An Uncertainty-Aware Quantum–Classical Digital Twin Framework for Predictive Maintenance of Rotating Machinery

Entropy-Gated Hybrid Inference, Innovation-Triggered Twin Synchronization and a Reproducible Simulation Study with Fault Injection

[1]Sonia Mohapatra, [2]Saswat Swain, [3]Chandrasekhar Panda
[1]Researcher, [2]Ph.D Scholar, [3]Researcher
[1,3]Research and Development Division,
Tekstein Solutions Private Limited, Bhubaneswar, Odisha, India
[2]KSOM — KIIT Deemed to be University,
Bhubaneswar, Odisha, India
[1]smctc98@gmail.com, [2]saswat.mck@gmail.com, [3]contact@chandrasekharpanda.com

Abstract*—Digital twins are increasingly used to monitor rotating machinery, and a growing literature proposes adding quantum machine-learning components to them. Whether a near-term quantum model improves a maintenance twin, and how it should be integrated so that it cannot degrade the twin when it does not, has not been tested in a controlled setting. This paper presents a framework in which a six-qubit variational quantum classifier (VQC) ensemble and a classical multilayer-perceptron ensemble are combined through an uncertainty-aware gate: calibrated predictive entropy decides whether the quantum circuit is executed at all (an escalation cascade) and how its output is weighted (entropy-weighted fusion). The twin state (speed, bearing temperature, defect severity and its growth rate) is tracked by an extended Kalman filter that is synchronized with the edge node only when the normalized innovation exceeds a threshold, and the severity belief drives a remaining-useful-life (RUL) estimate and a cost-based maintenance rule. A physics-based motor–pump–bearing simulator with McFadden–Smith defect impulses, motor-current sidebands and fault injection provides 2,500 labelled windows over 0–20 dB and 80 run-to-failure trajectories. Two results are reported without softening. First, the gate keeps the hybrid at the accuracy of the better model (98.0 % against 98.2 % for the classical ensemble and 88.8 % for the VQC) while executing the circuit on 4.4 % of windows, and holds 98.0 % when simulated hardware noise drives the VQC alone to 61.4 %. Second, the VQC did not outperform a parameter-matched 61-parameter classical model under any tested condition, so no quantum advantage is claimed. Innovation-triggered synchronization reduced uplink traffic by 87 % at a 13 % increase in severity-tracking error, and the twin-based maintenance rule matched the best oracle-tuned periodic interval without knowledge of the fleet life distribution. All code, data generators and results are released for reproduction.

***Index Terms*—Digital twin, predictive maintenance, variational quantum classifier, hybrid quantum–classical inference, uncertainty quantification, event-triggered estimation, rolling-element bearing, remaining useful life, Industry 5.0.**

I. Introduction

Rotating machinery accounts for a large share of unplanned downtime in process industries, and most of it originates in rolling-element bearings. A digital twin, understood here as a continuously synchronized state model of a specific asset [1], [2], allows the maintenance decision to be based on the estimated internal state (defect severity, its growth rate and the remaining useful life, RUL) rather than on threshold alarms. Reviews of twin-based predictive maintenance [3] find that the state-estimation and decision layers are usually the weakest part of published twins and that the cost of keeping a cloud-hosted twin synchronized with field sensors is rarely quantified; Industry 5.0 perspectives add that the resulting decisions must remain explainable to the operator [4].

In parallel, quantum machine learning (QML) has entered condition monitoring: parameterized quantum circuits have been trained on the Case Western Reserve University (CWRU) and Jiangnan University bearing data [8], [9], a quantum least-squares support vector machine has been applied to bearing diagnosis [10], and quantum components have been proposed inside manufacturing twins [5]–[7]. Two limitations motivate the present work. The quantum model is evaluated as a stand-alone classifier, so the question of how it should be placed inside a twin that must remain dependable when the quantum output is unreliable is not addressed; and the comparison is against a large classical network rather than a classical model of the same parameter count, which confounds the effect of the quantum feature map with the effect of capacity.

This paper treats the integration as an uncertainty-aware model-selection problem. Both models are trained as ensembles, calibrated by temperature scaling [18], and compared window

by window through their calibrated predictive entropy [19]. The entropy is used twice: as a gate that decides whether the quantum circuit is executed at all, so that the expensive resource is spent only on windows the classical model finds ambiguous, and as a weight that decides how much the quantum output contributes when it is. The same principle of acting only on information governs synchronization: the edge keeps a replica of the twin's predictor and transmits a measurement only when the normalized innovation exceeds a threshold, transferring event-based estimation [20], [21] into the twin-synchronization setting formalized in [22].

Validation uses a physics-based simulator rather than recorded data, because the two questions asked here (does the gate protect the twin when the quantum model fails, and does event triggering preserve state accuracy at reduced traffic) require fault injection at controlled severity, signal-to-noise ratio (SNR) and hardware noise. The simulator follows the McFadden–Smith defect model [24], envelope-analysis practice [25] and the stator-current sideband model of Schoen *et al.* [31]. Benchmarking on the CWRU and Paderborn data sets [26], [27] is identified as the necessary next step.

The contributions are: (i) a digital-twin architecture in which a VQC ensemble is integrated through calibrated-entropy gating and fusion, with a proposition showing that, under calibration, the selection rule cannot be worse than the better model; (ii) an innovation-triggered synchronization scheme for the twin's extended Kalman filter (EKF) with a deterministic bound on the twin–plant discrepancy between transmissions; (iii) an RUL belief that includes the first-passage variance of the degradation process, and a cost-based decision rule evaluated against periodic and run-to-failure policies; (iv) a controlled empirical comparison in which the VQC is matched against a classical model of equal parameter count, reported in full, including the negative finding that no quantum advantage was observed; and (v) a reproducibility bundle containing the simulator, the training code and every number in this paper.

II. Related Work

Table I summarizes the closest studies. Twin-based maintenance reviews [2], [3] establish the architecture and identify RUL estimation and synchronization cost as open issues; bearing twins [28], [30] combine physics and data but transmit at a fixed rate. Maior *et al.* [8] is the closest QML precedent: a parameterized circuit followed by an MLP classifies CWRU and JNU data with accuracy comparable to classical models, without a parameter-matched baseline or twin integration. Ud Din *et al.* [5] and Villalba-Diez *et al.* [7] place quantum components in manufacturing twins for risk estimation and process visualization, without a condition-monitoring state model. Rehman *et al.* [33] route uncertain samples from a classical to a quantum head, the same cascade pattern adopted here, but without calibration, twin integration or hardware-noise evaluation. Event-triggered estimation [20], [21] underpins the synchronization scheme; recent twin-synchronization work [22], [23] treats update timing as resource allocation but does not use the estimator's own innovation as the trigger.

To the authors' knowledge, no peer-reviewed study integrates a quantum classifier into a maintenance twin through an uncertainty-aware gate, evaluates the gate under simulated hardware noise, and compares the quantum model against a parameter-matched classical model on the same fault-injected data. These are the gaps addressed below.

TABLE I. Representative prior work and the gap addressed

Ref., year	Scope	Method	Limitation relevant here
[3] 2022	Twin-based PdM review	Systematic review	Sync cost and UQ rarely quantified
[28] 2022	Bearing twin	Physics + data fusion	Fixed-rate data transfer
[8] 2023	QML for bearings	PQC + MLP on CWRU/JNU	No parameter-matched baseline; no twin
[10] 2022	QML for bearings	Quantum LS-SVM	Small data; no uncertainty
[5] 2025	Quantum twin	Quantum + GAN	No fault-state model
[33] 2026	Hybrid routing	Confidence-gated classical → quantum	No calibration; no noise study
[20] 2014	Event-based estimation	Variance-triggered KF	Not applied to twins
[22] 2024	Twin synchronization	Optimization framework	No estimator-driven trigger
This work	Quantum–classical twin	Entropy-gated VQC + innovation-triggered EKF	Simulation only

III. Problem Statement

Consider an induction motor driving a centrifugal pump, with the drive-end bearing developing a single localized defect whose severity $s(t) \in [0, 1]$ grows over time. The edge node samples vibration $x(t)$ and phase current $i(t)$ at 12 kHz in windows of 0.5 s, together with shaft speed ω and bearing-housing temperature T_b once per minute. The twin service must (a) identify the fault mode $y \in$ {healthy, outer race, inner race, ball, unbalance} from each window with a calibrated confidence, (b) track the slow state $\mathbf{x} = [\omega, T_b, s, \theta]^T$, where θ is the degradation rate, (c) produce an RUL belief, and (d) recommend a maintenance time, subject to three constraints. The uplink budget is limited, so the twin cannot receive every measurement; the quantum inference resource is scarce and noisy, so it must be invoked selectively and its output must not be trusted blindly; and the decision must be robust to the fact that neither model is perfectly calibrated. The questions studied are therefore: how should two classifiers of very different character be combined so that the combination is never worse than the better one; how far can the uplink rate be reduced before state tracking and onset detection degrade; and does the resulting RUL belief support a decision rule that is competitive with the best periodic policy.

IV. Proposed Framework

Three design decisions distinguish Fig. 1 from a conventional twin. First, inference is performed by two

ensembles with a common probabilistic interface, so that every downstream component consumes a calibrated class distribution rather than a label. Second, the quantum ensemble is behind a gate: it is executed only when the classical ensemble's calibrated entropy exceeds a threshold τ, which bounds the quantum workload, and its output is fused with an entropy-dependent weight, which bounds its influence. Third, the twin's EKF is replicated at the edge, so that the edge can evaluate the innovation that the twin would compute and decide locally whether a transmission is informative.

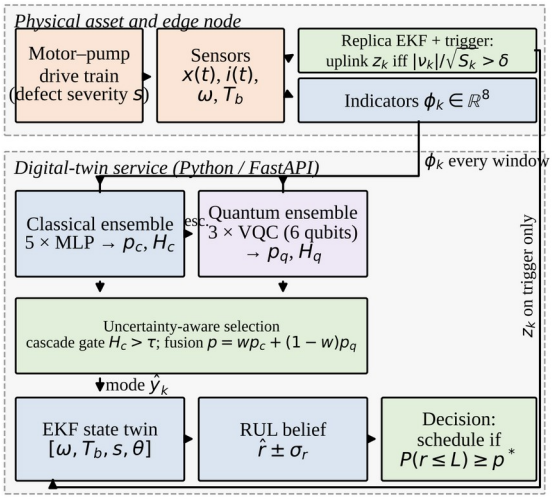


Fig. 1. Framework architecture. The edge node extracts an eight-dimensional indicator vector φ_k from every window and runs a replica of the twin's EKF predictor; measurements z_k are uplinked only when the normalized innovation exceeds δ. The twin service holds the classical and quantum ensembles, the entropy-based selection logic, the EKF state twin, the RUL belief and the decision rule.

A. Indicator extraction

Each vibration window is band-pass filtered around the dominant structural resonance (0.6–1.4 f_n), demodulated by the Hilbert transform and transformed to an envelope spectrum [25]. The indicator vector $\varphi \in \mathbb{R}^8$ contains the envelope-spectrum peaks at the ball-pass frequencies of the outer race (BPFO) and inner race (BPFI) and at twice the ball-spin frequency (2BSF), each normalized by the median envelope level in 10–500 Hz; the raw-spectrum shaft-order (1×) peak normalized by the median level in 10–2000 Hz; the kurtosis, crest factor and RMS of the raw window; and the stator-current sideband ratio, i.e. the sum of current-spectrum amplitudes at $f_e \pm f_{def}$ for the three bearing frequencies and the shaft frequency, divided by the fundamental [31]. The five ratio indicators are log-compressed. Eight dimensions are retained deliberately so that the vector can be angle-encoded on six qubits without a dimensionality-reduction stage that would obscure which physics the quantum model receives.

B. Classical and quantum ensembles

The classical ensemble consists of five MLPs with two hidden layers of 32 units (1,509 parameters each), trained from different initializations; ensembling provides the epistemic uncertainty used by the gate [17]. The quantum ensemble consists of three VQCs with 64 parameters each. Each VQC encodes the standardized indicators as rotation angles, applies three strongly entangling layers with data re-uploading between layers [16], and reads out the Pauli-Z expectation of five qubits, which a trainable affine map converts to class logits [11]. A parameter-matched classical ensemble (five MLPs with one hidden layer of four units, 61 parameters each) is trained as the control that separates the effect of the quantum feature map from the effect of capacity.

C. Uncertainty-aware selection

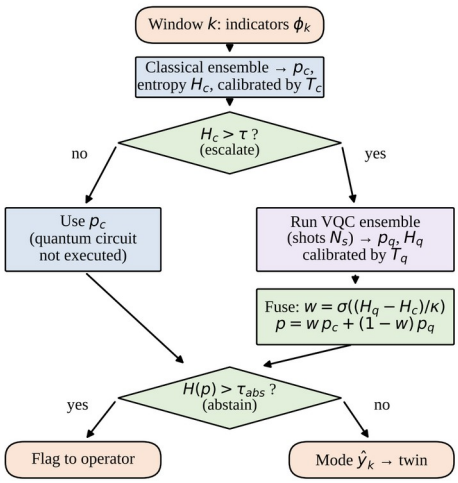


Fig. 2. Per-window decision logic. The quantum ensemble is executed only when the classical entropy exceeds τ; the fused distribution is abstained from (flagged to an operator) when its entropy exceeds τ_{abs}.

Fig. 2 shows the per-window logic. The calibrated distributions p_c and p_q and their normalized entropies H_c, H_q are computed as in Section V-C. Three rules are studied: maximum-confidence selection, which forwards whichever distribution has the larger calibrated maximum probability; entropy-weighted fusion, which mixes the two with a weight that favours the lower-entropy model; and the cascade, which executes the quantum ensemble only when $H_c > \tau$ and otherwise forwards p_c. Selective prediction is supported by abstaining when the final entropy exceeds τ_{abs}, in which case the window is flagged to an operator rather than passed to the twin.

D. State twin, synchronization and decision

The slow state is tracked by an EKF on the plant model of Section V-D. The edge runs an identical predictor and transmits z_k only when the largest normalized innovation component exceeds δ; because both sides apply the same deterministic update on the same transmitted data, the edge replica and the cloud twin remain identical without additional messages. The fault mode $\hat{y}_k$ from the selection logic chooses which envelope-spectrum channel (BPFO, BPFI or 2BSF) is used as the severity measurement. The RUL belief is derived from the posterior on (s, θ), and maintenance is scheduled when the probability that the RUL is shorter than the lead time L exceeds a threshold p^*.

V. Mathematical Modelling

A. Bearing kinematics and vibration model

For a bearing with n rolling elements of diameter d, pitch diameter D, contact angle α and shaft frequency f_r, the characteristic frequencies are

$$f_{BPFO} = (n/2)\, f_r\,(1 - r), \quad f_{BPFI} = (n/2)\, f_r\,(1 + r), \tag{1}$$

$$f_{BSF} = (D/2d)\, f_r\,(1 - r^2), \quad f_{FTF} = (f_r/2)(1 - r), \tag{2}$$

with $r = (d/D)\cos\alpha$. For the SKF 6205 geometry used ($n = 9$, $d = 7.94$ mm, $D = 39.04$ mm, $\alpha = 0$) at $f_r = 29.53$ Hz, Eq. 1 and Eq. 2 give 105.9, 159.9, 69.6 and 11.8 Hz. Following McFadden and Smith [24], a localized defect produces a train of impulses at the defect frequency, each exciting a damped structural resonance,

$$x_d(t) = s \sum_k A(t_k)\, h(t - t_k), \quad h(t) = e^{-\beta t}\sin(2\pi f_n t), \tag{3}$$

where $t_{k+1} - t_k = (1 + 0.01\varepsilon_k)/f_{def}$ with $\varepsilon_k \sim N(0, 1)$ represents slip jitter, $\beta = 2\pi\zeta f_n$ with $\zeta = 0.05$ and $f_n = 3.2$ kHz, and $A(t) = 1 + 0.6\cos(2\pi f_m t)$ is the load-zone modulation with $f_m = f_r$ for an inner-race defect, $f_m = f_{FTF}$ for a rolling-element defect (impacting at $2f_{BSF}$) and no modulation for an outer-race defect. The healthy baseline contains 1×, 2× and vane-pass (5×) components; unbalance adds a 1× component proportional to s. White Gaussian noise is added at the specified SNR relative to the deterministic content. The stator current carries sidebands at $f_e \pm kf_{def}$ ($k = 1, 2$) with amplitude $0.05s/k$ relative to the fundamental [31]. The simulator is thus a signal model, not a dynamic model of the bearing; its purpose is to produce defect signatures with the correct kinematic structure under controlled SNR.

B. Variational quantum classifier

Standardized indicators z_j are mapped to bounded angles $\vartheta_j = \pi\tanh(z_j/2)$. One VQC member realizes the state

$$|\psi(\boldsymbol{\vartheta}, \mathbf{W})\rangle = \Pi_{l=1}^{L}\, U_{ent}(\mathbf{W}_l)\, S(\boldsymbol{\vartheta})\, |0\rangle^{\otimes 6}, \tag{4}$$

where $S(\boldsymbol{\vartheta})$ applies $R_Y(\vartheta_j)$ on qubit $j = 0\ldots5$ and $R_Z(\vartheta_6)$, $R_Z(\vartheta_7)$ on qubits 0 and 1, and U_{ent} is a strongly entangling layer of general single-qubit rotations followed by a ring of CNOT gates. Re-uploading the data in every layer increases the number of Fourier frequencies the model can represent in the inputs [16]; the circuit is a standard variational ansatz in the sense of [13]. Class logits are

$$\ell_c = a_c\, \langle\psi|\, Z_c\, |\psi\rangle + b_c, \quad c = 1\ldots5, \tag{5}$$

with trainable scale a_c and bias b_c; $p(y = c \mid \boldsymbol{\vartheta}) = \mathrm{softmax}(\ell)_c$. With $L = 3$ the circuit has 54 rotation parameters, 10 readout parameters, 60 gates (24 encoding rotations, 18 general rotations, 18 CNOTs) and depth 27. Training minimizes the cross-entropy with Adam; gradients are obtained by back-propagation through the state-vector simulator, which is exact and avoids the shot noise of parameter-shift estimates during training.

C. Calibration, uncertainty and selection rules

Let $\mathbf{p}^{(m)}$ be the output of ensemble member $m \in \{1\ldots M\}$ and $\bar{\mathbf{p}}$ their mean. A single temperature T per ensemble is fitted on the validation split by minimizing the negative log-likelihood (NLL) of the mean logits [18]. The normalized predictive entropy and the mutual information (epistemic part) are

$$H(\bar{\mathbf{p}}) = -(1/\ln C)\, \Sigma_c\, \bar{p}_c \ln \bar{p}_c, \tag{6}$$

with $C = 5$ so that $H \in [0, 1]$, and $I = H(\bar{\mathbf{p}}) - (1/M)\, \Sigma_m\, H(\mathbf{p}^{(m)})$. Entropy-weighted fusion mixes the calibrated classical and quantum means as

$$\mathbf{p} = w\, \mathbf{p}_c + (1 - w)\, \mathbf{p}_q, \quad w = \sigma((H_q - H_c)/\kappa), \tag{7}$$

where σ is the logistic function and $\kappa = 0.15$ sets how sharply the weight follows the entropy difference; for $H_q = H_c$ the rule reduces to equal averaging. The cascade forwards $\mathbf{p}_c$ when $H_c \le \tau$ and Eq. 7 otherwise. Calibration is measured by the expected calibration error (ECE) over ten confidence bins, and selective prediction by the area under the risk–coverage curve (AURC).

Proposition 1. *Let two classifiers be calibrated in the sense that, for every window, the probability of being correct equals the reported confidence c_c, c_q (the maximum calibrated class probability). Then maximum-confidence selection has expected accuracy* $\mathrm{E}[\max(c_c, c_q)] \ge \max(\mathrm{E}[c_c], \mathrm{E}[c_q])$, *i.e., it is never worse than the better classifier alone.*

Proof. Under calibration the expected accuracy of any rule that forwards the classifier with confidence c is $\mathrm{E}[c]$. Selection forwards the maximum, and $\max(c_c, c_q) \ge c_c$ and $\ge c_q$ pointwise; taking expectations gives the result. ■ The proposition is conditional: with imperfect calibration the inequality can fail by an amount governed by the calibration error of the weaker model, which Section VIII quantifies (a 0.2-percentage-point loss against the classical ensemble, whose ECE is 0.007, caused by the VQC, whose ECE after temperature scaling remains 0.128).

D. Plant model, EKF and innovation-triggered synchronization

At the twin time step $\Delta t = 60$ s the mechanical time constant $J/(2k_p\omega + b) \approx 0.4$ s is negligible, so speed is treated as a slowly varying operating point. Bearing temperature follows a first-order thermal balance and severity follows an exponential degradation law in the log domain [29]:

$$T_{b,k+1} = T_{b,k} + (\Delta t/C_{th})\,[P_{loss}(s_k, \omega_k) - (T_{b,k} - T_{amb})/R_{th}], \tag{8}$$

$$\ln s_{k+1} = \ln s_k + \theta\,\Delta t + \sigma_s\sqrt{\Delta t}\,\xi_k, \quad \xi_k \sim N(0, 1), \tag{9}$$

with $P_{loss} = P_0 + P_s s_k + 0.15\, u\omega_k$, u the electromagnetic torque, $C_{th} = 420$ J/K, $R_{th} = 0.55$ K/W, $P_0 = 35$ W, $P_s = 110$ W, $\sigma_s = 3 \times 10^{-3}$ min$^{-1/2}$, and θ unknown to the twin and modelled as a random walk. The measurement model is $\mathbf{z}_k = [\omega_k,\ T_{b,k},\ c_0 + c_1 s_k]^T + \mathbf{v}_k$, where the third component is the log-compressed envelope peak of the mode-specific channel; $c_0 = 2.36$, $c_1 = 1.83$ and the residual standard deviation 0.48 were obtained by regression on 300 generated windows at 10 dB. The EKF uses the Jacobian of Eq. 8–9 with $\partial s_{k+1}/\partial\theta = s_k\Delta t\, e^{\theta\Delta t}$. With predicted state $\hat{\mathbf{x}}^-_k$ and covariance $\mathbf{P}^-_k$, the innovation and its covariance are

$$\mathbf{v}_k = \mathbf{z}_k - h(\hat{\mathbf{x}}^-_k), \quad \mathbf{S}_k = \mathbf{H}_k\mathbf{P}^-_k\mathbf{H}_k^T + \mathbf{R}, \tag{10}$$

and the edge transmits $\mathbf{z}_k$ if and only if $\max_i |v_{k,i}|/\sqrt{S_{k,ii}} > \delta$. Both sides then perform the Joseph-form update; otherwise both propagate the prediction only.

Proposition 2. *Between transmissions the twin's predicted measurement satisfies* $|z_{k,i} - h_i(\hat{\mathbf{x}}^-_k)| \le \delta\ \sqrt{S_{k,ii}}$ *for every component i, irrespective of model mismatch; and any*

mismatch that grows without bound forces a transmission in finite time.

Proof. The first statement is the negation of the trigger condition, which the edge evaluates on the same $\hat{\mathbf{x}}^-_k$, $\mathbf{P}^-_k$ as the twin because the two run identical deterministic recursions on identical transmitted data. For the second, $\mathbf{S}_k$ is bounded above by $\mathbf{H}\mathbf{P}_\infty\mathbf{H}^T + \mathbf{R}$, where $\mathbf{P}_\infty$ is the open-loop covariance limit (finite here because the only unstable mode, severity, is clipped by the failure threshold), whereas an unbounded mismatch makes $|\nu_{k,i}|$ unbounded. ■ Unlike the variance-based trigger of [20], which transmits when the prediction variance exceeds a bound regardless of the data, the innovation trigger transmits only when the data disagree with the twin, so traffic is low while severity is constant and rises when the indicator starts to move.

E. RUL belief and decision rule

From the posterior mean $(\hat{s}, \hat{\theta})$ the RUL point estimate to the failure threshold $s_f = 0.85$ is $\hat{r} = \ln(s_f/\hat{s})/\hat{\theta}$. Its variance has two parts: the first-order propagation of the posterior covariance through the gradient $\mathbf{g} = \partial\hat{r}/\partial\mathbf{x} = [0, 0, -1/(\hat{s}\hat{\theta}), -\hat{r}/\hat{\theta}]^T$, and the variance of the first-passage time of the drifted random walk in Eq. 9, which is $\sigma_s^2\hat{r}/\hat{\theta}^2$ under the inverse-Gaussian approximation:

$$\sigma_r^2 = \mathbf{g}^T\mathbf{P}_k\mathbf{g} + \sigma_s^2\,\hat{r}/\hat{\theta}^2. \tag{11}$$

The second term is essential: without it the 90 % interval covered only 46 % of true RUL values in a preliminary run, because the future process noise, not the estimation error, dominates late in life. Maintenance is scheduled at the first k at which $\Phi((L - \hat{r}_k)/\sigma_{r,k}) \geq p^*$, with lead time $L = 240$ min; the rule is armed only once $\sigma_r < \hat{r}$, a condition met at the first step on both fleets. The cost of a trajectory is $C_p + c_w(t_f - t_m)$ if maintenance at t_m precedes failure at t_f, and C_f otherwise, with $C_p = 1$, $C_f = 8$ and $c_w = 0.04$ per hour of wasted life. For this cost structure the risk-neutral threshold is $p^* = C_p/(C_p + C_f) \approx 0.11$, which the sweep in Section VIII confirms.

VI. Algorithm and Complexity

Algorithm 1: Gated hybrid inference and triggered synchronization (one window *k*)

```
Edge: compute φ_k from x, i; read ω_k, T_b,k; form z_k
Edge: predict (x̂⁻, P⁻) with the replica EKF; compute ν_k, S_k by Eq. 10
if max_i |ν_k,i|/√S_k,ii > δ then transmit z_k and φ_k; else transmit φ_k only
Twin: p_c ← mean of calibrated MLP members; H_c ← Eq. 6
if H_c > τ then
    p_q ← mean of calibrated VQC members (N_s shots each); H_q ← Eq. 6
    p ← entropy-weighted fusion, Eq. 7
else p ← p_c
if H(p) > τ_abs then flag window to operator; else ŷ_k ← arg max p
Twin: predict; if z_k received then Joseph-form update using channel(ŷ_k)
RUL: r̂, σ_r by Eq. 11; if Φ((L − r̂)/σ_r) ≥ p* then schedule maintenance at k + L
```

Per window, indicator extraction costs $O(N \log N)$ for $N = 6{,}000$ samples; the classical ensemble costs $O(M_c n_p)$ multiply–adds (7.5×10^3); the EKF costs $O(n_x^3)$ with $n_x = 4$. On a state-vector simulator one VQC evaluation costs $O(G\,2^n)$ with $G = 60$ gates and $n = 6$ qubits; on hardware the wall time per window is approximately $M_q N_s(D\,t_g + t_{ro})$ for circuit depth $D = 27$, gate time t_g and readout time t_{ro}. Taking $t_g = 0.1$ µs, $t_{ro} = 1$ µs and $N_s = 2{,}048$ gives roughly 8 ms per member and 23 ms for the ensemble, before queueing and classical post-processing; this is an estimate, not a measurement. The cascade multiplies that cost by the escalation fraction, which Section VIII shows to be 4.4 % at $\tau = 0.3$.

VII. Simulation Setup

The study was implemented in Python 3.13 with NumPy 2.5, SciPy 1.18, scikit-learn 1.9 and PennyLane 0.45 [15] (default.qubit for training and finite-shot evaluation, default.mixed for depolarizing noise) on two virtual CPU cores. The plant equations of Section V-D map one-to-one onto Simulink integrator and gain blocks, and a MATLAB script that constructs the equivalent Simulink model programmatically is included in the bundle; it has not been executed in this study, because no MATLAB licence was available in the build environment, and every reported number comes from the Python implementation. A FastAPI service exposing the gated inference, the triggered synchronization and the decision rule is included and was exercised through its test client, but not benchmarked under load. All random draws pass through one NumPy generator seeded with 20261026. Table II lists the parameters.

TABLE II. Simulation parameters

Parameter	Value
Sampling rate; window	12 kHz; 0.5 s (6,000 samples)
Shaft frequency; supply	29.53 Hz (±2 % per window); 50 Hz
Resonance f_n; damping ζ	3.2 kHz; 0.05
Classes	healthy, outer race, inner race, ball, unbalance
Severity s (faulty windows)	uniform 0.15–1.0
SNR levels	0, 5, 10, 15, 20 dB; 100 windows/class/SNR
Split	1,500 train / 500 validation / 500 test (stratified)
Classical ensemble	5 × MLP (32, 32), 1,509 parameters each
Parameter-matched control	5 × MLP (4), 61 parameters each
Quantum ensemble	3 × VQC, 6 qubits, $L = 3$, 64 parameters each
VQC training	Adam, lr 0.05, batch 64, 40 epochs, backprop gradients
Fusion; cascade	$\kappa = 0.15$; $\tau \in [0, 1]$ (21 values)
Twin step; horizon	60 s; up to 3,600 min per trajectory
Fleet A / Fleet B	40 trajectories each; $\theta \in$ [1.1, 2.4] / [0.6, 3.6] × 10^{-3} min^{-1}
Measurement noise	$\sigma_\omega = 0.3$ rad/s, $\sigma_T = 0.4$ K, $\sigma_{HI} = 0.48$
Sync policies	innovation $\delta \in \{0, 1, 1.5, 2, 2.5, 3\}$; periodic every 2, 5, 10, 20, 40 min
Decision	$L = 240$ min; $p^* \in \{0.05, 0.1, 0.2, 0.3, 0.5\}$; $C_p = 1$, $C_f = 8$, $c_w = 0.04$/h

Test scenarios comprise (S1) the standard test split per SNR; (S2) escalation-threshold sweep; (S3) finite-shot evaluation of

the trained VQCs at 128, 512 and 2,048 shots; (S4) depolarizing channels of probability 0.005, 0.02 and 0.05 on every qubit after each encoding and entangling block; (S5) harder and out-of-distribution conditions generated after training: −5 and −10 dB, shaft speed shifted by ±8 % (outside the ±2 % training range), and incipient severities 0.05–0.25 at 5 and 0 dB; (S6) synchronization and RUL on both fleets; and (S7) circuit-depth ablation. In the inference experiments the SNR is applied to vibration; the current channel is 20 dB cleaner, reflecting typical sensor quality.

VIII. Results and Discussion

A. Fault classification and the effect of the gate

Fig. 3 confirms that the simulator produces the intended signatures: the envelope spectra show isolated peaks at the kinematic frequencies with the expected harmonics and, for the inner-race case, shaft-frequency sidebands. Table III and Fig. 4(a) summarize the classification results on the 500-window test split. The classical ensemble reaches 98.2 % accuracy with an ECE of 0.007; the VQC ensemble reaches 88.8 % with an ECE of 0.128 even after temperature scaling (T_q = 0.59, i.e. the raw VQC output is under-confident). The parameter-matched classical control, with three fewer parameters than the VQC, reaches 96.8 %. The VQC is therefore not competitive at equal capacity on this task, and its deficit is concentrated at 0 dB (78.0 % against 97.0 %) and in the unbalance and ball classes, which it confuses with the healthy and inner-race classes respectively.

The gate behaves as Proposition 1 predicts, up to calibration. Maximum-confidence selection and entropy fusion both give 98.0 %, 0.2 percentage points below the classical ensemble, whereas equal averaging, which ignores uncertainty, falls to 97.6 % and inherits the VQC's miscalibration (ECE 0.119 against 0.013 for fusion). The oracle that picks the correct model whenever either is right reaches 98.6 %, so the headroom available to any selection rule on this data is only 0.4 points: the two models' errors are positively correlated (phi coefficient 0.29) and 0.4 % of windows are misclassified by the classical model alone. The selection rule chose the quantum output on 4.6 % of windows, and the mean fusion weight on the classical model was 0.81.

TABLE III. Test-split performance (500 windows, all SNRs). Parameters are per ensemble member.

Model / rule	Params	Acc. (%)	Macro -F1	NLL	ECE
Logistic regression	45	94.6	0.947	0.193	0.042
Tiny MLP ensemble (control)	61	96.8	0.968	0.109	0.028
VQC ensemble	64	88.8	0.887	0.418	0.128
MLP ensemble	1,509	98.2	0.982	0.057	0.007
Equal average (no gate)	—	97.6	0.976	0.195	0.119
Max-confidence selection	—	98.0	0.980	0.067	0.009
Entropy fusion (Eq. 7)	—	98.0	0.980	0.074	0.013
Cascade, τ = 0.3 (4.4 % escalated)	—	98.0	0.980	0.061	0.006
Oracle (either correct)	—	98.6	—	—	—

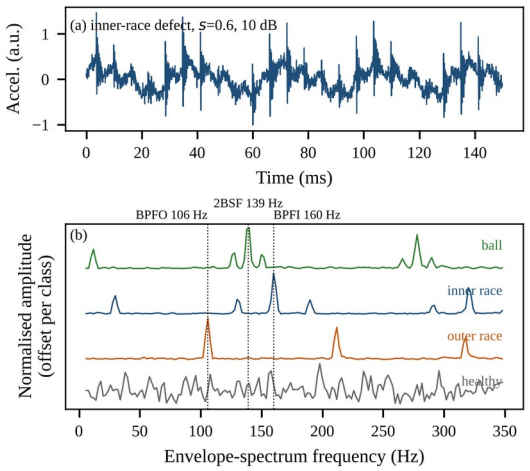


Fig. 3. Simulator output at 10 dB and s = 0.6. (a) Inner-race vibration window showing resonance bursts at the BPFI spacing with load-zone modulation. (b) Normalized envelope spectra of the four bearing conditions, offset vertically; dotted lines mark the kinematic frequencies of Eq. 1 and Eq. 2.

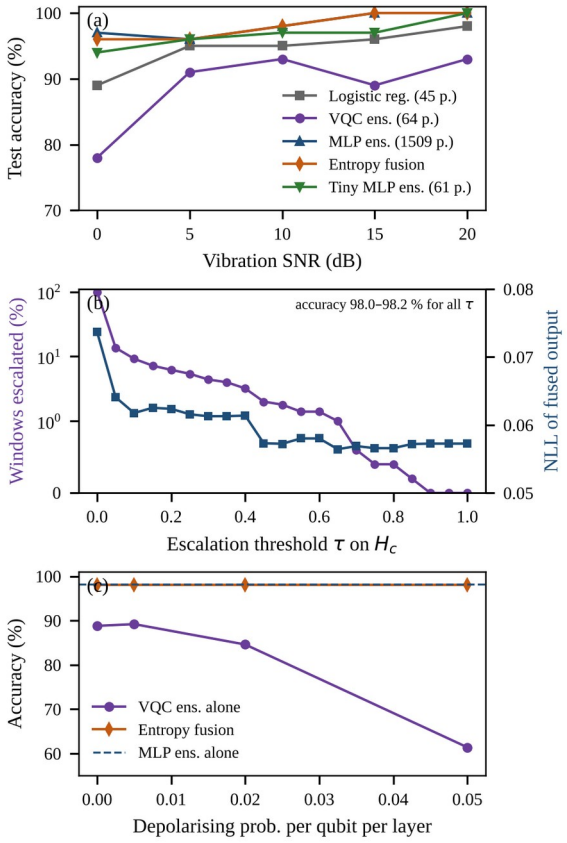


Fig. 4. Classification results on the test split. (a) Accuracy against vibration SNR; parameter counts are per member. (b) Escalation sweep of the cascade: fraction of windows on which the quantum ensemble is executed (left axis, symmetric-log) and NLL of the output (right axis) against the threshold τ. (c) Accuracy of the VQC ensemble alone and of entropy fusion under depolarizing noise, with the classical ensemble as reference.

B. Cascade: how often the circuit runs

Fig. 4(b) shows the escalation sweep. Accuracy stays within 98.0–98.2 % for every τ, so the choice of threshold is governed

by cost and calibration, not accuracy. At $\tau = 0.3$ the quantum ensemble is executed on 4.4 % of windows and the cascade achieves the lowest ECE of all rules (0.0055); at $\tau = 0.1$ the fraction rises to 9.2 % with no benefit. The escalated windows are concentrated at 0 dB (14.0 % of 0-dB windows were selected by the max-confidence rule against 0.0 % at 20 dB), i.e. the circuit is consulted where the classical model is uncertain. On the state-vector simulator the three-member VQC ensemble costs 0.23 ms per window against 2.5 µs for the classical ensemble, a ratio of about 90; the cascade reduces the mean quantum cost per window by the escalation fraction.

C. Robustness to shot noise and hardware noise

Table IV reports scenarios S3 and S4. Finite sampling degrades the VQC alone from 88.8 % (exact expectations) to 80.0 % at 128 shots, while the fused output moves only from 98.0 % to 97.6 %. Depolarizing noise is more damaging: at probability 0.05 per qubit per block, which is of the order of current two-qubit error rates accumulated over the 27-layer circuit, the VQC alone collapses to 61.4 % while its mean entropy rises from 0.30 to 0.67. Because the gate reads that entropy, the fused accuracy remains at 98.0 % (Fig. 4(c)). This is the central engineering result: the uncertainty-aware integration makes the twin insensitive to the failure of its quantum component, a necessary property for any deployment on noisy intermediate-scale devices [32].

TABLE IV. VQC UNDER FINITE SHOTS AND DEPOLARIZING NOISE (TEST SPLIT)

Condition	VQC acc. (%)	VQC ECE	Fusion acc. (%)	Fusion NLL
Exact expectations	88.8	0.128	98.0	0.074
128 shots	80.0	0.125	97.6	0.070
512 shots	86.8	0.120	98.0	0.072
2048 shots	88.6	0.126	98.0	0.073
Depolarizing $p = 0.005$	89.2	0.149	98.0	0.069
Depolarizing $p = 0.02$	84.6	0.212	98.0	0.065
Depolarizing $p = 0.05$	61.4	0.177	98.0	0.066

D. Calibration and selective prediction

Fig. 5(a) shows reliability diagrams. The classical ensemble and the fused output lie on the diagonal; the VQC is under-confident below 0.8 confidence even after scaling, which is why Proposition 1 does not hold exactly. Fig. 5(b) shows the risk–coverage curves: the AURC is 0.0005 for the classical ensemble, 0.0015 for fusion and 0.0348 for the VQC. At 90 % coverage the fused rule makes 0.4 % errors, so an abstention threshold that flags one window in ten to an operator removes almost all residual errors.

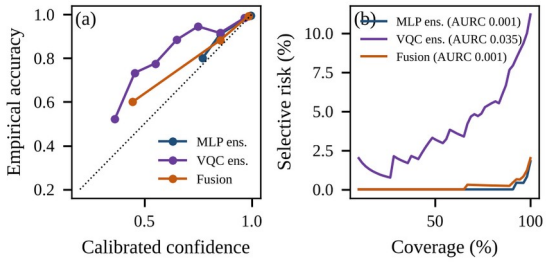


Fig. 5. (a) Reliability diagrams after temperature scaling. (b) Risk–coverage curves; AURC in the legend.

E. Harder and out-of-distribution conditions

Table V reports scenario S5, run after all models were frozen. At −5 dB every model degrades sharply, the parameter-matched control is best (48.2 %), the large MLP overfits (34.4 %) and the VQC is at chance (20.8 % for five classes); at −10 dB all models are at chance and the classical model is confidently wrong (mean entropy 0.01), the one condition in which the gate has no signal to act on. Under a −8 % speed shift the control again beats the large MLP (91.8 % against 82.2 %), so the large model has learned operating-point-specific structure; fusion with the control (89.2 %) follows the control, not the VQC (78.6 %). For incipient defects at 0 dB the gate escalates 34.8 % of windows to the VQC (34.0 %) without improving on the classical 55.8 %. In none of the six conditions did the VQC exceed the parameter-matched control, and in none did fusion exceed the better classical model by more than noise; the hypothesis that bounded angle encoding confers robustness under distribution shift is not supported.

TABLE V. ACCURACY (%) UNDER HARDER AND OUT-OF-DISTRIBUTION CONDITIONS (500 WINDOWS EACH, MODELS FROZEN)

Condition	MLP	Tiny MLP	VQC	Fusion	Oracle
−5 dB	34.4	48.2	20.8	32.4	40.8
−10 dB	18.8	19.6	18.8	18.8	29.4
Speed +8 %, 10 dB	99.6	99.0	92.8	99.2	99.6
Speed −8 %, 10 dB	82.2	91.8	78.6	82.0	88.6
$s \in [0.05, 0.25]$, 5 dB	62.4	65.2	47.8	62.6	67.2
$s \in [0.05, 0.25]$, 0 dB	55.8	46.8	34.0	52.0	59.8

F. Circuit-depth ablation

Scenario S7 trained single VQC members of depth $L = 1, 2, 3$ and 5 (28, 46, 64 and 100 parameters) for 50 epochs from one seed. Best validation accuracies were 48.6, 84.2, 72.2 and 66.8 %: capacity beyond two layers did not help and the deepest circuit trained worst, consistent with the flattening of the loss landscape with depth [14]. Training is also seed-sensitive: the three $L = 3$ members of the deployed ensemble ended at 83.4, 73.6, 81.8 % validation accuracy, and a 120-epoch run plateaued at the same level as 40 epochs. Ensembling is therefore not optional for the VQC; it is what makes its entropy informative enough for the gate to read.

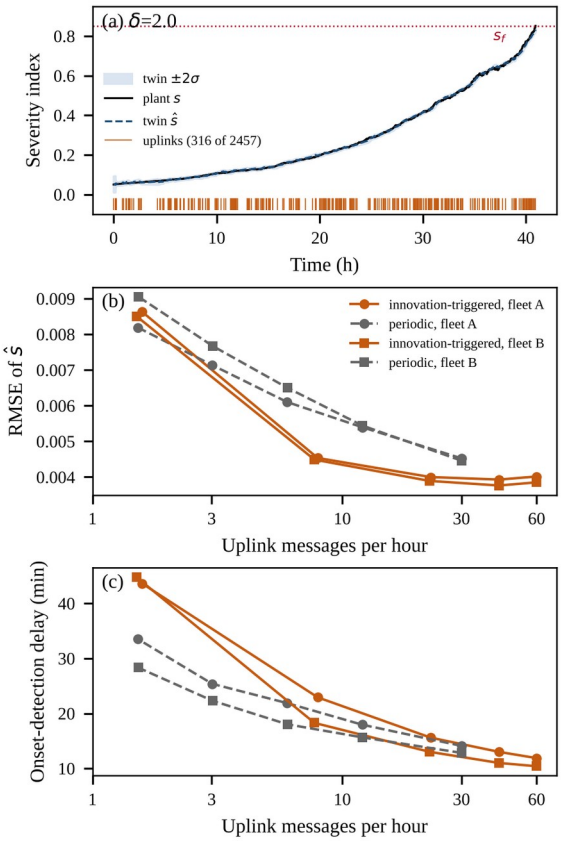


Fig. 6. Twin synchronization. (a) Fleet A trajectory under the innovation trigger with $\delta = 2$: plant severity, twin estimate with 2σ band, and uplink instants (ticks). (b) Severity RMSE and (c) onset-detection delay against uplink rate for innovation-triggered and periodic policies on both fleets; the $\delta = 3$ point is omitted for scale.

G. Twin synchronization

Fig. 6 and Table VI report scenario S6 on Fleet A (life 1164–2518 min, coefficient of variation 0.23). Fig. 6(a) shows one trajectory under $\delta = 2$: the twin tracks the plant severity within its 2σ band using 316 uplinks out of 2457 measurements, and the uplinks cluster in the final hours when severity accelerates, as Proposition 2 implies. Across the fleet, $\delta = 2$ reduces the uplink rate from 60 to 8.0 messages per hour (−87 %) while the severity RMSE rises from 0.0040 to 0.0045 (+13 %); the periodic policy at a comparable rate (every 10 min, 6.0 messages per hour) gives 0.0061. Fig. 6(b) shows that the innovation trigger dominates the periodic policy at every matched rate above about 2 messages per hour on both fleets. Onset-detection delay (first time the 2σ lower bound on $\hat{s}$ exceeds 0.2) does not share this advantage: it is 22.9 min at $\delta = 2$ against 21.9 min for the periodic policy at a similar rate, and below about 5 messages per hour the trigger is slower than periodic sampling (Fig. 6(c)), because the delay is set by the number of measurements accumulated after onset and the trigger withholds exactly the early, small-innovation samples. The trigger therefore improves tracking accuracy per message, not reaction time. Two limits are visible and are reported as such: at $\delta = 2.5$ the rate falls to 1.6 per hour but the onset delay doubles, and at $\delta = 3$ the trigger almost never fires (0.22 per hour) and tracking collapses (RMSE 0.051), because the innovation covariance grows with the open-loop prediction covariance and the threshold is never crossed. A threshold of $\delta \approx 2$ is therefore the operating point on this plant; a formal threshold-selection rule is left as future work.

TABLE VI. SYNCHRONIZATION POLICIES, FLEET A (40 TRAJECTORIES). RUL ERROR IS $|\hat{R} - R|/R$ AT THE STATED FRACTION OF LIFE; A-Λ IS THE FRACTION OF ESTIMATES WITHIN ±20 % AFTER 30 % OF LIFE; COVERAGE IS OF THE NOMINAL 90 % INTERVAL.

Policy	Msg/h	RMSE $\hat{s}$	Onset delay (min)	RUL err. 50 % / 90 %	α-λ (%)	Cov. (%)
Full rate ($\delta = 0$)	60.0	0.0040	11.9 (p90 19)	8.6 / 14.5	87	86
Innovation $\delta = 1.5$	22.6	0.0040	15.6 (p90 23)	8.7 / 14.4	87	87
Innovation $\delta = 2$	8.0	0.0045	22.9 (p90 36)	9.0 / 14.1	86	87
Innovation $\delta = 2.5$	1.6	0.0086	43.6 (p90 68)	10.6 / 12.9	81	90
Innovation $\delta = 3$	0.2	0.0510	433.0 (p90 894)	19.6 / 61.2	49	96
Periodic 10 min	6.0	0.0061	21.9 (p90 35)	8.0 / 13.3	87	87
Periodic 40 min	1.5	0.0082	33.5 (p90 51)	8.1 / 12.8	84	88

H. RUL and decision support

At full rate the RUL point estimate on Fleet A has a mean relative error of 8.6 % at half life and 14.5 % at 90 % of life, 87 % of estimates after 30 % of life fall inside the ±20 % α-λ cone, and the nominal 90 % interval covers 86 % of true values, i.e. the belief is slightly over-confident; on Fleet B (coefficient of variation 0.47, two trajectories censored at the 3,600 min horizon) the figures are 10.7 %, 18.4 % and 79 % coverage. The growth of relative error late in life is a property of the degradation process, not of the estimator: by Eq. 11 the first-passage standard deviation at a true RUL of 170 min is about 30 % of the RUL. Event triggering at $\delta = 2$ leaves all RUL metrics within one percentage point of full rate (Table VI). Fig. 7(a) shows a representative trajectory.

Fig. 7(b) and Table VII give the decision-support result, which is reported without embellishment. The twin-based rule is strongly sensitive to p^*: the risk-neutral value 0.1 predicted in Section V-E is the best or near-best setting on both fleets, whereas $p^* = 0.5$ fails 45–50 % of assets because the rule waits for a median RUL below the lead time. At $p^* = 0.1$ and $\delta = 2$ the mean cost is 1.43 on Fleet A (5.0 % failures) against 1.33 for the best periodic interval, and 1.71 on Fleet B against 1.51; with $p^* = 0.05$ the twin rule has zero failures on both fleets at a cost of 1.81 and 1.78, the extra cost being life wasted by scheduling while the early-life belief is still wide. The twin rule therefore matches, but does not beat, a periodic policy whose interval was tuned with knowledge of the realized fleet life distribution. Its advantage is one of information requirements: the periodic policy is fragile when that knowledge is wrong (on Fleet B, an interval of 0.6 of the mean life already fails 27.5 % of assets and costs 3.35), whereas the twin rule uses only the asset's own

measurements. The assumed c_w is small, and the conclusion should not be extrapolated beyond this cost structure.

TABLE VII. MEAN MAINTENANCE COST PER ASSET (FAILURE FRACTION IN PARENTHESES)

Policy	Fleet A	Fleet B
Run to failure	8.00 (100 %)	7.60 (100 %)
Periodic, 0.5 × mean life	1.55 (0 %)	1.51 (0 %)
Periodic, 0.7 × mean life	1.33 (0 %)	4.15 (40 %)
Twin, full rate, $p^* = 0.1$	1.59 (7.5 %)	1.76 (10.0 %)
Twin, $\delta = 2$, $p^* = 0.1$	1.43 (5.0 %)	1.71 (10.0 %)
Twin, $\delta = 2$, $p^* = 0.05$	1.81 (0.0 %)	1.78 (0.0 %)

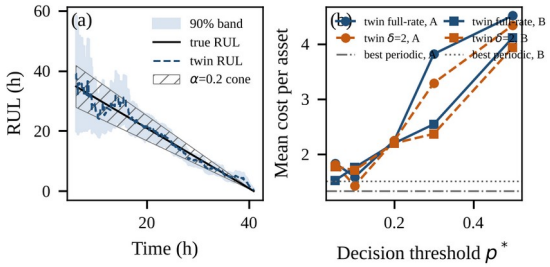


Fig. 7. (a) RUL belief on the Fleet A trajectory of Fig. 6(a): point estimate, 90 % band from Eq. 11 and the ±20 % α-λ cone. (b) Mean maintenance cost against the decision threshold p^* for the twin rule at full rate and at $\delta = 2$, with the best oracle-tuned periodic interval of each fleet as reference.

IX. COMPARATIVE ANALYSIS

TABLE VIII. STRUCTURAL COMPARISON WITH PRIOR APPROACHES

Criterion	[8]	[5]	[28]	[33]	This work
Fault-state model (twin)	No	Risk only	Yes	No	Yes (EKF)
Matched classical control	No	No	n/a	No	Yes (61 vs 64)
Calibrated uncertainty	No	No	No	Conf. only	Entropy, ECE
Hardware-noise evaluation	No	No	n/a	No	Shots, depol.
Sync traffic accounted	n/a	No	No	n/a	−87 % msgs
RUL with interval	No	No	Yes	No	Yes, coverage
Decision cost evaluated	No	No	No	No	Yes, vs periodic
Data	CWRU, JNU	Ind. IoT	Test rig	Benchmarks	Simulation

Table VIII positions the framework against the closest published approaches ([8] PQC + MLP, [5] quantum twin, [28] bearing twin, [33] gated routing) on the criteria of the problem statement; numerical comparison with [8]–[10] is not possible because they use recorded data and different class sets, so the comparison is structural. The decisive differences are that the present framework reports the quantum model against a parameter-matched control and finds no advantage, where prior QML studies compare against larger networks and report parity; makes the twin's dependence on the quantum output conditional on calibrated uncertainty and demonstrates the consequence under hardware noise; and couples inference to a synchronized estimator whose traffic is reduced by 87 % at a stated accuracy cost. On scalability, the classical path costs microseconds per window and the EKF is four-dimensional, so one core serves thousands of assets, while the quantum path is bounded by the escalation fraction; on sustainability, the uplink reduction translates directly into radio energy at the edge.

X. KEY OUTCOMES

(1) A hybrid inference rule whose accuracy equals the better of its two models to within 0.2 points and whose calibration is better than naive averaging, with a proposition stating when it cannot lose. (2) Evidence that the rule neutralizes a failing quantum component: 98.0 % fused accuracy when the VQC alone is at 61.4 % under 5 % depolarizing noise. (3) An innovation-triggered synchronization scheme with a deterministic discrepancy bound that cuts uplink traffic by 87 % for a 13 % increase in severity-tracking error. (4) An RUL belief whose nominal 90 % interval achieves 80–87 % coverage once the first-passage variance is included, and a decision rule that matches an oracle-tuned periodic policy without fleet-level life statistics. (5) The negative result that a 64-parameter six-qubit VQC did not outperform a 61-parameter classical model in any of eleven test conditions. The practical message is that a quantum component can be added to a maintenance twin without risk to the twin's dependability if it is integrated through calibrated uncertainty; whether it adds value is a separate question, answered negatively here for the present circuit family and data scale.

XI. CONCLUSION AND FUTURE SCOPE

An uncertainty-aware quantum–classical digital-twin framework for predictive maintenance of rotating machinery was formulated, analysed and validated in simulation. The entropy-gated cascade and fusion rules keep the twin at classical accuracy while executing the quantum circuit on 4.4 % of windows and while the circuit is subjected to shot and depolarizing noise; the innovation-triggered EKF synchronization reduces uplink traffic by 87 % with a provable bound on twin–plant discrepancy; and the RUL belief with first-passage variance supports a cost-based decision rule that is competitive with the best periodic policy. The VQC itself provided no advantage over a parameter-matched classical model.

The limitations define the future scope. First, all results are simulation-based; the signal model reproduces the kinematic structure of bearing defects but not full bearing dynamics, and the immediate next step is to repeat scenarios S1–S5 on the CWRU and Paderborn data sets [26], [27]. Second, the quantum family was limited to angle-encoded strongly entangling circuits on six qubits; quantum kernels [12] and amplitude encodings with larger feature sets may behave differently and should be tested under the same matched-control protocol. Third, the VQC's residual miscalibration (ECE 0.128 after temperature

scaling) weakens Proposition 1 in practice; per-class or conformal calibration of the quantum output is a direct improvement. Fourth, the threshold δ was chosen by sweep; a rule derived from the open-loop covariance would make the scheme self-tuning, and the implicit information carried by non-transmission [21] was not exploited. Fifth, the maintenance-cost conclusion depends on the assumed cost structure, in particular the small wasted-life cost. Finally, execution on real quantum hardware, including queueing latency and calibration drift [34], must be evaluated before any deployment claim.

ACKNOWLEDGMENT

This work was carried out within the Research and Development Division of Tekstein Solutions Private Limited. No external funding was received.

REFERENCES

[1] F. Tao, H. Zhang, A. Liu, and A. Y. C. Nee, "Digital twin in industry: State-of-the-art," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2405–2415, 2019, doi: 10.1109/TII.2018.2873186.

[2] I. Errandonea, S. Beltrán, and S. Arrizabalaga, "Digital twin for maintenance: A literature review," *Comput. Ind.*, vol. 123, Art. no. 103316, 2020, doi: 10.1016/j.compind.2020.103316.

[3] R. van Dinter, B. Tekinerdogan, and C. Catal, "Predictive maintenance using digital twins: A systematic literature review," *Inf. Softw. Technol.*, vol. 151, Art. no. 107008, 2022, doi: 10.1016/j.infsof.2022.107008.

[4] L. Bukowski and S. Werbinska-Wojciechowska, "Towards Maintenance 5.0: Resilience-based maintenance in AI-driven sustainable and human-centric industrial systems," *Sensors*, vol. 25, no. 16, Art. no. 5100, 2025, doi: 10.3390/s25165100.

[5] I. Ud Din, I. Taj, K. A. Awan, A. Almogren, and A. Altameem, "Quantum and GAN-driven digital twin approach for IoT-based consumer electronics manufacturing," *IEEE Internet Things J.*, vol. 12, no. 4, pp. 3734–3741, 2025, doi: 10.1109/JIOT.2024.3499737.

[6] W. Maass *et al.*, "QUASIM: Quantum computing enhanced service ecosystem for simulation in manufacturing," *KI – Künstliche Intelligenz*, vol. 38, no. 4, pp. 361–370, 2024, doi: 10.1007/s13218-024-00860-x.

[7] J. Villalba-Diez, M. Gutierrez, M. Grijalvo Martín, T. Sterkenburgh, J. C. Losada, and R. M. Benito, "Quantum JIDOKA. Integration of quantum simulation on a CNC machine for in-process control visualization," *Sensors*, vol. 21, no. 15, Art. no. 5031, 2021, doi: 10.3390/s21155031.

[8] C. B. S. Maior, L. M. M. Araújo, I. D. Lins, M. das Chagas Moura, and E. López Droguett, "Prognostics and health management of rotating machinery via quantum machine learning," *IEEE Access*, vol. 11, pp. 25132–25151, 2023, doi: 10.1109/ACCESS.2023.3255417.

[9] L. M. M. Araújo, I. D. Lins, C. B. S. Maior, M. J. das Chagas Moura, and E. López Droguett, "Rotating machinery health state diagnosis through quantum machine learning," in *Proc. 33rd Eur. Safety Rel. Conf. (ESREL)*, Southampton, U.K., 2023, doi: 10.3850/978-981-18-8071-1_P228-cd.

[10] Y. Li, L. Song, Q. Sun, H. Xu, X. Li, Z. Fang, and W. Yao, "Rolling bearing fault diagnosis based on quantum LS-SVM," *EPJ Quantum Technol.*, vol. 9, Art. no. 18, 2022, doi: 10.1140/epjqt/s40507-022-00137-y.

[11] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, "Circuit-centric quantum classifiers," *Phys. Rev. A*, vol. 101, no. 3, Art. no. 032308, 2020, doi: 10.1103/PhysRevA.101.032308.

[12] V. Havlíček *et al.*, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, no. 7747, pp. 209–212, 2019, doi: 10.1038/s41586-019-0980-2.

[13] M. Cerezo *et al.*, "Variational quantum algorithms," *Nat. Rev. Phys.*, vol. 3, no. 9, pp. 625–644, 2021, doi: 10.1038/s42254-021-00348-9.

[14] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nat. Commun.*, vol. 9, Art. no. 4812, 2018, doi: 10.1038/s41467-018-07090-4.

[15] V. Bergholm *et al.*, "PennyLane: Automatic differentiation of hybrid quantum-classical computations," arXiv:1811.04968, 2018.

[16] M. Schuld, R. Sweke, and J. J. Meyer, "Effect of data encoding on the expressive power of variational quantum-machine-learning models," *Phys. Rev. A*, vol. 103, no. 3, Art. no. 032430, 2021, doi: 10.1103/PhysRevA.103.032430.

[17] B. Lakshminarayanan, A. Pritzel, and C. Blundell, "Simple and scalable predictive uncertainty estimation using deep ensembles," in *Adv. Neural Inf. Process. Syst. 30*, Long Beach, CA, USA, 2017.

[18] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, PMLR vol. 70, Sydney, Australia, 2017, pp. 1321–1330.

[19] M. Abdar *et al.*, "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Inf. Fusion*, vol. 76, pp. 243–297, 2021, doi: 10.1016/j.inffus.2021.05.008.

[20] S. Trimpe and R. D'Andrea, "Event-based state estimation with variance-based triggering," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3266–3281, 2014, doi: 10.1109/TAC.2014.2351951.

[21] D. Han, Y. Mo, J. Wu, S. Weerakkody, B. Sinopoli, and L. Shi, "Stochastic event-triggered sensor schedule for remote state estimation," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2661–2675, 2015, doi: 10.1109/TAC.2015.2406975.

[22] B. Tan and A. Matta, "The digital twin synchronization problem: Framework, formulations, and analysis," *IISE Trans.*, vol. 56, no. 6, pp. 652–665, 2024, doi: 10.1080/24725854.2023.2253869.

[23] L. V. Çakır, S. Al-Shareeda, S. F. Oktuğ, M. Özdem, M. Broadbent, and B. Canberk, "How to synchronize digital twins? A communication performance analysis," in *Proc. IEEE 28th Int. Workshop Comput. Aided Model. Design Commun. Links Netw. (CAMAD)*, Edinburgh, U.K., 2023, pp. 123–127, doi: 10.1109/CAMAD59638.2023.10478422.

[24] P. D. McFadden and J. D. Smith, "Model for the vibration produced by a single point defect in a rolling element bearing," *J. Sound Vib.*, vol. 96, no. 1, pp. 69–82, 1984, doi: 10.1016/0022-460X(84)90595-9.

[25] R. B. Randall and J. Antoni, "Rolling element bearing diagnostics—A tutorial," *Mech. Syst. Signal Process.*, vol. 25, no. 2, pp. 485–520, 2011, doi: 10.1016/j.ymssp.2010.07.017.

[26] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study," *Mech. Syst. Signal Process.*, vols. 64–65, pp. 100–131, 2015, doi: 10.1016/j.ymssp.2015.04.021.

[27] C. Lessmeier, J. K. Kimotho, D. Zimmer, and W. Sextro, "Condition monitoring of bearing damage in electromechanical drive systems by using motor current signals of electric motors: A benchmark data set for data-driven classification," in *Proc. Eur. Conf. PHM Soc.*, vol. 3, no. 1, Bilbao, Spain, 2016, doi: 10.36001/phme.2016.v3i1.1577.

[28] Y. Qin, X. Wu, and J. Luo, "Data-model combined driven digital twin of life-cycle rolling bearing," *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 1530–1540, 2022, doi: 10.1109/TII.2021.3089340.

[29] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mech. Syst. Signal Process.*, vol. 104, pp. 799–834, 2018, doi: 10.1016/j.ymssp.2017.11.016.

[30] W. Zhao, C. Zhang, J. Wang, S. Wang, D. Lv, and F. Qin, "Research on digital twin driven rolling bearing model-data fusion life prediction method," *IEEE Access*, vol. 11, pp. 48611–48627, 2023, doi: 10.1109/ACCESS.2023.3277040.

[31] R. R. Schoen, T. G. Habetler, F. Kamran, and R. G. Bartheld, "Motor bearing damage detection using stator current monitoring," *IEEE Trans. Ind. Appl.*, vol. 31, no. 6, pp. 1274–1279, 1995, doi: 10.1109/28.475697.

[32] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018, doi: 10.22331/q-2018-08-06-79.

[33] S. Rehman, N. Innan, and M. Shafique, "QSTAR: Quantum selective transfer with adaptive routing," arXiv:2607.21411, 2026.

[34] E. Bautra, M. Dimitrijevs, and A. Yakaryilmaz, "Evaluating calibration-based digital twins for IBM quantum hardware simulation," arXiv:2603.14607, 2026.