

Continuous-variable photonic quantum extreme learning machines for fast collider-data selection

Benedikt Maier ^a **Michael Spannowsky** ^b **Simon Williams** ^b

^a*Blackett Laboratory, Imperial College, Prince Consort Road, London, SW7 2AZ, United Kingdom*

^b*Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, UK*

E-mail: benedikt.maier@cern.ch, michael.spannowsky@durham.ac.uk,
simon.j.williams@durham.ac.uk

ABSTRACT: We study continuous-variable photonic quantum extreme learning machines as fast, low-overhead front-ends for collider data processing. Data is encoded in photonic modes through quadrature displacements and propagated through a fixed-time Gaussian quantum substrate. The final readout occurs through Gaussian-compatible measurements to produce a high-dimensional random feature map. Only a linear classifier is trained, using a single linear solve, so retraining is fast, and the optical path and detector response set the analytical and inference latency. We evaluate this architecture on two representative classification tasks, top-jet tagging and Higgs-boson identification, with parameter-matched multi-layer perceptron (MLP) baselines. Using standard public datasets and identical train, validation, and test splits, the photonic Quantum Extreme Learning Machine (QELM) outperforms an MLP with two hidden units for all considered training sizes, and matches or exceeds an MLP with ten hidden units at large sample sizes, while training only the linear readout. These results indicate that Gaussian photonic extreme-learning machines can provide compact and expressive random features at fixed latency. The combination of deterministic timing, rapid retraining, low optical power, and room temperature operation makes photonic QELMs a credible building block for online data selection and even first-stage trigger integration at future collider experiments.

Contents

1	Introduction	1
2	Quantum Extreme Learning Machines	4
3	Continuous-variable quantum extreme learning machines	6
4	Results	10
4.1	Top jet tagging: QCD vs Top	11
4.2	Higgs boson identification	13
5	Summary and Conclusions	15

1 Introduction

Next-generation collider experiments will demand unprecedented computational power and performance in online data selection and reduction, large-scale simulation and reconstruction, and physics interpretation. Consequently, this has been driving interest in computational paradigms that go beyond conventional computing architectures. Quantum algorithms are emerging as a potential route to efficient, low-latency computation that exploit quantum behaviour to access computational regimes beyond classical computing. Recent studies have demonstrated the feasibility of quantum approaches to collider-physics tasks such as charged-particle track finding [1, 2], anomaly detection [3, 4], parton-shower evolution [5–7], and perturbative QCD methods [8–10]. In parallel, progress in simulating quantum field theories on near-term quantum devices [11–18] highlights a growing convergence between quantum information science and high-energy physics, with long-term implications for both theory and experiment.

In particular, continuous-variable quantum computing (CVQC), which encodes information in the quadrature variables of quantum oscillators called *qumodes*, offers a natural route to fast, highly parallel pre-processing. Information is carried and transformed optically, enabling nanosecond-scale operation governed by device transit times and optical control rather than firmware compile cycles. The underlying theory and toolbox for CV quantum information and computation are mature, including both Gaussian operations, such as displacement, squeezing, linear optics, and non-Gaussian resources needed for universality [19–23]. Already, CVQC has seen applications in high-energy physics, including event classification [24] and the simulation of scalar field theories [25–27].

From a machine learning perspective, extreme learning machines (ELMs) [28–30] provide a particularly attractive model class for low-latency inference and ultra-fast retraining. ELMs

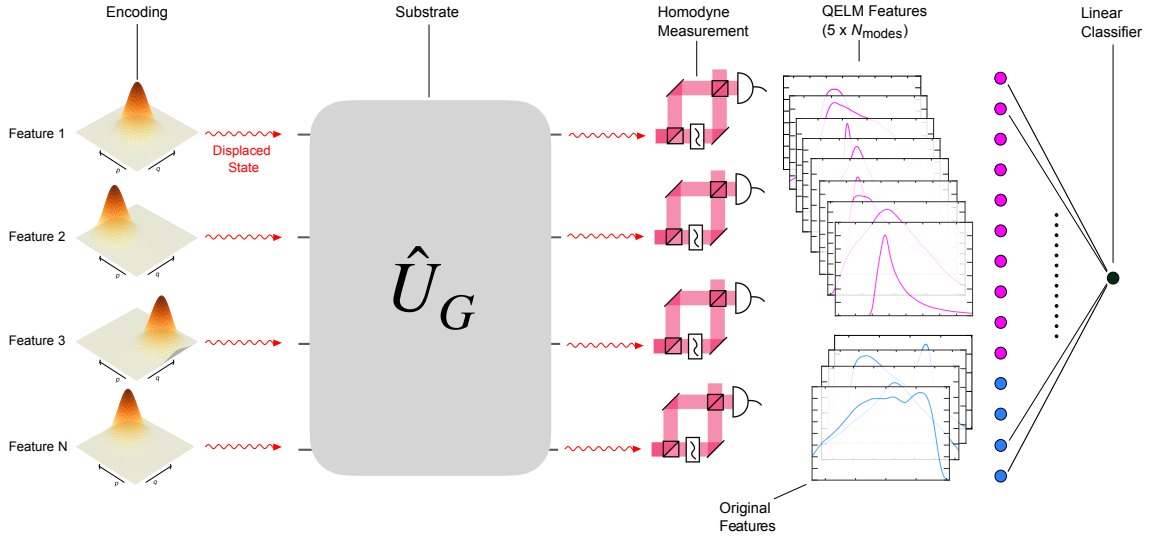


Figure 1: Visualisation of the QELM. The features of the data get encoded onto qumodes via displacement operations. A fixed-time Gaussian quantum substrate $\hat{U}_G$ followed by Gaussian-compatible measurements creates a random, high-dimensional feature map, which is used, together with the original data, in a linear classifier.

fix a random hidden layer and learn only a linear readout by solving a single regression, thereby eliminating back-propagation and its associated compute and tuning overhead [28]. When the random feature map is provided by a physical device, the approach inherits the device’s speed, parallelism, and analog bandwidth. Photonic platforms are especially compelling in this role, with extensive progress in nanophotonic neural primitives and integrated photonic reservoir computing that demonstrate low-power, high-throughput inference compatible with on-edge deployment [31–33].

This connects naturally to the quantum extreme learning machine (QELM) framework [34, 35], where a fixed, possibly uncalibrated quantum circuit produces a high-dimensional quantum random feature map that is read out and linearly trained. In this sense, QELMs instantiate the broader idea of quantum feature maps underpinning quantum kernel methods [36, 37]. Theory has clarified both potential and limits of QELMs, including concise descriptions via effective measurements, expressivity analyses, and guidance on circuit design to avoid concentration effects [35, 38, 39]. First applications already demonstrate state estimation and entanglement witnessing with photonic implementations, highlighting the practicality of the train-by-regression paradigm on near-term devices [40–42].

Within high-energy physics, fast AI on specialised hardware has matured from concept to deployed tools for triggers and real-time reconstruction, with the `hls4ml` dataset [43] providing a pathway from trained models to FPGA and even ASIC implementations under strict

latency budgets [44, 45]. In parallel, photonic computing has demonstrated integrated reservoirs and coherent optical neural networks with competitive accuracy and orders-of-magnitude speed and energy benefits in the lab [31, 33]. On the quantum side, QELMs and quantum reservoir computing synthesise these strands by using the inherent dynamics and large Hilbert spaces of quantum circuits to implement powerful random feature maps with minimal training, a direction surveyed in [38] and formalised in [35, 39]. Recent photonic QELM experiments underscore the feasibility for quantum state tasks [40–42]. Our work brings this emerging paradigm into the collider context, proposing and evaluating a continuous-variable photonic QELM as an ultra-fast, reconfigurable front-end for pattern recognition at the detector edge.

We present a continuous-variable quantum extreme learning machine (CV-QELM) whose quantum circuit is a fixed-time Gaussian photonic reservoir, called a quantum substrate, drawn from a reproducible ensemble. Inputs are encoded as quadrature displacements across M optical modes, qumodes, which propagate through layers of two-mode Gaussian operations that generate entanglement and mode mixing. The qumodes are then measured using either photon-number-resolving (PNR) or homodyne detection to produce a random feature vector, $\Phi_{\text{(hom)}}(\mathbf{x})$, of dimension R , such that the quantum substrate generates a high-dimensional random feature map,

$$\mathbf{x} \in \mathbb{R}^d \longmapsto \Phi(\mathbf{x}) \in \mathbb{R}^R. \quad (1.1)$$

Training on the linear readout is then performed via either logistic regression or ridge regression, providing analytic, closed-form updates that make retraining extremely fast. A schematic of the setup is shown in Figure 1. This design matches the requirements of front-end data processing, where latency is set by the optical path and detector response, power consumption is dominated by low optical powers and photodiodes, and reconfiguration is achieved by adjusting homodyne phases and pump settings rather than recompiling firmware.

The approach is motivated by two observations. First, ELM-style training is attractive under strict latency and retraining constraints, since only the readout is learned, allowing rapid adaptation to changing detector conditions and domain shifts in the data stream. Second, Gaussian photonic circuits supply a physically native, massively parallel random feature map with large effective bandwidth. Although classically simulable, such circuits act as analog random projections that approximate families of smooth kernels while operating at optical line rates. The photonic implementation therefore targets speed, stability, and parallelism rather than quantum advantage, which aligns with on-detector use cases.

Relative to the state of the art, this work focuses on establishing a controlled and reproducible evaluation of CV-QELMs for collider applications. The photonic substrate is designed to operate at fixed optical depth, ensuring that latency is determined solely by path length and detector response, consistent with Level-1 trigger constraints. The random feature map is drawn from a well-defined Gaussian ensemble, providing a stable and physically interpretable benchmark for quantum random features. To isolate the contribution of these photonic features from pure model capacity, we compare against parameter-matched MLPs trained on identical data splits and quantify accuracy as a function of training sample size. This frame-

work provides a clear, reproducible baseline for future work that will extend to systematic sweeps over substrate parameters, direct latency measurement, and hardware implementation studies.

Section 2 reviews ELMs and QELMs and sets our notation. Section 3 details the CV photonic substrate, data embedding, and readout. Section 4 presents results for top-jet tagging and Higgs event identification, including accuracy versus sample size and parameter-matched comparisons to classical baselines. Section 5 summarizes findings and outlines next steps.

2 Quantum Extreme Learning Machines

We begin with the classical extreme learning machine (ELM) formalism [28–30] before transitioning to its quantum analogue. Let $\{(\mathbf{x}_i, y_i)\}_{i=1}^N \subset \mathbb{R}^d \times \mathbb{R}$ be a training set, or more generally $y_i \in \mathbb{R}^C$ for classification into C classes. An ELM is a single-hidden-layer feedforward model whose output takes the form

$$f(\mathbf{x}) = \sum_{j=1}^L \beta_j h_j(\mathbf{a}_j, \mathbf{x}) = \boldsymbol{\beta}^\top \mathbf{h}(\mathbf{x}) , \quad (2.1)$$

where $\mathbf{h}(\mathbf{x}) = (h_1(\mathbf{a}_1, \mathbf{x}), \dots, h_L(\mathbf{a}_L, \mathbf{x}))^\top$ is a vector of nonlinearly transformed features. In the ELM paradigm, the hidden mapping parameters $\{\mathbf{a}_j\}_{j=1}^L$ are chosen randomly, e.g. from a Gaussian or uniform distribution, and held fixed. Only the output weights $\boldsymbol{\beta}$ are trained. Collecting the hidden outputs over the training set into the matrix,

$$H_{ij} = h_j(\mathbf{a}_j, \mathbf{x}_i), \quad i = 1, \dots, N, \quad j = 1, \dots, L , \quad (2.2)$$

and putting the target vector(s) in a matrix $T \in \mathbb{R}^{N \times C}$, the training reduces to solving,

$$H\boldsymbol{\beta} = T , \quad (2.3)$$

or its regularised version. A standard choice is ridge regression,

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta}} \|H\boldsymbol{\beta} - T\|_2^2 + \lambda \|\boldsymbol{\beta}\|_2^2 = (H^\top H + \lambda I)^{-1} H^\top T . \quad (2.4)$$

As only $\boldsymbol{\beta}$ is learned, ELMs avoid iterative gradient descent, backpropagation, learning-rate tuning, and convergence criteria. Under suitable assumptions on the random basis h_j , such as bounded activation functions and independent random weight draws, ELMs satisfy universal approximation theorems. They can achieve good generalisation performance for sufficiently large hidden-layer width L [28–30, 46]. Rigorous analyses have established that, with probability one, randomly generated hidden nodes span dense function spaces in L_2 and $C(X)$ norms [29, 30], and that the ELM output converges to the target function as $L \rightarrow \infty$. Moreover, kernel formulations show that ELMs implicitly implement a random feature approximation to certain reproducing-kernel Hilbert spaces, linking their generalisation bounds

to those of random-feature models [29, 46, 47]. These results formalise why, despite their conceptual simplicity, ELMs can approach the performance of deeper networks while retaining analytic and hardware efficiency.

The QELM generalises this architecture by replacing the fixed non-linear feature map $\mathbf{h}(\cdot)$ with a quantum random feature map produced by a quantum substrate or circuit. Concretely, one defines a parametric quantum encoding map

$$\mathbf{x} \mapsto \rho(\mathbf{x}) = U_{\text{enc}}(\mathbf{x}) \rho_0 U_{\text{enc}}^\dagger(\mathbf{x}) , \quad (2.5)$$

evolves under a fixed (random) quantum reservoir unitary U_{res} , and then measures a set of observables $\{O_k\}_{k=1}^M$. The predicted output is

$$f_{\boldsymbol{\eta}}(\mathbf{x}) = \sum_{k=1}^M \eta_k \text{Tr}[O_k U_{\text{res}} \rho(\mathbf{x}) U_{\text{res}}^\dagger] = \boldsymbol{\eta}^\top \mathbf{o}(\mathbf{x}) , \quad (2.6)$$

where the quantum feature vector is $\mathbf{o}(\mathbf{x}) = (\text{Tr}[O_1 \rho_{\text{res}}(\mathbf{x})], \dots, \text{Tr}[O_M \rho_{\text{res}}(\mathbf{x})])^\top$. The coefficients $\boldsymbol{\eta} \in \mathbb{R}^M$ are obtained by a linear solve akin to ridge regression:

$$\hat{\boldsymbol{\eta}} = (O^\top O + \lambda I)^{-1} O^\top T , \quad (2.7)$$

where $O_{ik} = \text{Tr}[O_k \rho_{\text{res}}(\mathbf{x}_i)]$. Thus, QELMs preserve the ‘train-only-the-output-layer’ philosophy of ELMs, but now the feature map is determined by quantum dynamics.

One principal advantage of QELMs is the ability of the quantum substrate to implicitly explore extremely large Hilbert spaces, and to embed non-trivial correlations, interference, and, in non-Gaussian or non-linear settings, entanglement structure, which may yield more expressive features than classical random projections for certain tasks. Of course, these advantages bring caveats. Recent work [35] develops a Fourier-expressivity analysis showing that the prediction $f_{\boldsymbol{\eta}}$ can be expanded in a Fourier series whose accessible frequencies are determined by the encoding scheme, while the effective number of independent coefficients is bounded by $\min(M, |\Omega|)$, where Ω is the frequency set induced by data encoding. This places an intrinsic ceiling on complexity when M , the number of observables, is small. Moreover, concentration phenomena in large quantum systems, due to randomness, noise, entanglement, or global measurements, can degrade independence of the quantum features, effectively collapsing the model to an input-agnostic constant [35, 39]. Complementary work uses Krylov-dimension or spread complexity to bound effective expressivity in QELMs and quantum reservoir computers, suggesting that beyond a certain system depth or randomness, marginal expressivity gains saturate [34].

In view of these theoretical constraints, a practical QELM must carefully balance the choice of encoding scheme, the number of observables measured, the depth and mixing of the substrate, and regularisation. In this paper we adopt a continuous variable photonic implementation, embedding classical inputs via quadrature displacements, evolving under a fixed Gaussian circuit, and performing Gaussian compatible measurement techniques to define $\{O_k\}$. We test the quantum substrate’s expressivity and robustness by comparing the CV-QELM performance against parameter-matched MLPs on collider classification tasks.

3 Continuous-variable quantum extreme learning machines

Continuous-variable quantum computing (CVQC) provides a natural substrate for quantum extreme learning machines (QELMS), as it encodes information in the continuous quadratures of quantum oscillators, so-called *qumodes*, whose Hilbert space is infinite-dimensional. Here we focus on a quantum optical platform as an experimentally viable realisation of CVQC, where information is carried by photonic modes and the quadrature operators are defined via the creation and annihilation operators,

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger), \quad \hat{p} = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}), \quad (3.1)$$

which obey the canonical commutation relation $[\hat{x}_i, \hat{p}_j] = i\delta_{ij}$ (where here we adopt natural units $\hbar = \omega = m = 1$). Gaussian operations, generated by Hamiltonians that are at most quadratic in the quadrature operators $\hat{x}$ and $\hat{p}$, act linearly on the quadrature phase space and include displacement, squeezing, and passive linear optics. Their general unitary form can be written as,

$$\hat{U}_G = \exp\left(-\frac{i}{2}\hat{\mathbf{r}}^\top G \hat{\mathbf{r}} + i\mathbf{d}^\top \hat{\mathbf{r}}\right), \quad (3.2)$$

where $\hat{\mathbf{r}} = (\hat{x}_1, \hat{p}_1, \dots, \hat{x}_M, \hat{p}_M)^\top$, G is a real symmetric matrix encoding quadratic couplings, and $\mathbf{d}$ a displacement vector. The corresponding symplectic transformation acts on first and second moments as $\mathbf{r} \rightarrow S\mathbf{r} + \mathbf{d}$ with $S \in \text{Sp}(2M, \mathbb{R})$ [20–22].

Single-mode Gaussian gates include displacement $\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$ and squeeze $\hat{S}(z) = \exp[\frac{1}{2}(z^*\hat{a}^2 - z(\hat{a}^\dagger)^2)]$ operations, with complex parameters $\alpha = re^{i\phi}$ and $z = r'e^{i\phi'}$. Two-mode Gaussian operations generate entanglement, such as the controlled-addition gate

$$\hat{C}_X(s) = \exp(-is\hat{x}_1 \otimes \hat{p}_2), \quad (3.3)$$

which can be decomposed into a series of squeeze operations and beam splitters, such that

$$\hat{C}_X(s) = \hat{B}(\frac{\pi}{2} + \theta, 0) [\hat{S}(r) \otimes \hat{S}(-r)] \hat{B}(\theta, 0), \quad (3.4)$$

where $(\sinh r = -s/2)$ and $(\cos(2\theta) = -\tanh r)$ with the beam splitter defined as $\hat{B}(\theta, \phi) = \exp[\theta(e^{i\phi}\hat{a}_1\hat{a}_2^\dagger - e^{-i\phi}\hat{a}_1^\dagger\hat{a}_2)]$. These gates, along with phase shifters and interferometers, form the building blocks of Gaussian photonic circuits that are experimentally accessible with current integrated optics technology. Non-Gaussian gates, generated by higher-than-quadratic Hamiltonians, such as cubic-phase or Kerr interactions, are required for universal quantum computation [19, 23], but are experimentally demanding due to weak optical non-linearities. In practice, measurement-induced or hybrid machine-learning-based schemes are often employed to emulate them [25–27, 48, 49], although these are not necessary for the QELMS considered here.

In a continuous-variable QELM, classical data $\mathbf{x} \in \mathbb{R}^d$ are embedded onto the optical device via displacements $\hat{D}(\alpha(\mathbf{x}))$ across M -qumodes. The state then evolves under the

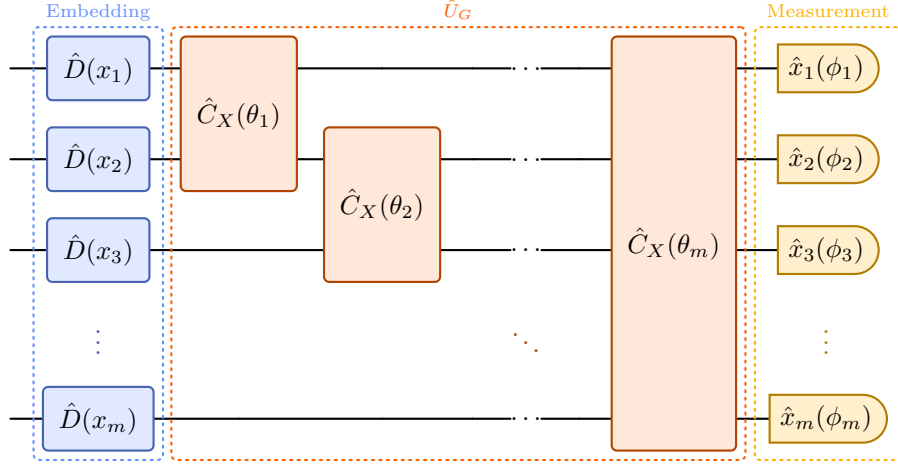


Figure 2: Schematic of the photonic substrate for the quantum extreme learning machine (QELM). Classical data are first embedded into optical modes via displacement operations before being propagated through a Gaussian circuit composed of cascading controlled-addition ($\hat{C}_X$) gates. Homodyne measurements are then performed on the output qumodes to generate the feature vector, which serves as input for downstream classical classification.

Gaussian circuit,

$$\hat{U}_G = \prod_{m=1}^M \left[\hat{C}_{X_{m, \overline{m+1}}}(\theta_m) \right] , \quad (3.5)$$

comprising a cascade of CX -gates with gate parameters θ_m which are randomly drawn from experimentally realisable ranges and then fixed. Here,

$$\overline{m+1} = (m+1) \bmod M . \quad (3.6)$$

denotes the periodicity of the gate structure that provides both mixing and entanglement. A schematic of the quantum substrate is shown in Figure 2. For a given input $\mathbf{x}$, the resulting output density operator is

$$\rho_{\text{out}}(\mathbf{x}) = \hat{U}_G \hat{D}(\alpha(\mathbf{x})) \rho_0 \hat{D}^\dagger(\alpha(\mathbf{x})) \hat{U}_G^\dagger , \quad (3.7)$$

where $\rho_0 = |0\rangle\langle 0|$ is the initial state of the photonic system where all qumodes are initialised in the vacuum state.

Finally, the output modes can be measured in two ways. The first method uses photon-number-resolving (PNR) measurements on M output modes to determine the occupation of each mode in the photon number basis, yielding discrete outcomes $\mathbf{n} \in \mathbb{N}^M$ sampled from

$$P(\mathbf{n} | \mathbf{x}) = \text{Tr}[\rho_{\text{out}}(\mathbf{x}) \Pi_{\mathbf{n}}] , \quad (3.8)$$

where,

$$\Pi_{\mathbf{n}} = \bigotimes_{m=0}^{M-1} |n_m\rangle\langle n_m| , \quad (3.9)$$

are ideal number-resolving measurements. From repeated shots of the circuit, the first and second-order photon-number moments can be obtained as,

$$\Phi(\mathbf{x}) = \text{Tr}[\rho_{\text{out}}(\mathbf{x}) \hat{\mathbf{O}}] , \quad \hat{\mathbf{O}} = (\hat{n}_1, \dots, \hat{n}_M, (\Delta\hat{n}_1)^2, \dots, (\Delta\hat{n}_M)^2)^\top , \quad (3.10)$$

where $\hat{n}_m = \hat{a}_m^\dagger \hat{a}_m$ and $(\Delta\hat{n}_m)^2 = \hat{n}_m^2 - \langle \hat{n}_m \rangle^2$. The resulting expectation values yield the mean photon numbers, $\bar{n}_m = \langle \hat{n}_m \rangle$, and variances, $\text{Var}(\hat{n}_m) = \langle \hat{n}_m^2 \rangle - \langle \hat{n}_m \rangle^2$, for each mode, which together form the classical feature vector $\Phi(\mathbf{x})$, which can then be used downstream for classification.

The second measurement method employs homodyne detection of selected quadratures $\{\hat{x}_m(\phi_m)\}_{m=0}^{M-1}$ on the output modes, yielding continuous-valued outcomes sampled from the probability distribution,

$$P(\mathbf{x}_{\text{hom}} | \mathbf{x}) = \text{Tr}[\rho_{\text{out}}(\mathbf{x}) \Pi_{\mathbf{x}_{\text{hom}}}] , \quad (3.11)$$

where,

$$\Pi_{\mathbf{x}_{\text{hom}}} = \bigotimes_{m=0}^{M-1} |x_m(\phi_m)\rangle\langle x_m(\phi_m)| , \quad (3.12)$$

denotes the operation corresponding to simultaneous quadrature measurements with phases $\{\phi_m\}^*$. From these measurements, the first and second-order quadrature moments are obtained as

$$\begin{aligned} \Phi_{\text{hom}}(\mathbf{x}) = \text{Tr}[\rho_{\text{out}}(\mathbf{x}) \hat{\mathbf{O}}_{\text{hom}}] , \quad \hat{\mathbf{O}}_{\text{hom}} = & (\hat{x}_0, \dots, \hat{x}_{M-1}, \hat{p}_0, \dots, \hat{p}_{M-1}, \\ & \hat{x}_0^2, \dots, \hat{x}_{M-1}^2, \hat{p}_0^2, \dots, \hat{p}_{M-1}^2, \\ & \frac{1}{2}\{\hat{x}_0, \hat{p}_0\}, \dots, \frac{1}{2}\{\hat{x}_{M-1}, \hat{p}_{M-1}\})^\top . \end{aligned} \quad (3.13)$$

This construction captures the local quadrature statistics for each qumode, including both mean values and second-order correlations between conjugate quadratures. The resulting expectation values define the continuous-variable feature vector $\Phi_{\text{hom}}(\mathbf{x})$, which serves as input to the downstream classification.

Compared to photon-number-resolving (PNR) detection, homodyne detection offers a highly efficient and scalable means of extracting continuous-valued features from Gaussian

*The measurement phase determines which quadrature, or which linear combination of quadratures, is measured:

$$\hat{x}(\phi) = \cos \phi \hat{x} + \sin \phi \hat{p} .$$

Note, here we consider measurements of the pure quadratures only, corresponding to $\phi_m = \{0, \frac{\pi}{2}\}$.

photonic circuits. Its high detection efficiency and access to quadrature correlations make it well suited to near-term implementations of QELMs, where the underlying states remain approximately Gaussian. PNR detection, by contrast, can capture genuinely non-Gaussian statistics at the cost of greater experimental complexity. Homodyne detection also offers an advantage in measurement speed, crucial for the collider applications discussed in this paper. Continuous quadrature measurements can be performed at high bandwidth using standard photodiodes, providing rapid acquisition of statistical moments [21, 50–52]. In contrast, PNR detection requires slower single-photon counting with limited repetition rates [53]. Moreover, for Gaussian photonic circuits, homodyne expectation values can be computed efficiently from covariance matrices, making this approach particularly attractive for high-throughput computing.

Once the data have been propagated through the quantum substrate, classification is performed using a simple classical linear classifier applied to the feature matrix $\Phi_{(\text{hom})}(\mathbf{x})$ obtained from the quantum measurements. Here, we consider two approaches: logistic regression and ridge regression. The former obtains the model parameters $\boldsymbol{\eta}$ by maximising the regularised log-likelihood,

$$\mathcal{L}(\boldsymbol{\eta}) = \sum_{i=1}^N \left[T_i \log \sigma(\boldsymbol{\eta}^\top \Phi(\mathbf{x}_i)) + (1 - T_i) \log(1 - \sigma(\boldsymbol{\eta}^\top \Phi(\mathbf{x}_i))) \right] - \frac{\lambda}{2} \|\boldsymbol{\eta}\|^2, \quad (3.14)$$

where N is the number of samples, $T_i \in \{0, 1\}$ is the target label, $\sigma(z) = 1/(1 + e^{-z})$ is the logistic sigmoid, and λ is the regularisation strength. The objective is convex in $\boldsymbol{\eta}$, ensuring a unique global optimum that can be efficiently reached via standard gradient-based solvers.

In contrast, ridge regression determines the model parameters $\boldsymbol{\eta}$ analytically, avoiding iterative gradient descent and thereby reducing the computational latency of the classification process. This is achieved by minimising the regularised least-squares objective,

$$\mathcal{L}(\boldsymbol{\eta}) = \sum_{i=1}^N \left(T'_i - \boldsymbol{\eta}^\top \Phi(\mathbf{x}_i) \right)^2 + \lambda \|\boldsymbol{\eta}\|^2, \quad (3.15)$$

where $T'_i \in \{-1, +1\}$ denotes the class label for the i -th training sample and $\lambda > 0$ once again controls the strength of the L2 regularisation. This objective defines a convex quadratic optimisation problem that admits a closed-form solution,

$$\boldsymbol{\eta} = \left(\Phi^\top \Phi + \lambda I \right)^{-1} \Phi^\top \mathbf{T}, \quad (3.16)$$

found directly via linear algebra rather than iterative gradient descent. The predicted class for a new input $\mathbf{x}$ is then given by

$$\hat{T} = \text{sign}(\boldsymbol{\eta}^\top \Phi(\mathbf{x})). \quad (3.17)$$

In both formulations, the Gaussian photonic circuit provides a fixed, random non-linear embedding of the data, while the regression models constitute the learnable output layer.

The use of Gaussian circuits for QELMs offers several advantages. They are experimentally stable, with operations that can be easily reconfigured via phase and pump control. Moreover, they operate on continuous-variable optical modes, enabling natural parallel encoding of real-valued input features. Even though purely Gaussian dynamics are classically simulable, their random projections act as efficient high-dimensional analogue feature maps that approximate smooth kernel families [31, 32, 54]. For real-time or on-detector use cases, this physical implementation emphasises determinism and latency over quantum advantage: the optical computation proceeds at the speed of light, and retraining requires simple linear regression, providing several orders of magnitude faster adaptation than iterative gradient-based neural networks. Compared to digital accelerators, the photonic substrate introduces negligible thermal load and can operate at room temperature with low optical power, fulfilling key criteria for deployment close to detectors.

Thus, our CV-QELM architecture consists of three stages: (i) input embedding by multi-mode displacement, (ii) propagation through a fixed Gaussian interferometric circuit that provides random mixing and entanglement via cascading $\hat{C}_X(s)$ -gates, and (iii) readout through Gaussian compatible detection to yield the feature vector $\Phi_{(\text{hom})}(\mathbf{x})$. This design balances expressivity and trainability while maintaining fixed optical latency, making continuous-variable photonics a practical platform for scalable quantum extreme learning at the front-end of high-energy physics data acquisition chains.

4 Results

In this section we evaluate the performance of the continuous-variable quantum extreme learning machine (CV-QELM) on representative classification tasks from collider physics. We first consider top jet tagging, a benchmark problem that distinguishes boosted hadronic top quarks from QCD background jets, followed by Higgs boson identification. These tasks span different levels of feature complexity and class separability, allowing us to assess the robustness and data efficiency of the QELM relative to compact classical baselines.

Let F denote the number of input features, H the number of nodes in the hidden layer of the MLP baseline, and M the number of optical modes in the quantum substrate[†]. As outlined in Section 3, we consider two measurement schemes. In the first, photon-number-resolving (PNR) detectors measure the observable vector $\hat{\mathbf{O}}(\mathbf{x})$ defined in Equation 3.10, yielding a feature vector $\Phi(\mathbf{x})$ of dimension $R = 2M$. In the second, homodyne measurements evaluate the observable vector $\hat{\mathbf{O}}_{\text{hom}}(\mathbf{x})$ from Equation 3.13, producing a feature vector $\Phi_{\text{hom}}(\mathbf{x})$ of dimension $R = 5M$. In both cases, $\Phi_{(\text{hom})}(\mathbf{x}) \in \mathbb{R}^R$ denotes the resulting random feature vector which is then used downstream for classification. For binary classification, we use a single output logit, so we write $C_{\text{out}} = 1$.

The CV-QELM trains only a linear readout on top of the concatenation of the fixed random features $\Phi_{(\text{hom})}(\mathbf{x})$ and the input features $\mathbf{x}$ of dimension F . With one logit, the

[†]Note, for the studies conducted in this paper, we will use $F = M$ and embed a single feature on each qumode. In general, one could use a substrate with more or fewer optical modes than F .

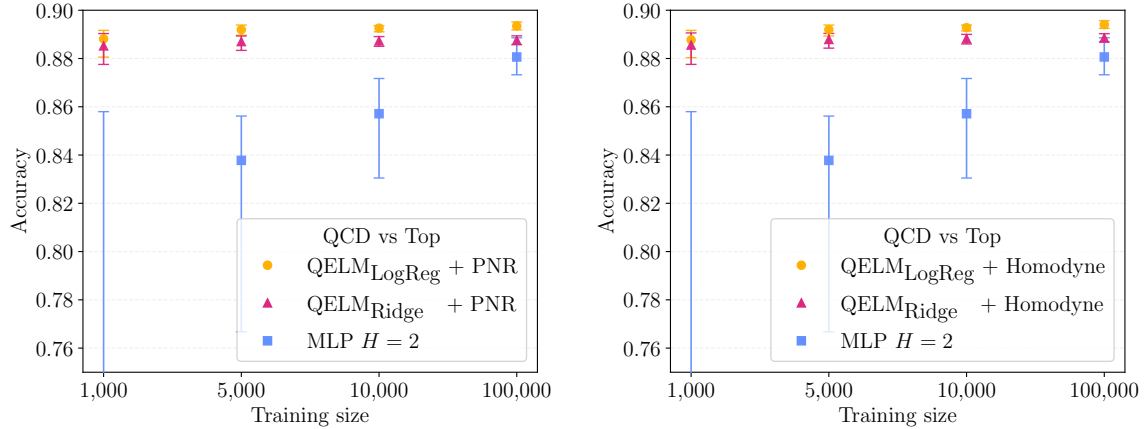


Figure 3: Scaling of test accuracy with training set size for the QCD versus Top jet classification task with $F = 16$ features [43]. Mean test accuracy is shown for the CV-QELM with (left) PNR and (right) homodyne measurements using both logistic and ridge regression readouts, compared against an MLP baseline with $H = 2$. Error bars denote the standard deviation across repeated runs. The CV-QELM consistently outperforms the MLP across all training sizes, maintaining high accuracy and low variance even in the low-statistics regime.

number of trainable parameters is therefore

$$n_{\text{train}}^{\text{QELM}} = F + R + 1 \quad (\text{weights plus bias}) . \quad (4.1)$$

The MLP baseline, trained on the classical features only, has one hidden layer of width H . With F inputs and one logit, the number of trainable parameters is

$$\begin{aligned} n_{\text{train}}^{\text{MLP}}(H) &= \underbrace{FH}_{\text{input} \rightarrow \text{hidden}} + \underbrace{H}_{\text{hidden biases}} + \underbrace{H}_{\text{hidden} \rightarrow \text{output}} + \underbrace{1}_{\text{output bias}} \\ &= H(F + 2) + 1 . \end{aligned} \quad (4.2)$$

In the following studies, H is chosen to test the CV-QELM’s performance against MLPs with comparable and larger numbers of trainable parameters.

4.1 Top jet tagging: QCD vs Top

We benchmark the CV-QELM on a standard collider classification task, distinguishing top jets from QCD background jets. The study uses the public HLS4ML LHC jet dataset [43], in which each jet is represented by $F = 16$ engineered high-level observables that encapsulating the jet’s substructure. To provide classical baselines, we train single-hidden-layer MLPs with widths $H \in \{2, 10\}$. From Equation 4.2, these values correspond respectively to parameter counts comparable to and exceeding those of the CV-QELM readout.

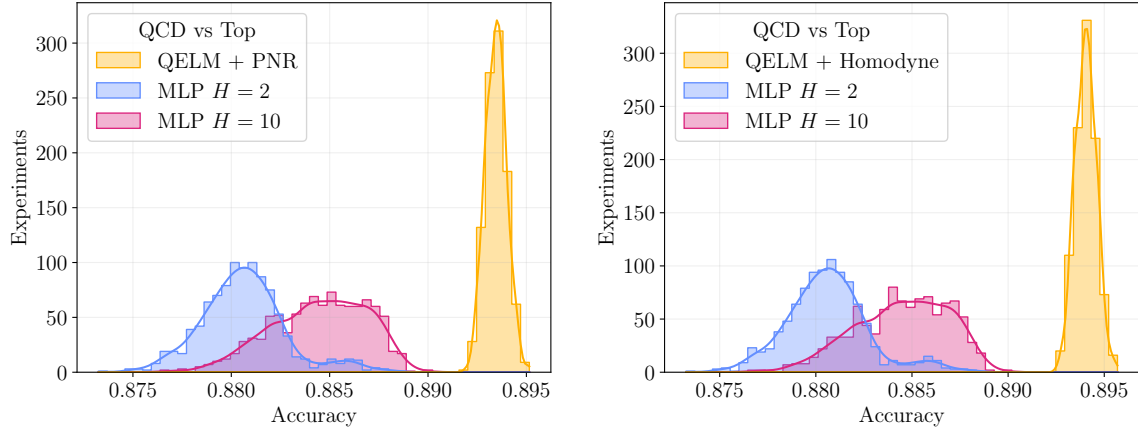


Figure 4: QCD versus Top jet classification with $F = 16$ features and $n_{\text{samples}} = 10^5$ from the HLS4ML dataset [43] using logistic regression on the CV-QELM outputs. Accuracy distributions are shown for the CV-QELM with (left) PNR and (right) homodyne measurements, compared against MLP baselines with $H = 2$ and $H = 10$. The CV-QELM achieves higher mean accuracy and substantially lower variance than both MLP baselines, demonstrating the stability and expressiveness of the Gaussian photonic random feature representation.

For the CV-QELM, as outlined in Section 3, the input features are embedded as quadrature displacements on the qumode device, such that $M = 16$. They are then propagated through a fixed randomly drawn Gaussian circuit, $\hat{U}_G$, and measured using either PNR or homodyne detection to produce $\Phi_{(\text{hom})}(\mathbf{x})$. The feature vector is then classified using either logistic regression or ridge regression. The MLP baselines are trained with cross-entropy loss, early stopping, and a small grid search over weight decay. Identical training, validation, and test splits are used across models.

To assess data efficiency, we vary the number of training examples $n_{\text{samples}} \in \{10^3, 5 \times 10^3, 10^4, 10^5\}$ and report test metrics averaged over independent initialisations. Figure 3 shows the scaling of test accuracy with training set size for both measurement schemes (PNR and homodyne) and both regression methods (logistic and ridge), alongside an MLP baseline with a single hidden layer of width $H = 2$.

Across all training sizes, the CV-QELM consistently outperforms the classical MLP baseline. The CV-QELM also exhibits remarkable stability, maintaining high and nearly constant accuracy with minimal variance across independent runs. Logistic and ridge regression yield comparable performance, with logistic regression marginally higher on average. This demonstrates that ridge regression provides competitive accuracy while retaining its analytical simplicity and low training cost. At larger training sizes, the homodyne measurement scheme slightly outperforms the PNR-based variant, indicating that homodyne detection combined with ridge regression offers an especially fast and efficient low-latency classification routine.

Figures 4 and 5 further illustrate this performance gain for $n_{\text{samples}} = 10^5$, showing

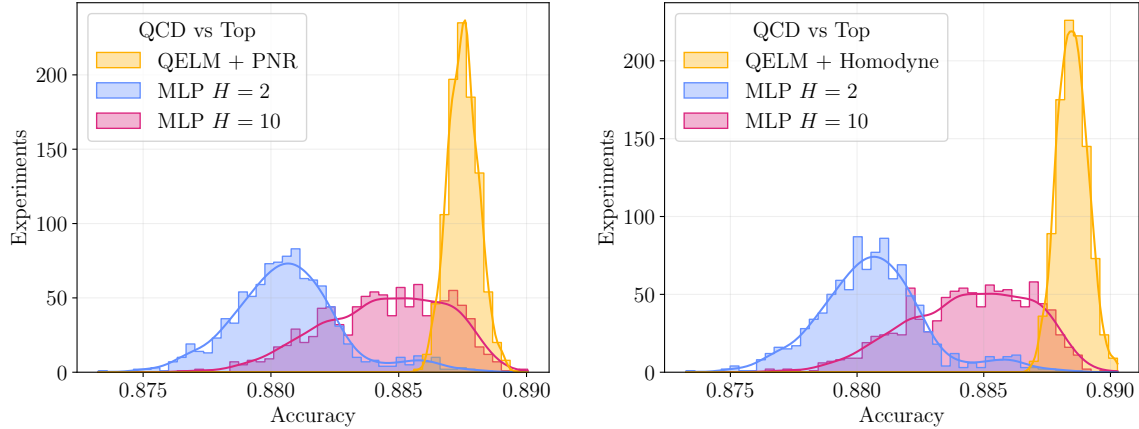


Figure 5: QCD versus Top jet classification with $F = 16$ features and $n_{\text{samples}} = 10^5$ from the HLS4ML dataset [43] using ridge regression on the CV-QELM outputs. Accuracy distributions are shown for the CV-QELM with (left) PNR and (right) homodyne measurements, compared against MLP baselines with $H = 2$ and $H = 10$. The CV-QELM achieves higher mean accuracy and markedly lower variance than both MLP baselines, highlighting the robustness and efficiency of the analytic ridge regression readout in combination with Gaussian photonic random features.

logistic and ridge regression on the CV-QELM outputs, respectively, compared with MLP baselines of $H = 2$ and $H = 10$. The CV-QELM surpasses both baselines, including the wider $H = 10$ network, which from Equation 4.2 contains 181 trainable parameters. The accuracy distributions demonstrate the robustness of the CV-QELM, which exhibits substantially smaller variance across repeated runs than the MLPs due to its fixed feature map and analytic readout fit. This combination of accuracy, predictable training cost, and fixed, nanosecond-scale inference latency supports the use of CV-QELMs as low-overhead front-end classifiers in collider data acquisition.

4.2 Higgs boson identification

We evaluate the continuous-variable QELM on the HIGGS BOSON MACHINE LEARNING CHALLENGE dataset [55], a binary classification task distinguishing simulated Higgs signal from background. We use the curated subset of $F = 10$ engineered features to enable direct comparison with MLP baselines and to mirror the setup in Sec. 4.1. Following the notation introduced above, the CV-QELM, which is constructed from $M = 10$ qumodes, produces a random feature vector $\Phi_{(\text{hom})}(\mathbf{x}) \in \mathbb{R}^R$ and trains a single linear readout with $n_{\text{train}}^{\text{QELM}} = F + R + 1$, while a single-hidden-layer MLP with width H has $n_{\text{train}}^{\text{MLP}}(H) = H(F + 2) + 1$. We again consider $H \in \{2, 10\}$ to provide a parameter-matched small baseline and a moderately larger one. Training, validation, and test splits are identical across models.

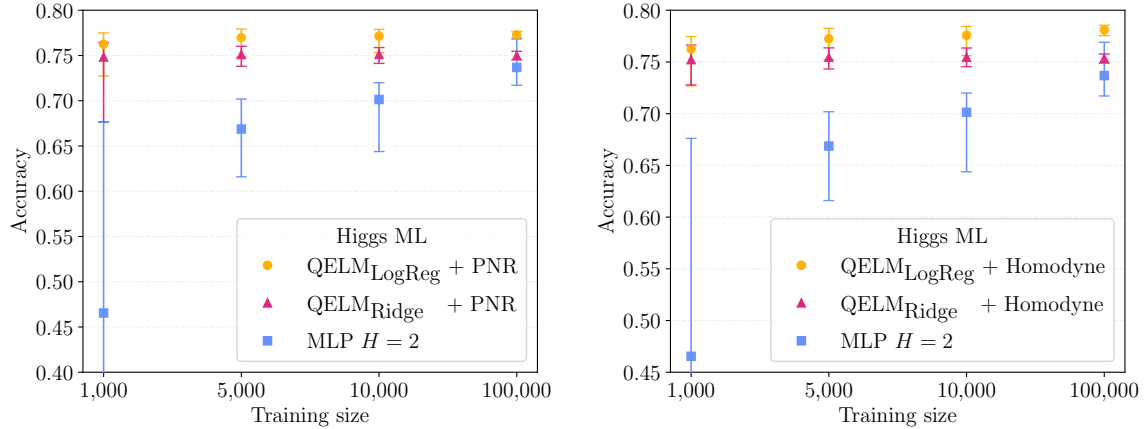


Figure 6: Scaling with training set size on the $F = 10$ Higgs task [55]. Mean test accuracy versus n_{samples} for the CV QELM with (left) PNR and (right) Homodyne measurements and the MLP with $H = 2$. Shaded bands show the standard deviation across initialisations. The QELM maintains a consistent advantage from 10^3 to 10^5 examples, mirroring the trends observed in Sec. 4.1.

The readout from the quantum substrate is trained using either logistic regression or ridge regression, while the MLPs are trained with cross-entropy loss, early stopping, and a small grid search over weight decay. To assess data efficiency, we vary the number of training examples $n_{\text{samples}} \in \{10^3, 5 \times 10^3, 10^4, 10^5\}$ and quantify performance by reporting the mean test accuracy averaged over independent runs. Figure 6 shows the scaling of test accuracy with training set size for both measurement schemes (PNR and homodyne) and both regression methods (logistic and ridge), alongside an MLP baseline with a single hidden layer of width $H = 2$. The CV-QELM consistently outperforms the parameter-matched MLP across all training sizes, with the margin most pronounced in the low-data regime, reflecting the stability of the analytic readout fit and the absence of stochastic optimisation variance. Absolute accuracies are lower than in the QCD versus Top task, consistent with the weaker class separability of the Higgs dataset; however, the relative advantage of the QELM persists. Training time for the QELM remains dominated by a single linear solve that scales predictably with n_{samples} , whereas the MLP requires iterative optimisation and hyperparameter tuning.

To further probe the QELM framework, we compare its performance against an MLP with a wider hidden layer of width $H = 10$, which from Equation 4.2 has 121 trainable parameters. Figures 7 and 8 show results for a training size of $n_{\text{samples}} = 10^5$ using logistic and ridge regression on the CV-QELM outputs, respectively, compared against MLP baselines with $H = 2$ and $H = 10$. In the logistic case, the CV-QELM achieves comparable performance to the $H = 10$ MLP under PNR detection and slightly surpasses it under homodyne detection. For ridge regression, both measurement schemes yield accuracies marginally below the $H = 10$ MLP but consistently above the $H = 2$ baseline. In all cases, the CV-QELM exhibits markedly

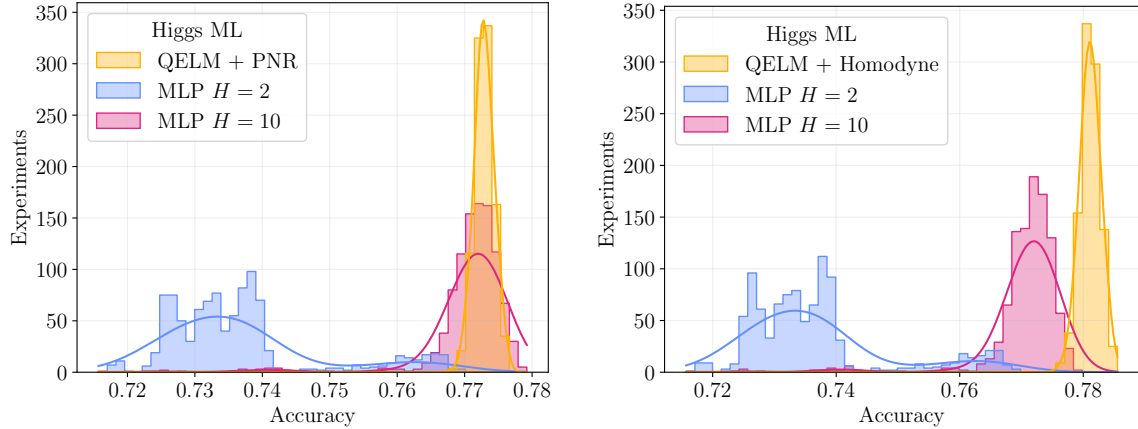


Figure 7: Higgs versus background classification with $F = 10$ features and $n_{\text{samples}} = 10^5$ of the HIGGS ML dataset [55] using logistic regression on the CV-QELM outputs. Accuracy distributions are shown for the CV-QELM with (left) PNR and (right) homodyne measurements, compared against MLP baselines with $H = 2$ and $H = 10$. The CV-QELM matches the wider $H = 10$ MLP under PNR detection and slightly surpasses it with homodyne detection, while maintaining substantially lower variance across runs.

smaller variance across runs, reflecting the reproducibility of the analytic readout training and the stability of the photonic random feature map.

5 Summary and Conclusions

We investigated continuous-variable quantum extreme learning machines (CV-QELMs) as fast, low-overhead front-ends for collider data processing. The architecture employs displacement based input embedding across M optical modes, a fixed-time Gaussian photonic quantum substrate, $\hat{U}_G$, that mixes and entangles the modes, and photon-number-resolving (PNR) or homodyne readout to generate a random feature vector, $\Phi_{(\text{hom})}(\mathbf{x})$, of dimension R . Only a linear readout layer is trained via a single regularised solve, making retraining extremely fast. In this work, we considered both logistic regression and ridge regression as the linear models used to train the output of the quantum substrate. The optical path and detector response set both the analytic and inference latency. The design emphasises predictable, non-stochastic behaviour, straightforward reconfigurability, and low-power operation, all of which are relevant for future Level-1 trigger integration or on-detector data reduction.

We established a parameter-matched evaluation protocol against classical baselines. For a multilayer perceptron (MLP) with one hidden layer of width H and F input features, the number of trainable parameters is $H(F + 2) + 1$. For the CV-QELM, only the linear readout parameters, $R + 1$, are trained. Using identical data splits and averaging over independent

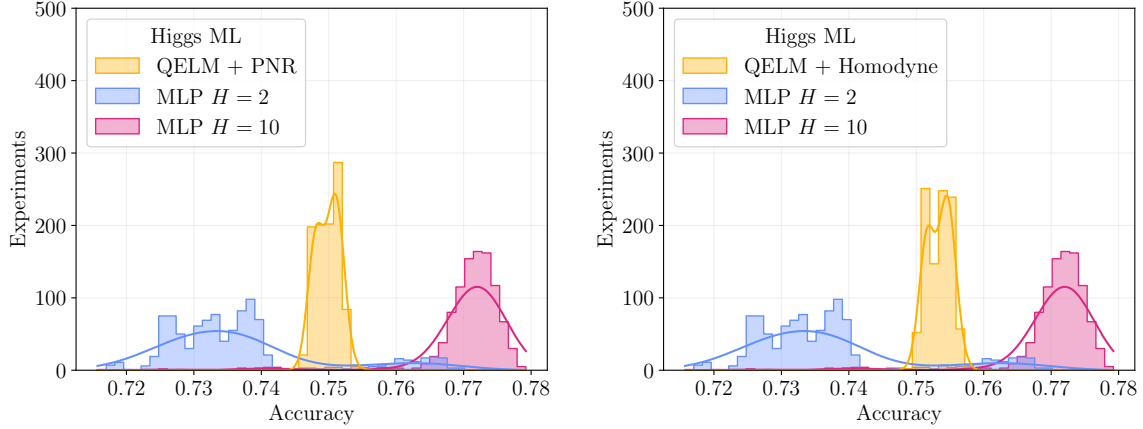


Figure 8: Higgs versus background classification with $F = 10$ features and $n_{\text{samples}} = 10^5$ of the HIGGS ML dataset [55] using ridge regression on the CV-QELM outputs. Accuracy distributions are shown for the CV-QELM with (left) PNR and (right) homodyne measurements, compared against MLP baselines with $H = 2$ and $H = 10$. Both QELM variants outperform the parameter-matched $H = 2$ MLP but fall slightly short of the $H = 10$ model, while maintaining far lower run-to-run variance due to the deterministic analytic readout.

runs, we benchmarked two collider-relevant tasks: top jet tagging and Higgs boson identification.

For the HLS4ML LHC jet dataset with $F = 16$ engineered features, the CV-QELM consistently outperformed the $H = 2$ MLP across all training sizes, with the largest margin in the low-statistics regime. At $n_{\text{samples}} = 10^5$, the CV-QELM also surpasses the wider $H = 10$ MLP, demonstrating that the Gaussian photonic random feature map provides a compact and expressive representation. For the HIGGS BOSON MACHINE LEARNING CHALLENGE dataset with $F = 10$ features, the CV-QELM again exceeded the $H = 2$ baseline across all sample sizes. With logistic regression, it achieved accuracy comparable to the $H = 10$ MLP under PNR detection and slightly higher under homodyne detection, while ridge regression yielded results marginally below the $H = 10$ model but above the $H = 2$ baseline. In all cases, the CV-QELM displayed significantly lower variance across runs, reflecting the reproducibility of the analytic readout training and the stability of the photonic feature map.

Beyond accuracy, we quantified the practical advantages that follow from analytic regression based training. Since only a linear readout is fitted through a single regularised regression, training scales predictably with sample size and avoids optimisation variance, learning rate tuning, or early stopping heuristics. Inference latency is fixed and extremely low, on the order of nanoseconds for homodyne detection, and independent of event complexity within the designed optical depth. Combined with the low optical power and room-temperature operation of integrated photonics, these properties support deployment close to the detector.

In summary, we demonstrated that a fixed Gaussian photonic quantum substrate cou-

pled to an analytic readout can deliver competitive or superior classification performance to classical networks with similar or larger numbers of trainable parameters, while offering deterministic, low-latency inference and negligible training overhead. These results establish CV-QELMs as a promising building block for real-time pattern recognition in future collider experiments. A natural next step is hardware co-design and on-bench validation of latency, power, and noise characteristics, including systematic sweeps over substrate depth, squeezing strength, and readout to extract design scaling laws. The approach can also be extended to multi-class tagging, calibration-sensitive regression, and robustness studies under domain shift.

References

- [1] P. Duckett, G. Facini, M. Jastrzebski, S. Malik, T. Scanlon, and S. Rettie, “Reconstructing charged particle track segments with a quantum-enhanced support vector machine,” *Phys. Rev. D* **109** (Mar, 2024) 052002. <https://link.aps.org/doi/10.1103/PhysRevD.109.052002>.
- [2] C. Brown, M. Spannowsky, A. Tapper, S. Williams, and I. Xiotidis, “Quantum pathways for charged track finding in high-energy collisions,” *Front. Artif. Intell.* **7** (2024) 1339785, [[arXiv:2311.00766 \[hep-ph\]](https://arxiv.org/abs/2311.00766)].
- [3] V. S. Ngairangbam, M. Spannowsky, and M. Takeuchi, “Anomaly detection in high-energy physics using a quantum autoencoder,” *Phys. Rev. D* **105** (May, 2022) 095004. <https://link.aps.org/doi/10.1103/PhysRevD.105.095004>.
- [4] A. Bal, M. Klute, B. Maier, M. Oughton, E. Pezone, and M. Spannowsky, “One particle - one qubit: Particle physics data encoding for quantum machine learning,” *Phys. Rev. D* **112** (Oct, 2025) 076004. <https://link.aps.org/doi/10.1103/18y2-87vq>.
- [5] K. Bepari, S. Malik, M. Spannowsky, and S. Williams, “Towards a quantum computing algorithm for helicity amplitudes and parton showers,” *Phys. Rev. D* **103** no. 7, (2021) 076020, [[arXiv:2010.00046 \[hep-ph\]](https://arxiv.org/abs/2010.00046)].
- [6] K. Bepari, S. Malik, M. Spannowsky, and S. Williams, “Quantum walk approach to simulating parton showers,” *Phys. Rev. D* **106** no. 5, (2022) 056002, [[arXiv:2109.13975 \[hep-ph\]](https://arxiv.org/abs/2109.13975)].
- [7] G. Gustafson, S. Prestel, M. Spannowsky, and S. Williams, “Collider events on a quantum computer,” *JHEP* **11** (2022) 035, [[arXiv:2207.10694 \[hep-ph\]](https://arxiv.org/abs/2207.10694)].
- [8] S. Ramírez-Urbe, A. E. Rentería-Olivo, G. Rodrigo, G. F. R. Sborlini, and L. Vale Silva, “Quantum algorithm for Feynman loop integrals,” *JHEP* **05** (2022) 100, [[arXiv:2105.08703 \[hep-ph\]](https://arxiv.org/abs/2105.08703)].
- [9] H. A. Chawdhry, M. Pellen, and S. Williams, “Quantum simulation of scattering amplitudes and interferences in perturbative QCD,” [[arXiv:2507.07194 \[hep-ph\]](https://arxiv.org/abs/2507.07194)].
- [10] E. Bashore, S. Moretti, and T. Vitos, “A quantum algorithm for the n-gluon MHV scattering amplitude,” [[arXiv:2507.14252 \[hep-ph\]](https://arxiv.org/abs/2507.14252)].
- [11] S. P. Jordan, K. S. M. Lee, and J. Preskill, “Quantum Algorithms for Quantum Field Theories,” *Science* **336** (2012) 1130–1133, [[arXiv:1111.3633 \[quant-ph\]](https://arxiv.org/abs/1111.3633)].

- [12] M. C. Bañuls *et al.*, “Simulating Lattice Gauge Theories within Quantum Technologies,” *Eur. Phys. J. D* **74** no. 8, (2020) 165, [arXiv:1911.00003 [quant-ph]].
- [13] C. W. Bauer, Z. Davoudi, A. B. Balantekin, T. Bhattacharya, M. Carena, W. A. de Jong, P. Draper, A. El-Khadra, N. Gemelke, M. Hanada, D. Kharzeev, H. Lamm, Y.-Y. Li, J. Liu, M. Lukin, Y. Meurice, C. Monroe, B. Nachman, G. Pagano, J. Preskill, E. Rinaldi, A. Roggero, D. I. Santiago, M. J. Savage, I. Siddiqi, G. Siopsis, D. Van Zanten, N. Wiebe, Y. Yamauchi, K. Yeter-Aydeniz, and S. Zorzetti, “Quantum simulation for high-energy physics,” *PRX Quantum* **4** (May, 2023) 027001. <https://link.aps.org/doi/10.1103/PRXQuantum.4.027001>.
- [14] J. Ingoldby, M. Spannowsky, T. Sypchenko, and S. Williams, “Enhancing quantum field theory simulations on NISQ devices with Hamiltonian truncation,” *Phys. Rev. D* **110** no. 9, (2024) 096016, [arXiv:2407.19022 [quant-ph]].
- [15] A. Di Meglio, K. Jansen, I. Tavernelli, C. Alexandrou, S. Arunachalam, C. W. Bauer, K. Borras, S. Carrazza, A. Crippa, V. Croft, R. de Putter, A. Delgado, V. Dunjko, D. J. Egger, E. Fernández-Combarro, E. Fuchs, L. Funcke, D. González-Cuadra, M. Grossi, J. C. Halimeh, Z. Holmes, S. Kühn, D. Lacroix, R. Lewis, D. Lucchesi, M. L. Martinez, F. Meloni, A. Mezzacapo, S. Montangero, L. Nagano, V. R. Pascuzzi, V. Radescu, E. R. Ortega, A. Roggero, J. Schuhmacher, J. Seixas, P. Silvi, P. Spentzouris, F. Tacchino, K. Temme, K. Terashi, J. Tura, C. Tüysüz, S. Vallecorsa, U.-J. Wiese, S. Yoo, and J. Zhang, “Quantum computing for high-energy physics: State of the art and challenges,” *PRX Quantum* **5** (Aug, 2024) 037001. <https://link.aps.org/doi/10.1103/PRXQuantum.5.037001>.
- [16] S. Williams, “Simulating quantum field theories on near-term quantum devices,” *Int. J. Mod. Phys. A* **40** no. 26, (2025) 2530010.
- [17] J. Ingoldby, M. Spannowsky, T. Sypchenko, S. Williams, and M. Wingate, “Real-Time Scattering on Quantum Computers via Hamiltonian Truncation,” [arXiv:2505.03878 [quant-ph]].
- [18] J. C. Halimeh, N. Mueller, J. Knolle, Z. Papić, and Z. Davoudi, “Quantum simulation of out-of-equilibrium dynamics in gauge theories,” [arXiv:2509.03586 [quant-ph]].
- [19] S. Lloyd and S. L. Braunstein, “Quantum computation over continuous variables,” *Phys. Rev. Lett.* **82** (Feb, 1999) 1784–1787. <https://link.aps.org/doi/10.1103/PhysRevLett.82.1784>.
- [20] S. L. Braunstein and P. van Loock, “Quantum information with continuous variables,” *Rev. Mod. Phys.* **77** (Jun, 2005) 513–577. <https://link.aps.org/doi/10.1103/RevModPhys.77.513>.
- [21] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, “Gaussian quantum information,” *Rev. Mod. Phys.* **84** no. 2, (2012) 621–669, [arXiv:1110.3234 [quant-ph]].
- [22] G. Adesso, S. Ragy, and A. R. Lee, “Continuous variable quantum information: Gaussian states and beyond,” *Open Systems & Information Dynamics* **21** no. 01n02, (2014) 1440001, [arXiv:1401.4679].
- [23] N. C. Menicucci, P. van Loock, M. Gu, C. Weedbrook, T. C. Ralph, and M. A. Nielsen, “Universal quantum computation with continuous-variable cluster states,” *Phys. Rev. Lett.* **97** no. 11, (2006) 110501, [arXiv:quant-ph/0605198].

- [24] A. Blance and M. Spannowsky, “Unsupervised event classification with graphs on classical and photonic quantum computers,” *JHEP* **21** (2020) 170, [[arXiv:2103.03897 \[hep-ph\]](#)].
- [25] S. Abel, M. Spannowsky, and S. Williams, “Simulating quantum field theories on continuous-variable quantum computers,” *Phys. Rev. A* **110** no. 1, (2024) 012607, [[arXiv:2403.10619 \[quant-ph\]](#)].
- [26] S. Abel, M. Spannowsky, and S. Williams, “Real-time scattering processes with continuous-variable quantum computers,” *Phys. Rev. A* **112** no. 1, (2025) 012614, [[arXiv:2502.01767 \[quant-ph\]](#)].
- [27] S. Abel, M. Spannowsky, and S. Williams, “Qumode Tensor Networks for False Vacuum Decay in Quantum Field Theory,” [[arXiv:2506.17388 \[quant-ph\]](#)].
- [28] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” *Neurocomputing* **70** no. 1-3, (2006) 489–501.
- [29] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, “Extreme learning machine for regression and multiclass classification,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* **42** no. 2, (2012) 513–529.
- [30] T. Guo, G.-B. Huang, and Y. Chen, “A comprehensive theoretical framework for the random feature mapping in extreme learning machines,” *Neurocomputing* **168** (2015) 35–48.
- [31] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, “Deep learning with coherent nanophotonic circuits,” *Nature Photonics* **11** (2017) 441–446, [[arXiv:1610.02365](#)].
- [32] M. Henderson, V. Srikumar, and S. Kais, “Quantum reservoir computing with continuous variables,” *Phys. Rev. Lett.* **125** no. 26, (2020) 260503, [[arXiv:2008.06653](#)].
- [33] K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, “Experimental demonstration of reservoir computing on a silicon photonics chip,” *Nature Communications* **5** (2014) 3541.
- [34] S. Čindrak, L. Jaurigue, and K. Lüdge, “Krylov expressivity in quantum reservoir computing and quantum extreme learning,” [[arXiv:2409.12079](#)].
- [35] W. Xiong, G. Facelli, M. Sahebi, O. Agnel, T. Chotibut, S. Thanasilp, and Z. Holmes, “On fundamental aspects of quantum extreme learning machines,” *Quantum Machine Intelligence* **7** no. 1, (2025) 20, [[arXiv:2312.15124](#)].
- [36] M. Schuld and N. Killoran, “Quantum Machine Learning in Feature Hilbert Spaces,” *Phys. Rev. Lett.* **122** (Feb, 2019) 040504. <https://link.aps.org/doi/10.1103/PhysRevLett.122.040504>.
- [37] 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** (2019) 209–212, [[arXiv:1804.11326 \[quant-ph\]](#)].
- [38] P. Mujal, R. Martínez-Peña, J. Nokkala, J. García-Bení, G. L. Giorgi, M. C. Soriano, and R. Zambrini, “Opportunities in quantum reservoir computing and extreme learning machines,” *Advanced Quantum Technologies* **4** no. 8, (2021) 2100027, [[arXiv:2102.11831](#)].
- [39] L. Innocenti, S. Lorenzo, I. Palmisano, A. Ferraro, M. Paternostro, and G. M. Palma,

- “Potential and limitations of quantum extreme learning machines,” *Communications Physics* **6** no. 1, (2023) 118, [arXiv:2210.00780].
- [40] M. Vetrano, G. Lo Monaco, L. Innocenti, S. Lorenzo, and G. M. Palma, “State estimation with quantum extreme learning machines beyond the scrambling time,” *npj Quantum Information* **11** no. 1, (2025) 20, [arXiv:2409.06782].
- [41] D. Zia, L. Innocenti, G. Minati, S. Lorenzo, A. Suprano, R. Di Bartolo, N. Spagnolo, T. Giordani, V. Cimini, G. M. Palma, A. Ferraro, F. Sciarrino, and M. Paternostro, “Quantum extreme learning machines for photonic entanglement witnessing,” [arXiv:2502.18361 [quant-ph]].
- [42] H. Assil, A. El Allati, and G. L. Giorgi, “Entanglement estimation of werner states with a quantum extreme learning machine,” *Phys. Rev. A* **111** (Feb, 2025) 022412. <https://link.aps.org/doi/10.1103/PhysRevA.111.022412>.
- [43] M. Pierini, J. M. Duarte, N. Tran, and M. Freytsis, “HLS4ML LHC Jet dataset (150 particles),” Jan., 2020. <https://doi.org/10.5281/zenodo.3602260>.
- [44] J. Duarte, S. Han, P. Harris, S. Jindariani, E. Kreinar, B. Kreis, J. Ngadiuba, M. Pierini, R. Rivera, N. Tran, and Z. Wu, “Fast inference of deep neural networks in FPGAs for particle physics,” *JINST* **13** no. 07, (2018) P07027, [arXiv:1804.06913].
- [45] T. Aarrestad *et al.*, “Fast convolutional neural networks on FPGAs with hls4ml,” *Machine Learning: Science and Technology* **2** no. 4, (2021) 045015, [arXiv:2101.05108].
- [46] S. Scardapane and D. Wang, “Randomness in neural networks: An overview,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **7** no. 2, (2017) e1200.
- [47] A. Rahimi and B. Recht, “Random features for large-scale kernel machines,” in *Advances in Neural Information Processing Systems*, vol. 20, pp. 1177–1184. 2008. [arXiv:0709.4483 [cs.LG]].
- [48] K. K. Sabapathy, H. Qi, J. Izaac, and C. Weedbrook, “Production of photonic universal quantum gates enhanced by machine learning,” *Phys. Rev. A* **100** (Jul, 2019) 012326. <https://link.aps.org/doi/10.1103/PhysRevA.100.012326>.
- [49] D. Su, C. R. Myers, and K. K. Sabapathy, “Conversion of gaussian states to non-gaussian states using photon-number-resolving detectors,” *Phys. Rev. A* **100** (Nov, 2019) 052301. <https://link.aps.org/doi/10.1103/PhysRevA.100.052301>.
- [50] S. L. Braunstein, “Homodyne statistics,” *Phys. Rev. A* **42** (Jul, 1990) 474–481. <https://link.aps.org/doi/10.1103/PhysRevA.42.474>.
- [51] S. L. Braunstein and D. D. Crouch, “Fundamental limits to observations of squeezing via balanced homodyne detection,” *Phys. Rev. A* **43** (Jan, 1991) 330–337. <https://link.aps.org/doi/10.1103/PhysRevA.43.330>.
- [52] D. Huang, D. Lin, C. Wang, W. Liu, S. Fang, J. Peng, P. Huang, and G. Zeng, “Continuous-variable quantum key distribution with 1 mbps secure key rate,” *Opt. Express* **23** no. 13, (Jun, 2015) 17511–17519. <https://opg.optica.org/oe/abstract.cfm?URI=oe-23-13-17511>.
- [53] R. H. Hadfield, J. Leach, F. Fleming, D. J. Paul, C. H. Tan, J. S. Ng, R. K. Henderson, and G. S. Buller, “Single-photon detection for long-range imaging and sensing,” *Optica* **10** no. 9,

(Sep, 2023) 1124–1141.

<https://opg.optica.org/optica/abstract.cfm?URI=optica-10-9-1124>.

- [54] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, “Continuous-variable quantum neural networks,” *Phys. Rev. Research* **1** no. 3, (2019) 033063, [[arXiv:1906.07682](https://arxiv.org/abs/1906.07682)].
- [55] C. Adam-Bourdarios, G. Cowan, C. Germain, I. Guyon, B. Kégl, and D. Rousseau, “The Higgs boson machine learning challenge,” in *Proceedings of the NIPS 2014 Workshop on High-energy Physics and Machine Learning*, G. Cowan, C. Germain, I. Guyon, B. Kégl, and D. Rousseau, eds., vol. 42 of *Proceedings of Machine Learning Research*, pp. 19–55. PMLR, Montreal, Canada, 13 dec, 2015. <https://proceedings.mlr.press/v42/cowa14.html>.