space through non-classical operations, enabling more efficient separation of signal and background events.

Although the present study is not explicitly optimized for measuring the trilinear Higgs boson self-coupling modifier, κ_λ , a one-dimensional profile-likelihood scan is performed to

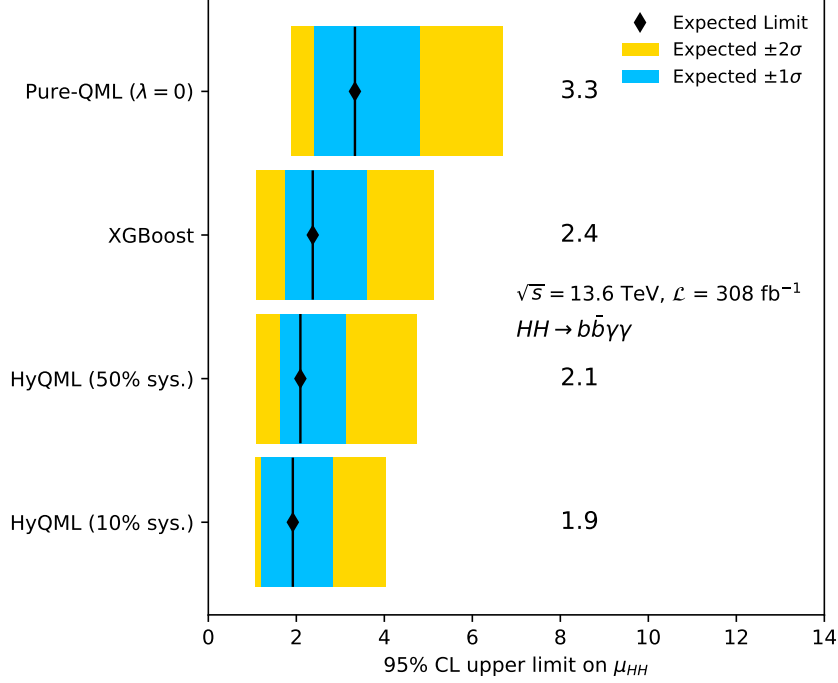


FIG. 7: Expected 95% CL upper limits on the signal strength μ_{HH} under the 10%, 50% background normalization uncertainties conditions and pure quantum model. The shaded bands represent the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty intervals.

estimate the 68% and 95% confidence intervals in both scenarios. In this procedure, only κ_λ is treated as a free parameter while all other couplings are fixed to their SM values. The scan accounts only for the dependence of the total production cross-section on κ_λ as defined by Equation 1, without including potential kinematic shape variations. Figure 8(a) shows the negative log-likelihood profile obtained with the 10%, 50% background normalization uncertainties and the pure quantum model. From the likelihood scan, the HyQML analysis yields an expected 95% (68%) confidence interval of approximately

$$\kappa_\lambda \in [-0.4, 4.9] \quad ([0.3, 4.2]),$$

which is consistent with the latest experimental constraints of $[-1.7, 6.6]$ (95%) and $[-0.4, 5.1]$ (68%) [9] and demonstrates a significant improvement over both the pure quantum architecture and classical XGBoost model. The comparison underscores the potential of quantum-enhanced methods to deliver precision measurements of the Higgs boson self-coupling in future high-luminosity LHC analyses and beyond.

A similar one-dimensional scan is performed for the κ_{2V} coupling, which parameterizes

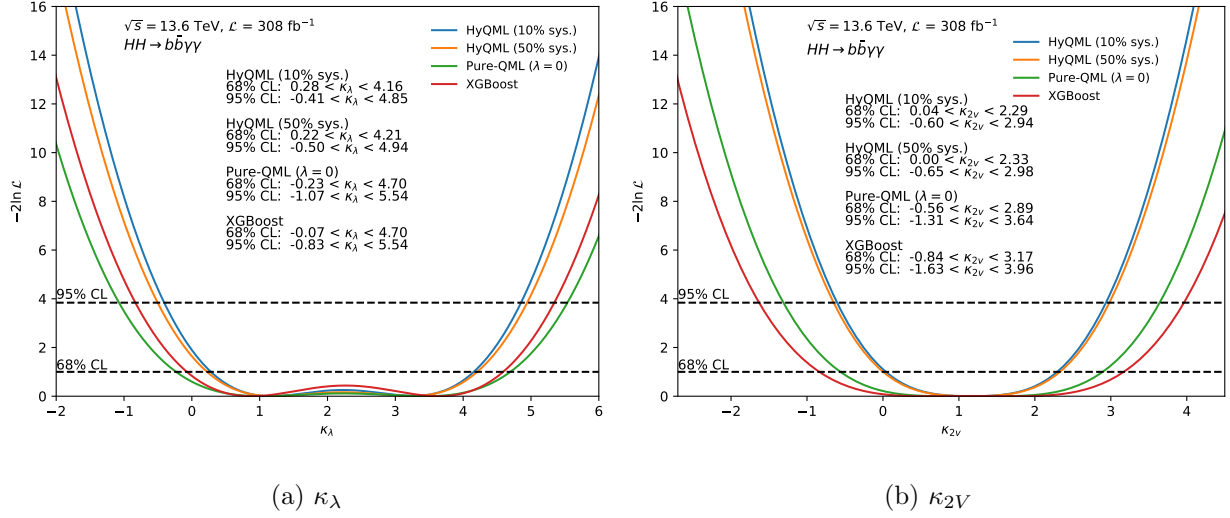


FIG. 8: Negative log-likelihood as a function of the trilinear Higgs self-coupling modifier κ_λ (a) and κ_{2V} (b) for the 10% (blue), 50% (orange) background uncertainties, the pure-QML (green) and the XGBoost model (red). Dashed lines indicate the 68% and 95% confidence level intervals.

deviations in the quartic $VVHH$ interaction entering the VBF Higgs pair production process (Equation 2). As shown in Figure 8(b), the HyQML model displays a well-defined likelihood minimum near the SM expectation $\kappa_{2V} = 1$, with symmetric confidence intervals at 68% and 95% CL. The expected sensitivity obtained from this analysis corresponds to

$$\kappa_{2V} \in [-0.6, 2.9] \quad ([0.0, 2.3]),$$

We must stress that this analysis does not include a dedicated VBF-category resulting of the lower sensitivity. These results confirm that the hybrid quantum framework maintains robust performance across different coupling hypotheses, providing consistent constraints on both κ_λ and κ_{2V} without the need for explicit retraining under new coupling scenarios. In addition, a two-dimensional likelihood fit is performed allowing both κ_λ and κ_{2V} to vary. Figure 9 shows the 95% and 68% CL contours with the best-fit values of κ_λ and κ_{2V} .

Overall, the HyQML training procedure demonstrates a significant improvement over both a pure-QML with $\lambda = 0$ and a classical XGBoost model. By learning event topology directly from multidimensional kinematic correlations, the quantum-assisted model retains strong discriminating power while generalizing effectively to BSM coupling variations. These results strongly highlight the potential of quantum-enhanced methods as competitive tools

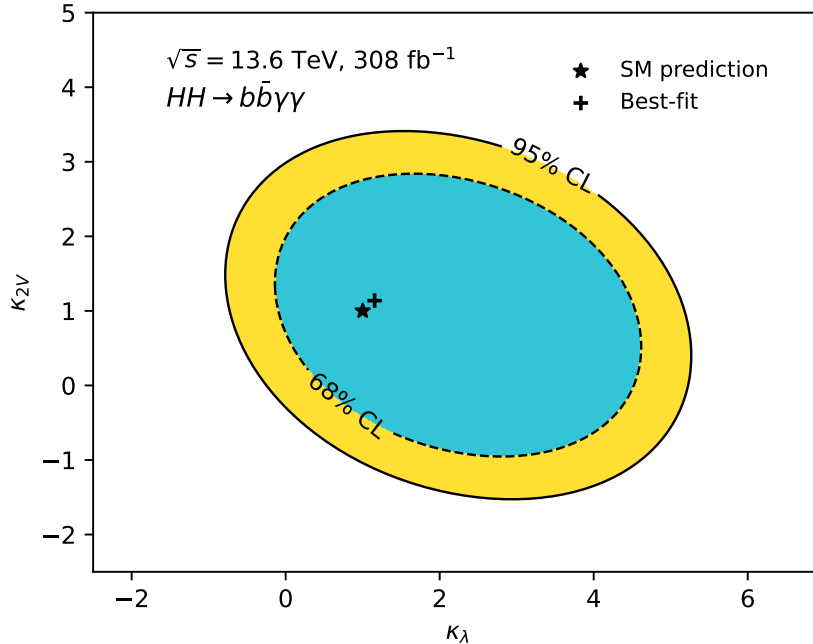


FIG. 9: Contours at 68% CL (dashed line) and 95% CL (solid line) in the $(\kappa_\lambda, \kappa_{2V})$ parameter space. The SM prediction ($\kappa_\lambda = 1, \kappa_{2V} = 1$) is indicated by a star, while the $+$ corresponds to the best-fit values.

for precision Higgs-sector measurements at the LHC and future high-luminosity upgrades.

VII. CONCLUSION

In this work, we proposed a HyQML algorithm to enhance the sensitivity of double Higgs boson searches at the LHC. The analysis focuses on the $HH \rightarrow b\bar{b}\gamma\gamma$ final state, which, despite its small branching ratio, offers a clean experimental signature and excellent mass resolution. The proposed framework combines parameterized quantum circuits with a classical neural meta-model that dynamically conditions quantum parameters on event-level features, thereby integrating quantum feature representations with the optimization stability of classical learning. This hybrid design bridges the gap between classical and quantum approaches, exploiting high-dimensional correlations that are difficult to capture with traditional models.

The hybrid model outperforms a purely quantum architecture, achieving an improvement of approximately 27% in the area under the ROC curve (AUC). Furthermore, the HyQML

framework yields an expected 95% CL upper limit on the non-resonant di-Higgs production cross-section of $1.9 \times \sigma_{\text{SM}}$ assuming a 10% background normalization uncertainty, and $2.1 \times \sigma_{\text{SM}}$ when the uncertainty is increased to 50% leading to almost a factor-of-two improvement compared to the pure-QML model and a 21% improvement with respect to an XGBoost model. These results demonstrate the robustness of the hybrid quantum–classical approach, maintaining stable performance even under large systematic variations and limited training statistics. Likelihood scans of the Higgs self-coupling modifier κ_λ and the quartic vector-boson–Higgs coupling κ_{2V} show confidence intervals consistent with recent ATLAS measurements, confirming the model’s reliability across different coupling hypotheses.

Overall, this study demonstrates that quantum-assisted learning can achieve strong discriminating power in realistic collider analyses, and with the rapid progress in quantum hardware—particularly improvements in qubit fidelity, circuit depth, and noise mitigation—hybrid quantum models hold significant promise for future searches of new physics at the LHC.

ACKNOWLEDGMENTS

This work is supported by the United Arab Emirates University (UAEU) Start-Up Grant No 12S157. The authors gratefully thank the AI and Robotics Lab of United Arab Emirates University for offering computing facilities including HPC and DGX1 for MC simulation and ML training.

DATA AND CODE AVAILABILITY

The datasets and code used in this analysis can be provided by the corresponding author upon reasonable request.

-
- [1] G. Aad *et al.* (ATLAS), Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, *Phys. Lett. B* **716**, 1 (2012), arXiv:1207.7214 [hep-ex].

- [2] S. Chatrchyan *et al.* (CMS), Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC, Phys. Lett. B **716**, 30 (2012), arXiv:1207.7235 [hep-ex].
- [3] G. Aad *et al.* (ATLAS), Combined Measurement of the Higgs Boson Mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ Decay Channels with the ATLAS Detector Using $\sqrt{s}=7, 8$, and 13 TeV pp Collision Data, Phys. Rev. Lett. **131**, 251802 (2023), arXiv:2308.04775 [hep-ex].
- [4] A. Hayrapetyan *et al.* (CMS), Measurement of the Higgs boson mass and width using the four-lepton final state in proton-proton collisions at $\sqrt{s}=13$ TeV, Phys. Rev. D **111**, 092014 (2025), arXiv:2409.13663 [hep-ex].
- [5] R. L. Workman *et al.* (Particle Data Group), Review of Particle Physics, PTEP **2022**, 083C01 (2022).
- [6] G. Aad *et al.* (ATLAS), Constraints on the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using pp collisions at $\sqrt{s}=13$ TeV, Phys. Lett. B **843**, 137745 (2023), arXiv:2211.01216 [hep-ex].
- [7] A. Hayrapetyan *et al.* (CMS), Constraints on the Higgs boson self-coupling from the combination of single and double Higgs boson production in proton-proton collisions at $\sqrt{s}=13$ TeV, Phys. Lett. B **861**, 139210 (2025), arXiv:2407.13554 [hep-ex].
- [8] D. de Florian *et al.* (LHC Higgs Cross Section Working Group), Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, CERN Yellow Rep. Monogr. **2**, 1 (2017), arXiv:1610.07922 [hep-ph].
- [9] G. Aad *et al.* (ATLAS), Study of Higgs boson pair production in the $HH \rightarrow b\bar{b}\gamma\gamma$ final state with 308 fb $^{-1}$ of data collected at $\sqrt{s} = 13$ TeV and 13.6 TeV by the ATLAS experiment, CERN-EP-2025-140 (2025), arXiv:2507.03495 [hep-ex].
- [10] HEP ML Community, A Living Review of Machine Learning for Particle Physics, online.
- [11] J. M. Duarte, Novel machine learning applications at the LHC, PoS **ICHEP2024**, 012 (2025), arXiv:2409.20413 [hep-ex].
- [12] M. Belfkir, M. A. Loualidi, and S. Nasri, Boosting Sensitivity to $HH \rightarrow b\bar{b}\gamma\gamma$ with Graph Neural Networks and XGBoost, pre-print (2025), arXiv:2508.01449 [hep-ph].
- [13] W. Guan, G. Perdue, A. Pesah, M. Schuld, K. Terashi, S. Vallecorsa, and J.-R. Vlimant, Quantum Machine Learning in High Energy Physics, Mach. Learn. Sci. Tech. **2**, 011003 (2021), arXiv:2005.08582 [quant-ph].

- [14] V. Havlicek, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, Supervised learning with quantum-enhanced feature spaces, *Nature* **567**, 209 (2019), arXiv:1804.11326 [quant-ph].
- [15] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, Circuit-centric quantum classifiers, *Phys. Rev. A* **101**, 032308 (2020), arXiv:1804.00633 [quant-ph].
- [16] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019), arXiv:1910.11333 [quant-ph].
- [17] S. L. Wu *et al.*, Application of quantum machine learning using the quantum kernel algorithm on high energy physics analysis at the LHC, *Phys. Rev. Res.* **3**, 033221 (2021), arXiv:2104.05059 [quant-ph].
- [18] A. Fadol, Q. Sha, Y. Fang, Z. Li, S. Qian, Y. Xiao, Y. Zhang, and C. Zhou, Application of quantum machine learning in a Higgs physics study at the CEPC, *Int. J. Mod. Phys. A* **39**, 2450007 (2024), arXiv:2209.12788 [hep-ex].
- [19] F. A. Dreyer, A. Karlberg, J.-N. Lang, and M. Pellen, Precise predictions for double-Higgs production via vector-boson fusion, *Eur. Phys. J. C* **80**, 1037 (2020), arXiv:2005.13341 [hep-ph].
- [20] F. A. Dreyer and A. Karlberg, Vector-Boson Fusion Higgs Pair Production at N³LO, *Phys. Rev. D* **98**, 114016 (2018), arXiv:1811.07906 [hep-ph].
- [21] G. C. Branco, P. M. Ferreira, L. Lavoura, M. N. Rebelo, M. Sher, and J. P. Silva, Theory and phenomenology of two-Higgs-doublet models, *Phys. Rept.* **516**, 1 (2012), arXiv:1106.0034 [hep-ph].
- [22] U. Ellwanger, C. Hugonie, and A. M. Teixeira, The Next-to-Minimal Supersymmetric Standard Model, *Phys. Rept.* **496**, 1 (2010), arXiv:0910.1785 [hep-ph].
- [23] T. Robens, T. Stefaniak, and J. Wittbrodt, Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios, *Eur. Phys. J. C* **80**, 151 (2020), arXiv:1908.08554 [hep-ph].
- [24] M. Grazzini, G. Heinrich, S. Jones, S. Kallweit, M. Kerner, J. M. Lindert, and J. Mazzei, Higgs boson pair production at NNLO with top quark mass effects, *JHEP* **05**, 059, arXiv:1803.02463 [hep-ph].
- [25] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, *JHEP* **06**, 043, arXiv:1002.2581

- [hep-ph].
- [26] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, *JHEP* **07**, 079, arXiv:1405.0301 [hep-ph].
 - [27] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of PYTHIA 8.3, *SciPost Phys. Codeb.* **2022**, 8 (2022), arXiv:2203.11601 [hep-ph].
 - [28] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3), DELPHES 3, A modular framework for fast simulation of a generic collider experiment, *JHEP* **02**, 057, arXiv:1307.6346 [hep-ex].
 - [29] G. Aad *et al.* (ATLAS), Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data, *JINST* **14** (12), P12006, arXiv:1908.00005 [hep-ex].
 - [30] G. Aad *et al.* (ATLAS), Measurement of photon identification efficiency using radiative z decays using 2022-2024 collision data at the atlas experiment ().
 - [31] M. Cacciari, G. P. Salam, and G. Soyez, The anti- k_t jet clustering algorithm, *JHEP* **04**, 063, arXiv:0802.1189 [hep-ph].
 - [32] M. Cacciari, G. P. Salam, and G. Soyez, FastJet User Manual, *Eur. Phys. J. C* **72**, 1896 (2012), arXiv:1111.6097 [hep-ph].
 - [33] G. Aad *et al.* (ATLAS), Transforming jet flavour tagging at ATLAS, CERN-EP-2025-103 (2025), arXiv:2505.19689 [hep-ex].
 - [34] G. Aad *et al.* (ATLAS), The atlas trigger system for lhc run 3 and trigger performance in 2022 ().
 - [35] G. Aad *et al.* (ATLAS), Studies of new Higgs boson interactions through nonresonant HH production in the $b\bar{b}\gamma\gamma$ final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *JHEP* **01**, 066, arXiv:2310.12301 [hep-ex].
 - [36] G. Aad *et al.* (ATLAS), Search for Higgs boson pair production in the two bottom quarks plus two photons final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, *Phys. Rev. D* **106**, 052001 (2022), arXiv:2112.11876 [hep-ex].
 - [37] G. Aad *et al.* (ATLAS), Measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels with the ATLAS detector using 25 fb $^{-1}$ of pp collision data, *Phys. Rev.*

- D **90**, 052004 (2014), arXiv:1406.3827 [hep-ex].
- [38] G. C. Strong, On the impact of selected modern deep-learning techniques to the performance and celerity of classification models in an experimental high-energy physics use case, *Mach. Learn. Sci. Tech.* **1**, 045006 (2020), arXiv:2002.01427 [physics.data-an].
 - [39] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, <https://arxiv.org/abs/1502.03167> (2015), arXiv:1502.03167 [cs.LG].
 - [40] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* **12**, 2825 (2011).
 - [41] M. Ait Haddou and M. Bennai, Sculpting Quantum Landscapes: Fubini-Study Metric Conditioning for Geometry Aware Learning in Parameterized Quantum Circuits, arXiv (2025), arXiv:2506.21940 [cs.LG].
 - [42] T. Haug and M. Kim, Natural parametrized quantum circuit, *Physical Review A* **106**, 052611 (2022).
 - [43] V. Bergholm *et al.*, PennyLane: Automatic differentiation of hybrid quantum-classical computations, pre-print (2018), arXiv:1811.04968 [quant-ph].
 - [44] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, Pytorch: An imperative style, high-performance deep learning library (2019), arXiv:1912.01703 [cs.LG].
 - [45] A. F. Agarap, Deep learning using rectified linear units (relu), <https://arxiv.org/abs/1803.08375> (2019), arXiv:1803.08375 [cs.NE].
 - [46] M. Bataille, Quantum circuits of CNOT gates, arXiv (2020), arXiv:2009.13247 [quant-ph].
 - [47] T. Hospedales, A. Antoniou, P. Micaelli, and A. Storkey, Meta-learning in neural networks: A survey, *IEEE transactions on pattern analysis and machine intelligence* **44**, 5149 (2021).
 - [48] M. Larocca, S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo, Barren plateaus in variational quantum computing, *Nature Rev. Phys.* **7**, 174 (2025), arXiv:2405.00781 [quant-ph].

- [49] A. Maćkiewicz and W. Ratajczak, Principal components analysis (pca), Computers & Geosciences **19**, 303 (1993).
- [50] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, Asymptotic formulae for likelihood-based tests of new physics, Eur. Phys. J. C **71**, 1554 (2011), [Erratum: Eur.Phys.J.C 73, 2501 (2013)], arXiv:1007.1727 [physics.data-an].
- [51] A. L. Read, Presentation of search results: The CL_s technique, J. Phys. G **28**, 2693 (2002).
- [52] M. Feickert, L. Heinrich, and G. Stark, pyhf: a pure-Python statistical fitting library with tensors and automatic differentiation, PoS **ICHEP2022**, 245 (2022), arXiv:2211.15838 [hep-ex].