RESOURCE-CONSTRAINED QUANTUM MODULATION OF LIQUID TEMPORAL DYNAMICS

Anonymous authors
Paper under double-blind review

ABSTRACT

What functional role can a small parameterised quantum circuit play when embedded inside a classical dynamical model? We study this question through resource-constrained quantum modulation of irregular temporal dynamics, where a low-qubit circuit modulates the effective time scale and nonlinear drift of a liquid recurrent update rather than acting as a stand-alone predictor. We characterise the induced function class and its resource sensitivity. Without entanglement, local measurements factorise into bounded nonlinear ridge functions of learned input projections; repeated data re-uploading can expand the potentially accessible local Fourier support under the stated angle-encoding assumptions, while entangling gates permit mixed-coordinate spectral interactions. We further derive a finite-shot trajectory perturbation bound showing that $O(S^{-1/2})$ measurement error is filtered through temporal spacing and recurrent state sensitivity. Experiments on a controlled irregularised Room Occupancy benchmark support this mechanistic distinction. Removing entanglement while retaining re-uploading improves the classical liquid backbone from 0.3065 to 0.4018 accuracy and also improves macro-F1, negative log-likelihood, Brier score, and expected calibration error relative to the full circuit where available, whereas removing re-uploading worsens most reported metrics. Strong MLP and GRU references remain substantially better overall predictors, so the results do not support a claim of predictive quantum advantage. Finite-shot evaluation produces non-monotonic operating-point changes, motivating measurement budget as an explicit property of recurrent quantum models. Overall, the results support viewing low-qubit quantum circuits as structured dynamical modulators whose usefulness depends on the function class required by the downstream temporal system.

1 INTRODUCTION

Hybrid quantum machine learning is increasingly moving away from broad claims of unconditional quantum advantage towards a more operational question: *what functional structure can a small parameterised quantum circuit contribute when embedded inside an otherwise classical learning system?* This question is particularly relevant in the noisy intermediate-scale quantum (NISQ) regime, where useful models must operate with limited qubit counts, shallow circuits, finite measurement budgets, and non-trivial trainability constraints (Biamonte et al., 2017; Cerezo et al., 2021). Under these constraints, the usefulness of a quantum component cannot be assessed by circuit expressivity alone (Abbas et al., 2021; Sim et al., 2019) It also depends on the inductive bias induced by the circuit, its interaction with the surrounding classical dynamics, and the measurement resources required to estimate its outputs. Recent learning-theoretic analyses similarly emphasise that the usefulness of a quantum learning model depends on the relation between the data, the target function, and the inductive bias encoded by its quantum representation (Huang et al., 2021; Kübler et al., 2021).

Irregular temporal learning provides a useful setting in which to study these questions. Real-world sequential observations may be missing, asynchronous, or separated by non-uniform time intervals, and both observation availability and elapsed time can carry predictive information (Che et al., 2018). Continuous-time and irregular-sequence models therefore allow latent dynamics to depend explicitly on temporal spacing rather than assuming uniform recurrence (Chen et al., 2018;

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Rubanova et al., 2019; Kidger et al., 2020; Hasani et al., 2021; 2022).Complementary approaches include event-gated recurrence, continuous-time gated recurrent dynamics, learned time-attention mechanisms, and continuous-time Transformers (Neil et al., 2016; De Brouwer et al., 2019; Shukla & Marlin, 2021; Chen et al., 2023). We adopt a liquid-style recurrent backbone in which each temporal input contains the observed features, an observation mask, and the elapsed time since the preceding retained observation. Rather than replacing this temporal model with a quantum predictor, we use a low-qubit parameterised quantum circuit (PQC) to modulate the time-scale and drift terms governing hidden-state evolution. Figure 1 provides an overview of this design and the structural view developed in the paper. Panel (a) shows how the two quantum branches enter the liquid dynamics, while panel (b) previews the distinct roles of factorisation, data re-uploading, entanglement, and finite-shot measurement.

This formulation turns the PQC into a structured perturbation of a recurrent vector field and permits a more precise analysis of what different circuit components contribute. We show that, with local Pauli-Z measurements and without entanglement, the quantum modulator factorises into bounded nonlinear functions of learned one-dimensional input projections. For the angle-encoding architecture considered here, repeated data re-uploading enlarges the accessible Fourier support of these local functions, whereas entangling operations enable dependence on multiple projected coordinates through mixed-frequency terms. Thus, re-uploading and entanglement do not simply represent weaker and stronger versions of the same mechanism: they modify different aspects of the induced function class. The corresponding structural expressions are given in Appendix A, with the formal characterisation developed in Section 3.3.

Quantum measurement introduces a second source of structure. Expectation values estimated from S shots incur $O(S^{-1/2})$ sampling error, but in a recurrent model this error does not remain local: a perturbation to a quantum feature changes the hidden state supplied to all subsequent temporal transitions. We therefore analyse finite-shot inference dynamically rather than as independent output noise. Under standard local Lipschitz assumptions, the resulting trajectory perturbation depends jointly on the measurement budget, the temporal spacing of observations, and the sensitivity of subsequent recurrent transitions. Details of this decomposition are provided in Appendix A.5, and the corresponding trajectory bound is formalised in Section 3.5.

These structural results motivate three controlled comparisons. We contrast the full entangled circuit with an entanglement-free variant that preserves data re-uploading and with a no-reuploading variant that retains the remaining variational structure. We additionally evaluate the full model under explicit finite-shot budgets. Experiments on an irregularized Room Occupancy benchmark show that strong classical MLP and GRU references remain substantially better overall predictors, so our results do not constitute evidence of predictive quantum advantage. Within the fixed liquid backbone, however, the circuit structure has a clear effect: removing entanglement while retaining re-uploading improves accuracy, macro-F1, negative log-likelihood, and expected calibration error relative to both the classical liquid model and the full quantum circuit, whereas removing re-uploading degrades most metrics. Finite-shot evaluation further changes the recurrent operating point, although the available point estimates do not support a claim that measurement noise is intrinsically beneficial.

Contributions. We make three contributions. First, we formulate a resource-constrained quantum modulation architecture in which low-qubit PQCs alter the drift and effective time scale of irregular liquid dynamics rather than directly predicting class labels. Second, we characterise the induced modulation function, separating local spectral enrichment from re-uploading and cross-coordinate interaction from entanglement, and connect finite-shot measurement error to recurrent trajectory sensitivity. Third, we evaluate these mechanisms through circuit-component ablations, calibration-sensitive metrics, finite-shot inference, and controlled temporal sparsification. Together, the analysis supports a mechanism-oriented view of hybrid quantum sequence modelling: under tight quantum resources, the relevant design question is not whether a circuit is maximally expressive, but whether its induced function class is appropriate for the dynamics it modulates.

2 Related Work

Irregular temporal dynamics. Irregular temporal learning addresses sequences with missing observations, asynchronous measurements, or non-uniform time intervals. Recurrent models such as GRU-D explicitly incorporate observation masks and elapsed-time information, while time-aware gating and attention mechanisms provide alternative treatments of irregular sampling (Che et al., 2018). Continuous-time approaches instead model latent evolution directly as a time-dependent dynamical system. Neural ODEs, Latent ODEs, ODE–RNNs, and neural controlled differential equations provide general continuous-time formulations for irregular observations (Chen et al., 2018; Rubanova et al., 2019; Kidger et al., 2020). Neural rough differential equations further extend this path-based continuous-time perspective to long and potentially irregular sequences (Morrill et al., 2021). Liquid Time-Constant Networks and Closed-form Continuous-time Networks further introduce state-dependent temporal dynamics in recurrent architectures (Hasani et al., 2021; 2022). Our setting builds on this family but does not propose a new general-purpose temporal backbone: the quantum component instead modulates the time scale and drift of an existing liquid recurrent update.

Quantum function classes and finite-shot hybrid learning. Parameterised quantum circuits (PQCs) are commonly used as trainable nonlinear feature maps in hybrid quantum–classical models (Havlicek et al., 2019; Schuld & Killoran, 2019). Their induced function class depends on the data encoding, variational gates, measurement observables, and circuit connectivity. In particular, data re-uploading repeatedly injects classical inputs into shallow circuits and can enlarge the accessible input-dependent spectral family (Perez-Salinas et al., 2020; Schuld et al., 2021). A complementary function-class perspective places data re-uploading within a broader hierarchy of quantum learning models and highlights distinct qubit and data-resource trade-offs (Jerbi et al., 2023). Entangling operations alter a different structural property by coupling different encoded coordinates. More generally, increased circuit expressivity does not automatically imply improved optimisation or task-level performance, and variational models may exhibit substantial trainability and measurement costs (Sim et al., 2019; Abbas et al., 2021; McClean et al., 2018; Cerezo et al., 2021). Hardware noise can introduce an additional optimisation bottleneck through noise-induced barren plateaus, further strengthening the case for shallow and resource-conscious circuit designs (Wang et al., 2021).

These considerations become especially important when a PQC is evaluated repeatedly inside a recurrent model. Quantum observables estimated from a finite number of shots incur sampling error, and parameter-shift or related gradient estimators likewise require repeated circuit executions . In a static predictor, such error perturbs an individual feature evaluation; in a temporal model, it can alter a hidden state and thereby propagate through later transitions. Our work therefore differs from prior static uses of PQCs in two respects: we treat the circuit as a *dynamical modulator* rather than as a stand-alone predictor, and we analyse re-uploading, entanglement, and finite-shot measurement as distinct mechanisms affecting the induced temporal dynamics. A more detailed discussion of these connections and the positioning of our work is provided in Appendix B.

3 Methodology

3.1 Problem setup and liquid dynamics

We consider irregular temporal classification from windows

$$\mathcal{S} = \{(\mathbf{x}_t, \mathbf{m}_t, \Delta t_t)\}_{t=1}^{L},$$

where $\mathbf{x}_t \in \mathbb{R}^d$ denotes the observed or imputed features, $\mathbf{m}_t \in \{0,1\}^d$ is the observation mask, and $\Delta t_t \in \mathbb{R}_+$ is the elapsed time since the preceding retained observation. At each temporal position the model receives

$$\mathbf{z}_t = [\mathbf{x}_t \,\|\, \mathbf{m}_t \,\|\, \Delta t_t] \in \mathbb{R}^{2d+1}. \tag{1}$$

Classical liquid reference. Let $\mathbf{h}_t \in \mathbb{R}^H$ denote the hidden state. The liquid backbone uses a state- and input-dependent time scale and an explicit time-aware update,

$$\begin{aligned} \boldsymbol{\tau}_t &= \sigma\left(\mathbf{W}_{\tau h}\mathbf{h}_{t-1} + \mathbf{W}_{\tau z}\mathbf{z}_t + \mathbf{b}_\tau\right) + \epsilon\mathbf{1}_H, \\ \mathbf{h}_t &= \mathbf{h}_{t-1} + \Delta t_t\left[-\mathbf{h}_{t-1} \oslash \boldsymbol{\tau}_t + \tanh\left(\mathbf{W}_h\mathbf{h}_{t-1} + \mathbf{W}_z\mathbf{z}_t + \mathbf{b}_g\right)\right]. \end{aligned} \tag{2}$$

Here $\sigma(\cdot)$ is the element-wise sigmoid, $\oslash$ denotes element-wise division, and $\epsilon > 0$ prevents division by zero. The factor Δt_t explicitly scales the state change by the local temporal spacing and will also determine how quantum-feature perturbations enter the trajectory.

3.2 Quantum modulation architecture

The temporal input is first projected into a q-dimensional quantum input space,

$$\mathbf{u}_t = \mathbf{P}\mathbf{z}_t + \mathbf{b}_P \in \mathbb{R}^q, \tag{3}$$

where q is the number of qubits. This projection forms an explicit low-dimensional bottleneck between the irregular temporal input and the quantum circuit.

For each projected coordinate, we use angle encoding $R_z(u_{t,j})R_y(u_{t,j})$. The compact circuit definition is

$$\begin{aligned} E(\mathbf{u}_t) &= \bigotimes_{j=1}^{q} R_z(u_{t,j})R_y(u_{t,j}), \\ V_\ell &= \bigotimes_{j=1}^{q} R_z(\theta_{\ell,j,3})R_y(\theta_{\ell,j,2})R_x(\theta_{\ell,j,1}), \\ U_{\mathrm{ent}} &= \prod_{j=1}^{q-1} \mathrm{CNOT}_{j,j+1}, \\ U_\ell^{\mathrm{full}}(\mathbf{u}_t) &= U_{\mathrm{ent}} V_\ell E(\mathbf{u}_t), \end{aligned} \tag{4}$$

where the data-encoding map is re-applied in each of the D variational layers. Starting from an initial encoding, the resulting state and local readout are

$$\begin{aligned} |\psi_t\rangle &= U_D^{\mathrm{full}}(\mathbf{u}_t)\cdots U_1^{\mathrm{full}}(\mathbf{u}_t)E(\mathbf{u}_t)|0\rangle^{\otimes q}, \\ q_{t,j} &= \langle\psi_t|Z_j|\psi_t\rangle, \qquad \mathbf{q}_t = [q_{t,1},\ldots,q_{t,q}]^\top \in [-1,1]^q. \end{aligned} \tag{5}$$

The complete gate-level specification is provided in Appendix D.

We use two independent circuits with the same structural template and separate trainable parameters. Their outputs modulate the effective time scale and nonlinear drift,

$$\boldsymbol{\phi}_{q,\tau}(\mathbf{z}_t) = \mathbf{V}_\tau \mathbf{q}_{t,\tau} + \mathbf{c}_\tau, \qquad \boldsymbol{\phi}_{q,g}(\mathbf{z}_t) = \mathbf{V}_g \mathbf{q}_{t,g} + \mathbf{c}_g. \tag{6}$$

The resulting recurrent dynamics are

$$\begin{aligned} \boldsymbol{\tau}_t^{(q)} &= \sigma\left(\mathbf{W}_{\tau h}\mathbf{h}_{t-1} + \boldsymbol{\phi}_{q,\tau}(\mathbf{z}_t)\right) + \epsilon\mathbf{1}_H, \\ \mathbf{h}_t &= \mathbf{h}_{t-1} + \Delta t_t \left[-\mathbf{h}_{t-1} \oslash \boldsymbol{\tau}_t^{(q)} + \tanh\left(\mathbf{W}_h\mathbf{h}_{t-1} + \boldsymbol{\phi}_{q,g}(\mathbf{z}_t)\right)\right]. \end{aligned} \tag{7}$$

Thus, the PQC does not directly predict the class label. It parameterises a low-dimensional nonlinear perturbation of the vector field governing hidden-state evolution.

Figure 1(a) summarises the resulting architecture. The irregular temporal input is first compressed into a low-dimensional quantum input space, after which two independent parameterised quantum circuits produce modulation features for the effective time scale and nonlinear drift. The quantum outputs therefore enter the recurrent dynamics through structured modulation terms rather than through a direct prediction head.

3.3 Structural characterization of the quantum modulator

The circuit variants differ not merely in gate count, but in the structure of the input-dependent functions they can induce.

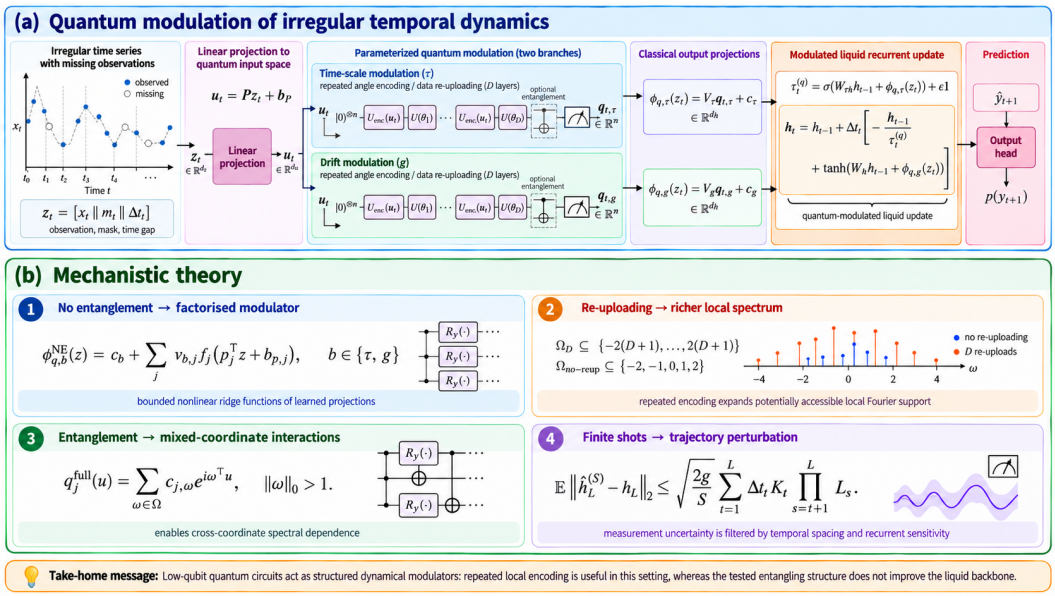

Figure 1: **Overview of the proposed quantum temporal modulation framework and its mechanistic interpretation. (a)** An irregular temporal observation $\mathbf{z}_t = [\mathbf{x}_t \| \mathbf{m}_t \| \Delta t_t]$ is projected into a low-dimensional quantum input space. Two independent low-qubit parameterised quantum circuits then generate features that modulate the effective time scale and nonlinear drift of the liquid recurrent update, rather than directly predicting the class label. **(b)** The structural analysis separates four mechanisms: removing entanglement yields a factorised bank of bounded nonlinear ridge functions; repeated data re-uploading can expand the potentially accessible local Fourier support; entanglement permits mixed-coordinate spectral interactions; and finite-shot measurement perturbations propagate through temporal spacing and recurrent state sensitivity. Detailed definitions and proofs are given in Sections 3.2–3.5 and Appendices C–D.

Proposition 1 (Factorised modulator). If all multi-qubit gates are removed while retaining angle encoding, re-uploading, trainable single-qubit rotations, and local Pauli-Z measurements, then

$$q_j^{\mathrm{NE}}(\mathbf{u}) = f_j(u_j), \qquad \phi_{q,b}^{\mathrm{NE}}(\mathbf{z}) = \mathbf{c}_b + \sum_{j=1}^{q} \mathbf{v}_{b,j} f_j\left(\mathbf{p}_j^\top \mathbf{z} + b_{P,j}\right), \quad b \in \{\tau, g\}, \tag{8}$$

for bounded functions $f_j : \mathbb{R} \to [-1, 1]$. The entanglement-free quantum module is therefore a finite bank of bounded nonlinear ridge functions applied to learned projections of the temporal input. The proof is given in Appendix C.1.

Proposition 2 (Re-uploading expands potentially accessible spectral support). For the repeated $R_z(u)R_y(u)$ encoding used here, each local expectation value is a finite trigonometric polynomial. Under the encoding-count argument in Appendix C.2, the accessible support satisfies

$$\Omega_D \subseteq \{-2(D+1), \ldots, 2(D+1)\}, \qquad \Omega_{\text{no-reup}} \subseteq \{-2, -1, 0, 1, 2\}. \tag{9}$$

These are upper bounds on accessible support: particular frequencies may vanish because of circuit symmetries or parameter choices. Re-uploading therefore expands the set of harmonics that can potentially contribute to each local feature without increasing the number of qubits.

Proposition 3 (Entanglement enables mixed-coordinate dependence). For the full circuit, local observables admit a multivariate Fourier representation

$$q_j^{\mathrm{full}}(\mathbf{u}) = \sum_{\boldsymbol{\omega} \in \Omega} c_{j,\boldsymbol{\omega}} e^{i \boldsymbol{\omega}^\top \mathbf{u}}, \qquad \boldsymbol{\omega} \in \mathbb{Z}^q. \tag{10}$$

Entangling operations can permit terms with $\|\boldsymbol{\omega}\|_0 > 1$, whereas without entanglement and under local readout the j-th feature has support only along its own coordinate, $\boldsymbol{\omega} = k\mathbf{e}_j$. Entangle-

ment can therefore introduce cross-coordinate frequency interactions, whereas re-uploading primarily changes the potentially accessible local spectral support. A proof of this structural distinction is given in Appendix C.3.

3.4 Circuit variants and quantum resources

The theoretical decomposition motivates three controlled circuit variants, all embedded in the same liquid backbone:

$$\begin{aligned} U_\ell^{\text{full}}(\mathbf{u}_t) &= U_{\text{ent}} V_\ell E(\mathbf{u}_t), \\ U_\ell^{\text{no-ent}}(\mathbf{u}_t) &= V_\ell E(\mathbf{u}_t), \\ U_\ell^{\text{no-reup}} &= U_{\text{ent}} V_\ell. \end{aligned} \tag{11}$$

The no-entanglement variant preserves repeated local encoding but removes multi-qubit interactions. The no-reuploading variant retains the variational and entangling layers but injects the classical input only through the initial encoding. The comparison therefore separates two mechanisms: *local spectral enrichment through repeated encoding* and *cross-coordinate interaction through entanglement*.

These variants also have different logical resource requirements. For one branch and temporal step, removing entanglement eliminates $D(q-1)$ CNOT gates, whereas removing re-uploading reduces the number of data-dependent rotations from $2q(D+1)$ to $2q$. Because two independent quantum branches are evaluated across L temporal positions, analytic inference requires $2L$ circuit evaluations and S-shot inference requires

$$N_{\text{prep}} = 2LS \tag{12}$$

state preparations per sequence. These are nominal logical counts rather than hardware-transpiled depth or latency; full accounting is given in Appendix D.5.

3.5 Finite-shot inference and recurrent stability

At inference time, analytic Pauli expectations can be replaced by empirical estimates obtained from S measurements. For local Pauli readout, the estimator is unbiased with componentwise variance at most $1/S$, yielding a characteristic feature-estimation error of order $O(S^{-1/2})$. Training uses analytic expectation values, while finite-shot sampling is introduced only at evaluation time with all learned parameters fixed. We evaluate

$$S \in \{10, 50, 100, 500, \infty\}, \tag{13}$$

where $S = \infty$ denotes analytic evaluation. The estimator derivation is given in Appendix C.4.

Finite-shot error is qualitatively different in a recurrent model from a static predictor because a perturbation introduced at time t changes the state supplied to every later transition. Define

$$\boldsymbol{\xi}_t = [\mathbf{q}_{t,\tau} \| \mathbf{q}_{t,g}] \in \mathbb{R}^{2q}, \qquad \mathbf{h}_t = \mathcal{F}_t(\mathbf{h}_{t-1}, \boldsymbol{\xi}_t).$$

Assume that, over the region traversed by the model,

$$\|\mathcal{F}_t(\mathbf{h}, \boldsymbol{\xi}) - \mathcal{F}_t(\mathbf{h}', \boldsymbol{\xi}')\|_2 \le L_t \|\mathbf{h} - \mathbf{h}'\|_2 + \Delta t_t K_t \|\boldsymbol{\xi} - \boldsymbol{\xi}'\|_2, \tag{14}$$

where L_t measures local state sensitivity and K_t measures sensitivity to the quantum modulation features.

Proposition 4 (Finite-shot trajectory perturbation). Let $\mathbf{h}_L$ denote the analytic trajectory and $\widehat{\mathbf{h}}_L^{(S)}$ the trajectory obtained using S-shot quantum features. Under equation 14,

$$\mathbb{E}\left[\left\|\widehat{\mathbf{h}}_L^{(S)} - \mathbf{h}_L\right\|_2\right] \le \sqrt{\frac{2q}{S}} \sum_{t=1}^{L} \Delta t_t K_t \prod_{s=t+1}^{L} L_s. \tag{15}$$

The proof is given in Appendix C.5.

Equation (15) separates finite-shot sensitivity into three factors: measurement uncertainty $S^{-1/2}$, local temporal spacing Δt_t, and dynamical propagation through subsequent sensitivities. An early perturbation may therefore be attenuated by contractive later transitions or amplified by locally sensitive dynamics. This motivates treating measurement budget as an explicit property of recurrent quantum inference rather than assuming that analytic expectations fully characterise deployment.

3.6 EVALUATION SCOPE

We compare the quantum-modulated models with an MLP, a GRU, and the classical liquid backbone. The MLP and GRU contextualise performance on the overall prediction task, whereas the controlled mechanistic comparison is between the classical liquid model and its quantum-modulated variants. Models are trained with cross-entropy and evaluated using accuracy, macro-F1, negative log-likelihood, Brier score where available, and expected calibration error. Preprocessing, optimisation, and metric definitions are provided in Appendix E.

4 EXPERIMENTS

4.1 EXPERIMENTAL PROTOCOL

We evaluate the proposed framework on the Room Occupancy Estimation dataset from the UCI Machine Learning Repository (Candanedo & Feldheim, 2016; UCI Machine Learning Repository, 2023). The dataset contains $N = 10{,}129$ timestamped observations with $d = 16$ sensor features and a four-class occupancy target. Because the source observations are regularly timestamped, we construct an *irregularised* benchmark through controlled observation thinning rather than treating the original dataset as naturally irregular. Chronologically ordered windows have length $L = 12$ and stride 3; intermediate observations are removed with base probability $\rho = 0.2$, while the final observation is retained. The resulting split contains 2022 training, 500 validation, and 672 held-out test windows. Further construction details are given in Appendix E.

We compare an MLP operating on flattened temporal windows, a GRU processing the augmented irregular sequence, the classical liquid backbone, and three quantum-liquid variants: the full circuit, a no-entanglement circuit, and a no-reuploading circuit. The comparison is intentionally hierarchical. The MLP and GRU contextualise the overall prediction task, whereas the classical liquid model provides the controlled reference for assessing quantum modulation. Quantum models are trained using analytic expectation values. At test time, the trained full quantum model is additionally evaluated with $S \in \{10, 50, 100, 500, \infty\}$, where $S = \infty$ denotes analytic evaluation. We report accuracy, macro-F1, negative log-likelihood (NLL), Brier score where available, and expected calibration error (ECE).

4.2 MAIN RESULTS AND CIRCUIT-COMPONENT ABLATIONS

Table 1 combines the classical reference comparison and the circuit-component ablations. The first result defines the scope of the quantum comparison. The GRU reaches 0.8795 accuracy, NLL 0.5246, and ECE 0.0986, while the MLP reaches 0.8765 accuracy and the largest reported macro-F1, 0.3612. By contrast, the classical liquid backbone reaches 0.3065 accuracy, and the full Quantum Liquid ODE reaches 0.3021. Thus, the full quantum circuit does not repair the weak operating point of the liquid backbone, and the results do not support a claim of predictive quantum advantage over the strong classical references.

Table 1: Main results and circuit-component ablations on the irregularised Room Occupancy benchmark under analytic inference. MLP and GRU contextualise the overall task; the liquid-family rows provide the controlled modulation comparison. Brier score is reported where available. Higher is better for Accuracy and Macro-F1; lower is better for NLL, Brier, and ECE.

Model	Params	Acc. ↑	F1 ↑	NLL ↓	Brier ↓	ECE ↓
GRU	5124	**0.8795**	0.3064	**0.5246**	–	**0.0986**
MLP	18308	0.8765	**0.3612**	0.6340	–	0.1119
Classical Liquid	3396	0.3065	0.1787	10.3855	–	0.6745
Quantum no entanglement	4652	0.4018	0.2093	9.2952	**1.1883**	0.5921
Quantum Liquid ODE, full	4652	0.3021	0.1563	10.4476	1.3732	0.6832
Quantum no re-uploading	4652	0.2649	0.1727	11.4689	1.4595	0.7296

Within the fixed liquid backbone, however, circuit structure has a clear component-level effect. Removing entanglement while retaining data re-uploading increases accuracy from 0.3021 for the full

circuit to 0.4018, and from 0.3065 for the classical liquid model. Macro-F1 increases to 0.2093, while NLL decreases to 9.2952. The probabilistic metrics follow the same direction: relative to the full circuit, Brier score decreases from 1.3732 to 1.1883 and ECE from 0.6832 to 0.5921. Consequently, the observed improvement is not restricted to classification accuracy. Nevertheless, the liquid-family calibration remains substantially worse than that of the GRU and MLP.

Removing re-uploading produces the opposite pattern. The model reaches 0.2649 accuracy, 0.1727 macro-F1, NLL 11.4689, Brier score 1.4595, and ECE 0.7296, worsening most reported metrics relative to the full circuit. This ordering is consistent with the structural characterisation in Section 3.3: the strongest quantum-liquid operating point preserves repeated local encoding while removing the tested entanglement-induced cross-coordinate structure.

This result is task- and architecture-specific. It does not establish that entanglement is generically detrimental, nor that the improvement is uniquely quantum. Rather, it identifies which component of the tested circuit coincides with the improved operating point. The same variant also removes $D(q-1)$ CNOT gates per branch and temporal step; the corresponding resource accounting is given in Section 3.4 and Appendix D.5.

4.3 Finite-shot sensitivity

We next evaluate how replacing analytic quantum expectations with finite-shot estimates changes the full Quantum Liquid ODE while holding the trained parameters fixed. Table 2 reports the complete sweep.

The observed response to shot count is non-monotonic. Accuracy changes from 0.2976 at $S = 10$ to 0.2842 at $S = 50$, 0.3021 at $S = 100$, and 0.3125 at $S = 500$, compared with 0.3021 under analytic evaluation. Macro-F1, NLL, and ECE also vary across the finite-shot operating points.

These differences should be interpreted as *resource sensitivity*, not as evidence that a particular finite shot budget improves generalisation. The experiment does not contain repeated training seeds and independent shot realisations sufficient to estimate uncertainty around these numerical differences. In particular, the stronger point estimates at $S = 500$ do not establish finite-shot regularisation.

The result is consistent with Proposition 4 in a narrower sense. Although the variance of each local quantum estimator decreases as $O(S^{-1})$, its downstream effect is transformed by the recurrent trajectory. Measurement perturbations enter at different temporal positions, are weighted by Δt_t, and can be amplified or attenuated by subsequent state transitions. Analytic and finite-shot inference should therefore be regarded as distinct operating regimes for a recurrent quantum model.

4.4 Robustness to stronger temporal sparsification

Finally, we test whether making the observation process increasingly sparse qualitatively changes the comparison. Test sequences are reconstructed at $\rho \in \{0.0, 0.2, 0.4, 0.6\}$, and the full Quantum Liquid ODE is evaluated with both analytic expectations and 100-shot estimates.

The GRU remains numerically stable across the sweep, with accuracy between 0.8780 and 0.8854 and ECE between approximately 0.094 and 0.101. The full quantum model remains substantially weaker under every sparsification level and under both measurement regimes. No monotonic advantage for the full circuit emerges as intermediate observations are removed.

This experiment therefore does not support a simple explanation in which the full quantum architecture becomes advantageous merely because the temporal sequence is made more sparse. The scope of this conclusion is limited: the sparsification sweep evaluates the full circuit rather than all circuit ablations, so it does not establish how the no-entanglement variant behaves under increasing sparsity.

5 Conclusion

We studied low-qubit quantum circuits as structured modulators of liquid temporal dynamics rather than as stand-alone sequence predictors. The analysis shows that different circuit components induce qualitatively different function classes: without entanglement, local readout reduces the quantum

Table 2: Finite-shot inference for the full Quantum Liquid ODE. $S = \infty$ denotes analytic expectation evaluation. Values are point estimates and should not be interpreted as statistically ordered without repeated independent measurement realisations.

	$S = 10$	$S = 50$	$S = 100$	$S = 500$	$S = \infty$
Accuracy ↑	0.2976	0.2842	0.3021	0.3125	0.3021
Macro-F1 ↑	0.1430	0.1603	0.1658	0.1678	0.1563
NLL ↓	10.8704	10.8090	10.5782	10.4414	10.4476
ECE ↓	0.6903	0.6950	0.6869	0.6763	0.6832

Table 3: Robustness under controlled temporal sparsification. The drop probability ρ controls the removal of intermediate observations. The full quantum model is evaluated analytically and with 100 shots.

Model / Metric	$\rho = 0.0$	$\rho = 0.2$	$\rho = 0.4$	$\rho = 0.6$
GRU Accuracy	0.8824	0.8780	0.8854	0.8839
Quantum Accuracy, analytic	0.2887	0.3438	0.2768	0.2902
Quantum Accuracy, 100 shots	0.3006	0.3378	0.2723	0.2976
GRU ECE	0.0958	0.1008	0.0944	0.0945
Quantum ECE, analytic	0.6205	0.6468	0.7219	0.7098
Quantum ECE, 100 shots	0.6153	0.6479	0.7262	0.7024

modulator to bounded nonlinear ridge functions of learned input projections; data re-uploading increases the potentially accessible local spectral support; and entanglement enables mixed-coordinate interactions. Finite-shot measurement introduces a complementary resource effect, with local $O(S^{-1/2})$ estimation perturbations propagated through temporal spacing and recurrent state sensitivity.

The experiments are consistent with this mechanistic decomposition. Strong MLP and GRU references remain substantially better overall predictors on the irregularised Room Occupancy benchmark, so the results do not support a claim of predictive quantum advantage. Within the fixed liquid backbone, however, removing entanglement while retaining data re-uploading produces the strongest quantum-family operating point and improves accuracy, macro-F1, NLL, and ECE relative to both the classical liquid model and the full entangled circuit. Removing re-uploading worsens most metrics. Finite-shot inference additionally changes the recurrent operating point, but the available point estimates are best interpreted as resource sensitivity rather than evidence that measurement noise improves generalisation.

These results support a simple design principle: the value of a low-qubit quantum component depends on the alignment between its induced function class, the surrounding classical dynamics, and the available measurement resources, rather than on circuit complexity alone. The present evidence is limited to a single synthetically irregularised benchmark, a comparatively weak liquid backbone, and point-estimate evaluation without repeated training and measurement realisations or device-level noise. Extended discussion, limitations, and future directions are provided in Appendix G.

AI USE STATEMENT

In this work, we used generative AI tools for proofreading the manuscript. We have not used generative AI tools for any other tasks. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REFERENCES

Amira Abbas, David Sutter, Christa Zoufal, Aurelien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1:403–409, 2021. doi: 10.1038/s43588-021-00084-1.

Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549:195–202, 2017. doi: 10.1038/nature23474.

Luis M. Candanedo and Veronique Feldheim. Accurate occupancy detection of an office room from light, temperature, humidity and co2 measurements using statistical learning models. *Energy and Buildings*, 112:28–39, 2016. doi: 10.1016/j.enbuild.2015.11.071.

M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3 (9):625–644, 2021. doi: 10.1038/s42254-021-00348-9.

Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1):6085, 2018. doi: 10.1038/s41598-018-24271-9.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems 31*, pp. 6571–6583, 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/hash/69386f6bb1dfed68692a24c8686939b9-Abstract.html.

Yuqi Chen, Kan Ren, Yansen Wang, Yuchen Fang, Weiwei Sun, and Dongsheng Li. Contiformer: Continuous-time transformer for irregular time series modeling. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-2042. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/9328208f88ec69420031647e6ff97727-Abstract-Conference.html.

Edward De Brouwer, Jaak Simm, Adam Arany, and Yves Moreau. GRU-ODE-Bayes: Continuous modeling of sporadically-observed time series. In *Advances in Neural Information Processing Systems*, volume 32, pp. 7377–7388, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/455cb2657aaa59e32fad80cb0b65b9dc-Abstract.html.

Ramin Hasani, Mathias Lechner, Alexander Amini, Daniela Rus, and Radu Grosu. Liquid time-constant networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 7657–7666, 2021. doi: 10.1609/aaai.v35i9.16936.

Ramin Hasani, Mathias Lechner, Alexander Amini, Lucas Liebenwein, Aaron Ray, Max Tschaikowski, Gerald Teschl, and Daniela Rus. Closed-form continuous-time neural networks. *Nature Machine Intelligence*, 4(11):992–1003, 2022. doi: 10.1038/s42256-022-00556-7.

Vojtech Havlicek, Antonio D. Corcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212, 2019. doi: 10.1038/s41586-019-0980-2.

Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R. McClean. Power of data in quantum machine learning. *Nature Communications*, 12:2631, 2021. doi: 10.1038/s41467-021-22539-9.

Sofiene Jerbi, Lukas J. Fiderer, Hendrik Poulsen Nautrup, Jonas M. Kübler, Hans J. Briegel, and Vedran Dunjko. Quantum machine learning beyond kernel methods. *Nature Communications*, 14:517, 2023. doi: 10.1038/s41467-023-36159-y.

Patrick Kidger, James Morrill, James Foster, and Terry Lyons. Neural controlled differential equations for irregular time series. In *Advances in Neural Information Processing Systems 33*, pp. 6696–6707, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/4a5876b450b45371f6cfe5047ac8cd45-Abstract.html.

Jonas M. Kübler, Simon Buchholz, and Bernhard Schölkopf. The inductive bias of quantum kernels. In *Advances in Neural Information Processing Systems*, volume 34, pp. 12661–12673, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/69adc1e107f7f7d035d7baf04342e1ca-Abstract.html.

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9:4812, 2018. doi: 10.1038/s41467-018-07090-4.

James Morrill, Cristopher Salvi, Patrick Kidger, and James Foster. Neural rough differential equations for long time series. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 7829–7838. PMLR, 2021. URL https://proceedings.mlr.press/v139/morrill21b.html.

Daniel Neil, Michael Pfeiffer, and Shih-Chii Liu. Phased lstm: Accelerating recurrent network training for long or event-based sequences. In *Advances in Neural Information Processing Systems*, volume 29, pp. 3882–3890, 2016. URL https://proceedings.neurips.cc/paper/2016/hash/5bce843dd76db8c939d5323dd3e54ec9-Abstract.html.

Adrian Perez-Salinas, Alba Cervera-Lierta, Elies Gil-Fuster, and Jose I. Latorre. Data re-uploading for a universal quantum classifier. *Quantum*, 4:226, 2020. doi: 10.22331/q-2020-02-06-226.

Yulia Rubanova, Ricky T. Q. Chen, and David K. Duvenaud. Latent ordinary differential equations for irregularly-sampled time series. In *Advances in Neural Information Processing Systems 32*, pp. 5321–5331, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/42a6845a557bef704ad8ac9cb4461d43-Abstract.html.

Maria Schuld and Nathan Killoran. Quantum machine learning in feature hilbert spaces. *Physical Review Letters*, 122(4):040504, 2019. doi: 10.1103/PhysRevLett.122.040504.

Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3):032430, 2021. doi: 10.1103/PhysRevA.103.032430.

Satya Narayan Shukla and Benjamin M. Marlin. Multi-time attention networks for irregularly sampled time series. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=4c0D2lgzZK.

Sukin Sim, Peter D. Johnson, and Alan Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070, 2019. doi: 10.1002/qute.201900070.

UCI Machine Learning Repository. Room occupancy estimation. UCI Machine Learning Repository, 2023. URL https://archive.ics.uci.edu/dataset/864/room+occupancy+estimation. Accessed 2026-05-06.

Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J. Coles. Noise-induced barren plateaus in variational quantum algorithms. *Nature Communications*, 12:6961, 2021. doi: 10.1038/s41467-021-27045-6.

APPENDIX

A STRUCTURAL DETAILS FOR QUANTUM TEMPORAL MODULATION

This appendix expands the structural perspective introduced in Section 1. The purpose is to make explicit the functional forms that motivate the circuit comparisons while keeping the main-text Introduction compact. Formal propositions and their use in the model analysis are developed in Section 3.3.

A.1 IRREGULAR TEMPORAL INPUT

At temporal position t, the model receives the augmented input

$$\mathbf{z}_t = [\mathbf{x}_t \,\|\, \mathbf{m}_t \,\|\, \Delta t_t] \in \mathbb{R}^{2d+1}, \tag{16}$$

where $\mathbf{x}_t \in \mathbb{R}^d$ denotes the observed or imputed feature vector, $\mathbf{m}_t \in \{0,1\}^d$ is the corresponding observation mask, and $\Delta t_t \in \mathbb{R}_+$ is the elapsed time since the preceding retained observation. This representation exposes both observed values and the temporal structure of observation availability to the recurrent dynamics.

The classical input is projected into the q-dimensional quantum input space according to

$$\mathbf{u} = \mathbf{P}\mathbf{z} + \mathbf{b}_P \in \mathbb{R}^q, \tag{17}$$

where the j-th projected coordinate is

$$u_j = \mathbf{p}_j^\top \mathbf{z} + b_{P,j}. \tag{18}$$

Here, $\mathbf{p}_j^\top$ denotes the j-th row of $\mathbf{P}$.

A.2 Factorised ridge-function representation

Consider the entanglement-free circuit with local Pauli-Z readout. Since all circuit operations are tensor products of single-qubit operations, the resulting state remains separable:

$$\left|\psi^{\mathrm{NE}}(\mathbf{u})\right\rangle = \bigotimes_{j=1}^{q} \left|\psi_j(u_j)\right\rangle . \tag{19}$$

Consequently, the local expectation value measured on qubit j depends only on the corresponding scalar coordinate,

$$q_j^{\mathrm{NE}}(\mathbf{u}) = \langle \psi_j(u_j)| \, Z \, |\psi_j(u_j)\rangle = f_j(u_j), \tag{20}$$

where $f_j : \mathbb{R} \to [-1, 1]$ is the bounded nonlinear function induced by the j-th single-qubit circuit.

After the classical output projection, a modulation branch $b \in \{\tau, g\}$ therefore takes the form

$$\boldsymbol{\phi}_{q,b}^{\mathrm{NE}}(\mathbf{z}) = \mathbf{c}_b + \sum_{j=1}^{q} \mathbf{v}_{b,j} f_j \left(\mathbf{p}_j^\top \mathbf{z} + b_{P,j}\right), \qquad b \in \{\tau, g\}, \tag{21}$$

where $\mathbf{v}_{b,j}$ is the j-th column of the output projection matrix $\mathbf{V}_b$.

Equation (21) makes explicit that, without entanglement, the quantum temporal modulator is a finite sum of bounded nonlinear ridge functions applied to learned projections of the irregular input.

A.3 Fourier interpretation of data re-uploading

For the angle encoding used in this work, a projected scalar input enters through Pauli-generated rotation gates. The corresponding expectation value can therefore be expressed as a finite trigonometric polynomial,

$$f_j(u) = \sum_{k \in \Omega_D} c_{j,k} e^{iku}, \tag{22}$$

or, equivalently,

$$f_j(u) = a_{j,0} + \sum_{k \in \Omega_D^+} \left[a_{j,k} \cos(ku) + b_{j,k} \sin(ku)\right]. \tag{23}$$

For the circuit considered in this work, the scalar input occurs in two rotation gates in the initial encoding and in two additional rotation gates in each of the D re-uploading layers. The accessible frequency set therefore satisfies the upper bound

$$\Omega_D \subseteq \{-2(D+1), \ldots, 2(D+1)\}. \tag{24}$$

If data are encoded only once, the corresponding support is restricted to

$$\Omega_{\text{no-reup}} \subseteq \{-2, -1, 0, 1, 2\}. \tag{25}$$

The exact realised subset of frequencies depends on the trainable parameters and circuit symmetries. Equations equation 24 and equation 25 therefore describe accessible rather than necessarily non-zero frequency components. Their structural implication is that repeated input injection enlarges the family of harmonics that can contribute to each local quantum feature without increasing the qubit count.

A.4 MIXED-COORDINATE STRUCTURE INDUCED BY ENTANGLEMENT

Entangling gates qualitatively change the input dependence of local measurements. Once different encoded qubits interact, the j-th local expectation value can depend jointly on several coordinates of $\mathbf{u}$. Its input dependence can be written in the multivariate form

$$q_j^{\text{full}}(\mathbf{u}) = \sum_{\boldsymbol{\omega}\in\Omega} c_{j,\boldsymbol{\omega}} e^{i\boldsymbol{\omega}^\top \mathbf{u}}, \qquad \boldsymbol{\omega} \in \mathbb{Z}^q. \tag{26}$$

The entangling circuit can contain terms whose frequency vectors satisfy

$$\|\boldsymbol{\omega}\|_0 > 1, \tag{27}$$

so that a single measured observable depends on more than one projected input coordinate.

By contrast, in the entanglement-free circuit with local readout, the frequency support of the j-th observable is restricted to vectors of the form

$$\boldsymbol{\omega} = k\mathbf{e}_j, \tag{28}$$

where $\mathbf{e}_j$ is the j-th coordinate basis vector.

This decomposition separates the two circuit mechanisms considered in the main paper. Data re-uploading enlarges the accessible frequency content of local coordinate functions, whereas entanglement permits mixed-coordinate interactions in the induced quantum feature map.

A.5 FINITE-SHOT PERTURBATIONS IN RECURRENT DYNAMICS

Let

$$\widehat{q}_{t,j}^{(S)} = \frac{1}{S}\sum_{s=1}^{S} z_{t,j}^{(s)}, \qquad z_{t,j}^{(s)} \in \{-1,+1\}, \tag{29}$$

denote the S-shot estimator of a Pauli-Z_j expectation value. The estimator satisfies

$$\mathbb{E}\left[\widehat{q}_{t,j}^{(S)}\right] = q_{t,j}, \tag{30}$$

and

$$\mathrm{Var}\left[\widehat{q}_{t,j}^{(S)}\right] = \frac{1-q_{t,j}^2}{S} \le \frac{1}{S}. \tag{31}$$

Thus, the characteristic feature-estimation error decreases as $O(S^{-1/2})$. In a recurrent model, however, a perturbation introduced at time t alters the state used by all subsequent updates. To express this dependence, let

$$\boldsymbol{\xi}_t = [\mathbf{q}_{t,\tau} \,\|\, \mathbf{q}_{t,g}] \in \mathbb{R}^{2q} \tag{32}$$

collect the two quantum modulation branches, and write one recurrent transition abstractly as

$$\mathbf{h}_t = \mathcal{F}_t\left(\mathbf{h}_{t-1}, \boldsymbol{\xi}_t\right). \tag{33}$$

Suppose the transition is locally Lipschitz in both arguments,

$$\|\mathcal{F}_t(\mathbf{h},\boldsymbol{\xi}) - \mathcal{F}_t(\mathbf{h}',\boldsymbol{\xi}')\|_2 \le L_t \|\mathbf{h}-\mathbf{h}'\|_2 + \Delta t_t K_t \|\boldsymbol{\xi}-\boldsymbol{\xi}'\|_2, \tag{34}$$

where L_t characterises local state sensitivity and K_t characterises sensitivity to the quantum modulation feature.

Under this assumption, the analytic trajectory $\mathbf{h}_L$ and the finite-shot trajectory $\widehat{\mathbf{h}}_L^{(S)}$ satisfy the bound

$$\mathbb{E}\left[\left\|\widehat{\mathbf{h}}_L^{(S)} - \mathbf{h}_L\right\|_2\right] \le \sqrt{\frac{2q}{S}} \sum_{t=1}^{L} \Delta t_t K_t \prod_{r=t+1}^{L} L_r. \tag{35}$$

Equation (35) separates three factors governing finite-shot sensitivity:

$$\underbrace{S^{-1/2}}_{\text{measurement uncertainty}} \qquad \underbrace{\Delta t_t}_{\text{temporal spacing}} \qquad \underbrace{\prod_{r=t+1}^{L} L_r}_{\text{dynamical propagation}}. \tag{36}$$

Consequently, the shot budget alone does not determine the downstream effect of quantum measurement noise. A perturbation can be amplified or attenuated depending on the subsequent recurrent dynamics, while irregular temporal spacing directly scales the instantaneous effect of the perturbed drift. A formal statement and proof of the trajectory perturbation result are given in Section 3.5 and its corresponding proof appendix.

B Extended Related Work and Positioning

This appendix provides a more detailed discussion of the literature underlying the compact review in Section 2. We separate prior work on irregular temporal learning, quantum feature maps, circuit expressivity, finite-shot learning, and hybrid quantum architectures before positioning the present study relative to these directions.

B.1 Irregular temporal learning and continuous-time dynamics

Irregular temporal learning concerns sequential data for which observations may be missing, asynchronous, or separated by non-uniform time intervals. In such settings, the observation process itself can be informative: both missingness patterns and elapsed times may correlate with the latent state being modelled. Recurrent approaches such as GRU-D explicitly incorporate observation masks and time gaps into the state update, while related architectures use time-dependent gating or attention mechanisms to account for irregular sampling (Che et al., 2018). These methods motivate representing an observation jointly with its availability indicator and temporal spacing rather than treating missingness as a preprocessing problem.

A complementary family replaces fixed-step recurrence with explicitly continuous-time latent dynamics. Neural ODEs parameterise hidden trajectories through learned differential equations, while Latent ODEs and ODE–RNN constructions adapt this framework to partially observed and irregularly sampled sequences (Chen et al., 2018; Rubanova et al., 2019). Neural controlled differential equations instead represent sequential data as continuous paths that drive the latent dynamics (Kidger et al., 2020). These formulations are particularly natural when the intervals between measurements carry predictive information.

Liquid neural architectures provide another continuous-time perspective. Liquid Time-Constant Networks introduce state- and input-dependent time constants controlling how rapidly hidden units respond to new information, while Closed-form Continuous-time Networks derive efficient recurrent updates that retain continuous-time inductive structure (Hasani et al., 2021; 2022). The separation between an effective time scale and a nonlinear drift term makes this family particularly suitable for modulation by an auxiliary function.

The present work uses this liquid-style structure as a controlled dynamical backbone. Our objective is not to propose a replacement for existing continuous-time models, but to study how an auxiliary low-qubit quantum function changes the local vector field and effective response rate of the backbone.

B.2 Parameterized quantum circuits as structured function classes

Parameterized quantum circuits are widely studied as nonlinear components of hybrid quantum–classical learning systems. Quantum feature-map methods encode classical data into quantum states whose measurements or kernel evaluations provide nonlinear representations (Havlicek et al., 2019). Trainable circuit models combine input-dependent gates, variational parameters, and differentiable measurements to produce learned quantum transformations.

From this perspective, a PQC is not characterised only by its number of parameters. Its induced function class is jointly determined by the input encoding, the arrangement of variational gates, the circuit connectivity, and the observables used for readout. This functional viewpoint is particularly important in low-qubit settings, where architectural structure rather than system width determines much of the available input dependence.

Data re-uploading repeatedly inserts classical variables into a quantum circuit, allowing a shallow low-qubit model to generate a richer family of nonlinear input transformations (Perez-Salinas et al.,

2020). Fourier analyses of quantum data-encoding circuits further show that expectation values can be expressed as finite trigonometric expansions whose accessible frequencies are determined by the encoding generators and their repetition (Schuld et al., 2021). Repeated encoding can therefore be interpreted as modifying the spectral structure of the induced function rather than merely increasing circuit depth.

The present work uses this perspective to distinguish two circuit mechanisms. In the entanglement-free architecture, local readout produces independent nonlinear functions of learned projected coordinates, whereas entangling gates can couple these coordinates and introduce multivariate dependence. The formal characterisation used in our analysis is given in Section 3.3 and Appendix A.

B.3 Expressivity, entanglement, and trainability

Variational quantum learning has long highlighted a tension between expressive capacity and practical trainability. More expressive ansätze can represent larger families of quantum states or functions, but additional expressivity does not imply improved downstream prediction (Sim et al., 2019; Abbas et al., 2021). Entangling layers, in particular, increase the possible multi-qubit correlations while simultaneously affecting circuit depth, optimisation geometry, and hardware requirements.

Trainability introduces an additional constraint. Parameterized circuits can exhibit barren plateaus or gradient concentration under certain depth, width, initialisation, and architectural regimes (McClean et al., 2018; Cerezo et al., 2021). Analytic differentiation and parameter-shift techniques make PQCs compatible with gradient-based optimisation, but the resulting derivatives may still require repeated quantum circuit evaluations.

These observations motivate separating *expressive structure* from *useful structure*. In our setting, entanglement and re-uploading change different aspects of the modulation function and are therefore evaluated as distinct architectural interventions rather than as a simple ordering from less expressive to more expressive circuits.

B.4 Finite-shot estimation and resource-aware hybrid learning

The resource requirements of a variational quantum model extend beyond its trainable parameter count. Practical cost depends on the number of qubits, circuit depth, entangling operations, circuit executions, and the number of measurements required to estimate observables (Biamonte et al., 2017; Cerezo et al., 2021).

For an observable estimated from S independent measurements, the characteristic statistical error decreases at the Monte Carlo rate $O(S^{-1/2})$. In a static hybrid model, this uncertainty affects a single feature or output estimate. When the same quantum module is repeatedly evaluated inside a recurrent architecture, however, an estimation error can modify the hidden state supplied to all subsequent transitions.

The resulting resource question is therefore dynamical. Measurement uncertainty must be considered together with the local sensitivity of the recurrent system and the temporal spacing at which quantum perturbations are introduced. Our finite-shot analysis formalises this coupling in Section 3.5, with supporting details in Appendix A.5.

We consequently treat shot count as an explicit inference resource rather than assuming that analytic expectation values fully characterise deployment. The finite-shot experiments are interpreted as resource-sensitivity measurements, not as evidence that sampling noise is intrinsically beneficial.

B.5 Hybrid quantum architectures as dynamical modulators

Hybrid quantum–classical models commonly embed a quantum circuit as a submodule inside a larger classical computation graph. Depending on the architecture, the quantum component may act as a feature extractor, nonlinear bottleneck, kernel-like map, or intermediate trainable representation (Havlicek et al., 2019). Such modular designs are attractive under resource constraints because only a small part of the learning pipeline needs to be implemented quantum mechanically.

Our setting differs in the role assigned to this submodule. The circuit output is not directly interpreted as a class score or static latent embedding. Instead, it changes the effective time scale and

nonlinear drift of a recurrent continuous-time update. The PQC therefore acts on the *dynamics* of the classical model.

This distinction changes how circuit structure should be evaluated. A local nonlinear perturbation may influence an entire later trajectory, while a richer interaction term may be unhelpful if it is misaligned with the structure required by the temporal task. Circuit usefulness is therefore determined jointly by the induced function class and by the dynamical system into which that function is inserted.

B.6 Positioning of the present work

The present study lies at the intersection of irregular continuous-time sequence modelling, variational quantum function approximation, and resource-aware hybrid learning. Prior work has developed temporal models for irregular observations (Che et al., 2018; Chen et al., 2018; Rubanova et al., 2019; Kidger et al., 2020; Hasani et al., 2021; 2022), established PQCs as trainable nonlinear feature maps (Havlicek et al., 2019), and characterised how encoding, expressivity, trainability, and measurement constraints shape variational quantum models (Perez-Salinas et al., 2020; Schuld et al., 2021; Sim et al., 2019; Cerezo et al., 2021).

The combination studied here differs primarily in the role played by the quantum circuit. A low-qubit PQC repeatedly modulates the vector field of an irregular recurrent dynamical system. This creates three coupled questions: which function class is induced by local re-uploading and entangling structure, how that structure affects a repeatedly applied temporal update, and how finite-shot uncertainty propagates through the resulting trajectory.

Our empirical comparisons are interpreted through this mechanistic lens. The goal is neither to establish unconditional quantum advantage nor to treat circuit complexity as an end in itself. Instead, the analysis asks which quantum-induced function structure is useful for the fixed temporal backbone and at what measurement and circuit-resource cost.

C Proofs and Theoretical Details

This appendix provides the proofs and supporting derivations for Propositions 1–4 in Section 3. Gate-level circuit definitions are collected separately in Appendix D.

C.1 Proof of Proposition 1: factorised modulator

In the absence of entangling operations, every circuit layer is a tensor product of single-qubit unitaries. Writing the complete single-qubit operation on qubit j as $U_j(u_j)$, the joint unitary has the form

$$U^{\mathrm{NE}}(\mathbf{u}) = \bigotimes_{j=1}^{q} U_j(u_j). \tag{37}$$

Starting from $|0\rangle^{\otimes q}$, the final state therefore remains separable:

$$|\psi^{\mathrm{NE}}(\mathbf{u})\rangle = \bigotimes_{j=1}^{q} |\psi_j(u_j)\rangle, \qquad |\psi_j(u_j)\rangle = U_j(u_j)|0\rangle. \tag{38}$$

For the local observable $Z_j = I^{\otimes(j-1)} \otimes Z \otimes I^{\otimes(q-j)}$,

$$\begin{aligned} q_j^{\mathrm{NE}}(\mathbf{u}) &= \langle \psi^{\mathrm{NE}}(\mathbf{u}) | Z_j | \psi^{\mathrm{NE}}(\mathbf{u}) \rangle \\ &= \langle \psi_j(u_j) | Z | \psi_j(u_j) \rangle = f_j(u_j). \end{aligned} \tag{39}$$

Because Z has eigenvalues $\{-1, +1\}$,

$$|f_j(u_j)| \le 1. \tag{40}$$

Using

$$u_j = \mathbf{p}_j^\top \mathbf{z} + b_{P,j} \tag{41}$$

and $\boldsymbol{\phi}_{q,b} = \mathbf{V}_b\mathbf{q} + \mathbf{c}_b$, we obtain

$$\boldsymbol{\phi}^{\mathrm{NE}}_{q,b}(\mathbf{z}) = \mathbf{c}_b + \sum_{j=1}^{q} \mathbf{v}_{b,j} f_j\left(\mathbf{p}_j^\top \mathbf{z} + b_{P,j}\right), \qquad b \in \{\tau, g\}, \tag{42}$$

which proves Proposition 1.

C.2 Derivation for Proposition 2: accessible Fourier support

For the convention

$$R_\alpha(u) = \exp\left(-\frac{i}{2} u \sigma_\alpha\right), \qquad \alpha \in \{x, y, z\}, \tag{43}$$

the matrix elements of an encoding rotation are linear combinations of $e^{iu/2}$ and $e^{-iu/2}$. Repeated appearances of the same scalar input therefore generate amplitudes whose exponents are sums of the corresponding generator spectral gaps. Expectation values are bilinear in these amplitudes, so the resulting observable is a finite trigonometric polynomial,

$$f(u) = \sum_{k \in \Omega} c_k e^{iku}. \tag{44}$$

For the present encoding,

$$E(u) = R_z(u) R_y(u), \tag{45}$$

the scalar u appears twice per encoding stage. With one initial encoding and D re-uploading stages, the rotation-count argument gives the upper bound

$$\Omega_D \subseteq \{-2(D+1), \ldots, 2(D+1)\}. \tag{46}$$

If the input is encoded only once,

$$\Omega_{\text{no-reup}} \subseteq \{-2, -1, 0, 1, 2\}. \tag{47}$$

These expressions characterise potentially accessible support rather than guaranteeing a non-zero coefficient at every listed frequency. Circuit symmetries and parameter choices can remove individual components. Under the stated encoding convention, repeated input injection therefore expands the candidate harmonic support available to the local feature.

C.3 Proof of Proposition 3: mixed-coordinate dependence

Without entanglement, Proposition 1 gives

$$q_j^{\mathrm{NE}}(\mathbf{u}) = f_j(u_j) = \sum_k c_{j,k} e^{iku_j}, \tag{48}$$

so the multivariate frequency vectors are restricted to

$$\boldsymbol{\omega} = k\mathbf{e}_j. \tag{49}$$

With entangling operations, conjugating a local observable through the circuit can produce Pauli strings with support on multiple qubits. Schematically,

$$U_{\mathrm{ent}}^\dagger Z_j U_{\mathrm{ent}} = \sum_{\boldsymbol{\alpha}} \gamma_{\boldsymbol{\alpha}} P_{\alpha_1} \otimes \cdots \otimes P_{\alpha_q}, \tag{50}$$

where $P_{\alpha_k} \in \{I, X, Y, Z\}$. Expectation values of multi-qubit terms can contain products of coordinate-dependent trigonometric functions,

$$\prod_{r=1}^{q} f_r(u_r) = \sum_{\boldsymbol{\omega}} c_{\boldsymbol{\omega}} e^{i\boldsymbol{\omega}^\top \mathbf{u}}, \tag{51}$$

and hence may contain

$$\|\boldsymbol{\omega}\|_0 > 1. \tag{52}$$

Entanglement therefore permits mixed-coordinate spectral terms that are absent from the factorised local-readout circuit. It does not imply that such terms must have non-zero coefficients for every entangled parameter setting.

C.4 FINITE-SHOT ESTIMATOR

Let

$$z_{t,j}^{(s)} \in \{-1, +1\} \tag{53}$$

denote the outcome of shot s for Pauli-Z_j. The S-shot estimator is

$$\widehat{q}_{t,j}^{(S)} = \frac{1}{S} \sum_{s=1}^{S} z_{t,j}^{(s)}. \tag{54}$$

Since $\mathbb{E}[z_{t,j}^{(s)}] = q_{t,j}$,

$$\mathbb{E}\left[\widehat{q}_{t,j}^{(S)}\right] = q_{t,j}, \tag{55}$$

and

$$\mathrm{Var}\left[\widehat{q}_{t,j}^{(S)}\right] = \frac{1 - q_{t,j}^2}{S} \leq \frac{1}{S}. \tag{56}$$

For one q-dimensional branch,

$$\mathbb{E}\left[\left\|\widehat{\mathbf{q}}_t^{(S)} - \mathbf{q}_t\right\|_2^2\right] \leq \frac{q}{S}. \tag{57}$$

C.5 PROOF OF PROPOSITION 4: FINITE-SHOT TRAJECTORY PERTURBATION

Let

$$\delta_t = \left\|\widehat{\mathbf{h}}_t^{(S)} - \mathbf{h}_t\right\|_2. \tag{58}$$

Applying the local sensitivity condition in equation 14 gives

$$\delta_t \leq L_t \delta_{t-1} + \Delta t_t K_t \left\|\widehat{\boldsymbol{\xi}}_t^{(S)} - \boldsymbol{\xi}_t\right\|_2. \tag{59}$$

Since $\delta_0 = 0$, recursive expansion yields

$$\delta_L \leq \sum_{t=1}^{L} \Delta t_t K_t \left(\prod_{s=t+1}^{L} L_s\right) \left\|\widehat{\boldsymbol{\xi}}_t^{(S)} - \boldsymbol{\xi}_t\right\|_2. \tag{60}$$

Because $\boldsymbol{\xi}_t$ concatenates two q-dimensional measurement branches,

$$\mathbb{E}\left[\left\|\widehat{\boldsymbol{\xi}}_t^{(S)} - \boldsymbol{\xi}_t\right\|_2^2\right] \leq \frac{2q}{S}. \tag{61}$$

Jensen's inequality then gives

$$\mathbb{E}\left[\left\|\widehat{\boldsymbol{\xi}}_t^{(S)} - \boldsymbol{\xi}_t\right\|_2\right] \leq \sqrt{\frac{2q}{S}}. \tag{62}$$

Taking expectations in equation 60 therefore produces

$$\mathbb{E}\left[\left\|\widehat{\mathbf{h}}_L^{(S)} - \mathbf{h}_L\right\|_2\right] \leq \sqrt{\frac{2q}{S}} \sum_{t=1}^{L} \Delta t_t K_t \prod_{s=t+1}^{L} L_s, \tag{63}$$

which proves Proposition 4.

D FULL ARCHITECTURE AND RESOURCE ACCOUNTING

This appendix provides the gate-level specification corresponding to the compact circuit definition in Section 3.2 and the resource accounting used in Section 3.4.

D.1 Input projection and encoding

The augmented temporal input is projected as

$$\mathbf{u}_t = \mathbf{P}\mathbf{z}_t + \mathbf{b}_P \in \mathbb{R}^q. \tag{64}$$

For qubit j, the local encoding is

$$E_j(u_{t,j}) = R_z(u_{t,j})R_y(u_{t,j}), \tag{65}$$

and the product encoding is

$$E(\mathbf{u}_t) = \bigotimes_{j=1}^{q} E_j(u_{t,j}). \tag{66}$$

The same encoding map is re-applied within each variational layer for the re-uploading variants.

D.2 Variational and entangling layers

The trainable single-qubit operation is

$$V_{\ell,j} = R_z(\theta_{\ell,j,3})R_y(\theta_{\ell,j,2})R_x(\theta_{\ell,j,1}), \tag{67}$$

with

$$V_\ell = \bigotimes_{j=1}^{q} V_{\ell,j}. \tag{68}$$

The full circuit uses nearest-neighbour entanglement,

$$U_{\text{ent}} = \prod_{j=1}^{q-1} \text{CNOT}_{j,j+1}, \tag{69}$$

and one full layer is

$$U_\ell^{\text{full}}(\mathbf{u}_t) = U_{\text{ent}} V_\ell E(\mathbf{u}_t). \tag{70}$$

After D layers,

$$|\psi_t\rangle = U_D^{\text{full}}(\mathbf{u}_t) \cdots U_1^{\text{full}}(\mathbf{u}_t) E(\mathbf{u}_t)|0\rangle^{\otimes q}. \tag{71}$$

D.3 Measurement and recurrent modulation

The measured feature on qubit j is

$$q_{t,j} = \langle\psi_t|Z_j|\psi_t\rangle, \qquad j = 1, \ldots, q, \tag{72}$$

and

$$\mathbf{q}_t = [q_{t,1}, \ldots, q_{t,q}]^\top. \tag{73}$$

Two independent circuit instances produce

$$\boldsymbol{\phi}_{q,\tau}(\mathbf{z}_t) = \mathbf{V}_\tau \mathbf{q}_{t,\tau} + \mathbf{c}_\tau, \tag{74}$$

$$\boldsymbol{\phi}_{q,g}(\mathbf{z}_t) = \mathbf{V}_g \mathbf{q}_{t,g} + \mathbf{c}_g. \tag{75}$$

The quantum-modulated recurrent quantities are

$$\boldsymbol{\tau}_t^{(q)} = \sigma\left(\mathbf{W}_{\tau h}\mathbf{h}_{t-1} + \boldsymbol{\phi}_{q,\tau}(\mathbf{z}_t)\right) + \epsilon\mathbf{1}_H, \tag{76}$$

$$\mathbf{g}_t^{(q)} = -\mathbf{h}_{t-1} \oslash \boldsymbol{\tau}_t^{(q)} + \tanh\left(\mathbf{W}_h\mathbf{h}_{t-1} + \boldsymbol{\phi}_{q,g}(\mathbf{z}_t)\right), \tag{77}$$

$$\mathbf{h}_t = \mathbf{h}_{t-1} + \Delta t_t \mathbf{g}_t^{(q)}. \tag{78}$$

After the final temporal position,

$$\mathbf{o} = \mathbf{W}_{\text{cls}}\mathbf{h}_L + \mathbf{b}_{\text{cls}}, \tag{79}$$

with class probabilities

$$p_\theta(y = c \mid \mathcal{S}) = \frac{\exp(o_c)}{\sum_{k=1}^{C}\exp(o_k)}. \tag{80}$$

D.4 CIRCUIT-COMPONENT VARIANTS

The three evaluated circuits are

$$U_\ell^{\text{full}}(\mathbf{u}_t) = U_{\text{ent}} V_\ell E(\mathbf{u}_t), \tag{81}$$

$$U_\ell^{\text{no-ent}}(\mathbf{u}_t) = V_\ell E(\mathbf{u}_t), \tag{82}$$

$$U_\ell^{\text{no-reup}} = U_{\text{ent}} V_\ell. \tag{83}$$

The initial encoding $E(\mathbf{u}_t)$ remains present before the variational stack in the no-reuploading model; only repeated data injection inside the subsequent layers is removed.

D.5 CIRCUIT-RESOURCE ACCOUNTING

For one quantum branch at one temporal position, the nominal logical gate counts are shown in Table 4.

Table 4: Nominal logical resources per quantum branch and temporal step. Counts exclude state preparation and measurement.

Variant	Data rotations	Trainable 1-qubit rotations	CNOTs
Full	$2q(D+1)$	$3qD$	$D(q-1)$
No entanglement	$2q(D+1)$	$3qD$	0
No re-uploading	$2q$	$3qD$	$D(q-1)$

Because the model uses two independent quantum branches at each of L temporal positions, analytic inference requires

$$N_{\text{circ}} = 2L \tag{84}$$

circuit evaluations per sequence. Under an S-shot measurement budget, the number of state preparations is

$$N_{\text{prep}} = 2LS. \tag{85}$$

All local Pauli-Z observables within one branch are diagonal in the same computational basis and can therefore be estimated from the same set of shots. Increasing q consequently increases circuit width and gate count but does not require a distinct measurement basis for each Z_j.

Removing entanglement eliminates

$$D(q-1) \tag{86}$$

CNOT gates per branch and temporal step, or

$$2LD(q-1) \tag{87}$$

across both branches of an L-step sequence. These counts describe nominal algorithmic resources only; they do not include connectivity-dependent transpilation, hardware gate decomposition, execution latency, readout error, or other device-level overheads.

E EXPERIMENTAL DETAILS

E.1 DATASET AND CHRONOLOGICAL CONSTRUCTION

The Room Occupancy Estimation dataset contains $N = 10{,}129$ timestamped observations with $d = 16$ sensor variables and a four-class occupancy label,

$$y \in \{0, 1, 2, 3\}. \tag{88}$$

All records are sorted chronologically before window construction and dataset splitting.

Because the original benchmark is regularly timestamped, temporal irregularity is introduced through controlled stochastic observation thinning. The chronological sequence is converted into sliding windows with

$$L = 12, \qquad \text{stride} = 3. \tag{89}$$

Intermediate observations are removed with base probability

$$\rho = 0.2, \tag{90}$$

while the final observation is always retained so that the target remains associated with the terminal temporal state.

Missing observations are represented through the feature vector, observation mask, and elapsed time defined by equation 1. For $d = 16$, the resulting augmented input has dimension

$$2d + 1 = 33. \tag{91}$$

The final benchmark contains

$$N_{\text{train}} = 2022, \qquad N_{\text{val}} = 500, \qquad N_{\text{test}} = 672. \tag{92}$$

E.2 Model-comparison hierarchy

The MLP operates on flattened temporal windows and provides a non-recurrent classical reference. The GRU processes the augmented irregular sequence and provides a strong recurrent reference. The classical liquid model uses the temporal backbone of Section 3.1 without quantum modulation.

The quantum-liquid family contains the full circuit, the no-entanglement variant, and the no-reuploading variant defined in Section 3.4. The MLP and GRU are used to contextualise the overall difficulty of the classification task, whereas the controlled mechanistic comparison is between the classical liquid backbone and its quantum-modulated variants.

All three quantum circuit variants contain 4652 trainable parameters. Their logical gate requirements differ as detailed in Appendix D.5.

E.3 Analytic training and finite-shot evaluation

Quantum models are trained using analytic expectation values. During the finite-shot study, the learned parameters are held fixed and measurement sampling is introduced only at test time.

The evaluated budgets are

$$S \in \{10, 50, 100, 500, \infty\}, \tag{93}$$

where $S = \infty$ denotes analytic expectation evaluation.

This protocol separates inference-time measurement effects from additional stochasticity that would arise if finite-shot estimates were also introduced during optimisation.

E.4 Reported metrics

All models are evaluated on the same held-out test windows. We report accuracy, macro-F1, negative log-likelihood, Brier score where available, and expected calibration error. These metrics respectively provide complementary views of overall classification performance, class-balanced discrimination, probability assignment, squared probabilistic error, and confidence calibration.

F Additional Experimental Interpretation

F.1 Scope of the circuit-component comparison

The main ablation identifies a clear ordering within the fixed liquid architecture: retaining reuploading while removing the tested entangling layer gives the strongest reported quantum-liquid operating point across accuracy, macro-F1, NLL, Brier score, and ECE.

This result has two important interpretive limits. First, it does not imply that entanglement is generically harmful. Proposition 3 characterises entanglement as enabling additional cross-coordinate structure, but whether that structure is useful depends on the modulation function required by the task. Second, the ablation does not establish that the improvement is uniquely quantum. The no-entanglement circuit has the bounded ridge-function structure described by Proposition 1, and parameter-matched classical nonlinear modulators are not included in the present comparison.

F.2 Scope of the calibration result

Calibration follows the same component-level ordering as the predictive metrics. Removing entanglement reduces ECE from 0.6832 to 0.5921, while removing re-uploading increases it to 0.7296. Brier score similarly changes from 1.3732 for the full model to 1.1883 without entanglement and 1.4595 without re-uploading.

These improvements are relative to the liquid family only. The GRU and MLP remain substantially better calibrated, with ECE values of 0.0986 and 0.1119, respectively. The ablation therefore supports a component-level calibration claim, not a claim that the quantum-liquid models are well calibrated in absolute terms.

F.3 Scope of the finite-shot comparison

The finite-shot sweep reports point estimates for one trained model under the evaluated measurement budgets. The experiment does not provide repeated independent measurement realisations or repeated training runs from which uncertainty intervals for the shot-specific differences could be estimated.

Accordingly, the fact that several metrics are numerically strongest at $S = 500$ should not be interpreted as evidence that finite-shot sampling is beneficial. The defensible conclusion is narrower: changing the measurement budget changes the operating point of the recurrent quantum model, and the effect is mediated by the downstream temporal dynamics described in Proposition 4.

F.4 Scope of the sparsification comparison

The controlled sparsification study evaluates the GRU and the full quantum circuit over $\rho \in \{0.0, 0.2, 0.4, 0.6\}$. It shows that stronger thinning does not produce a monotonic advantage for the full quantum architecture.

The experiment does not repeat the complete circuit-component ablation at each sparsification level. It therefore cannot establish whether the no-entanglement variant retains its base-setting advantage as temporal sparsity increases.

G Extended Discussion, Limitations, and Future Directions

This appendix expands on the interpretation and scope of the results summarised in Section 5. The discussion separates implications of the circuit structure from limitations of the current empirical setting and from directions requiring additional investigation.

G.1 Interpreting the circuit-component results

The main circuit ablation should not be interpreted simply as evidence that "quantum modulation helps" or that entanglement is detrimental. The more specific observation is that, within the fixed liquid backbone, the strongest quantum-family operating point is obtained when repeated local input encoding is retained and the tested entangling layer is removed.

This ordering is consistent with the structural analysis in Section 3.3. The entanglement-free circuit implements a finite bank of bounded nonlinear ridge functions, while repeated data encoding can increase the potentially accessible local frequency support. Entangling operations instead permit mixed-coordinate dependence. These mechanisms are therefore qualitatively different rather than simply ordered by generic circuit complexity.

The empirical ordering suggests that the local nonlinear structure retained by re-uploading is better aligned with the tested temporal modulation problem than the additional cross-coordinate structure introduced by the chosen entangling circuit. This conclusion is specific to the architecture and task studied here. It does not imply that entanglement is generically harmful, nor does it establish that the observed gain is uniquely quantum.

The resource comparison strengthens the architectural interpretation. The no-entanglement variant removes all CNOT gates from each quantum branch and temporal step while improving the reported

predictive and calibration-sensitive metrics relative to the full circuit. Detailed resource counts are given in Appendix D.5. Thus, in the present setting, the empirically preferred quantum variant is also the simpler one in two-qubit-gate requirements.

G.2 Finite-shot measurement as a dynamical resource

Finite-shot inference introduces a different form of resource constraint. For local Pauli measurements, the standard estimation error decreases with measurement budget, but in a recurrent architecture the corresponding perturbation is repeatedly transformed by subsequent state transitions.

Proposition 4 formalises this effect by coupling three quantities: measurement uncertainty, temporal spacing, and downstream dynamical sensitivity. Consequently, identical local measurement errors can have different effects depending on where they enter the sequence and whether later transitions amplify or contract the resulting state perturbation.

The finite-shot experiment in Section 4.3 illustrates this distinction. The reported operating points vary non-monotonically across $S \in \{10, 50, 100, 500, \infty\}$, even though the expected magnitude of the local measurement estimator error decreases with S. Because the experiment does not contain repeated independent training and measurement realisations, the numerically stronger point estimate observed at $S = 500$ should not be interpreted as evidence of stochastic regularisation or a generally preferable measurement budget.

The more conservative implication is that analytic expectation evaluation and finite-shot deployment should be regarded as distinct operating regimes for recurrently applied quantum modules. Measurement budget is therefore part of the deployed dynamical system rather than only a circuit-level implementation detail.

G.3 Limitations

Single benchmark and synthetic irregularity. The empirical study uses one relatively small and imbalanced Room Occupancy benchmark. Moreover, the original observations are regularly timestamped and irregularity is introduced through controlled stochastic thinning. The results therefore characterise the tested architecture under a controlled irregularisation protocol rather than a naturally asynchronous data-generating process.

Backbone dependence. The classical liquid backbone performs substantially worse than the MLP and GRU references. Improvements within this backbone should therefore not be interpreted as competitiveness with strong classical sequence models. The central empirical contribution is a component-level comparison of modulation mechanisms rather than state-of-the-art predictive performance.

Statistical uncertainty. The reported experiments do not include sufficient independent training seeds, thinning realisations, and finite-shot measurement realisations to estimate uncertainty intervals for the observed differences. In particular, the finite-shot sweep should be interpreted as point-estimate sensitivity rather than evidence of a systematic benefit or cost at any particular value of S.

Classical modulation controls. The current comparison does not include parameter-matched classical nonlinear modulators designed to reproduce the function class suggested by the entanglement-free analysis. Classical sinusoidal features, learned Fourier features, or compact nonlinear networks would provide important controls for separating the usefulness of the modulation structure itself from any specifically quantum implementation of that structure.

Hardware realism. Finite-shot inference captures sampling uncertainty but not the complete error process of current quantum hardware. Gate noise, readout error, restricted connectivity, transpilation overhead, and hardware-dependent latency are not modelled. Similarly, the circuit-resource accounting in Appendix D.5 reports nominal logical resources rather than compiled hardware depth or end-to-end device cost.

Scope of the sparsification study. The stronger-sparsification experiment in Appendix **??** evaluates the full quantum circuit rather than all circuit-component variants. It therefore shows that greater temporal sparsity does not by itself reveal an advantage for the full model, but it does not establish robustness of the no-entanglement variant under increasing sparsification.

G.4 Future directions

Several directions follow directly from the structural interpretation developed in this work.

Interaction-controlled temporal tasks. The distinction between local spectral structure and mixed-coordinate interaction suggests evaluating quantum modulators on controlled dynamical problems for which the required interaction structure is known. Tasks can be constructed to distinguish targets dominated by nonlinear functions of individual projected coordinates from targets requiring genuine cross-coordinate interactions. Such comparisons would provide a more direct test of when the additional terms enabled by entanglement are useful.

Naturally irregular temporal data. Future studies should extend the evaluation to naturally asynchronous settings, including clinical event streams, asynchronous sensor systems, and other event-driven datasets. Such benchmarks would test whether the circuit-component ordering observed here persists when temporal irregularity arises from the data generation process rather than controlled thinning.

Repeated training and measurement realisations. A stronger statistical evaluation should separately quantify variability from model optimisation, temporal thinning, and quantum measurement. Repeated initialisations and thinning realisations would characterise training and data variability, while repeated finite-shot evaluations of fixed trained models would isolate measurement-induced uncertainty.

Structure-aware circuit design. The circuit design space can be explored through the structural quantities identified by the analysis rather than through gate count alone. Relevant directions include varying re-uploading depth, using sparse or adaptive entanglement patterns, changing measurement observables, modifying the low-dimensional input projection, and evaluating quantum modules only at selected temporal positions.

Adaptive measurement allocation. Proposition 4 also suggests that a fixed total measurement budget need not be distributed uniformly across temporal positions. If a_t summarises the downstream influence of the quantum feature evaluated at time t, a sensitivity-aware allocation could take the schematic form

$$S_t \propto \sqrt{a_t}. \tag{94}$$

This is not part of the present method; it is a possible future consequence of the trajectory-sensitivity analysis. Developing such a rule would turn shot count from a single global hyperparameter into a temporal resource-allocation problem.

Parameter-matched classical modulators. The factorised analysis suggests a particularly informative class of future controls. Since the no-entanglement quantum circuit behaves as a collection of bounded spectral ridge functions, parameter-matched classical modulators based on sinusoidal features, learned Fourier maps, or compact nonlinear networks could be compared against the same liquid backbone. Such experiments would address a sharper question: whether a low-qubit circuit provides a useful or resource-efficient implementation of a modulation structure that is itself beneficial.

G.5 Broader design implication

Taken together, the theoretical and empirical results motivate evaluating resource-constrained quantum temporal models through functional structure and dynamical interaction rather than circuit complexity in isolation. In the regime studied here, repeated local encoding coincides with the most useful quantum modulation signal, whereas the tested entangling structure increases two-qubit-gate requirements without a corresponding empirical improvement.

A useful design principle for subsequent hybrid temporal models is therefore to use the smallest quantum module whose induced function class matches the required dynamics, and to treat measurement resources according to how strongly quantum perturbations influence the downstream trajectory.

APPENDIX REFERENCES

Amira Abbas, David Sutter, Christa Zoufal, Aurelien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1:403–409, 2021. doi: 10.1038/s43588-021-00084-1.

Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549:195–202, 2017. doi: 10.1038/nature23474.

M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles. Variational quantum algorithms. *Nature Reviews Physics*, 3 (9):625–644, 2021. doi: 10.1038/s42254-021-00348-9.

Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1):6085, 2018. doi: 10.1038/s41598-018-24271-9.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David K. Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems 31*, pp. 6571–6583, 2018. URL https://proceedings.neurips.cc/paper_files/paper/2018/hash/69386f6bb1dfed68692a24c8686939b9-Abstract.html.

Ramin Hasani, Mathias Lechner, Alexander Amini, Daniela Rus, and Radu Grosu. Liquid time-constant networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 7657–7666, 2021. doi: 10.1609/aaai.v35i9.16936.

Ramin Hasani, Mathias Lechner, Alexander Amini, Lucas Liebenwein, Aaron Ray, Max Tschaikowski, Gerald Teschl, and Daniela Rus. Closed-form continuous-time neural networks. *Nature Machine Intelligence*, 4(11):992–1003, 2022. doi: 10.1038/s42256-022-00556-7.

Vojtech Havlicek, Antonio D. Corcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212, 2019. doi: 10.1038/s41586-019-0980-2.

Patrick Kidger, James Morrill, James Foster, and Terry Lyons. Neural controlled differential equations for irregular time series. In *Advances in Neural Information Processing Systems 33*, pp. 6696–6707, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/4a5876b450b45371f6cfe5047ac8cd45-Abstract.html.

Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9:4812, 2018. doi: 10.1038/s41467-018-07090-4.

Adrian Perez-Salinas, Alba Cervera-Lierta, Elies Gil-Fuster, and Jose I. Latorre. Data re-uploading for a universal quantum classifier. *Quantum*, 4:226, 2020. doi: 10.22331/q-2020-02-06-226.

Yulia Rubanova, Ricky T. Q. Chen, and David K. Duvenaud. Latent ordinary differential equations for irregularly-sampled time series. In *Advances in Neural Information Processing Systems 32*, pp. 5321–5331, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/42a6845a557bef704ad8ac9cb4461d43-Abstract.html.

Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Physical Review A*, 103(3):032430, 2021. doi: 10.1103/PhysRevA.103.032430.

Sukin Sim, Peter D. Johnson, and Alan Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070, 2019. doi: 10.1002/qute.201900070.