Astronomy and Computing 57 (2026) 101168

Contents lists available at ScienceDirect

Astronomy and Computing

journal homepage: www.elsevier.com/locate/ascom

Full length paper

Brain-inspired quantum machine learning architecture and its benchmarking on astronomical classification

Subhash S. Pandey [a], Mousam Mondal [b], Xing Liang [c], Christos Kalyvas [a], Dimitrios Makris [c], Hongwei Wu [a,*]

[a] School of Physics, Engineering and Computer Science, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB, UK
[b] Centre for Astrophysics Research, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB, UK
[c] School of Computer Science and Mathematics (CSM), Kingston University London, UK

ARTICLE INFO

Dataset link: https://www.sdss4.org/dr14/, https://www.sdss4.org/dr17/

Keywords:
Brain-inspired learning
Quantum machine learning
Spiking neural networks
Variational quantum circuits
Astronomical classification
Sloan digital sky survey
Cross-domain generalization

ABSTRACT

Automated classification of stars, galaxies, and quasi-stellar objects (QSOs) is essential for large-scale astronomical surveys, where reliable classifiers must remain stable across survey releases, class imbalance, and incomplete observational inputs. In this work, we present a Brain-Inspired Quantum Machine Learning (BIQML) framework for SDSS star–galaxy–QSO classification as a step toward robust source classification for next-generation photometric pipelines. The model combines a spiking Leaky Integrate-and-Fire (LIF) controller with a classically simulated variational quantum circuit (VQC), in which the controller sequentially selects quantum gates using measurement feedback; all circuits are classically simulated, and no hardware-level quantum advantage is claimed. The framework is trained and internally tested on SDSS DR17, with SDSS DR14 used as an external cross-release test set, and is benchmarked against a broad suite of baselines including Random Forest, Extra Trees, HistGradientBoosting, Logistic Regression, Linear Support Vector Machine, and Multilayer Perceptron classifiers under a repeated ten-split protocol with paired statistical significance testing. In the full-feature setting, BIQML is statistically on par with the strongest classical baselines, reaching $97.36 \pm 0.05\%$ accuracy (96.90% balanced) on SDSS DR17 and $98.76 \pm 0.09\%$ accuracy (98.41% balanced) on the external SDSS DR14 test set, where the leading tree ensembles span 97.3–97.8% and 98.8–99.4%, respectively. To probe robustness when key observational channels are unavailable, we introduce a low-sensor (g, r, i) no-redshift setting in which spectroscopic redshift, the u- and z-bands, and metadata are removed. Under this reduced-feature condition, BIQML remains among the top-performing models on a single controlled split (0.805 balanced accuracy, versus 0.64–0.80 for the classical baselines), achieving 85.58% accuracy on SDSS DR17 and 93.44% accuracy on SDSS DR14, suggesting that it remains competitive, degrading no more than the strongest ensembles, when discriminative information is scarce. A component-level ablation isolates the contributions of the VQC, the spiking controller, and the measurement-feedback loop, while both an analytic noise screening and a finite-shot, IBM-calibrated device-noise benchmark show that BIQML predictions remain stable under representative depolarizing, amplitude-damping, readout, and realistic NISQ-style noise channels. Together, these results position BIQML as a competitive and robust quantum-inspired architecture for astronomical source classification under cross-release shift and limited observational information.

1. Introduction

Modern ground-based and space-based observatories have produced astronomical datasets of unprecedented scale, making automated source classification an essential component of observational astronomy. In particular, reliable astronomical object classification, such as the separation of stars, galaxies, and quasi-stellar objects (QSOs), is a foundational task for large spectroscopic surveys including the Sloan Digital Sky Survey (SDSS) (Abolfathi et al., 2018; Abdurro'uf et al., 2022). Over the past decade, numerous studies demonstrated that this problem could be effectively addressed using supervised, unsupervised, and ensemble-based machine-learning approaches applied to SDSS data (Logan and Fotopoulou, 2020; Cunha and Humphrey, 2022; Wen et al., 2023; Whiting et al., 2020; Peruzzi et al., 2021; Wang et al., 2023; Portillo et al., 2020; Fraix-Burnet et al., 2021; Chatterjee and Ghosh,

* Corresponding author.
E-mail addresses: subhashshankarpandeyy@gmail.com (S.S. Pandey), mousamphysics137@gmail.com (M. Mondal), X.Liang@kingston.ac.uk (X. Liang), c.kalyvas@herts.ac.uk (C. Kalyvas), D.Makris@kingston.ac.uk (D. Makris), h.wu6@herts.ac.uk (H. Wu).

https://doi.org/10.1016/j.ascom.2026.101168
Received 7 February 2026; Accepted 12 July 2026
Available online 18 July 2026

2025). While these methods achieved high accuracy when trained and tested on data drawn from the same survey release, their robustness across different survey versions, sky coverage, and data distributions remained an active area of investigation. A further and less frequently examined challenge is robustness to *incomplete* observational inputs, where spectroscopic redshift or specific photometric bands may be unavailable, expensive to obtain, or deliberately excluded in fast photometric pipelines. In this work, we address these robustness challenges by introducing a Brain-Inspired Quantum Machine Learning (BIQML) framework for star–galaxy–QSO classification and benchmarking it against a broad suite of classical tabular baselines under repeated-split and external cross-release evaluation protocols.

Brain-Inspired Machine Learning (BIML) draws motivation from biological neural systems, which process information through sparse, event-driven, and temporally adaptive dynamics rather than continuous activations (Wang et al., 2026; Irfan et al., 2021; Singh et al., 2025; Atashgahi et al., 2022). BIML models commonly employ spiking neuron formulations (Kasabov, 2010), such as the Leaky Integrate-and-Fire (LIF) neuron, enabling temporal integration (Ai et al., 2025), stateful processing, and sensitivity to feedback. These properties allow BIML systems to learn efficiently from limited data and to adapt rapidly under uncertainty. In the context of astronomical classification, such dynamics are advantageous for handling overlapping class boundaries and evolving survey characteristics. However, BIML models alone remain constrained by classical representational limits and do not explicitly exploit non-classical feature interactions (Schmidgall et al., 2024).

Quantum Machine Learning (QML) leverages quantum–mechanical principles to construct expressive models based on parameterized quantum circuits, often referred to as Variational Quantum Circuits (VQCs) (Martín-Guerrero and Lamata, 2022; Lu et al., 2024). In QML, classical data are encoded into quantum states and transformed through trainable unitary operations, enabling highly non-linear feature mappings via superposition and interference. Despite their representational power, practical QML models typically rely on fixed circuit ansätze and hybrid quantum–classical optimization, which can suffer from barren plateaus, sensitivity to initialization, and inefficient exploration of the circuit design space. These limitations restrict their robustness and scalability for real-world astronomical datasets (Bhavsar et al., 2023; Rahmaniar et al., 2024). Hybrid quantum–classical models have also been applied to astrophysical classification; for example, Rauf et al. (2026) proposed a hybrid quantum–classical convolutional neural network that amplitude-encodes raw astronomical *images* into a fixed, randomized quantum circuit used as a front-end feature extractor for a separately trained CNN. BIQML differs in three respects: it operates on tabular photometric and spectroscopic features rather than images; it treats the circuit structure as a learned, sequentially constructed object with measurement feedback rather than a fixed extractor decoupled from the classifier; and it targets robustness under cross-release shift and reduced observational information rather than in-distribution accuracy. Given these differences in data modality and evaluation objective, a direct numerical comparison is not meaningful, and we position BIQML relative to Rauf et al. (2026) on architectural grounds.

BIQML integrates the adaptive, feedback-driven dynamics of BIML with the expressive capacity of QML (Andrés et al., 2024; Nguyen et al., 2024; Duan et al., 2025). In this framework, spiking neural controllers sequentially construct and adapt a quantum circuit through measurement-driven feedback, forming a closed-loop hybrid learning system. Rather than optimizing a static quantum ansatz, BIQML treats quantum circuit design as a decision-making process guided by brain-inspired principles such as sparse activation, temporal accumulation, and iterative refinement. This combination is hypothesized to reduce the effective search complexity of quantum circuit construction while promoting robust, transferable representations, a hypothesis we examine empirically rather than assume. We state our scope explicitly. In this study, all quantum circuits are executed using classical state-vector simulation; we therefore do not claim, and our results do not require, any hardware-level quantum speed-up or quantum advantage. The variational quantum circuit is used as a structured, quantum-inspired nonlinear feature transformation with measurement feedback, and any performance reported here is attributable to this architecture under classical simulation rather than to quantum hardware execution. Accordingly, BIQML is presented as a quantum-inspired solver architecture for robust astronomical source classification, particularly for future large-scale photometric survey pipelines where reliable decisions may be required under cross-release shifts and incomplete observational features.

The central question of this work is whether brain-inspired, measurement-driven circuit construction yields a classifier that remains reliable both across survey releases and under reduced observational information. To address this, we make the following contributions:

- **Comprehensive benchmarking.** We evaluate BIQML against strong classical tabular baselines—Random Forest, Extra Trees, HistGradientBoosting, Logistic Regression, Linear SVM, and a Multilayer Perceptron—under a repeated multi-split protocol with paired statistical significance testing and external DR17→DR14 cross-release evaluation.
- **Low-sensor robustness stress test.** We introduce a reduced low-sensor g, r, i / no-redshift setting in which spectroscopic redshift, the u- and z-bands, and metadata are removed, isolating performance under limited observational information.
- **Component-level ablation.** We decompose BIQML to quantify the individual contributions of the VQC, the spiking controller, and the measurement-feedback loop, distinguishing genuine architectural benefit from raw model capacity.
- **Quantum-noise screening.** We assess the stability of BIQML predictions under representative simulated noise channels (depolarizing, amplitude-damping, and readout noise), characterizing reliability under non-ideal circuit execution.

Together, this structure allows the proposed architecture to be assessed not only in terms of accuracy, but also in terms of robustness, component contribution, and reliability under observational constraints.

The remainder of this paper is organized as follows. Section 2 describes the data and its preparation, Section 3 presents the Brain-Inspired Quantum Machine Learning architecture, Section 4 reports the results and discussion, and Section 5 concludes the paper.

2. Data description and preparation

The SDSS is a long-running astronomical survey that has mapped large regions of the sky using both imaging and spectroscopy with a dedicated 2.5-m optical telescope at Apache Point Observatory in New Mexico, USA. SDSS provides high-quality photometric and spectroscopic measurements for millions of objects, making it a key resource for studies of stars, galaxies, and quasars. The datasets used in this work were obtained from SDSS DR14[1] (Abolfathi et al., 2018; Rahman, 2018) and SDSS DR17[2] (Abdurro'uf et al., 2022; Ghosh, 2024). Each object in both datasets was described by a common set of photometric and spectroscopic measurements and a class label identifying it as a star, galaxy, or QSO. From the raw catalog columns, comprising the five photometric magnitudes (u, g, r, i, z), spectroscopic redshift, and associated observational metadata, a set of derived colour and summary features was constructed, yielding the full feature representation used by the models. The features were constructed by joining the SDSS photometric and spectroscopic catalogs, ensuring physically reliable class labels derived from spectroscopy.

SDSS DR17 (Abdurro'uf et al., 2022; Ghosh, 2024) originally contained approximately 100,000 objects whose classes (star, galaxy, or

[1] https://www.sdss4.org/dr14/
[2] https://www.sdss4.org/dr17/

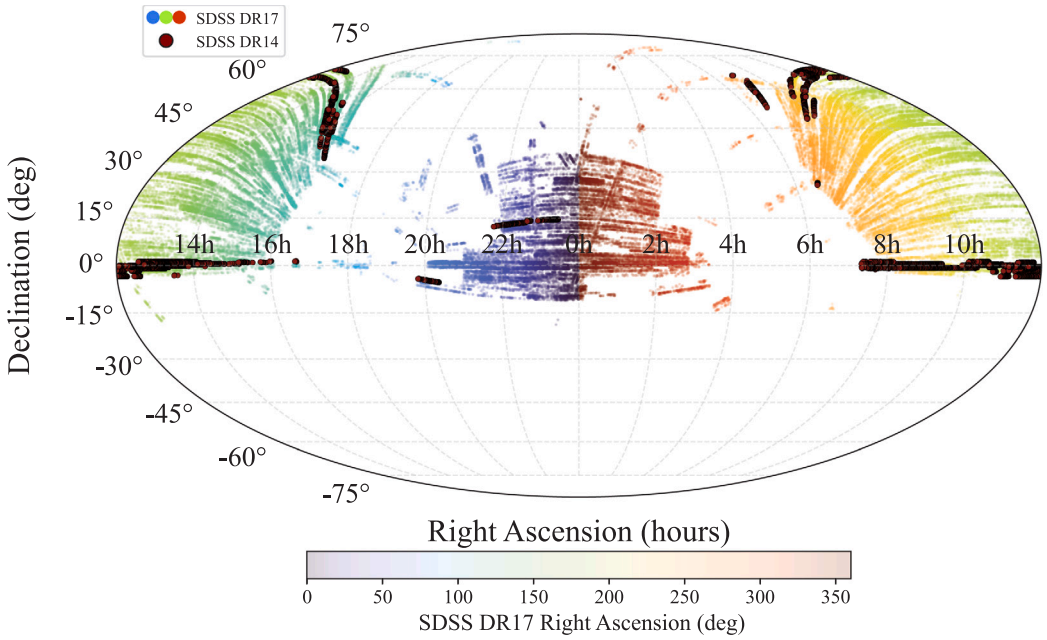


Fig. 1. Mollweide projection of the sky distribution of SDSS spectroscopic sources in equatorial coordinates. The horizontal axis shows right ascension (hours) and the vertical axis shows declination (degrees). The SDSS DR17 training sample (99,872 sources) is displayed using a continuous colour scale representing right ascension, illustrating the sky coverage of the survey. The SDSS DR14 validation subset (10,000 sources) is overplotted in red. The two samples occupy distinct regions on the sky with no positional overlap, ensuring spatial independence for cross-release evaluation.

quasar) were determined from their observed spectra, while the SDSS DR14 (Abolfathi et al., 2018; Rahman, 2018) subset used in this work contained approximately 10,000 objects. The two datasets were cross-matched using sky coordinates, specifically right ascension (RA) and declination (DEC), which identified 128 common sources. These overlapping objects were removed from the SDSS DR17 sample to prevent any potential data leakage. After this filtering, a total of 99,872 SDSS DR17 sources were retained and used for model training, while the full SDSS DR14 subset of 10,000 objects was reserved exclusively for external testing.

SDSS DR17 covered a significantly broader sky area ($0° \leq \text{RA} \leq 360°$, $-18.8° \leq \text{DEC} \leq 83.0°$) compared to the SDSS DR14 test subset ($8.2° \leq \text{RA} \leq 260.9°$, $-5.4° \leq \text{DEC} \leq 68.5°$). This difference provided an external cross-release evaluation setting for assessing model generalization across a wider and previously unseen sky footprint. The spatial distribution of the 99,872 SDSS DR17 training sources is shown in Fig. 1 using a colour scale based on right ascension, while the 10,000 SDSS DR14 validation sources are overplotted in red. The two samples occupied distinct regions on the sky, with no positional overlap within the adopted matching tolerance, confirming that the DR14 subset used in this work was spatially independent of the DR17 sample.

The primary observational features included the five SDSS photometric magnitudes (u, g, r, i, z), spanning ultraviolet to near-infrared wavelengths and encoding essential colour information for classification, together with spectroscopic redshift. The class distributions for both SDSS DR17 (training set) and SDSS DR14 (test set) were shown in Fig. 2.

In addition to this full-feature setting, we also considered a reduced low-sensor g, r, i no-redshift setting to evaluate model behaviour when key observational channels are unavailable. In this setting, spectroscopic redshift, the u-band, the z-band, and metadata columns were removed. Only the g, r, and i photometric bands were retained, together with derived colour and compact photometric summary features: $g - r$, $r - i$, $g - i$, $\overline{m}_{gri}$, σ_{gri}, and s_{gri}, where $\overline{m}_{gri}$, σ_{gri}, and s_{gri} denote the mean magnitude, magnitude standard deviation, and colour slope across the retained bands. This reduced-feature setting was designed to emulate incomplete or low-sensor observational conditions, where spectroscopic redshift or specific photometric channels may be unavailable, expensive, corrupted, or deliberately excluded.

Redshift was a key astrophysical parameter that encoded distance and cosmic expansion effects, and it served as a physically meaningful discriminator among different spectroscopic classes. As shown in Fig. 3, stars were predominantly nearby objects and therefore populated the lowest redshift regime. This behaviour was evident in the blue curves, where the dashed line corresponded to the SDSS DR17 sample and the solid line to the SDSS DR14 sample, both concentrated at $\lesssim 10^{-3}$. Galaxies, shown in green, occupied intermediate redshift ranges, reflecting their extragalactic but relatively local nature. QSOs, depicted in purple, extended to the highest redshifts, consistent with their origin as distant, luminous active galactic nuclei. The clear separation in redshift distributions across classes highlighted the strong physical relevance of redshift as a discriminative feature in SDSS-based classification.

The DR17 and DR14 samples also differed in class composition, with DR17 being more galaxy-dominated and DR14 containing a larger relative fraction of stars. Therefore, overall accuracy alone may not fully capture cross-release performance. For this reason, the experiments reported balanced accuracy, macro-averaged F1 score, class-wise precision and recall, and one-vs-rest AUC in addition to overall accuracy. This allowed the evaluation to distinguish genuine cross-release generalization from changes caused by different class priors.

The DR17–DR14 comparison was therefore treated as an explicit cross-release distribution-shift problem rather than only as an external test split. The distinct sky coverage in Fig. 1, the class-prior differences in Fig. 2, and the redshift-distribution differences in Fig. 3 provide the physical and statistical context for this shift. In the extended analysis, feature-wise distribution differences were quantified using the Kolmogorov–Smirnov statistic and Wasserstein distance.

3. Brain-inspired quantum machine learning architecture

3.1. Motivation and design philosophy

Biological neural systems are able to solve complex, high-dimensional tasks efficiently, even with limited training exposure.

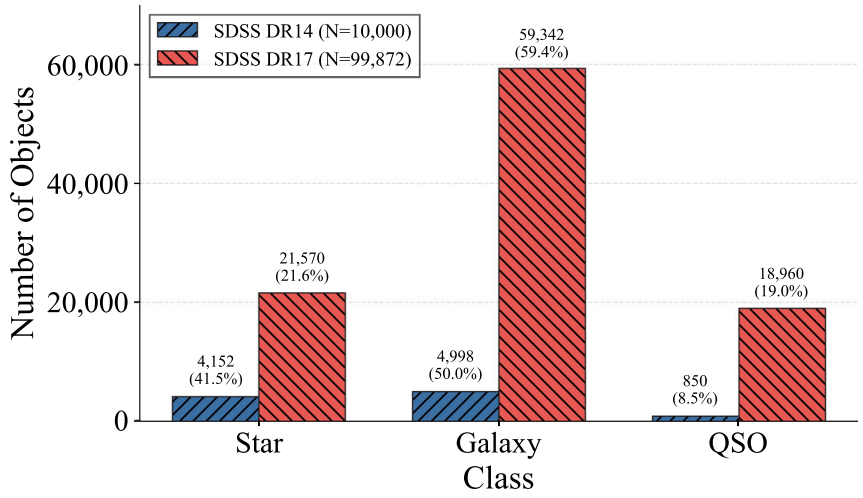


Fig. 2. Distribution of stars, galaxies, and quasars in the SDSS DR17 training set (Abdurro'uf et al., 2022; Ghosh, 2024) and the SDSS DR14 test set (Abolfathi et al., 2018; Rahman, 2018). The different class proportions motivate the use of balanced accuracy, macro-averaged metrics, and class-wise performance reporting in addition to overall accuracy.

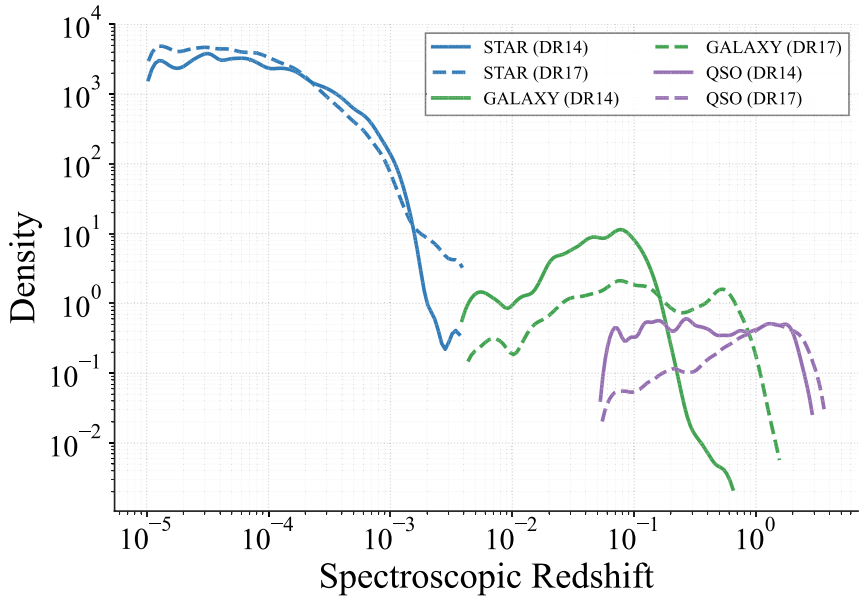


Fig. 3. Redshift distribution of stars, galaxies, and quasars in the SDSS DR17 training set (Abdurro'uf et al., 2022; Ghosh, 2024) and the SDSS DR14 test set (Abolfathi et al., 2018; Rahman, 2018). Redshift provides a strong discriminative feature in the full-feature setting and motivates the additional low-sensor no-redshift experiment.

Here, this biological analogy is used only as an architectural motivation rather than as a claim of direct biological equivalence. This ability is commonly attributed to three key principles: (i) recurrent, stateful processing, which allows information to be accumulated and refined over time; (ii) sparse, event-driven computation, which focuses learning signals on a small subset of active neurons; and (iii) closed-loop feedback, which enables continuous adaptation based on outcomes. Together, these mechanisms support rapid learning and robust generalization under uncertain conditions.

In contrast, variational quantum algorithms typically rely on fixed circuit ansätze, whose optimization is often impeded by highly non-convex loss landscapes, sensitivity to initialization, and the presence of barren plateaus. The discrete and combinatorial nature of quantum circuit design further complicates gradient-based optimization, particularly when the gate structure itself must be learned. The BIQML architecture is therefore motivated by the hypothesis that incorporating brain-inspired decision-making mechanisms into quantum circuit construction can reduce the effective search complexity and accelerate convergence toward task-relevant quantum representations. In the present work, this hypothesis is evaluated empirically through full-feature benchmarking, reduced-feature stress testing, and component-level ablation rather than being presented as a proof of quantum advantage.

3.2. Architecture overview and data flow

BIQML is formulated as a closed-loop hybrid system comprising a brain-inspired feature compressor, a spiking neural controller, and a variational quantum circuit (VQC) coupled through measurement feedback, as illustrated in Fig. 4. Rather than optimizing a static quantum circuit, the model learns to *sequentially construct* the circuit through a series of discrete gate-selection decisions informed by intermediate quantum measurements. In this implementation, the VQC is executed using a classical state-vector simulator (PennyLane

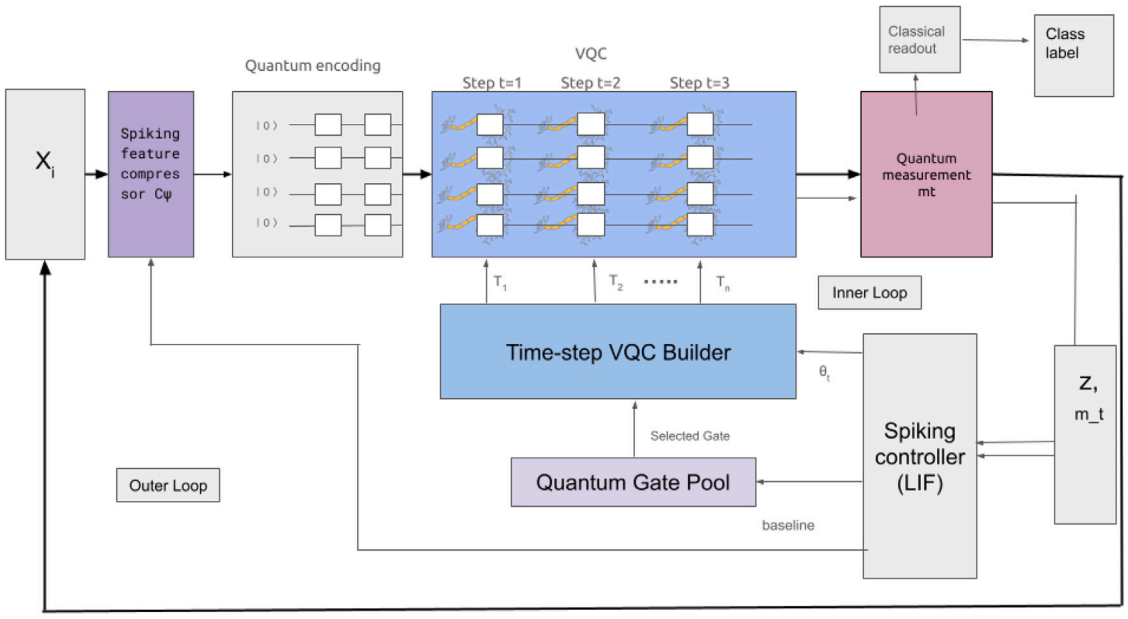


Fig. 4. Schematic of the BIQML architecture. Input features x_i are first mapped by a brain-inspired spiking feature compressor C_ψ into a qubit-aligned representation and encoded into a classically simulated variational quantum circuit (VQC). At each decision step $t = 1, \ldots, T$, the spiking LIF controller, conditioned on the compressed features and previous measurement $[z, m_{t-1}]$, selects one gate a_t per qubit from a finite quantum gate pool ("Selected Gate") and generates the corresponding rotation parameters θ_t, which the time-step VQC builder appends as the next circuit layer (inner loop). The resulting circuit is measured, and the intermediate Pauli-Z expectation values m_t are fed back to the controller to inform subsequent decisions (outer loop). The controller additionally produces a value "baseline" used for variance reduction during training. After the final step, the terminal measurement vector is passed to a classical readout layer for star–galaxy–QSO classification.

`lightning.qubit`); the quantum component is therefore treated as a quantum-inspired feature transformation and measurement-feedback mechanism rather than as evidence of hardware-level quantum advantage. This closed-loop interaction enables the controller to adaptively refine the circuit structure based on the evolving quantum measurement state.

Let $x \in \mathbb{R}^D$ denote a classical input sample. A brain-inspired compressor C_ψ (Section 3.3) maps x into a qubit-aligned angle vector $z \in [-\pi, \pi]^{n_q}$, where n_q is the number of qubits. In this work, $n_q = 6$ qubits are used. The initial quantum state is then prepared according to

$$|\psi_0(x)\rangle = \mathcal{E}(z)|0\rangle^{\otimes n_q}, \tag{1}$$

where $\mathcal{E}$ denotes a parameterized data-encoding operator, implemented as a single-rotation R_Y angle embedding, $\mathcal{E}(z) = \bigotimes_{q=1}^{n_q} R_Y(z_q)$. Eq. (1) defines the starting point of the sequential circuit-construction process depicted in Fig. 4. The controller then unrolls the VQC over $T = 5$ decision steps, selecting one gate per qubit at each step from a finite gate pool of size $K = 8$.

3.3. Brain-inspired feature compressor

Before any quantum operation, the classical input $x \in \mathbb{R}^D$ is mapped to the qubit-aligned angle vector $z \in [-\pi, \pi]^{n_q}$ by a brain-inspired compressor C_ψ that itself uses spiking dynamics. The input is first projected to a hidden current and gated,

$$c = \big(\mathrm{GELU}\,\mathrm{LN}\,W_{\mathrm{in}}x\big) \odot \sigma\big(\mathrm{LN}\,W_g x\big), \tag{2}$$

where LN denotes layer normalization, σ a sigmoid context gate, and $\odot$ elementwise multiplication. This current drives a recurrent LIF layer over S internal micro-steps with a learnable per-step temporal bias, producing spike and membrane traces $\{s^{(j)}, v^{(j)}\}_{j=1}^{S}$. A temporal-attention weighting $\beta_j = \mathrm{softmax}_j(W_a v^{(j)})$ aggregates the traces into a pooled spike code and membrane code, which are concatenated and passed through a bottleneck head with a tanh output,

$$z = \pi \tanh\Big(\mathrm{Head}\big[\textstyle\sum_j \beta_j s^{(j)};\ \sum_j \beta_j v^{(j)}\big]\Big) \in [-\pi, \pi]^{n_q}. \tag{3}$$

Algorithm 1 BIQML: brain-inspired compressor, spiking controller with probabilistic gate selection, and measurement feedback

Input: sample $x \in \mathbb{R}^D$, qubits $n_q = 6$, horizon $T = 5$, gate-pool size $K = 8$
Gate pool: $\mathcal{G} = \{R_X, R_Y, R_Z, \mathrm{PhaseShift}, H, S, T, I\}$
Output: class logits $\ell(x)$
Compress x to a qubit-aligned angle vector $z = C_\psi(x) \in [-\pi, \pi]^{n_q}$
Prepare the encoded state $|\psi_0(x)\rangle = \mathcal{E}(z)|0\rangle^{\otimes n_q}$ using R_Y embedding
Initialize feedback $m_0 \leftarrow \mathbf{0}$, membrane state $v_0 \leftarrow \mathbf{0}$, partial circuit $U_{1:0} \leftarrow I$
Initialize empty action list $\mathcal{A}$ and angle list Θ
for $t = 1$ to T **do**
  $(\ell_t, \theta_t, b_t, \hat{m}_t, v_t) \leftarrow \mathrm{LIFController}(z, m_{t-1}, v_{t-1})$
  for $q = 1$ to n_q **do**
    Sample gate action $a_{t,q} \sim \mathrm{Cat}\big(\mathrm{softmax}(\ell_{t,q})\big)$
    If the identity action is selected, apply the no-operation gate I
  end for
  Append a_t to $\mathcal{A}$ and θ_t to Θ
  Apply selected single-qubit gates, then the entangling block $U_t(a_t, \theta_t)$
  Update the partial circuit $U_{1:t} = U_t U_{1:t-1}$
  Evolve the encoded state: $|\psi_t\rangle \leftarrow U_{1:t}|\psi_0(x)\rangle$
  Measure local Pauli-Z expectations $m_t = \big(\langle Z_1\rangle_t, \ldots, \langle Z_{n_q}\rangle_t\big)$
  Feed m_t back to the controller for the next step
end for
$\ell(x) = W_o m_T + b_o$
return $\ell(x)$

During training, C_ψ is regularized by a spike-rate target term and a latent-decorrelation term that discourages redundancy among the n_q output channels. The compressor hidden width is 256, with $S = 5$ internal steps, leak coefficient $\alpha = 0.90$, and spike-rate target 0.10. The spiking compressor is retained in all component-level ablation variants; its independent contribution relative to a non-spiking feed-forward compressor is not isolated in the present study.

3.4. Spiking controller and gate-selection policy

The core of the BIQML framework is a brain-inspired spiking controller that operates as a recurrent decision-making system. The controller maintains an internal membrane state v_t and generates binary spike events s_t according to leaky integrate-and-fire (LIF) dynamics (Ahmadvand et al., 2025; Sepulchre, 2022):

$$v_{t+1} = \alpha v_t + W u_t - \vartheta s_t, \tag{4}$$

$$s_t = H(v_t - \vartheta), \tag{5}$$

where $v_t \in \mathbb{R}^{n_h}$ is the membrane potential at step t with n_h hidden units, $\alpha \in (0, 1)$ is the leak coefficient, W is a learned input weight matrix, and ϑ is the firing threshold. The Heaviside function $H(\cdot)$ produces binary spikes $s_t \in \{0, 1\}^{n_h}$, and the term $-\vartheta s_t$ implements a soft reset preventing unbounded membrane accumulation. During backpropagation, the non-differentiable Heaviside function is replaced by a sigmoid-based surrogate gradient, $\partial s_t/\partial v_t \approx \beta\, \sigma(\beta(v_t - \vartheta))\big(1 - \sigma(\beta(v_t - \vartheta))\big)$ with $\beta = 10$. The controller hidden dimension is $n_h = 256$.

The controller input is

$$u_t = [z, m_{t-1}], \tag{6}$$

where $z \in \mathbb{R}^{n_q}$ is the compressed input and $m_{t-1} \in [-1, 1]^{n_q}$ is the measurement vector from the previous step. The initial feedback vector is $m_0 = \mathbf{0}$, so each circuit-construction decision is conditioned on the input representation and the most recent quantum measurement, never on unavailable future measurements.

Based on the spike activity s_t, the controller produces four outputs: (i) logits $\ell_t \in \mathbb{R}^{n_q \times K}$ defining a categorical gate distribution per qubit; (ii) rotation angles $\theta_t \in \mathbb{R}^{n_q}$; (iii) a value estimate b_t used as a variance-reduction baseline; and (iv) a predicted measurement vector $\hat{m}_t$ used as an auxiliary predictive-feedback signal. Here $n_q = 6$ and $K = 8$, with the gate pool

$$\mathcal{G} = \{R_X, R_Y, R_Z, \mathrm{PhaseShift}, H, S, T, I\}, \tag{7}$$

where S and T are realized as fixed phase shifts of $\pi/2$ and $\pi/4$, respectively, and the identity gate I provides an explicit no-operation pathway. When the controller selects an inactive or spike-suppressed action for a qubit, the corresponding operation is the identity transformation rather than a forced rotation; this defines the spike-suppression edge case in the gate-selection process.

The policy over circuit actions is

$$\pi_\phi(a_t \mid z, m_{t-1}) = \prod_{q=1}^{n_q} \mathrm{Cat}\left(a_{t,q}; \mathrm{softmax}(\ell_{t,q})\right), \tag{8}$$

where $a_{t,q}$ selects the gate type for qubit q at step t and ϕ denotes the controller parameters. Gate selection is performed independently per qubit while remaining conditioned on the shared recurrent state and the previous measurement feedback. Across the $T = 5$ steps, the controller generates a temporally ordered action sequence $\mathcal{A} = \{a_1, \ldots, a_T\}$ and angle sequence $\Theta = \{\theta_1, \ldots, \theta_T\}$ specifying the VQC evaluated by the simulator. A temperature applied to the gate logits is annealed from 1.1 to 0.85 over training to sharpen the policy as learning proceeds.

3.5. Quantum circuit construction and measurement feedback

At each step t, the controller samples discrete gate actions $a_t = (a_{t,1}, \ldots, a_{t,n_q})$ and continuous rotation parameters $\theta_t = (\theta_{t,1}, \ldots, \theta_{t,n_q})$, selecting one operation per qubit from the gate pool $\mathcal{G}$. The circuit layer at step t is

$$U_t(a_t, \theta_t) = \mathcal{E}_{\mathrm{ent}}^{(t)} \left(\prod_{q=1}^{n_q} g_{a_{t,q}}(\theta_{t,q}) \right), \tag{9}$$

where $g_{a_{t,q}}(\theta_{t,q})$ is the controller-selected single-qubit operation on qubit q, and $\mathcal{E}_{\mathrm{ent}}^{(t)}$ is the entangling block. The adaptive part of the construction is the controller-driven selection of single-qubit gates and their rotation angles. The entangling block has a fixed topology with parameterized couplings: it applies a CNOT ring, an alternating layer of CZ gates whose pattern depends on the step parity $t \bmod 2$, and a ring of weak controlled-R_Y rotations $\mathrm{CRY}\left(\frac{1}{2}\theta_{t,q}\right)$ driven by the controller angles. Two-qubit gates are therefore not selected as independent discrete actions; instead, a fixed connectivity pattern is modulated by continuous, controller-generated coupling strengths.

Each decision step incrementally extends the circuit by appending a controller-generated layer, so that BIQML constructs the circuit as a temporally ordered sequence of $T = 5$ decisions rather than a single fixed ansatz. The same input sample thus influences the circuit repeatedly through the interaction between the compressor, the spiking controller, the selected gate sequence, and the measurement-feedback signal. After T steps the complete circuit is

$$U_{1:T} = U_T U_{T-1} \cdots U_1, \tag{10}$$

and the state at step t is

$$|\psi_t\rangle = U_{1:t} |\psi_0(x)\rangle. \tag{11}$$

All circuit evaluations use a classical state-vector simulator; the VQC is therefore a quantum-inspired nonlinear feature transformation and feedback source, not a demonstration of hardware-level quantum speed-up.

Measurement feedback is computed from local Pauli-Z expectation values,

$$m_t = \Big(\langle Z_1\rangle_t, \ldots, \langle Z_{n_q}\rangle_t\Big), \qquad \langle Z_q\rangle_t = \langle\psi_t | Z_q | \psi_t\rangle, \tag{12}$$

yielding $m_t \in [-1, 1]^{n_q}$, a compact state-dependent summary fed back to the controller at the next step. The controller input at step $t + 1$ thus depends on both the compressed astronomical features and the measurement response of the circuit built up to step t. This closed-loop construction is the main architectural distinction of BIQML relative to fixed-ansatz VQC models.

3.6. Final readout and prediction

After the final step T, the terminal measurement vector m_T is mapped to class logits by a classical readout layer,

$$\ell(x) = W_o m_T + b_o, \tag{13}$$

followed by a softmax,

$$\hat{y} = \mathrm{softmax}(\ell(x)), \tag{14}$$

with trainable parameters W_o, b_o. The VQC is not optimized as an independent standalone circuit; its sampled gate sequence is generated by the spiking controller and evaluated through the simulator, while the compressor, controller, and readout are trained jointly through the supervised and policy-gradient objective. The complete data flow is summarized in Fig. 5.

Table 1
BIQML hyperparameters. The same configuration is used in both the full-feature and low-sensor settings; the two differ only in the input feature set.

Group	Hyperparameter	Value
Quantum circuit	Qubits n_q	6
	Decision horizon T	5
	Gate-pool size K	8
	Data encoding	R_Y angle embedding
	Entangling block	CNOT ring + alt. CZ + weak CRY
	Simulator	PennyLane `lightning.qubit`
Compressor	Type	LIF spiking
	Hidden width	256
	Internal steps S	5
	Leak α	0.90
	Spike-rate target	0.10
	Decorrelation weight	0.01
	Dropout	0.05
Controller	Hidden units n_h	256
	Surrogate slope β	10
	Gate temperature (anneal)	$1.1 \rightarrow 0.85$
Optimization	Optimizer	AdamW
	Base learning rate	7×10^{-4}
	Weight decay	5×10^{-5}
	Scheduler	ReduceLROnPlateau (max val. acc., patience 3)
	Batch size	128
	Epochs	60
	Readout/controller/angle/value lr mult.	2.5/0.8/1.0/0.5
Loss weights	λ_{pg} (policy)	0.40
	λ_{v} (value)	0.20
	λ_{pc} (pred. consistency)	0.08
	λ_{s} (spike)	2×10^{-3}
	λ_{e} (energy)	2×10^{-5}
	λ_{c} (compressor)	0.02

3.7. Training objective and optimization

Training optimizes supervised classification performance together with the efficiency and stability of circuit construction. Because the gate choices a_t are discrete, the gate-selection component is optimized with a policy-gradient objective. The overall loss is

$$\mathcal{L} = \mathcal{L}_{\text{CE}} + \lambda_{\text{pg}}\mathcal{L}_{\text{pg}} + \lambda_{\text{v}}\mathcal{L}_{\text{v}} + \lambda_{\text{pc}}\mathcal{L}_{\text{pc}} + \lambda_{\text{s}}\mathcal{R}_{\text{spike}} + \lambda_{\text{e}}\mathcal{R}_{\text{energy}}, \tag{15}$$

where $\mathcal{L}_{\text{CE}}$ is the cross-entropy classification loss, $\mathcal{L}_{\text{pg}}$ the policy-gradient loss for the discrete gate policy, $\mathcal{L}_{\text{v}}$ the value-baseline loss, and $\mathcal{L}_{\text{pc}}$ a predictive-consistency loss penalizing disagreement between the controller-predicted $\hat{m}_t$ and the simulator-observed m_t. The terms $\mathcal{R}_{\text{spike}}$ and $\mathcal{R}_{\text{energy}}$ penalize excessive spiking activity and membrane-state energy. An additional compressor regularization term (spike-rate and latent-decorrelation, Section 3.3) is included with weight λ_{c}.

The model is trained with the AdamW optimizer using component-specific learning rates (the readout, controller, angle head, and value head use multipliers of 2.5, 0.8, 1.0, and 0.5 of the base rate, respectively) and a `ReduceLROnPlateau` schedule that monitors validation accuracy. All hyperparameters are identical across the full-feature and low-sensor settings and are listed in Table 1; the two settings differ only in the input feature set (18 versus 9 features, with redshift, u-, and z-bands removed in the low-sensor case). The discrete circuit structure is learned through the controller policy, while the VQC measurements provide simulator-based feedback used by the readout and auxiliary losses, keeping the procedure reproducible and compatible with classical simulation. The complete data flow, from feature compression through the closed-loop circuit construction to the final classical readout, is summarized in Fig. 5.

4. Results and discussion

4.1. Experimental protocol and baselines

All models were evaluated under a unified, controlled protocol. SDSS DR17 (99,872 sources, with the 128 DR14-overlapping objects removed) was used for training and internal testing, and the full SDSS DR14 subset (10,000 sources) served as a held-out external cross-release test set. To obtain stable estimates and enable statistical comparison, the full pipeline was repeated over ten random training/validation/test splits (seeds 42, 52, 62, 72, 82, 92, 102, 112, 122, 132); for every split, all models were trained and evaluated on identical data partitions and identical preprocessing. We report mean ± standard deviation across the ten splits for accuracy, balanced accuracy, macro-F_1, and macro one-vs-rest AUC.

BIQML was benchmarked against six classical baselines spanning linear models, ensembles, and neural networks: Logistic Regression, a Linear Support Vector Machine, Random Forest, Extra Trees, HistGradientBoosting, and a Multilayer Perceptron (MLP). Because the DR17 and DR14 class priors differ (Fig. 2), we report balanced accuracy, macro-averaged F1, and one-vs-rest (OvR) AUC alongside overall accuracy, so that performance is not dominated by the majority class. Statistical significance of the BIQML–baseline differences was assessed using paired per-split t-tests and Wilcoxon signed-rank tests across the ten seeds.

Three complementary experiments are reported: (i) a full-feature benchmark using the complete photometric and spectroscopic representation; (ii) a low-sensor g, r, i no-redshift stress test that removes redshift, the u- and z-bands, and metadata; and (iii) a component-level ablation that isolates the contributions of the VQC, the spiking controller, and the measurement-feedback loop. A simulated quantum-noise screening and a quantitative distribution-shift analysis complete the evaluation.

4.2. Full-feature benchmark

Table 2 reports the full-feature results on the DR17 internal test set and the DR14 external set, averaged over ten repeated splits and reported as mean ± standard deviation. For BIQML, the internal DR17 test performance was 97.36 ± 0.05 percent accuracy, 96.90 ± 0.08 percent balanced accuracy, 0.9697 ± 0.05 macro-F_1, and 0.9955 ± 0.0003 macro-OvR AUC. On the external DR14 set, BIQML reached 98.76 ± 0.09

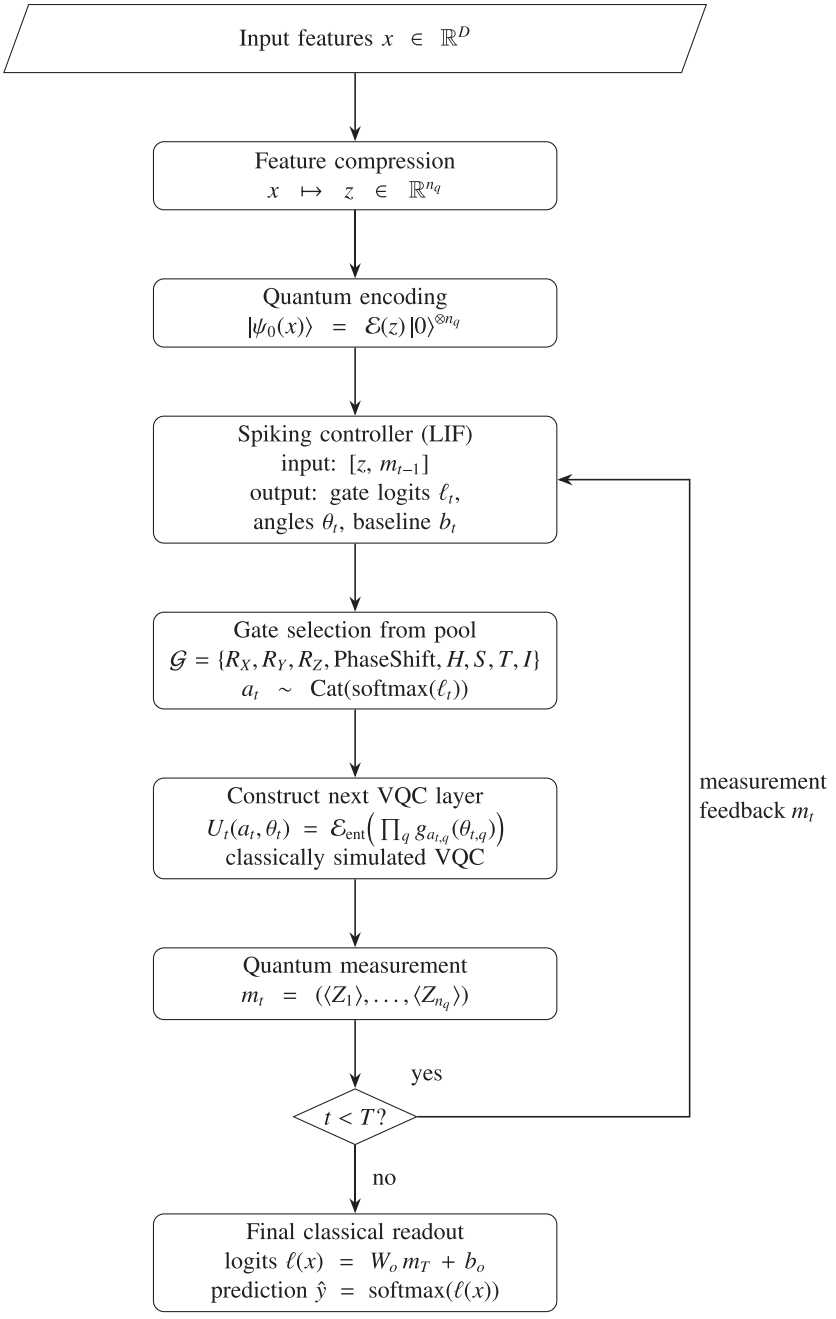


Fig. 5. Closed-loop data flow of the BIQML architecture. The classical input is compressed into a qubit-aligned representation and encoded into a classically simulated VQC. At each of the T decision steps, the spiking LIF controller receives the compressed representation together with the previous measurement vector m_{t-1}, selects one gate per qubit from the finite gate pool, and generates the corresponding rotation parameters; the resulting layer extends the circuit, which is measured to produce m_t. The measurement is fed back to the controller (right-hand loop), and after the final step the terminal measurement vector m_T is mapped to class logits by a classical readout layer. This sequential, measurement-driven construction, rather than a fixed ansatz, is the defining feature of BIQML.

percent accuracy, 98.41 ± 0.09 percent balanced accuracy, 0.9839 ± 0.11 macro-F_1, and 0.9980 ± 0.0002 macro-OvR AUC. Across both datasets, BIQML was competitive with the strongest classical ensembles and clearly outperformed the linear baselines.

The paired significance tests in Table 3 make this comparison precise. Throughout, Δ denotes the mean accuracy difference, BIQML minus baseline; therefore, a positive value favours BIQML and a negative value favours the baseline. On the external DR14 set, BIQML significantly outperformed Logistic Regression ($\Delta = +4.29$ percentage points, paired-t $p = 9.8 \times 10^{-15}$), Linear SVM ($\Delta = +2.41$ percentage points, $p = 8.5 \times 10^{-14}$), and MLP ($\Delta = +0.47$ percentage points, $p = 2.1 \times 10^{-5}$). Its difference from Extra Trees was not significant ($\Delta = -0.02$ percentage points, $p = 0.51$). Random Forest and HistGradientBoosting were marginally but significantly ahead of BIQML on DR14, with $\Delta = -0.61$ percentage points ($p = 9.3 \times 10^{-10}$) and $\Delta = -0.14$ percentage points ($p = 7.5 \times 10^{-3}$), respectively. The Wilcoxon signed-rank tests in Table 3 support the same overall interpretation. A similar ranking pattern was observed on the DR17 internal test set. We therefore do not claim that BIQML is the single best classifier in the full-feature regime. Rather, it is statistically competitive with the strongest tabular baselines, clearly stronger than the linear models, and

Table 2
Full-feature benchmark on SDSS DR17 (internal test) and SDSS DR14 (external cross-release test), reported as mean ± standard deviation over ten random splits. Best mean per column is shown in bold. AUC is one-vs-rest (OvR).

Dataset	Model	Accuracy	Balanced Acc.	AUC_{macro}	$F1_{macro}$
DR17 (internal)	Random Forest	**0.9779 ± 0.0009**	0.9717 ± 0.0012	0.9958 ± 0.0002	**0.9742 ± 0.0010**
	HistGradientBoosting	0.9758 ± 0.0011	**0.9743 ± 0.0013**	**0.9963 ± 0.0003**	0.9722 ± 0.0013
	BIQML	0.9736 ± 0.0005	0.9690 ± 0.0008	0.9955 ± 0.0003	0.9697 ± 0.0005
	Extra Trees	0.9726 ± 0.0010	0.9654 ± 0.0012	0.9947 ± 0.0003	0.9684 ± 0.0011
	MLP	0.9716 ± 0.0012	0.9662 ± 0.0013	0.9950 ± 0.0004	0.9675 ± 0.0013
	Logistic Regression	0.9430 ± 0.0017	0.9509 ± 0.0016	0.9882 ± 0.0005	0.9365 ± 0.0019
	Linear SVM	0.9372 ± 0.0014	0.9455 ± 0.0015	0.9859 ± 0.0006	0.9307 ± 0.0015
DR14 (external)	Random Forest	**0.9937 ± 0.0003**	**0.9851 ± 0.0007**	**0.9991 ± 0.0001**	**0.9880 ± 0.0005**
	HistGradientBoosting	0.9890 ± 0.0006	0.9840 ± 0.0006	0.9988 ± 0.0001	0.9828 ± 0.0008
	Extra Trees	0.9878 ± 0.0003	0.9808 ± 0.0007	0.9981 ± 0.0001	0.9828 ± 0.0004
	BIQML	0.9876 ± 0.0009	**0.9841 ± 0.0009**	0.9980 ± 0.0002	0.9839 ± 0.0011
	MLP	0.9829 ± 0.0014	0.9797 ± 0.0014	0.9970 ± 0.0003	0.9800 ± 0.0015
	Linear SVM	0.9635 ± 0.0006	0.9491 ± 0.0004	0.9909 ± 0.0001	0.9547 ± 0.0005
	Logistic Regression	0.9448 ± 0.0009	0.9306 ± 0.0006	0.9913 ± 0.0001	0.9411 ± 0.0006

Table 3
Paired significance tests of BIQML versus each baseline on the external DR14 set, across the ten splits. Δ is the mean accuracy difference (BIQML – baseline). Positive Δ favours BIQML.

Baseline	Δ (pp)	Paired-t p	Wilcoxon p
Logistic Regression	+4.29	9.8×10^{-15}	0.002
Linear SVM	+2.41	8.5×10^{-14}	0.002
MLP	+0.47	2.1×10^{-5}	0.002
Extra Trees	−0.02	0.51 (n.s.)	0.44
HistGradientBoosting	−0.14	7.5×10^{-3}	0.004
Random Forest	−0.61	9.3×10^{-10}	0.002

Table 4
BIQML per-class precision, recall, and F1 (mean over ten splits) in the full-feature setting.

Dataset	Class	Precision	Recall	F1
DR17 (internal)	GALAXY	0.9775	0.9786	0.9781
	QSO	0.9613	0.9354	0.9482
	STAR	0.9730	0.9930	0.9829
DR14 (external)	GALAXY	0.9912	0.9842	0.9877
	QSO	0.9737	0.9735	0.9736
	STAR	0.9862	0.9947	0.9904

is subsequently assessed under cross-release and reduced-information conditions.

Table 4 reports BIQML's class-wise performance, averaged over the ten splits. On DR14, BIQML maintained high and balanced per-class quality (GALAXY $F_1 = 0.988$, STAR $F_1 = 0.990$, QSO $F_1 = 0.974$), confirming that the strong overall accuracy is not driven by the majority class. The corresponding confusion matrices for the full-feature and low-sensor settings are shown in Fig. 6; these are computed on the held-out test partitions rather than the full samples of Fig. 2, so their row totals are smaller than the global class counts. The relatively lower QSO recall on the internal DR17 split (0.935; Fig. 6a) compared with DR14 (0.974; Fig. 6b) reflects the higher QSO fraction and the harder galaxy–QSO boundary in the DR17 sample.

4.3. Low-sensor g, r, i stress test

To probe behaviour when key observational channels are unavailable, all models were retrained and evaluated on the reduced nine-feature representation (the g, r, i magnitudes and the derived colour and summary features), with redshift, the u- and z-bands, and metadata removed. This regime is substantially harder: removing redshift eliminates the single most discriminative feature for separating quasars, so all models lose accuracy. The key question is which model degrades most gracefully. The training and internal optimization dynamics of BIQML in this low-sensor setting are shown in Fig. 7. The close agreement between the training and validation curves, together with the smooth evolution of the individual loss components, indicates stable convergence without evident overfitting.

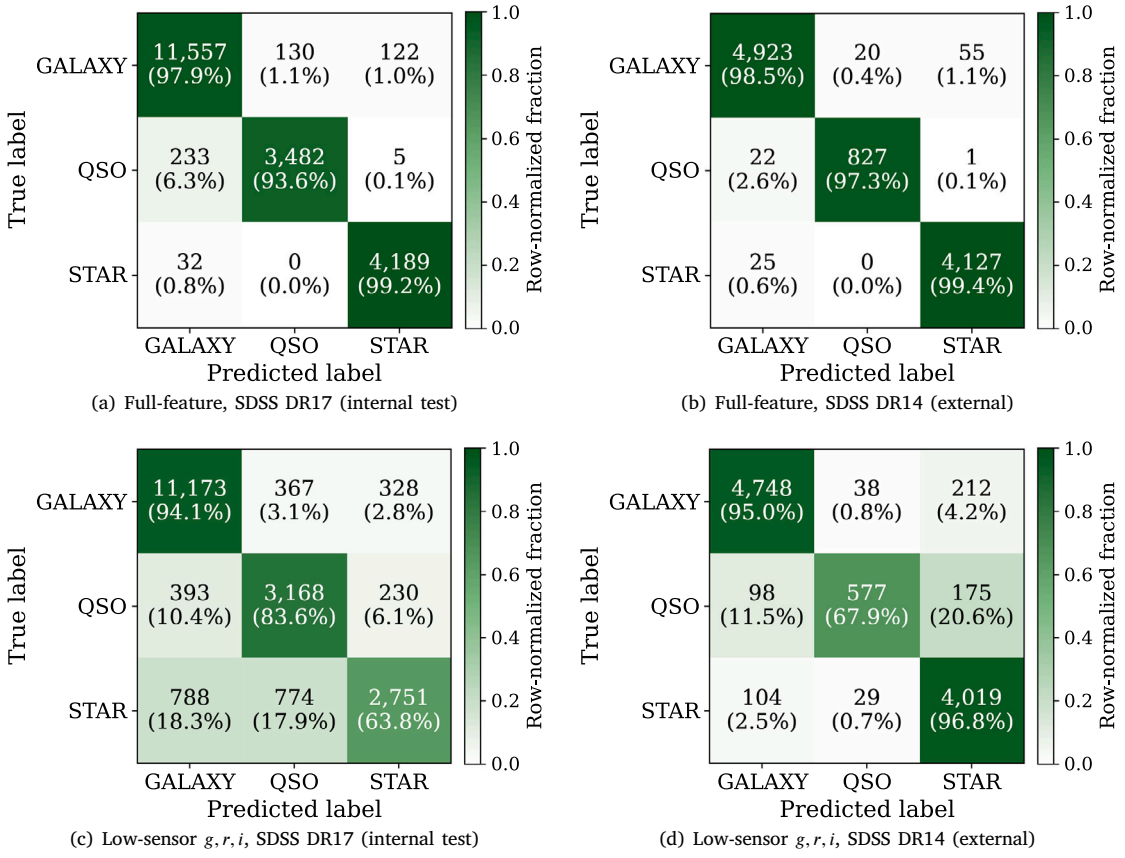
Table 5
Low-sensor g, r, i no-redshift benchmark (nine features), ranked by balanced accuracy on the DR17 internal test set, on a single controlled split. BIQML records the top balanced accuracy, overall accuracy, and macro AUC on this split.

Rank	Model	Acc.	Bal. Acc.	Macro F1	AUC
1	**BIQML**	**0.8558**	**0.8050**	**0.8089**	**0.9430**
2	HistGradientBoosting	0.8536	0.7986	0.8052	0.9398
3	MLP	0.8521	0.7949	0.8022	0.9412
4	Random Forest	0.8440	0.7873	0.7929	0.9335
5	Extra Trees	0.8322	0.7730	0.7785	0.9270
6	Logistic Regression	0.7538	0.6637	0.6540	0.8285
7	Linear SVM	0.7391	0.6386	0.6028	0.8245

Table 5 and Fig. 8 report the ranking. In this information-limited regime BIQML ranked among the top models on this single controlled split, reaching a balanced accuracy of 0.8050, marginally ahead of HistGradientBoosting (0.7986), the MLP (0.7949), and Random Forest (0.7873), with the linear models trailing far behind (0.64–0.66). On the same split BIQML also recorded the top overall accuracy (85.58 percent on DR17) and macro-OvR AUC (0.9430). On the external DR14 set, BIQML reached 93.44 percent accuracy (86.56 percent balanced accuracy). The per-class confusion matrices (Fig. 6c–d) localize the residual error: removing redshift leaves galaxy and star recall comparatively high (DR17 GALAXY 0.941, STAR 0.638; DR14 GALAXY 0.950, STAR 0.968) but reduces QSO recall to 0.836 on DR17 and 0.679 on DR14, confirming that the galaxy–QSO boundary is the dominant remaining source of error once redshift is unavailable. The ordering is nonetheless the reverse of the full-feature case: the gradient ensembles that narrowly led when redshift was available now fall behind BIQML once that feature is removed. This is consistent with the central hypothesis of this work—that the brain-inspired, measurement-driven representation remains competitive when discriminative information is scarce. We emphasize, however, that this low-sensor comparison is reported on a single controlled split, and we identify placing it on the full ten-split footing as the primary item of future work.

4.4. Component-level ablation

The full-feature SDSS task is close to saturated when spectroscopic redshift is available, making small performance differences difficult to



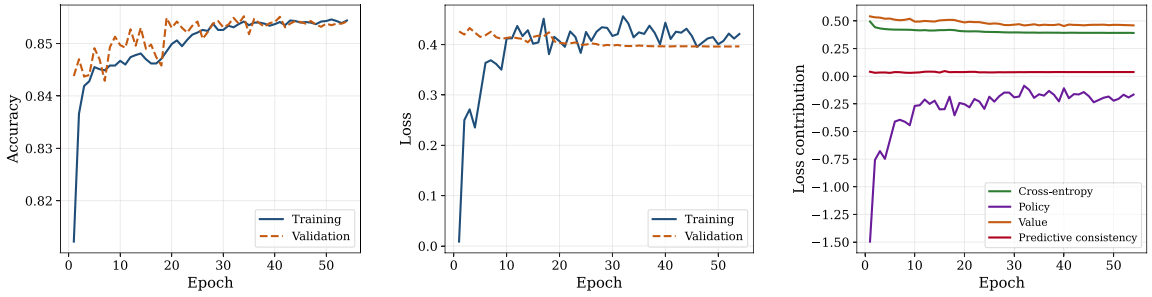
(a) Full-feature, SDSS DR17 (internal test)

(b) Full-feature, SDSS DR14 (external)

(c) Low-sensor g, r, i, SDSS DR17 (internal test)

(d) Low-sensor g, r, i, SDSS DR14 (external)

Fig. 6. Confusion matrices for BIQML star–galaxy–QSO classification. Panels (a) and (b) show the full-feature model on the SDSS DR17 internal test set and the external SDSS DR14 set; panels (c) and (d) show the low-sensor g, r, i no-redshift model on the same two evaluation sets. Removing redshift in the low-sensor setting increases galaxy–QSO confusion, consistent with the photometric similarity of compact galaxies and quasars at fixed g, r, i colours.



(a) Training and validation accuracy versus epoch.

(b) Training and validation loss versus epoch.

(c) Evolution of the individual loss components.

Fig. 7. Training and internal optimization dynamics of the BIQML architecture in the low-sensor setting. (a) Training and validation accuracy track each other closely, indicating stable convergence without overfitting. (b) The corresponding loss curves confirm stable optimization under the combined objective. (c) The individual loss components, comprising the cross-entropy, policy, value, and predictive-consistency terms, evolve smoothly, reflecting balanced multi-objective learning across the supervised and brain-inspired terms.

attribute reliably to individual architectural components. We therefore performed the component-level ablation in the more demanding low-sensor setting, using the same nine g, r, i features as the stress test above. The experiment was designed as a parameter-matched functional diagnostic rather than a repeated-split superiority study: Full BIQML and every ablated variant were retrained from scratch under an identical low-sensor input representation, DR17 train–validation–test partition, preprocessing, optimization schedule, loss weights, batch size, and temperature schedule, with each model containing exactly $164,039$ trainable parameters. This short-budget component analysis is separate from the main low-sensor benchmark in Table 5: the ablation re-runs Full BIQML and every variant under a reduced 15-epoch budget at seed 92, so its Full BIQML balanced accuracy (0.8016) is slightly below the main-benchmark value (0.8050). The two sets of numbers are therefore not directly comparable and should be read only within their respective tables.

The ablation isolates three elements of the sequential circuit-construction loop while retaining the BIML spiking compressor, the qubit-aligned representation, the temporal depth, and the readout in all variants. First, the binary spiking controller is replaced by a continuous

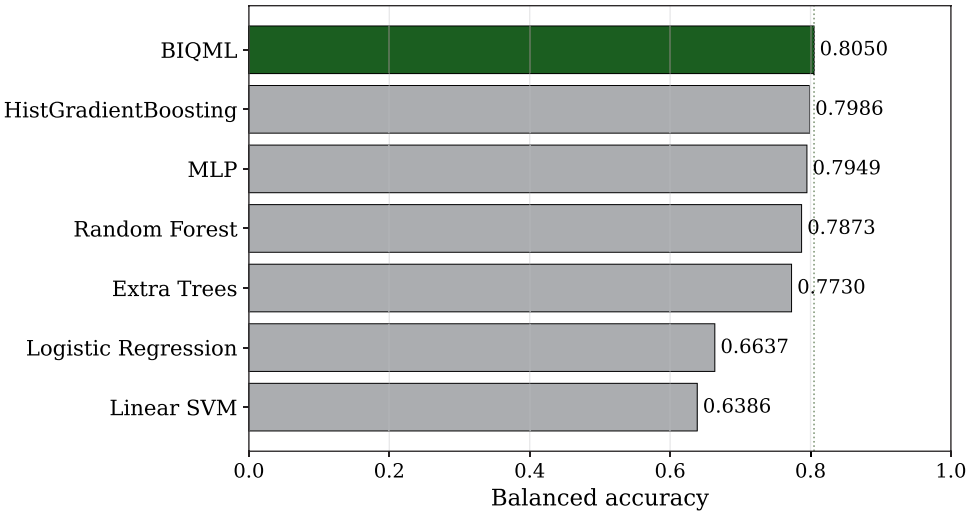


Fig. 8. Model ranking by balanced accuracy in the low-sensor g, r, i no-redshift setting, on a single controlled split. When spectroscopic redshift and the u- and z-bands are removed, BIQML records the top balanced accuracy on this split, marginally ahead of the gradient-boosted and tree ensembles that narrowly lead in the full-feature regime. The linear baselines degrade markedly, indicating that the discriminative signal in this regime is strongly non-linear.

rate-LIF controller with the same membrane update and parameterization. Second, the VQC state transition is replaced by a parameter-free classical state map receiving the same controller-generated gate identifiers, angles, depth, and nearest-neighbour coupling pattern. Third, measurement-to-controller feedback is removed by zeroing the incoming measurement vector while retaining circuit evaluation and predictive-state supervision. A final variant removes both binary spiking and the VQC transition; since the BIML compressor remains active throughout, this is not a fully classical model.

Table 6 and Fig. 9 summarize the results. On the primary low-sensor DR17 test set, Full BIQML achieves the highest balanced accuracy (0.8016) and the highest macro one-vs-rest AUC (0.9409). Replacing binary spiking with a rate-based controller reduces DR17 balanced accuracy by 0.57 percentage points, removing the VQC transition by 0.23, suppressing measurement feedback by 0.34, and the combined rate-controller plus classical-state-map baseline by 0.64. Although these changes are not additive because the components interact, their common direction on the internal test set indicates that the complete closed-loop configuration is preferred under limited observational information, and exact parameter matching shows this cannot be attributed to a larger model. The external DR14 comparison provides a qualification: Full BIQML attains 0.8630 balanced accuracy, while the rate-LIF, no-VQC, and combined variants are lower (0.8495, 0.8589, 0.8595), but the no-feedback variant reaches 0.8716, exceeding the full configuration by 0.87 points on this single external split. The diagnostic thus consistently supports the contribution of binary spiking and the VQC transition across both releases, whereas the contribution of measurement feedback is positive on DR17 but not uniformly so on external balanced accuracy. This mixed result is reported directly; repeating the parameter-matched ablation over multiple pre-specified seeds will be required before drawing a statistical component-level conclusion. The VQC contribution is evaluated through classical state-vector simulation and therefore carries substantial computational cost, with no hardware-level quantum advantage implied.

4.5. Computational cost

Table 7 reports wall-clock training time, inference time, and trainable-parameter counts for all models, measured on identical hardware over the ten splits. BIQML is, as expected, the most expensive model: training takes on the order of 1.8×10^4 s per split, versus seconds to a few minutes for the classical baselines, because every forward pass evaluates a multi-step quantum circuit on the simulator. We emphasize, however, that this comparison is not like-for-like and should not be read as a measure of intrinsic algorithmic efficiency. The classical baselines are executed through highly optimized, low-level numerical libraries (e.g. the compiled `scikit-learn` ensemble and linear solvers), whereas the dominant cost of BIQML is the *classical state-vector simulation* of its quantum circuits, which scales exponentially in qubit count and is incurred entirely by emulating quantum execution on a classical processor. This simulation overhead would not be present on native quantum hardware, where the circuit would be evaluated directly rather than emulated; the reported training time therefore reflects the cost of classical emulation, not a fundamental bottleneck of the architecture itself. Its $168,647$ trainable parameters (full-feature configuration) exceed the tree ensembles and the linear models but are comparable to the MLP, confirming that the runtime gap is driven by quantum-circuit simulation rather than by model size. The same architecture and hidden dimensions are used in the low-sensor regime, but the absolute parameter count differs because the input dimensionality differs: the full-feature configuration accepts 18 inputs and contains $168,647$ trainable parameters, whereas the nine-feature low-sensor g, r, i configuration contains $164,039$. The $4,608$-parameter reduction arises exclusively from the two input-facing BIML compressor projections; within each regime, all ablation variants are exactly parameter matched. These costs are an honest limitation of classically simulated quantum-inspired models and are reported transparently; they also motivate the deliberately shallow circuit ($T = 5$, $n_q = 6$) used throughout. For survey-scale deployment, the trained model's inference cost (of order 10^2 s for $\sim 2\times10^4$ objects) is the more relevant figure and is dominated by the simulator rather than by the classical components.

4.6. Quantum-noise screening

To assess the sensitivity of BIQML to representative quantum-noise channels, we performed a simulator-based noise screening using depolarizing, amplitude-damping, and readout bit-flip channels applied to the final circuit at strengths of 0.5 percent, 1.0 percent, and 2.0 percent. Each of the ten trained BIQML models was re-evaluated under all nine noisy conditions and the ideal noiseless baseline, on both the internal DR17 test set and the external DR14 set, using analytic (infinite-shot) expectation values. Across all conditions, the model showed no measurable change in accuracy, balanced accuracy, or macro-F1 relative to the ideal simulator. We note that, under analytic expectation values, a global depolarizing channel contracts each Pauli-Z expectation by a common multiplicative factor, to which the linear readout followed by

Table 6
Parameter-matched component-level ablation in the low-sensor g, r, i setting. All variants use identical low-sensor inputs, data partition, optimization protocol, and 164,039 trainable parameters. Δ is computed as Full BIQML minus the corresponding variant in percentage points; a positive value favours Full BIQML. Results are from one controlled diagnostic run and are not averages across repeated seeds.

Variant	SDSS DR17 internal test			SDSS DR14 external test	
	BAcc.	Δ_{DR17} (pp)	AUC_{macro}	BAcc.	Δ_{DR14} (pp)
Full BIQML	**0.8016**	—	**0.9409**	0.8630	–
Rate-LIF controller (no binary spikes)	0.7959	+0.57	0.9406	0.8495	+1.34
Fixed classical state map (no VQC)	0.7993	+0.23	0.9404	0.8589	+0.41
No measurement-to-controller feedback	0.7982	+0.34	**0.9413**	**0.8716**	−0.87
Rate-LIF + fixed state map	0.7952	+0.64	0.9403	0.8595	+0.35

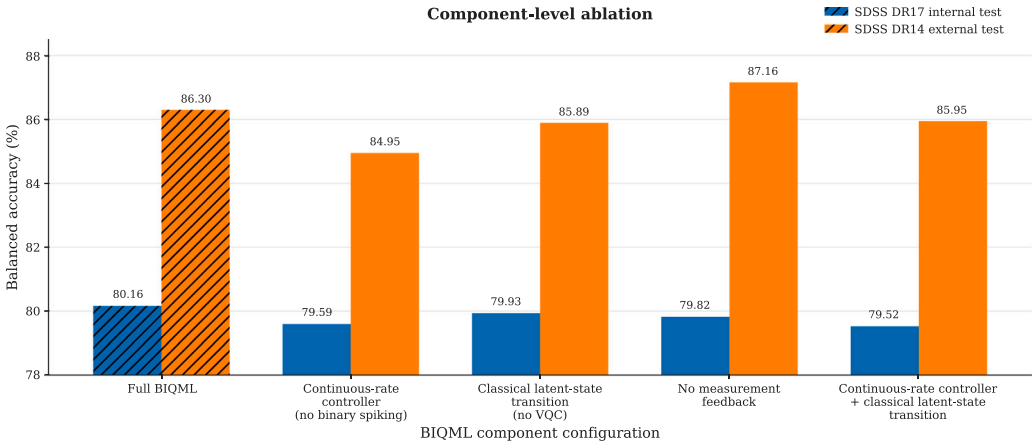


Fig. 9. Component-level ablation in the low-sensor (g,r,i) setting. Balanced accuracy is shown for the SDSS DR17 internal test and SDSS DR14 external test for Full BIQML and each parameter-matched component-removal variant. All variants use the same low-sensor representation, data partition, optimization protocol, and number of trainable parameters. The figure presents a controlled one-run diagnostic and does not display uncertainty intervals or imply statistical significance across random seeds.

Table 7
Computational cost per split (representative seed). Training and inference are wall-clock seconds; parameters are trainable counts.

Model	Train (s)	Infer DR14 (s)	Params
BIQML	~18,513	57.1	168,647
MLP	111.3	0.06	46,211
Random Forest	31.9	0.30	–
Extra Trees	5.0	0.34	–
HistGradientBoosting	3.7	0.73	–
Logistic Regression	8.0	0.004	57
Linear SVM	0.8	0.004	57

an arg max decision is largely invariant; the observed stability is therefore partly a property of this evaluation regime rather than evidence of hardware-level fault tolerance. Two architectural factors reinforce this stability: the VQC is intentionally shallow (a horizon of $T = 5$ decision steps), limiting the number of operations exposed to accumulated gate errors, and the spiking controller aggregates information across multiple decision steps rather than relying on a single isolated circuit output. We therefore interpret these results as indicating that BIQML predictions are stable under small, representative analytic noise channels, while a full finite-shot and hardware-noise characterization is left to future work.

4.7. Quantitative cross-release distribution shift

To verify that the DR17→DR14 evaluation constitutes a genuine distribution shift rather than only a class-prior change, we quantified per-feature differences between the two releases using the Kolmogorov–Smirnov (KS) statistic and the Wasserstein distance on the standardized feature space. The shift is substantial and statistically significant for the photometric features: the magnitude features exhibit KS statistics in the range 0.69–0.72 with $p < 10^{-3}$ and Wasserstein distances above 1.6 standard units, confirming a marked covariate shift between the two releases (Fig. 10). This establishes that BIQML's stable external performance reflects genuine cross-release generalization under measurable covariate shift, and that the reported metrics are not an artefact of class composition alone, which is separately controlled through the balanced-accuracy and macro-averaged reporting.

4.8. Discussion

Taken together, the three experiments support a nuanced and defensible conclusion. In the full-feature regime, where spectroscopic redshift nearly separates the three classes, BIQML is statistically on par with the best classical tabular models (Extra Trees, HistGradientBoosting, Random Forest) and clearly superior to linear baselines, but it does not establish a decisive accuracy advantage, and we do not claim one. The distinctive behaviour of BIQML emerges under stress: in the low-sensor g, r, i regime it records the top balanced accuracy on the reported single split (Fig. 8), and across the DR17→DR14 covariate shift it maintains high external accuracy. The component ablation (Fig. 9) is consistent with this being a property of the full architecture rather than of any single part or of raw capacity, since a larger fully classical recurrent model does not surpass it, though we read the low-sensor advantage as indicative pending repeated-split confirmation.

The class-wise results provide astrophysical context for where the models succeed and fail. Redshift cleanly isolates quasars, which occupy the high-redshift regime, while stars and galaxies overlap at low redshift (Fig. 3). When redshift is available, QSO recognition is strong for all models; when it is removed in the low-sensor setting, the galaxy–QSO boundary becomes the dominant source of error (Fig. 6c–d): QSO

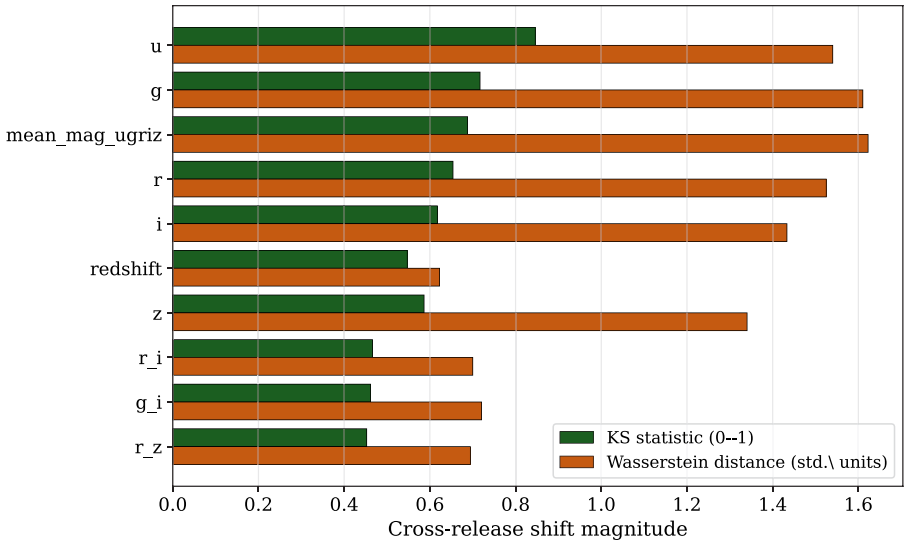


Fig. 10. Quantitative cross-release distribution shift between SDSS DR17 and DR14. For the most-shifted standardized features, the Kolmogorov–Smirnov statistic and Wasserstein distance confirm a substantial and statistically significant covariate shift, establishing that the DR17→DR14 evaluation tests genuine cross-release generalization rather than only a change in class priors.

Table 8
Inference-time finite-shot device-inspired noise robustness benchmark for BIQML at 256 shots. The same previously trained BIQML checkpoint was re-evaluated under the ideal finite-shot reference and four noisy device-inspired presets; no retraining, noise-aware fine-tuning, or parameter adaptation was performed. *Top:* balanced accuracy on balanced 1500-sample subsets of the internal DR17 test set and external DR14 test set. *Bottom:* balanced-accuracy drop relative to the shot-matched ideal reference. At the displayed precision, the trained model retains balanced accuracies of 0.9667 on DR17 and 0.9787 on DR14 across all tested noise presets, corresponding to a reported balanced-accuracy drop of 0.000. These results indicate decision-level stability of the closed-loop BIQML inference procedure under the tested finite-shot noise conditions.

Preset	1Q error (%)	2Q error (%)	Readout error (%)	T_1 (μs)	T_2 (μs)	Gate duration (1Q/2Q, ns)	Shots
Ideal finite-shot	0.0	0.0	0.0	–	–	–	256
IBM-light	0.1	1.0	1.0	100	80	50/300	256
IBM-medium	0.3	3.0	2.0	80	60	70/400	256
IBM-heavy	0.6	7.0	5.0	50	40	100/600	256
Readout-heavy	0.1	1.0	8.0	80	60	70/400	256

recall falls to 0.836 on DR17 and 0.679 on DR14, whereas galaxy and star recall remain high, because at fixed g, r, i colours compact galaxies and quasars are photometrically similar. BIQML's behaviour in this regime is consistent with its measurement-feedback representation capturing higher-order colour interactions that linear and even tree-based models do not exploit as effectively. Crucially, BIQML shows no systematic class-specific weakness on quasars: under the repeated-split protocol it maintains a QSO F1 of 0.974 on DR14, comparable to its galaxy and star performance and to the MLP baseline, confirming that its QSO recognition is robust once results are averaged over splits rather than read from a single confusion matrix.

5. Conclusion

We introduced BIQML, a brain-inspired quantum machine learning architecture in which a spiking leaky integrate-and-fire controller sequentially constructs a classically simulated variational quantum circuit through measurement feedback, rather than optimizing a fixed ansatz. Applied to SDSS star–galaxy–QSO classification, it was evaluated through a repeated ten-split full-feature benchmark with paired significance testing, a low-sensor g, r, i no-redshift stress test, and a component-level ablation, supported by analytic and finite-shot device-inspired noise screenings and an explicit cross-release distribution-shift analysis.

The picture is deliberately measured. In the full-feature regime, where spectroscopic redshift nearly separates the three classes, BIQML is statistically on par with the strongest classical tabular models and clearly outperforms the linear baselines, but we claim no decisive advantage, and the ablation confirms that no single component dominates near saturation. Its distinctive behaviour appears under stress: in the low-sensor regime it ranks among the top models on a single controlled split, degrading no more than the strongest ensembles, and across the DR17→DR14 covariate shift, quantified by Kolmogorov–Smirnov and Wasserstein diagnostics, it maintains high external accuracy. A larger fully classical recurrent baseline does not surpass it, though we treat this as indicative pending repeated-split confirmation.

BIQML also proved stable under non-ideal circuit execution. Across analytic screening and a finite-shot, IBM-calibrated benchmark at 256 shots, spanning depolarizing, amplitude-damping, readout, and thermal-relaxation presets (Table 8), the trained model retained its balanced accuracy to the reported precision (0.9667 on DR17, 0.9787 on DR14), with a 0.000 drop at every condition (Fig. 11). This decision-level stability is consistent with the deliberately shallow design ($n_q = 6$, $T = 5$) and the recurrent aggregation of intermediate measurements, suggesting BIQML may retain comparable accuracy on real NISQ hardware within similar noise and shot-budget regimes.

Limitations bound these claims. All circuits are evaluated by classical state-vector simulation, so no hardware-level quantum advantage is

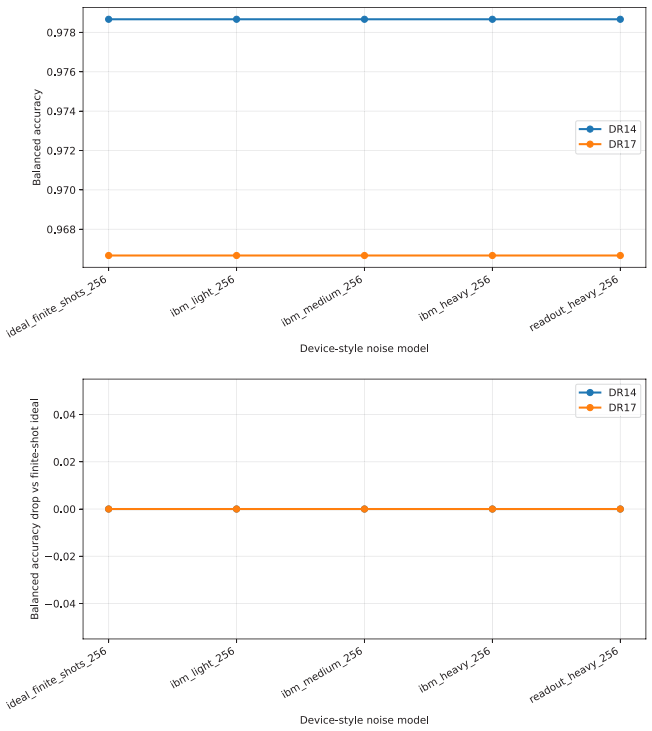


Fig. 11. Finite-shot device-style noise benchmark for BIQML at 256 shots. *Top:* balanced accuracy on the internal DR17 test set and external DR14 set across four IBM-calibrated noise presets and the finite-shot ideal reference; both curves are flat, at 0.9667 (DR17) and 0.9787 (DR14). *Bottom:* the corresponding balanced-accuracy drop relative to the finite-shot ideal, which is 0.000 at every noise level for both releases (the DR14 line lies exactly under DR17).

claimed and the training cost reflects that simulation rather than an intrinsic bottleneck. The analytic screening uses infinite-shot expectation values and does not by itself establish hardware fault tolerance, though the finite-shot benchmark adds complementary evidence under shot noise. The low-sensor stress test used a single controlled split, and the ablation compares a ten-seed full model against single-seed variants; we therefore frame BIQML's low-sensor standing as competitive parity with the strongest ensembles rather than a definitive ranking.

The cost of the approach is also reported transparently. BIQML is by far the most expensive model to train (of order 1.8×10^4 s per split against seconds to minutes for the classical baselines), but this gap is dominated by the classical state-vector simulation of its quantum circuits rather than by model size: at 168,647 trainable parameters it is comparable to the MLP, and the simulation overhead would not arise on native quantum hardware. For survey-scale deployment the trained model's inference cost (of order 10^2 s for $\sim 2 \times 10^4$ objects) is the more relevant figure, motivating the deliberately shallow circuit used throughout.

Several directions follow. A hardware-noise characterization on real NISQ devices is needed to test whether the finite-shot stability persists under device-specific error channels; extending the repeated-split protocol and ablation to the low-sensor regime would place that result on the same footing as the full-feature benchmark; scaling the qubit count and horizon with feature-importance analysis (e.g. SHAP) would clarify which photometric interactions BIQML exploits when redshift is absent; and evaluation across additional surveys, next-generation pipelines (e.g. LSST-era data), and the reverse DR14→DR17 direction would test broader survey-to-survey transfer.

CRediT authorship contribution statement

Subhash S. Pandey: Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Mousam Mondal:** Visualization, Software, Resources, Investigation, Data curation. **Xing Liang:** Writing – review & editing, Investigation, Formal analysis. **Christos Kalyvas:** Investigation, Formal analysis. **Dimitrios Makris:** Writing – review & editing, Investigation. **Hongwei Wu:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The SDSS DR14 and DR17 datasets were obtained from the official Sloan Digital Sky Survey website. SDSS DR14 is available at https://www.sdss4.org/dr14/, and SDSS DR17 is available at https://www.sdss4.org/dr17/. The code used in this study is available from the authors upon reasonable request.

References

Abdurro'uf, et al., 2022. The seventeenth data release of the Sloan Digital Sky Surveys: Complete release of MaNGA, MaStar, and APOGEE-2 data. Astrophys. J. Suppl. Ser. 259 (2), 35. doi:10.3847/1538-4365/ac4414.

Abolfathi, B., et al., 2018. The fourteenth data release of the Sloan Digital Sky Survey: First spectroscopic data from the extended baryon oscillation spectroscopic survey and from the second phase of the Apache Point Observatory Galactic Evolution Experiment. Astrophys. J. Suppl. Ser. 235 (2), 42. doi:10.3847/1538-4365/aa9e8a.

Ahmadvand, R., Sharif, S.S., Banad, Y.M., 2025. Neuromorphic robust framework for integrated estimation and control in dynamical systems using spiking neural networks. Sci. Rep. 15 (1), 44753.

Ai, Q., et al., 2025. A cross-layer residual spiking neural network with adaptive threshold leaky integrate-and-fire neuron and learnable surrogate gradient. Knowl.-Based Syst. 319, 113575.

Andrés, E., Cuéllar, M.P., Navarro, G., 2024. Brain-inspired agents for quantum reinforcement learning. Mathematics 12 (8), 1230.

Atashgahi, Z., et al., 2022. A brain-inspired algorithm for training highly sparse neural networks. Mach. Learn. 111 (12), 4411–4452.

Bhavsar, R., et al., 2023. Classification of potentially hazardous asteroids using supervised quantum machine learning. IEEE Access 11, 75829–75848.

Chatterjee, D., Ghosh, P., 2025. Redshift-agnostic machine learning classification: Unveiling peak performance in galaxy, star, and quasar classification using SDSS DR17. Astron. Nachr. 346 (5), e20240057. doi:10.1002/asna.20240057.

Cunha, P., Humphrey, A., 2022. Photometric redshift-aided classification using ensemble learning. Astron. Astrophys. 666, A87.

Duan, C., et al., 2025. A brain-inspired paradigm for scalable quantum vision. arxiv Preprint arXiv:2509.05919.

Fraix-Burnet, D., Bouveyron, C., Moultaka, J., 2021. Unsupervised classification of SDSS galaxy spectra. Astron. Astrophys. 649, A53. doi:10.1051/0004-6361/202040046.

Ghosh, P., 2024. From deep learning maze to neural network waltz: Unveiling peak performance in galaxy, star, and quasar classification using SDSS DR17. doi:10.7910/DVN/XNZDXA, Harvard Dataverse.

Irfan, M., et al., 2021. Brain-inspired lifelong learning model based on neural-based learning classifier system for underwater data classification. Expert Syst. Appl. 186, 115798.

Kasabov, N., 2010. To spike or not to spike: A probabilistic spiking neuron model. Neural Netw. 23 (1), 16–19.

Logan, C.H.A., Fotopoulou, S., 2020. Unsupervised star, galaxy, QSO classification: Application of HDBSCAN. Astron. Astrophys. 633, A154. doi:10.1051/0004-6361/201936648.

Lu, W., et al., 2024. Quantum machine learning: Classifications, challenges, and solutions. J. Ind. Inf. Integr. 42, 100736.

Martín-Guerrero, J.D., Lamata, L., 2022. Quantum machine learning: A tutorial. Neurocomputing 470, 457–461.

Nguyen, H.-Q., et al., 2024. Quantum-brain: Quantum-inspired neural network approach to vision-brain understanding. arxiv Preprint arXiv:2411.13378.

Peruzzi, T., et al., 2021. Interpreting automatic AGN classifiers with saliency maps. Astron. Astrophys. 652, A19. doi:10.1051/0004-6361/202038911.

Portillo, S.K.N., et al., 2020. Dimensionality reduction of SDSS spectra with variational autoencoders. Astron. J. 160 (1), 45. doi:10.3847/1538-3881/ab9644.
Rahman, F., 2018. Predicting star, galaxy and quasar with SVM. Available online, Kaggle Notebook.
Rahmaniar, W., et al., 2024. Deep learning and quantum algorithms approach to investigating the feasibility of wormholes: A review. Astron. Comput. 47, 100802.
Rauf, A., Amin, J., Nabi, J.-U., 2026. Hybrid quantum-classical convolutional neural network for astrophysical object classification. Phys. Rev. E 113 (1), 015302.
Schmidgall, S., et al., 2024. Brain-inspired learning in artificial neural networks: a review. APL Mach. Learn. 2 (2).
Sepulchre, R., 2022. Spiking control systems. Proc. IEEE 110 (5), 577–589.
Singh, S.P., et al., 2025. NeuroEvolve: A brain-inspired mutation optimization algorithm for enhancing intelligence in medical data analysis. Int. J. Cogn. Comput. Eng..
Wang, L.L., et al., 2023. Spectral classification of LAMOST emission-line galaxies based on machine learning methods. New Astron. 99, 101965.
Wang, D., et al., 2026. Optically-driven nine-state memory and pattern recognition in a lead-free tin-doped $Cs_2AgBiBr_6$ memristor for neuromorphic visual system. Mater. Today Commun. 51, 114742.
Wen, X.-Q., Jiang, Y.-Z., Liu, F.-H., Mi, J.-L., Li, C.-X., Hu, J., Shi, X.-P., Dong, X.-W., 2023. The classification for the sources in SDSS DR18: Searching for QSOs by machine learning. preprint org.
Whiting, M.T., et al., 2020. Astronomical data analysis software and systems XXVII. In: Astronomical Society of the Pacific Conference Series, vol. 522, p. 469.