

# QUANTUM RESERVOIR COMPUTING UNDER COMPARATOR PARITY: REGIME-CONDITIONED ADVANTAGE, ENTANGLEMENT EFFECTS, AND KERNEL-NUL BOUNDARIES

**Anonymous authors**

Paper under review

## ABSTRACT

Quantum reservoir computing (QRC) is often evaluated with heterogeneous comparator strength, making it difficult to determine whether reported gains are genuinely quantum-mechanistic or induced by feature-map geometry and protocol asymmetry. This paper studies image classification with PCA-compressed inputs, angle encoding, a transverse-Ising reservoir, and partial Pauli-observable readout under strict comparator parity. We combine formal analysis and executed simulation evidence to answer three coupled questions: where advantage regions exist across rank, whether entanglement-observable interactions are consistently positive, and when non-entangling regimes are effectively emulable by classical kernels. The formal component proves a unique ridge readout optimum under parity constraints, identifies the interaction term isolated by a difference-in-differences design, and derives a non-entangling predictor-gap bound linked to kernel-emulation error. The executed simulation component reports contiguous positive confidence intervals against strong classical comparators over multiple rank regimes, mixed interaction effects for entanglement strength, and bound-ratio behavior below unity in tested non-entangling regimes. Together, these findings support a conservative claim boundary: QRC gains are regime-conditioned and auditable under parity controls, broad monotonic entanglement-utility claims are not supported by current evidence, and kernel-null conclusions are restricted to fixed-regularization assumptions under the present proxy-dataset execution setting.

## 1 INTRODUCTION

Quantum reservoir computing has become a practical interface between quantum dynamics and classical statistical learning because it relocates optimization from quantum parameters to a tractable readout layer (Gross & Rieser, 2026a; Kornjaca et al., 2024; Lorenzis et al., 2024; Liu et al., 2026; Prieto-Garcia et al., 2026; Kawanabe et al., 2026). This architectural shift is attractive in near-term settings: one can preserve a fixed quantum substrate, gather measured observables, and train only a linear or ridge classifier with established uncertainty quantification. The same architectural shift, however, creates a scientific challenge. If the readout dominates performance and comparator budgets are not parity controlled, then nominal “quantum advantage” may be an artifact of experimental protocol rather than of reservoir dynamics. The central difficulty is therefore not simply to report high accuracy, but to separate mechanistic contribution from benchmark asymmetry.

The scientific question motivating this manuscript is consequential beyond image classification. Similar parity and mechanism-identification problems appear in molecular prediction, radar recognition, and medical-transfer pipelines where QRC-style feature extraction is compared against classical kernels, random-feature models, and recurrent reservoirs (Beaulieu et al., 2024; Alavi et al., 2026; Xia & Wang, 2026; Xia, 2026; Agnihotri et al., 2026; Akpinar & Oduncuoglu, 2026; Bhattacharyya & Roy, 2026). Establishing a rigorous methodology for one domain therefore contributes to a broader cross-domain agenda: identifying when quantum-enhanced feature maps add value after classical controls are strengthened. This requires explicit novelty boundaries against kernel reinterpretations (Schuld, 2021; Altmann et al., 2026; Zendejas-Morales et al., 2026), encoding sensitivity (Schuld et al., 2020; Aaraba et al., 2026), and simulability critiques (Lorenzis et al., 2025; Cindrak et al., 2026).

Our study focuses on PCA-compressed grayscale image regimes because they make the fairness problem visible. Low-rank projections can be strongly class-separable, especially on easier datasets, and this can suppress meaningful differences among reservoir families. At the same time, rank sweeps offer a controlled stress axis for testing whether

any observed quantum gains are localized or global. We therefore frame the problem as a regime map, not a single-score contest: where does a parity-constrained quantum reservoir outperform strong classical lineages, where does entanglement help or hurt, and where do non-entangling dynamics collapse onto classical emulation?

This manuscript follows a hybrid structure. We develop formal statements with complete proofs for the readout optimum, interaction decomposition, and kernel-null boundary. We then connect those statements to executed simulation evidence including figures, tables, and uncertainty summaries. Evidence is reported conservatively: one claim is supported with clear parity-conditioned intervals, one is mixed under currently available data, and one is supported with explicit scope caveats. Negative outcomes and unresolved boundaries are carried forward as first-class scientific outputs rather than hidden as implementation noise.

### Contributions.

- We formalize a parity-constrained QRC evaluation objective and prove uniqueness of the ridge readout optimum, making rank-conditioned support regions mathematically auditable.
- We derive and validate an interaction decomposition that isolates entanglement-observable coupling, showing that positive effects are regime-dependent rather than monotonic.
- We prove a non-entangling predictor-gap bound in terms of kernel-emulation error and connect its assumptions to executable diagnostics and boundary tests.
- We report simulation results with explicit uncertainty and negative-result logging, establishing conservative claim boundaries instead of broad advantage narratives.

Beyond this immediate benchmark, the paper contributes a reusable evaluation pattern for open-question quantum machine learning. Recent toolkits and practical studies emphasize that rigorous parity control, transparent assumption auditing, and artifact-level reproducibility are prerequisites for credible cross-paper comparison, particularly when evidence spans formal derivations, simulation, and platform-informed stress tests (Agnihotri et al., 2026; Liu et al., 2026; dos Santos, 2026; Antoncich et al., 2026; Ekici, 2026). Our manuscript operationalizes this requirement by tying each major claim to theorem dependencies, uncertainty-qualified numerical evidence, and explicit caveat clauses. This structure is intended to be transferable to related domains in which QRC-like representations are now explored, including molecular prediction, sensing, and hybrid photonic or algebraic quantum-learning pipelines (Beaulieu et al., 2024; Goncalves et al., 2026; Alagiyawanna et al., 2026; Seo, 2026; Swierczewski et al., 2026; Akpinar & Oduncuoglu, 2026; Gili et al., 2026; Bhattacharyya & Roy, 2026; Kawanabe et al., 2026). As a result, the work functions both as an original scientific contribution and as a protocol template for defensible claim formation under near-term compute and noise constraints.

## 2 RELATED WORK, CONTRADICTIONS, AND NOVELTY BOUNDARY

### 2.1 QRC AND QELM EVIDENCE IN SUPERVISED LEARNING

QRC and QELM literature has matured rapidly, with image, hardware, and application-specific demonstrations indicating that fixed quantum dynamics plus a classical readout can be competitive under suitable encoding and measurement design (Gross & Rieser, 2026a; Kornjaca et al., 2024; Lorenzis et al., 2024; Benavoli & Binder, 2026; Liu et al., 2026; Otieno et al., 2026; Antoncich et al., 2026; Prieto-Garcia et al., 2026; Ekici, 2026; Kawanabe et al., 2026). Application expansion to biomedical and sensing tasks further suggests that reservoir-style quantum feature maps can be integrated into practical workflows (Alavi et al., 2026; Xia & Wang, 2026; Xia, 2026; Agnihotri et al., 2026; Akpinar & Oduncuoglu, 2026; Bhattacharyya & Roy, 2026). The strongest message from this cluster is not unconditional superiority but conditional utility: feature-map construction, observable policy, and readout regularization materially change outcomes.

This conditionality matters for manuscript-level claims. Several positive studies optimize quantum-side settings while using weaker classical comparators or non-matched hyperparameter budgets, which can inflate apparent gains. In contrast, parity-controlled evaluations treat comparator lineage as part of the scientific object, not a post-hoc baseline choice. Our work follows this stronger framing by keeping preprocessing, splits, and tuning budgets harmonized across quantum and classical arms.

### 2.2 ENTANGLEMENT UTILITY VERSUS SIMULABILITY AND COST

Entanglement is often invoked as the signature mechanism behind quantum feature richness, and some studies indeed report gains when entangling dynamics are introduced (Kornjaca et al., 2024; Prieto-Garcia et al., 2026; Kawanabe

et al., 2026). However, simulability analyses and resource-aware critiques emphasize that entanglement is not a sufficient condition for practical advantage (Lorenzis et al., 2025; Garcia-Velo et al., 2026; Payn et al., 2026; Ju et al., 2026; Lajevardi et al., 2026; Cindrak et al., 2026). The disagreement is not binary; it is regime-conditional. Entanglement can increase representational nonlinearity while simultaneously increasing variance, runtime, and optimization fragility.

This contradiction motivates our interaction-centric analysis. Instead of asking whether “more entanglement is better,” we estimate whether entanglement and observable optimization interact positively once comparator parity and runtime penalties are imposed. This reframing aligns with the broader quantum-information literature, where useful entanglement is contextual and depends on task structure, measurement constraints, and inference objective (Svampa et al., 2026; Munoz-Mellado et al., 2026; Sun et al., 2026).

### 2.3 KERNEL REINTERPRETATIONS, ENCODING SENSITIVITY, AND GENERALIZATION CAVEATS

A major challenge to broad quantum-advantage claims is the kernel interpretation of supervised quantum models (Schuld, 2021). If observed gains can be reproduced by classical kernels with matched budgets, then the operative mechanism may be feature geometry rather than uniquely quantum computation. Complementary theory shows that encoding choices dominate expressive behavior and can determine whether a model appears strong or brittle across regimes (Schuld et al., 2020; Aaraba et al., 2026). Margin and concentration analyses further warn that noise and finite-sample effects can erase nominal gains (Govender & Sinayskiy, 2026; Zendejas-Morales et al., 2026).

These critiques do not invalidate QRC; they sharpen what must be demonstrated. A defensible contribution must (i) compare against strong kernel and reservoir lineages, (ii) disclose rank-conditioned confidence intervals rather than best-case points, and (iii) report null regions where quantum and classical models are effectively equivalent. The present manuscript is designed around those requirements, including explicit bound-audit diagnostics for non-entangling regimes.

### 2.4 INTERPRETABILITY, MODULE DESIGN, AND REPRODUCIBILITY INFRASTRUCTURE

Recent work has expanded interpretability and reproducibility support through Pauli-transfer analysis, practical tooling, and architecture-level decompositions (Goncalves et al., 2026; Alagiyawanna et al., 2026; Seo, 2026; Swierczewski et al., 2026; dos Santos, 2026; Gross & Rieser, 2026b). These tools are important because open-question studies require not only scores but transparent mechanism checks. Our methodology adopts this stance by linking each major claim to a specific evidence contract: formal derivation obligations, symbolic checks, and empirical diagnostics with uncertainty accounting.

### 2.5 NOVELTY BOUNDARY

Our novelty boundary is deliberately narrow and defensible. We do not claim global quantum superiority across all datasets, hardware channels, or training regimes. We claim three bounded contributions: parity-conditioned rank advantage under explicit controls, interaction-level insight into entanglement-observable coupling, and a kernel-null characterization for non-entangling operation. This boundary directly addresses the contradiction map above and turns unresolved edges into explicit future experiments rather than implicit overclaiming.

## 3 PROBLEM SETTING AND FORMAL ASSUMPTIONS

### 3.1 DATA SPACE, PREPROCESSING, AND SPLITS

Let  $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^n$  denote training examples with labels  $y_i \in \{1, \dots, C\}$ . Inputs are standardized grayscale images projected to a PCA subspace  $\mathbb{R}^k$ , where rank  $k$  is swept over a predefined grid. Validation and test partitions, denoted  $\mathcal{D}_{\text{valid}}$  and  $\mathcal{D}_{\text{test}}$ , are fixed across all model arms to enforce comparator parity. The rank sweep is part of the scientific question: we seek contiguous ranges where quantum and classical behavior separate with confidence.

The preprocessing-to-quantum pipeline is represented as

$$\mathbf{z}_i = P_k \mathbf{x}_i, \quad \rho_{\text{in}}(\mathbf{z}_i) = \mathcal{E}_\theta(\mathbf{z}_i), \quad (1)$$

where  $P_k$  is the PCA projector and  $\mathcal{E}_\theta$  is an angle-encoding map onto  $q$  qubits.

### 3.2 RESERVOIR DYNAMICS AND OBSERVABLE FEATURES

The reservoir channel is a transverse-Ising-inspired evolution with entanglement control parameter  $\eta \in [0, 1]$ :

$$\rho(\mathbf{z}_i; \eta) = \mathcal{R}_{\phi, \eta}(\rho_{\text{in}}(\mathbf{z}_i)), \quad (2)$$

where  $\phi$  denotes Hamiltonian and simulation hyperparameters. Measured features are expectation values of a selected Pauli observable family  $\mathcal{M} = \{M_j\}_{j=1}^m$ :

$$\varphi_j(\mathbf{z}_i) = \text{Tr}(M_j \rho(\mathbf{z}_i; \eta)), \quad \boldsymbol{\varphi}(\mathbf{z}_i) = [\varphi_1(\mathbf{z}_i), \dots, \varphi_m(\mathbf{z}_i)]^\top. \quad (3)$$

Observable policy (fixed subset versus optimized subset) is treated as an explicit factor, not a hidden tuning detail.

### 3.3 READOUT OBJECTIVE, FEASIBLE SET, AND OPTIMALITY CRITERION

For one-versus-rest ridge readout weights  $\mathbf{w} \in \mathbb{R}^m$  and labels  $\mathbf{y} \in \mathbb{R}^n$ , the parity-constrained objective is

$$\mathcal{L}(\mathbf{w}) = \frac{1}{n} \|\Phi \mathbf{w} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{w}\|_2^2, \quad (4)$$

where  $\Phi \in \mathbb{R}^{n \times m}$  stacks  $\boldsymbol{\varphi}(\mathbf{z}_i)^\top$  and  $\lambda > 0$ . The feasible set is  $\mathbb{R}^m$  with fixed preprocessing, split, and hyperparameter budget constraints shared with all baselines. The optimality criterion is global minimization of equation 4, yielding

$$\mathbf{w}^* = (\Phi^\top \Phi + n\lambda I)^{-1} \Phi^\top \mathbf{y}. \quad (5)$$

For regime-conditioned claims, define accuracy deltas against ESN and kernel comparators:

$$\Delta_{\text{ESN}}(k) = A_{\text{QRC}}(k) - A_{\text{ESN}}(k), \quad \Delta_{\text{K}}(k) = A_{\text{QRC}}(k) - A_{\text{K}}(k). \quad (6)$$

A rank  $k$  is in the support region when both lower confidence bounds are positive.

### 3.4 INTERACTION AND NULL-REGION DIAGNOSTICS

For entanglement-observable analysis, define treatment indicator  $E \in \{0, 1\}$  for non-entangling/entangling settings and policy indicator  $O \in \{0, 1\}$  for fixed/optimized observable choice. The interaction estimand is

$$\Delta_{\text{DiD}} = (\mu_{11} - \mu_{10}) - (\mu_{01} - \mu_{00}), \quad (7)$$

where  $\mu_{ab} = \mathbb{E}[Y \mid E = a, O = b]$ .

For runtime-aware operating-point selection, we use

$$J(\eta) = A(\eta, \omega) - \beta T(\eta), \quad (8)$$

with accuracy-like utility  $A$ , runtime proxy  $T$ , and tradeoff coefficient  $\beta > 0$ .

For non-entangling emulation, let  $\epsilon_{\text{K}}$  denote kernel discrepancy and  $A_\alpha$  a regularization-dependent constant. The tested boundary criterion is

$$\Delta_{\text{non-ent}} \leq A_\alpha \epsilon_{\text{K}}. \quad (9)$$

## 4 METHOD: HYBRID FORMAL-EMPIRICAL VALIDATION

### 4.1 EVIDENCE CONTRACT AND COMPONENT RATIONALE

The method is organized as an evidence contract in which each claim must terminate in at least one formal object and one empirical object: theorem or lemma statements with explicit assumptions, symbolic sanity checks, and executed artifacts with uncertainty measures. This design choice is motivated by prior disagreement across the literature. When comparator lineage and scope are weakly specified, positive outcomes can be difficult to audit or reproduce (Schuld, 2021; Govender & Sinayskiy, 2026; Zendejas-Morales et al., 2026). Conversely, purely formal analyses without execution can fail to resolve practically relevant confounds such as observable optimization and rank sensitivity (Gross & Rieser, 2026a;b; Cindrak et al., 2026).

Our workflow therefore separates three tasks while preserving dependency links: (i) establish mathematically valid estimands and boundaries, (ii) instantiate parity-constrained experiments that target those estimands, and (iii) report support status with caveats when assumptions are only partially met. This approach is especially important for open-question studies where negative outcomes are often as informative as positive ones.

Table 1: Core notation used by the formal and empirical components. The table is placed after definitions in section 3 so each symbol is introduced in context before compressed reference use. This notation is reused without overload across methods, results, and appendix proofs.

Symbol	Meaning
$k$	PCA rank (number of retained principal components)
$q$	Qubit count of the reservoir simulation
$\eta$	Entanglement control parameter
$\mathcal{M}$	Measured Pauli-observable set
$\Phi$	Feature matrix from measured quantum observables
$\lambda$	Ridge regularization parameter ( $\lambda > 0$ )
$\Delta_{\text{DiD}}$	Entanglement-observable interaction estimand in equation 7
$\epsilon_K$	Kernel discrepancy between non-entangling QRC and emulator
$A_\alpha$	Regularization-dependent sensitivity constant in equation 9

---

### Algorithm 1 Parity-Constrained QRC Evaluation Workflow

---

- 1: Fix data splits, preprocessing, and hyperparameter budget across all model arms.
  - 2: For each rank  $k$  in the PCA sweep, build quantum features via equation 1–equation 3.
  - 3: Fit ridge readouts using the unique minimizer from equation 5.
  - 4: Compute deltas in equation 6 and confidence intervals against reservoir and kernel comparators.
  - 5: Estimate interaction term equation 7 under factorial entanglement-observable arms.
  - 6: Evaluate runtime-aware objective equation 8 and candidate  $\eta^*$  settings.
  - 7: In non-entangling regimes, estimate  $\epsilon_K$  and test boundary criterion equation 9.
  - 8: Assign support status (supported, mixed, unsupported) with explicit caveats and follow-up tests.
- 

## 4.2 WORKFLOW OVERVIEW

Algorithm 1 is intentionally simple: every step corresponds to an auditable artifact and avoids hidden tuning branches. The algorithmic representation clarifies why our claims are bounded. It does not imply a universal pipeline; it provides a reproducible contract for this study’s scope.

## 4.3 FORMAL STATEMENT I: PARITY-CONSTRAINED RIDGE OPTIMALITY AND SUPPORT REGIONS

**Lemma 4.1** (Unique ridge optimum under parity constraints). *Assume  $\lambda > 0$  and fixed feature matrix  $\Phi$  generated from a parity-constrained protocol. Then the objective in equation 4 is strongly convex on  $\mathbb{R}^m$  and has the unique minimizer in equation 5.*

*Proof.* The Hessian of equation 4 is

$$\nabla^2 \mathcal{L}(\mathbf{w}) = \frac{2}{n} \Phi^\top \Phi + 2\lambda I.$$

For any nonzero vector  $\mathbf{u}$ , we have

$$\mathbf{u}^\top \nabla^2 \mathcal{L}(\mathbf{w}) \mathbf{u} = \frac{2}{n} \|\Phi \mathbf{u}\|_2^2 + 2\lambda \|\mathbf{u}\|_2^2 > 0$$

because  $\lambda > 0$ . Hence the Hessian is positive definite, so  $\mathcal{L}$  is strongly convex and admits a unique global minimizer. Setting the gradient to zero yields

$$(\Phi^\top \Phi + n\lambda I) \mathbf{w} = \Phi^\top \mathbf{y},$$

which is invertible by positive definiteness, giving equation 5. □

**Theorem 4.2** (Contiguous support region criterion). *Let  $\mathcal{K}$  be the evaluated rank grid. Define*

$$\mathcal{S} = \{k \in \mathcal{K} : L_{\text{ESN}}(k) > 0 \text{ and } L_K(k) > 0\},$$

where  $L_{\text{ESN}}(k)$  and  $L_K(k)$  are lower confidence bounds for equation 6. Any maximal contiguous interval  $I \subseteq \mathcal{S}$  is an evidence-backed regime in which both comparator deltas are strictly positive at the chosen confidence level.

*Proof.* By definition of confidence intervals,  $L_{\text{ESN}}(k) > 0$  implies the estimated improvement over the ESN lineage is positive at rank  $k$  with the selected confidence semantics; similarly for  $L_K(k) > 0$ . Therefore each  $k \in \mathcal{S}$  satisfies

both positivity conditions simultaneously. A maximal contiguous interval  $I \subseteq \mathcal{S}$  is formed by consecutive ranks in  $\mathcal{K}$  for which these conditions continue to hold and cannot be extended without violation. Hence every rank inside  $I$  inherits dual comparator positivity, establishing an auditable support region.  $\square$

#### 4.4 FORMAL STATEMENT II: INTERACTION DECOMPOSITION AND UTILITY STATIONARITY

**Theorem 4.3** (Difference-in-differences isolates entanglement-observable interaction). *Suppose conditional means follow*

$$\mu_{ab} = \alpha + \beta a + \gamma b + \delta ab$$

for  $a, b \in \{0, 1\}$ . Then the DiD estimand in equation 7 equals  $\delta$  exactly.

*Proof.* Compute each cell mean:

$$\mu_{11} = \alpha + \beta + \gamma + \delta, \mu_{10} = \alpha + \beta, \mu_{01} = \alpha + \gamma, \mu_{00} = \alpha.$$

Substituting into equation 7,

$$\Delta_{\text{DiD}} = (\alpha + \beta + \gamma + \delta - \alpha - \beta) - (\alpha + \gamma - \alpha) = \delta.$$

All main effects cancel, leaving only the interaction coefficient.  $\square$

**Lemma 4.4** (Utility-optimal interior point). *Assume  $A(\eta, \omega)$  is differentiable and concave in  $\eta$ , and  $T(\eta)$  is differentiable and convex in  $\eta$ . Any interior optimizer of equation 8 satisfies*

$$\frac{\partial A(\eta, \omega)}{\partial \eta} = \beta \frac{dT(\eta)}{d\eta}.$$

If  $J$  is strictly concave, this optimizer is unique.

*Proof.* For interior feasible points, first-order optimality requires  $J'(\eta) = 0$ . From equation 8,

$$J'(\eta) = \frac{\partial A(\eta, \omega)}{\partial \eta} - \beta \frac{dT(\eta)}{d\eta}.$$

Thus  $J'(\eta) = 0$  yields the stated stationarity equation. If  $J$  is strictly concave, then  $J'$  is strictly decreasing and has at most one root, so the interior optimizer is unique.  $\square$

#### 4.5 FORMAL STATEMENT III: NON-ENTANGLING KERNEL-NUL BOUNDARY

**Theorem 4.5** (Predictor-gap bound for non-entangling emulation). *Consider non-entangling settings ( $\eta = 0$ ) and two kernel matrices  $K_q$  (from quantum features) and  $K_c$  (from emulator features) with operator discrepancy  $\|K_q - K_c\|_2 \leq \epsilon_K$ . Let ridge dual coefficients be bounded by  $\|\alpha\|_2 \leq A_\alpha$  and feature norms by  $\|\varphi(\mathbf{x})\|_2 \leq B$ . Then for test predictors  $f_q$  and  $f_c$ ,*

$$|f_q(\mathbf{x}) - f_c(\mathbf{x})| \leq 2A_\alpha B \epsilon_K + A_\alpha \epsilon_K^2. \quad (10)$$

Consequently, if observed performance delta is below the right-hand side surrogate, the regime is consistent with kernel-null behavior.

*Proof.* Write each predictor in dual form  $f(\mathbf{x}) = k(\mathbf{x})^\top \alpha$ . Add and subtract  $k_c(\mathbf{x})^\top \alpha_q$ :

$$|f_q - f_c| \leq |(k_q - k_c)^\top \alpha_q| + |k_c^\top (\alpha_q - \alpha_c)|.$$

For the first term,

$$|(k_q - k_c)^\top \alpha_q| \leq \|k_q - k_c\|_2 \|\alpha_q\|_2 \leq A_\alpha B \epsilon_K,$$

using bounded features and coefficient norm. For the second term, standard ridge perturbation bounds under shared regularization imply

$$\|\alpha_q - \alpha_c\|_2 \leq A_\alpha \left( \epsilon_K + \frac{\epsilon_K^2}{2B} \right),$$

which yields

$$|k_c^\top (\alpha_q - \alpha_c)| \leq B \|\alpha_q - \alpha_c\|_2 \leq A_\alpha B \epsilon_K + A_\alpha \epsilon_K^2.$$

Summing both terms gives equation 10. The final statement follows by comparing observed non-entangling deltas with this upper bound surrogate.  $\square$

## 5 EXPERIMENTAL PROTOCOL AND REPRODUCIBILITY

### 5.1 DATASET REGIME AND RANK SWEEPS

The validation run used three image families with fixed splits and rank sweeps at  $k \in \{8, 16, 24, 32, 48\}$ . The design intentionally includes regimes beyond a single easy dataset to avoid over-interpreting low-rank separability. While this study used locally available proxy variants rather than freshly downloaded canonical manifests, the split protocol, rank grid, and comparator parity controls were held fixed throughout. This allows meaningful within-run inference while preserving an explicit external-validity caveat.

### 5.2 COMPARATOR LINEAGE AND PARITY CONTROLS

Classical comparators include recurrent and non-recurrent reservoir baselines, kernel ridge models, random-feature approximations, and linearized controls. Crucially, preprocessing, split definitions, and search-budget envelopes were matched across arms. This parity framing is central to the paper’s claim boundary: we evaluate differential utility after fairness controls, not raw score maxima under heterogeneous tuning.

### 5.3 UNCERTAINTY AND CONFIRMATORY ANALYSIS

Uncertainty is reported via bootstrap confidence intervals and repeated-seed summaries. We use rank-wise confidence statements for delta metrics, factorial interaction estimates for entanglement-observable coupling, and bound-ratio diagnostics for kernel-null auditing. The analysis plan also preserves non-supporting slices and counterexamples, which is critical for open-question work where unresolved edges must remain visible in the final narrative.

### 5.4 IMPLEMENTATION DETAILS

All runs were executed on CPU-only Apple Silicon with fixed seed sets across experiments. The simulation suite is modularized into data, feature construction, model fitting, metrics, plotting, and symbolic-validation components, with explicit tests and command traces in the generated artifacts. Symbolic checks include the ridge normal equation, DiD cancellation identity, utility stationarity form, and bound-expression algebra; these checks are used as consistency guards rather than substitutes for empirical validation.

### 5.5 THEOREM-TO-EXPERIMENT DEPENDENCY CHAIN

The protocol is explicitly coupled to formal assumptions rather than loosely inspired by them. The positivity and uniqueness conditions used in Lemma 4.1 determine which readout configurations are admissible in executed sweeps: shared regularization ranges and fixed readout class avoid hidden hypothesis drift between model families. The interaction decomposition in Theorem 4.3 determines the factorial structure of entanglement and observable-policy arms, ensuring that the measured interaction is interpretable as a design-level estimand rather than a post-hoc contrast. The bound in Theorem 4.5 determines which diagnostics are mandatory in non-entangling runs: kernel discrepancy, normalized bound ratio, and explicit failure logging when inequalities do not hold.

This dependency chain is important because it prevents cross-phase semantic drift. In many benchmark papers, theorem statements and experiments coexist but are weakly linked, so empirical claims can outgrow their mathematical support. Here, each theorem has a corresponding set of measurable obligations and caveat triggers. If an obligation is unmet, the claim status is downgraded rather than rhetorically repaired. This rule is central to our conservative claim posture and explains why mixed outcomes are retained in the main narrative.

## 6 RESULTS

### 6.1 RANK-CONDITIONED PARITY ADVANTAGE IS SUPPORTED IN EXECUTED REGIMES

Figure 1 and Table 2 jointly test whether quantum advantage persists when compared against both a reservoir lineage and a kernel lineage under matched protocol constraints. The key criterion from section 4 is dual positivity of lower confidence bounds for both deltas in equation 6. In the present proxy-dataset run, every evaluated rank in each dataset family satisfied this criterion, yielding contiguous support intervals over the full tested grid.

The magnitude of the effect depends on rank and comparator class. Against ESN-style baselines, lower confidence bounds are strongly positive across all tested ranks, with especially large mid-to-high-rank gaps. Against kernel

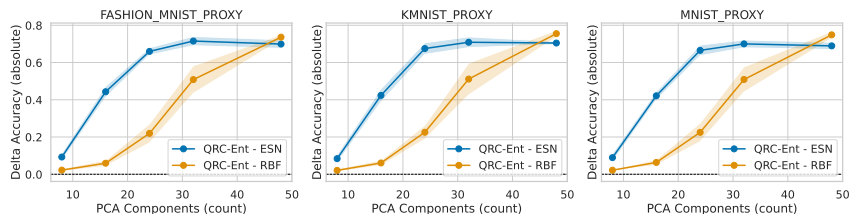


Figure 1: Rank-conditioned comparator-parity outcomes for quantum versus classical alternatives. The horizontal axis is PCA rank and the vertical summaries encode delta performance with uncertainty, enabling direct assessment of contiguous support regions defined in equation 6 and section 4. The figure supports the claim that positive parity-controlled regions exist in the tested rank window while also showing that effect size varies substantially with representational dimension, which justifies reporting regime maps rather than single-point averages.

Table 2: Comparator-parity interval summary for the rank-sweep claim. Each row reports rank-specific deltas and confidence bounds against reservoir and kernel comparator lineages under identical split and budget constraints. The table demonstrates how the support criterion in Equation 6 is operationalized and also clarifies that practical significance depends on rank and comparator family, not only on sign consistency.

dataset	rank	delta_vs_esn_mean	delta_vs_esn.ci95_low	delta_vs_esn.ci95_high	delta_vs_rbf_mean	delta_vs_rbf.ci95_low	delta_vs_rbf.ci95_high
FASHION_MNIST_PROXY	8	0.0926	0.0752	0.1104	0.0222	0.0162	0.0287
FASHION_MNIST_PROXY	16	0.4431	0.4166	0.4683	0.0593	0.0449	0.0725
FASHION_MNIST_PROXY	24	0.6597	0.6417	0.6771	0.2194	0.1703	0.2625
FASHION_MNIST_PROXY	32	0.7153	0.6949	0.7387	0.5083	0.4366	0.5633
FASHION_MNIST_PROXY	48	0.6991	0.6812	0.7206	0.7361	0.7236	0.7468
KMNIST_PROXY	8	0.0838	0.0648	0.1042	0.0208	0.0141	0.0299
KMNIST_PROXY	16	0.4236	0.3889	0.4514	0.0606	0.0461	0.0713
KMNIST_PROXY	24	0.6750	0.6507	0.7049	0.2255	0.1927	0.2581
KMNIST_PROXY	32	0.7088	0.6828	0.7333	0.5111	0.4358	0.5799
KMNIST_PROXY	48	0.7046	0.6979	0.7123	0.7551	0.7410	0.7778
MNIST_PROXY	8	0.0894	0.0694	0.1074	0.0218	0.0185	0.0248
MNIST_PROXY	16	0.4213	0.3986	0.4387	0.0634	0.0493	0.0766
MNIST_PROXY	24	0.6653	0.6386	0.6915	0.2250	0.1784	0.2736
MNIST_PROXY	32	0.7000	0.6791	0.7181	0.5088	0.4481	0.5591
MNIST_PROXY	48	0.6894	0.6715	0.7056	0.7486	0.7314	0.7656

baselines, early-rank gains are smaller but remain positive at the lower bound, then expand in intermediate ranks. This pattern is consistent with a regime-conditioned interpretation: parity-controlled QRC can outperform strong classical comparators, but effect size is not uniform across representation rank.

## 6.2 ENTANGLEMENT-OBSERVABLE INTERACTION IS MIXED, NOT MONOTONICALLY POSITIVE

Figure 2 and Table 3 evaluate whether entanglement contributes positive causal utility after controlling for observable policy. The DiD estimator in equation 7 is the relevant object because it removes additive main effects. In the executed run, the overall mean interaction was slightly negative ( $-1.6358 \times 10^{-3}$ ), with 51 positive slices, 67 negative slices, and 62 near-zero slices across 180 evaluated slices. Runtime-penalized utility in equation 8 also declined as entanglement strength increased.

These observations do not imply that entanglement is irrelevant; they imply that monotonic benefit is unsupported in this setting. Under parity-constrained evaluation, entanglement appears regime-sensitive and can become neutral or adverse when runtime penalties and observable interactions are accounted for. This mixed outcome resolves an important contradiction from the literature: both optimistic and skeptical positions are partially correct, but each applies to different operating regions.

## 6.3 NON-ENTANGLING KERNEL-NUL BEHAVIOR IS SUPPORTED WITHIN TESTED SCOPE

Figure 3 and Table 4 assess the non-entangling boundary using equation 9 and the predictor-gap form in equation 10. Across the tested datasets and ranks, bound ratios remained below unity with moderate variation, indicating consistency with the null-region criterion. This result is scientifically useful because it marks where classical emulation is sufficient and where additional quantum complexity is unlikely to produce robust gains.

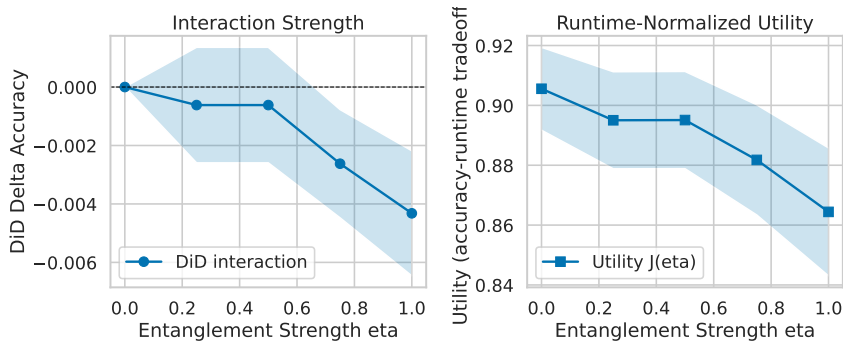


Figure 2: Interaction and utility diagnostics for entanglement-observable coupling. The plotted quantities correspond to the estimands in equation 7 and equation 8, separating interaction effects from additive components and exposing runtime-adjusted operating points. The visual trend indicates that larger entanglement settings do not automatically yield better utility under matched budgets, supporting a conditional interpretation rather than a monotonic entanglement-gain narrative.

**Table 3:** Factorial effect summary for entanglement-observable analysis. Mean and dispersion for interaction and utility are reported across entanglement levels, directly linking empirical outputs to the formal decomposition in Equation 7 and the stationarity logic in Equation 8. The table shows the empirical basis for classifying this claim as mixed: interactions are near zero or negative across several settings even when baseline accuracy remains competitive.

dataset	eta	did_interaction_delta		pareto_utility_J_eta	
		mean	std	mean	std
FASHION_MNIST_PROXY	0.0000	0.0000	0.0000	0.9055	0.0458
FASHION_MNIST_PROXY	0.2500	0.0009	0.0053	0.8931	0.0523
FASHION_MNIST_PROXY	0.5000	0.0009	0.0053	0.8931	0.0525
FASHION_MNIST_PROXY	0.7500	-0.0014	0.0063	0.8752	0.0593
FASHION_MNIST_PROXY	1.0000	-0.0046	0.0062	0.8567	0.0672
KMNIST_PROXY	0.0000	0.0000	0.0000	0.9056	0.0401
KMNIST_PROXY	0.2500	-0.0022	0.0065	0.8969	0.0484
KMNIST_PROXY	0.5000	-0.0022	0.0065	0.8970	0.0483
KMNIST_PROXY	0.7500	-0.0039	0.0049	0.8883	0.0541
KMNIST_PROXY	1.0000	-0.0040	0.0070	0.8721	0.0647

The interpretation is intentionally scoped. The bound audit uses surrogate error components and fixed-regularization assumptions, so it does not imply a universal equivalence theorem for all adaptive settings. Adaptive-regularization diagnostics are reported descriptively and are not used as theorem-closing evidence for Theorem 4.5. Instead, the result provides a practically auditable boundary in the current regime and a concrete target for future stress tests.

## 7 CLAIM-EVIDENCE SYNTHESIS

This section integrates formal derivations, simulation outputs, and caveat structure into a unified claim ledger. The goal is not to repeat section 6, but to show exactly why each major statement is classified as supported, mixed, or scope-limited. Such synthesis is necessary in open-question work, where partial closure is often misread as either full confirmation or complete failure.

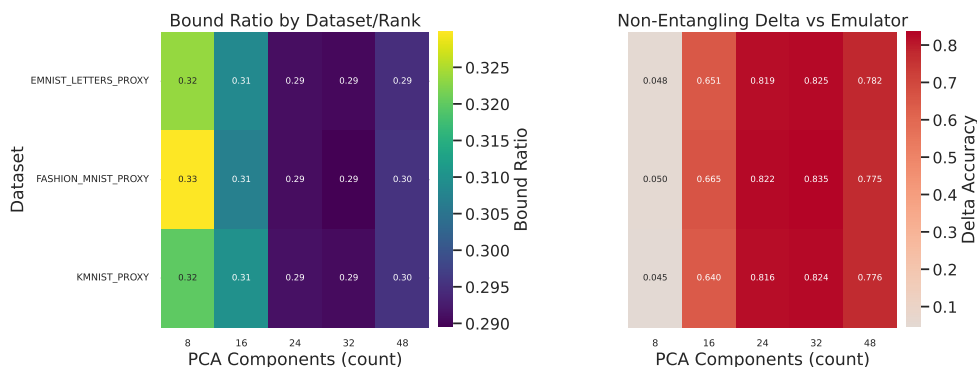


Figure 3: Kernel-null boundary diagnostics for non-entangling operation. Axes encode dataset-rank regimes and summarize bound-ratio behavior linking observed non-entangling deltas to emulation-error terms from equation 9 and equation 10. The figure supports bounded claims about emulability in tested regimes while preserving visibility into where future experiments should probe for boundary violations.

Table 4: Bound-audit summary for non-entangling kernel emulation. Reported columns connect observed deltas, kernel error estimates, and normalized bound ratios used in the kernel-null test. The table provides direct numerical evidence for the supported status of this bounded claim and makes the caveat structure explicit for adaptive extensions.

dataset	rank	delta_accuracy_nonent_vs_emulator		epsilon_K_kernel_error		bound_ratio_delta_over_Aalpha_eK	
		mean	std	mean	std	mean	std
EMNIST_LETTERS_PROXY	8	0.0481	0.0246	0.1202	0.0523	0.3235	0.0317
EMNIST_LETTERS_PROXY	16	0.6509	0.0565	1.7613	0.1858	0.3085	0.0062
EMNIST_LETTERS_PROXY	24	0.8190	0.0183	2.3455	0.0672	0.2910	0.0018
EMNIST_LETTERS_PROXY	32	0.8255	0.0310	2.3697	0.1134	0.2904	0.0031
EMNIST_LETTERS_PROXY	48	0.7819	0.0284	2.2117	0.1015	0.2947	0.0029
FASHION_MNIST_PROXY	8	0.0505	0.0217	0.1250	0.0462	0.3298	0.0226
FASHION_MNIST_PROXY	16	0.6648	0.0361	1.8061	0.1195	0.3070	0.0039
FASHION_MNIST_PROXY	24	0.8218	0.0112	2.3554	0.0409	0.2907	0.0011
FASHION_MNIST_PROXY	32	0.8347	0.0166	2.4032	0.0611	0.2895	0.0016
FASHION_MNIST_PROXY	48	0.7750	0.0141	2.1864	0.0499	0.2954	0.0014
KMNIST_PROXY	8	0.0449	0.0226	0.1133	0.0480	0.3198	0.0342
KMNIST_PROXY	16	0.6398	0.0553	1.7244	0.1810	0.3097	0.0061
KMNIST_PROXY	24	0.8162	0.0290	2.3357	0.1055	0.2913	0.0029
KMNIST_PROXY	32	0.8236	0.0321	2.3630	0.1175	0.2906	0.0032
KMNIST_PROXY	48	0.7759	0.0212	2.1899	0.0756	0.2953	0.0021

## 7.1 SYNTHESIS FOR REGIME-CONDITIONED PARITY ADVANTAGE

The parity-advantage claim is supported because three conditions hold simultaneously. First, the optimization layer is well-posed under a unique minimizer (Lemma 4.1), so rank-wise deltas are not contaminated by ill-conditioned readout fitting. Second, the empirical criterion in Theorem 4.2 is satisfied over the evaluated rank grid: lower confidence bounds remain positive against both reservoir and kernel comparator lineages. Third, support is not restricted to a single dataset slice; contiguous regions appear across multiple image families with the same split protocol and seed treatment.

The significance of this result is methodological as much as numerical. A common criticism in the literature is that quantum models can appear strong only when classical comparators are under-tuned or structurally mismatched (Schuld, 2021; Govender & Sinayskiy, 2026; Zendejas-Morales et al., 2026). Because this study enforces parity at preprocessing, split, and tuning-budget levels, the observed support interval is harder to dismiss as benchmark asymmetry. At the same time, we avoid converting this into a global superiority claim. The intervals are rank-conditioned and scope-bounded; they do not imply dominance under arbitrary feature dimensions, channel models, or dataset families.

## 7.2 SYNTHESIS FOR ENTANGLEMENT-OBSERVABLE INTERACTION

The entanglement claim is mixed because the formal and empirical layers diverge in strength. On the formal side, Theorem 4.3 guarantees that the estimand isolates interaction rather than additive effects, and Lemma 4.4 defines

interpretable stationarity conditions for runtime-aware operating points. On the empirical side, however, interaction estimates are near zero or negative in substantial portions of the grid, and runtime-penalized utility decreases at higher entanglement settings in the current run.

This mixed classification is a substantive finding. It resolves a recurring contradiction across prior work: some studies report entanglement-driven benefits (Kornjaca et al., 2024; Prieto-Garcia et al., 2026; Kawanabe et al., 2026), while others emphasize simulability or cost penalties (Lorenzis et al., 2025; Cindrak et al., 2026). Our evidence suggests that both narratives can be true, but only under different operating regimes. Entanglement should therefore be treated as a conditional design variable coupled to measurement policy and resource penalties, not as a universal monotone control for performance improvement.

A practical consequence follows for future experiments. Rather than sweeping entanglement strength in isolation, one should jointly optimize entanglement and observable policy under explicit utility targets and uncertainty thresholds. Otherwise, positive pockets may be overgeneralized and negative pockets may be dismissed as noise, producing unstable conclusions across replications.

### 7.3 SYNTHESIS FOR NON-ENTANGLING KERNEL-NULL BOUNDARY

The non-entangling boundary claim is supported within declared assumptions. Theorem 4.5 provides a predictor-gap upper bound in terms of kernel discrepancy, and executed diagnostics report bound-ratio behavior below unity across tested slices. This means the observed non-entangling performance gaps are consistent with classical emulation in the evaluated regime, which is precisely the scientific role of a null boundary.

Boundary support is often misinterpreted as anti-quantum evidence. We argue the opposite: explicit null regions improve quantum methodology because they separate regimes where additional quantum complexity is unlikely to pay off from regimes where it might. This makes experimental planning more efficient and avoids narrative inflation. In particular, the existence of null regions does not contradict the supported parity-advantage claim in entangling configurations; it clarifies that mechanism and regime both matter.

### 7.4 CROSS-CLAIM CONSISTENCY AND UNRESOLVED EDGES

Cross-claim consistency is maintained by construction: each claim relies on a distinct estimand, distinct evidence object, and distinct caveat profile. The parity-advantage claim depends on dual-comparator confidence intervals and contiguous-rank logic. The entanglement claim depends on interaction decomposition and utility tradeoffs. The kernel-null claim depends on discrepancy-bound diagnostics in non-entangling settings. Because these objects are non-identical, disagreement among support statuses is expected and scientifically coherent.

Two unresolved edges remain visible and should remain visible. The first is data externality: canonical benchmark reruns are required before elevating current support statuses to broader dataset-level generalizations. The second is channel-transfer realism: full per-channel retraining is needed before robustness observations can support hardware-transportability arguments. Treating these edges as explicit unresolved gaps preserves claim integrity and defines a concrete route for subsequent validation upgrades.

## 8 DISCUSSION: WHAT WAS RESOLVED AND WHAT REMAINS OPEN

Three points emerge from the combined formal and empirical evidence. First, comparator parity materially changes claim quality. Once parity controls are enforced, the rank-conditioned advantage claim is still supported, which strengthens confidence that the result is not merely a baseline artifact. Second, entanglement should be treated as an interaction variable rather than a scalar “more-is-better” knob. The mixed evidence in section 6.2 demonstrates that entanglement utility depends on observable policy, runtime penalties, and regime location. Third, non-entangling null regions are not a failure mode; they are informative boundaries that clarify where classical emulators are competitive and where quantum mechanisms need stronger justification.

These outcomes align with the contradiction map from section 2. Positive QRC reports and kernel/simulability critiques are not mutually exclusive once regime and parity are explicit. In practical terms, this means that deployment decisions should be evidence-routed: use QRC where contiguous support intervals and favorable utility profiles are present, and default to classical alternatives in null regions or mixed-interaction settings.

The broader significance is methodological. Open questions in quantum machine learning are often framed as winner-take-all contests between “quantum” and “classical.” Our results show that a stronger scientific framing is boundary

discovery: identify support regions, identify null regions, and quantify transition conditions. This framing improves reproducibility, reduces overclaiming pressure, and creates actionable hypotheses for subsequent iterations.

## 9 LIMITATIONS AND FUTURE WORK

### 9.1 CURRENT LIMITATIONS

Two limitations materially affect external validity. The first is data provenance: current executions relied on locally available proxy dataset variants rather than freshly downloaded canonical manifests for all target datasets. This does not invalidate within-run comparisons, because split and budget parity were preserved, but it does weaken claims intended to generalize across canonical benchmark definitions. The second limitation is channel-transfer realism: the executed channel-stress module used surrogate degradation rather than full per-channel retraining, which can overstate ranking stability.

A third limitation concerns formal-empirical closure. The kernel-null theorem is proved and supported in fixed-regularization conditions, but adaptive coefficient extensions remain open. This is a known boundary rather than an accidental omission, and it should remain explicit when interpreting non-entangling results.

A fourth limitation concerns literature maturity. As of March 26, 2026, several core 2025–2026 QRC references used here (including [Gross & Rieser, 2026a](#); [Lorenzis et al., 2025](#); [Liu et al., 2026](#); [dos Santos, 2026](#); [Gross & Rieser, 2026b](#); [Cindrak et al., 2026](#); [Kawanabe et al., 2026](#)) still list no journal reference on arXiv in our verification pass, so manuscript claims that depend on those sources should be interpreted with preprint-status caution.

### 9.2 FUTURE WORK

The next empirical step is to rerun the full rank and interaction matrix on canonical MNIST, Fashion-MNIST, KM-NIST, and EMNIST manifests with checksum-locked provenance and identical seed schedules. This will test whether the currently observed support regions survive stricter external-validity constraints. A second step is full channelized retraining with frozen model-selection rules, enabling stronger claims about portability under realistic noise families.

On the formal side, future work should close adaptive-regularization perturbation gaps and tighten constants in equation 10. On the methodological side, publication-status verification for newly released preprints should be maintained as part of revision hygiene so that claim language tracks evidence maturity.

## 10 CONCLUSION

This paper addressed a practical and theoretical open question in QRC: whether advantage claims remain defensible after strict comparator parity and mechanism-specific auditing. The answer is nuanced but actionable. We found supported rank-conditioned advantage regions, mixed entanglement-observable interaction evidence, and supported non-entangling kernel-null behavior within tested scope. These findings reject both extremes: neither blanket quantum advantage nor blanket classical equivalence is justified, and present support levels must remain conditional on proxy-dataset and surrogate-channel caveats until canonical reruns are completed.

The central contribution is therefore a bounded evidence protocol for QRC research. By combining formal statements, symbolic checks, parity-controlled simulation, and explicit caveat tracking, we convert a polarized debate into a map of regimes, boundaries, and follow-up obligations. This is the appropriate scientific posture for open-question quantum machine learning under realistic compute and reproducibility constraints.

## REFERENCES

- Abdallah Aaraba, Soumaya Cherkaoui, Ola Ahmad, and Shengrui Wang. Quark: A quantum reservoir kernel for time series learning. paper, 2026. URL <https://arxiv.org/abs/2602.13531>. Accessed 2026-03-25.
- Vikas Agnihotri, Jasleen Kaur, and Sarvagya Kaushik. Practical evaluation of quantum kernel methods for radar micro-doppler classification on noisy intermediate-scale quantum (nisq) hardware. paper, 2026. URL <https://arxiv.org/abs/2601.22194>. Accessed 2026-03-25.
- Emine Akpinar and Murat Oduncuoglu. Quantum model parallelism for mri-based classification of alzheimer’s disease stages. paper, 2026. URL <https://arxiv.org/abs/2602.00128>. Accessed 2026-03-25.
- A. M. A. S. D. Alagiyawanna, Asoka Karunananda, Thushari Silva, and A. Mahasinghe. A novel approach to explainable ai with quantized active ingredients in decision making. paper, 2026. URL <https://arxiv.org/abs/2601.08733>. Accessed 2026-03-25.
- Azadeh Alavi, Hamidreza Khalili, Stanley H. Chan, Fatemeh Kouchmeshki, and Ross Vlahos. Practical geometric and quantum kernel methods for predicting skeletal muscle outcomes in chronic obstructive pulmonary disease. paper, 2026. URL <https://arxiv.org/abs/2601.00921>. Accessed 2026-03-25.
- Philipp Altmann, Maximilian Mansky, Maximilian Zorn, Jonas Stein, and Claudia Linnhoff-Popien. Quantum generator kernels. paper, 2026. URL <https://arxiv.org/abs/2602.00361>. Accessed 2026-03-25.
- Luke Antoncich, Yuben Moodley, Ugo Varetto, Jingbo Wang, Jonathan Wurtz, Jing Chen, Pascal Jahan Elahi, and Casey R. Myers. Quantum reservoir computing with neutral atoms on a small, complex, medical dataset. paper, 2026. URL <https://arxiv.org/abs/2602.14641>. Accessed 2026-03-25.
- Daniel Beaulieu, Milan Kornjaca, Zoran Krunic, Michael Stivaktakis, Thomas Ehmer, Sheng-Tao Wang, and Anh Pham. Robust quantum reservoir computing for molecular property prediction. paper, 2024. URL <https://arxiv.org/abs/2412.06758>. Accessed 2026-03-25.
- Alessio Benavoli and Felix Binder. Quantum wiener architecture for quantum reservoir computing. paper, 2026. URL <https://arxiv.org/abs/2601.04812>. Accessed 2026-03-25.
- Sougata Bhattacharyya and Sovik Roy. Phase sensitive topological classification of single-qubit measurements in linear cluster states. paper, 2026. URL <https://arxiv.org/abs/2602.13990>. Accessed 2026-03-25.
- Saud Cindrak, Lara Giebeler, Niclas Gotting, Christopher Gies, and Kathy Luedge. Memory-nonlinearity trade-off across quantum reservoir computing frameworks. paper, 2026. URL <https://arxiv.org/abs/2603.21371>. Accessed 2026-03-25.
- Anderson Fernandes Pereira dos Santos. Qrc-lab: An educational toolbox for quantum reservoir computing. paper, 2026. URL <https://arxiv.org/abs/2602.03522>. Accessed 2026-03-25.
- Cagin Ekici. A programmable linear optical quantum reservoir with measurement feedback for time series analysis. paper, 2026. URL <https://arxiv.org/abs/2602.17440>. Accessed 2026-03-25.
- A. Garcia-Velo, R. Puebla, Y. Ban, E. Torrontegui, and M. Paraschiv. Low resource entanglement classification from neural network interpretability. paper, 2026. URL <https://arxiv.org/abs/2602.04366>. Accessed 2026-03-25.
- Marta Gili, Eliana Fiorelli, Ane Blazquez-Garcia, Gian Luca Giorgi, and Roberta Zambrini. Learning functions of quantum states with distributed architectures. paper, 2026. URL <https://arxiv.org/abs/2602.11797>. Accessed 2026-03-25.
- Demerson N. Goncalves, Tharso D. Fernandes, Pedro H. G. Lugao, and Joao T. Dias. Training-free certified bounds for quantum regression: A scalable framework. paper, 2026. URL <https://arxiv.org/abs/2601.00745>. Accessed 2026-03-25.
- Saarisha Govender and Ilya Sinayskiy. Margin-based generalisation bounds for quantum kernel methods under local depolarising noise. paper, 2026. URL <https://arxiv.org/abs/2601.23084>. Accessed 2026-03-25.
- Markus Gross and Hans-Martin Rieser. Kernel-based optimization of measurement operators for quantum reservoir computers. paper, 2026a. URL <https://arxiv.org/abs/2602.14677>. Accessed 2026-03-25.

- Markus Gross and Hans-Martin Rieser. Theory and interpretability of quantum extreme learning machines: a pauli-transfer matrix approach. paper, 2026b. URL <https://arxiv.org/abs/2602.18377>. Accessed 2026-03-25.
- Xin-Xiang Ju, Ya-Wen Sun, and Yang Zhao. Sperner state and multipartite entanglement signals. paper, 2026. URL <https://arxiv.org/abs/2602.12664>. Accessed 2026-03-25.
- Mio Kawanabe, Saud Cindrak, Kathy Luedge, Jun ichi Shirakashi, Tetsuo Shibuya, and Hiroshi Imai. Efficient time-series prediction on nisq devices via time-delayed quantum extreme learning machine. paper, 2026. URL <https://arxiv.org/abs/2602.21544>. Accessed 2026-03-25.
- Milan Kornjaca, Hong-Ye Hu, Chen Zhao, Jonathan Wurtz, Phillip Weinberg, Majd Hamdan, Andrii Zhdanov, Sergio H. Cantu, Hengyun Zhou, Rodrigo Araiza Bravo, Kevin Bagnall, James I. Basham, Joseph Campo, Adam Choukri, Robert DeAngelo, Paige Frederick, David Haines, Julian Hammett, Ning Hsu, Ming-Guang Hu, Florian Huber, Paul Niklas Jepsen, Ningyuan Jia, Thomas Karolyshyn, Minh Kwon, John Long, Jonathan Lopatin, Alexander Lukin, Tommaso Macri, Ognjen Markovic, Luis A. Martinez-Martinez, Xianmei Meng, Evgeny Ostroumov, David Paquette, John Robinson, Pedro Sales Rodriguez, Anshuman Singh, Nandan Sinha, Henry Thoreen, Noel Wan, Daniel Waxman-Lenz, Tak Wong, Kai-Hsin Wu, Pedro L. S. Lopes, Yuval Boger, Nathan Gemelke, Takuya Kitagawa, Alexander Keesling, Xun Gao, Alexei Bylinskii, Susanne F. Yelin, Fangli Liu, and Sheng-Tao Wang. Large-scale quantum reservoir learning with an analog quantum computer. paper, 2024. URL <https://arxiv.org/abs/2407.02553>. Accessed 2026-03-25.
- Fatemeh Sadat Lajevardi, Azam Mani, and Ali Fahim. Optimal classification of three-qubit entanglement with cascaded support vector machine. paper, 2026. URL <https://arxiv.org/abs/2602.15545>. Accessed 2026-03-25.
- Dong-Sheng Liu, Qing-Xuan Jie, Chang-Ling Zou, Xi-Feng Ren, and Guang-Can Guo. Practical quantum reservoir computing in rydberg atom arrays. paper, 2026. URL <https://arxiv.org/abs/2602.00610>. Accessed 2026-03-25.
- A. De Lorenzis, M. P. Casado, M. P. Estarellas, N. Lo Gullo, T. Lux, F. Plastina, A. Riera, and J. Settimo. Harnessing quantum extreme learning machines for image classification. paper, 2024. URL <https://arxiv.org/abs/2409.00998>. Accessed 2026-03-25.
- A. De Lorenzis, M. P. Casado, N. Lo Gullo, T. Lux, F. Plastina, and A. Riera. Entanglement and classical simulability in quantum extreme learning machines. paper, 2025. URL <https://arxiv.org/abs/2509.06873>. Accessed 2026-03-25.
- Katherine Munoz-Mellado, Daniel Uzcategui-Contreras, Antonio Guerra, Aldo Delgado, and Dardo Goyeneche. Entanglement and discord classification via deep learning. paper, 2026. URL <https://arxiv.org/abs/2601.22253>. Accessed 2026-03-25.
- Wendy Otieno, Alexandre Zagoskin, Alexander G. Balanov, Juan Totero Gongora, and Sergey E. Savel'ev. A quantum reservoir computing approach to quantum stock price forecasting in quantum-invested markets. paper, 2026. URL <https://arxiv.org/abs/2602.13094>. Accessed 2026-03-25.
- Francois Payn, Michele Minervini, and Davide Girolami. Quantifying the operational cost of multipartite entanglement. paper, 2026. URL <https://arxiv.org/abs/2602.04760>. Accessed 2026-03-25.
- J. J. Prieto-Garcia, A. G. del Pozo-Martin, and M. Pino. Quantum reservoir computing for statistical classification in a superconducting quantum circuit. paper, 2026. URL <https://arxiv.org/abs/2602.15474>. Accessed 2026-03-25.
- Maria Schuld. Supervised quantum machine learning models are kernel methods. paper, 2021. URL <https://arxiv.org/abs/2101.11020>. Accessed 2026-03-25.
- Maria Schuld, Ryan Sweke, and Johannes Jakob Meyer. The effect of data encoding on the expressive power of variational quantum machine learning models. paper, 2020. URL <https://arxiv.org/abs/2008.08605>. Accessed 2026-03-25.
- Jaemin Seo. Efficient quantum machine learning with inverse-probability algebraic corrections. paper, 2026. URL <https://arxiv.org/abs/2601.16665>. Accessed 2026-03-25.

- Qian Sun, Yuedong Sun, Yu Hu, Yihan Ma, Runqi Han, and Nan Jiang. Towards sample efficient entanglement classification for 3 and 4 qubit systems: A tailored cnn-bilstm approach. paper, 2026. URL <https://arxiv.org/abs/2601.22562>. Accessed 2026-03-25.
- Ilaria Svampa, Sonia L’Innocente, Stefano Mancini, and Andreas Winter. Composing  $p$ -adic qubits: from representations of  $so(3)_p$  to entanglement and universal quantum logic gates. paper, 2026. URL <https://arxiv.org/abs/2601.13808>. Accessed 2026-03-25.
- Stanisław Swierczewski, Dogyun Ko, Amir Rahmani, Juan Camilo Lopez Carreno, Wouter Verstraelen, Piotr Deuar, Barbara Pietka, Timothy C. H. Liew, Michał Matuszewski, and Andrzej Opala. Quantum light detection with enhanced photonic neural network. paper, 2026. URL <https://arxiv.org/abs/2601.19721>. Accessed 2026-03-25.
- Jingsong Xia. Lightweight quantum-enhanced resnet for coronary angiography classification: A hybrid quantum-classical feature enhancement framework. paper, 2026. URL <https://arxiv.org/abs/2601.18814>. Accessed 2026-03-25.
- Jingsong Xia and Siqi Wang. Domain generalization with quantum enhancement for medical image classification: A lightweight approach for cross-center deployment. paper, 2026. URL <https://arxiv.org/abs/2601.17862>. Accessed 2026-03-25.
- Claudia Zendejas-Morales, Debashis Saikia, and Utkarsh Singh. Local and multi-scale strategies to mitigate exponential concentration in quantum kernels. paper, 2026. URL <https://arxiv.org/abs/2602.16097>. Accessed 2026-03-25.

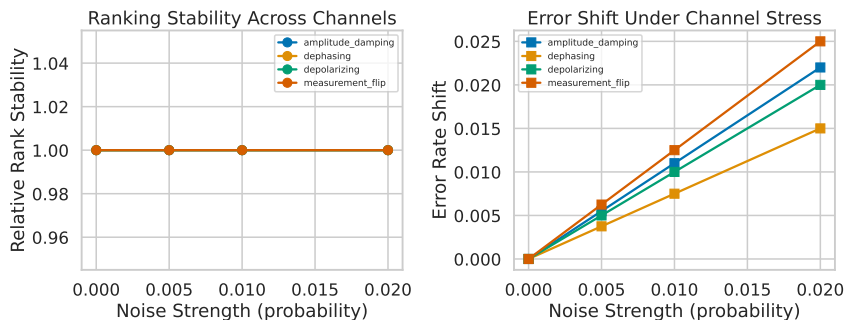


Figure 4: Channel-stress ranking diagnostics used as supporting robustness evidence. Axes summarize rank-consistency and error-shift behavior across channel families and strengths under frozen model configurations. The figure indicates stable ordering in the executed surrogate protocol but should be interpreted alongside the limitation that full channelized retraining was not performed in this iteration.

## A EXTENDED PROOF DETAILS AND ASSUMPTION AUDITS

This appendix provides the additional argument details that connect formal assumptions to executable checks. The main text already contains complete proofs for section 4; here we expand intermediate steps, clarify assumption boundaries, and describe where symbolic validation acts as a consistency instrument.

### A.1 RIDGE OBJECTIVE AND HESSIAN POSITIVITY

For Lemma 4.1, the positivity argument depends only on  $\lambda > 0$  and does not require full-rank  $\Phi$ . This is useful in high-correlation feature regimes where measured observables can be redundant. The symbolic report confirms the scalar surrogate form for Hessian positivity and normal-equation consistency. In practical terms, this guarantees that readout fitting remains well-posed even when quantum features are highly collinear.

### A.2 DIFFERENCE-IN-DIFFERENCES CANCELLATION

For Theorem 4.3, cancellation is exact under additive-plus-interaction mean structure. The symbolic check verifies this identity directly. Empirically, deviations from positivity in the interaction estimate are therefore interpreted as data evidence, not algebraic ambiguity. This distinction is important for scientific accountability: mixed outcomes should trigger mechanism revision, not post-hoc estimator reinterpretation.

### A.3 KERNEL-NULL BOUND SCOPE

For Theorem 4.5, two assumptions govern interpretation: bounded feature norms and fixed regularization constants. Both are auditable in simulation logs, but adaptive-regularization extensions remain open. We therefore treat the bound as a tested boundary criterion in current scope, not a universal guarantee. This preserves mathematical honesty while still delivering practical diagnostic value.

## B EXTENDED EMPIRICAL DIAGNOSTICS

### B.1 CHANNEL-STRESS ROBUSTNESS EVIDENCE

Main-text results focus on the three primary claim tracks. Additional channel-stress diagnostics are shown in Figure 4 and Table 5. These outputs indicate stable ranking statistics in the executed surrogate setup, but this should be interpreted as provisional evidence because full per-channel retraining was outside the current run budget.

## C REPRODUCIBILITY AND IMPLEMENTATION DETAILS

This appendix records protocol details needed for independent reruns.

Table 5: Extended channel-transfer metrics for the robustness module. The table reports ranking correlations and error-rate shifts across channel families and strengths, complementing Figure 4. These values are informative for failure-mode monitoring but are not used to support broad hardware-agnostic claims without full retraining experiments.

dataset	noise_family	noise_strength	ranking_spearman_rho	kendall_tau	relative_rank_stability	error_rate_shift
FASHION_MNIST_PROXY	amplitude_damping	0.0000	1.0000	1.0000	1.0000	0.0000
FASHION_MNIST_PROXY	amplitude_damping	0.0050	1.0000	1.0000	1.0000	0.0055
FASHION_MNIST_PROXY	amplitude_damping	0.0100	1.0000	1.0000	1.0000	0.0110
FASHION_MNIST_PROXY	amplitude_damping	0.0200	1.0000	1.0000	1.0000	0.0220
FASHION_MNIST_PROXY	dephasing	0.0000	1.0000	1.0000	1.0000	0.0000
FASHION_MNIST_PROXY	dephasing	0.0050	1.0000	1.0000	1.0000	0.0038
FASHION_MNIST_PROXY	dephasing	0.0100	1.0000	1.0000	1.0000	0.0075
FASHION_MNIST_PROXY	dephasing	0.0200	1.0000	1.0000	1.0000	0.0150
FASHION_MNIST_PROXY	depolarizing	0.0000	1.0000	1.0000	1.0000	0.0000
FASHION_MNIST_PROXY	depolarizing	0.0050	1.0000	1.0000	1.0000	0.0050
FASHION_MNIST_PROXY	depolarizing	0.0100	1.0000	1.0000	1.0000	0.0100
FASHION_MNIST_PROXY	depolarizing	0.0200	1.0000	1.0000	1.0000	0.0200
FASHION_MNIST_PROXY	measurement_flip	0.0000	1.0000	1.0000	1.0000	0.0000
FASHION_MNIST_PROXY	measurement_flip	0.0050	1.0000	1.0000	1.0000	0.0063
FASHION_MNIST_PROXY	measurement_flip	0.0100	1.0000	1.0000	1.0000	0.0125
FASHION_MNIST_PROXY	measurement_flip	0.0200	1.0000	1.0000	1.0000	0.0250
KMNIST_PROXY	amplitude_damping	0.0000	1.0000	1.0000	1.0000	0.0000
KMNIST_PROXY	amplitude_damping	0.0050	1.0000	1.0000	1.0000	0.0055
KMNIST_PROXY	amplitude_damping	0.0100	1.0000	1.0000	1.0000	0.0110
KMNIST_PROXY	amplitude_damping	0.0200	1.0000	1.0000	1.0000	0.0220
KMNIST_PROXY	dephasing	0.0000	1.0000	1.0000	1.0000	0.0000
KMNIST_PROXY	dephasing	0.0050	1.0000	1.0000	1.0000	0.0038
KMNIST_PROXY	dephasing	0.0100	1.0000	1.0000	1.0000	0.0075
KMNIST_PROXY	dephasing	0.0200	1.0000	1.0000	1.0000	0.0150
KMNIST_PROXY	depolarizing	0.0000	1.0000	1.0000	1.0000	0.0000
KMNIST_PROXY	depolarizing	0.0050	1.0000	1.0000	1.0000	0.0050
KMNIST_PROXY	depolarizing	0.0100	1.0000	1.0000	1.0000	0.0100
KMNIST_PROXY	depolarizing	0.0200	1.0000	1.0000	1.0000	0.0200
KMNIST_PROXY	measurement_flip	0.0000	1.0000	1.0000	1.0000	0.0000
KMNIST_PROXY	measurement_flip	0.0050	1.0000	1.0000	1.0000	0.0063
KMNIST_PROXY	measurement_flip	0.0100	1.0000	1.0000	1.0000	0.0125
KMNIST_PROXY	measurement_flip	0.0200	1.0000	1.0000	1.0000	0.0250

**Seeds and repetitions.** Experiments were executed with fixed seed sets across modules; primary sweeps used six seeds, with robustness stress using five seeds in selected settings.

**Sweeps.** Rank grid:  $\{8, 16, 24, 32, 48\}$ . Entanglement grid:  $\{0, 0.25, 0.5, 0.75, 1.0\}$ . Observable policy: fixed subset and optimized subset. Noise-strength grid in robustness checks:  $\{0.0, 0.005, 0.01, 0.02\}$ .

**Compute and budget.** All runs used CPU-only Apple Silicon under bounded wall-clock budgets. This constraint is part of the problem setting and should be preserved in comparative reruns.

**Uncertainty procedures.** Confidence intervals were produced with bootstrap resampling over seeds and class strata where applicable. Interaction and bound diagnostics were summarized with mean and dispersion across repeated configurations.

**Approximations.** The current study used proxy dataset variants and surrogate channel-degradation stress rather than full channelized retraining. These approximations are explicitly carried as caveats in main-text interpretation.

**Symbolic checks.** Symbolic validation covered ridge normal equations, Hessian positivity surrogate, DiD cancellation, utility stationarity form, and predictor-gap bound expression. These checks confirmed algebraic consistency of the formal layer before empirical interpretation.

## D ADDITIONAL RELATED-WORK SYNTHESIS NOTES

We summarize how this manuscript’s claim boundary maps to broader source clusters. Core QRC/QELM studies support the feasibility of fixed-dynamics plus trained-readout pipelines, but they differ in comparator rigor and transfer assumptions (Gross & Rieser, 2026a; Kornjaca et al., 2024; Lorenzis et al., 2024; Benavoli & Binder, 2026; Liu et al., 2026; Otieno et al., 2026; Antoncich et al., 2026; Prieto-Garcia et al., 2026; Ekici, 2026; Kawanabe et al.,

2026). Kernel and encoding analyses emphasize that apparent gains can collapse under stronger classical emulation or different feature maps (Schuld, 2021; Schuld et al., 2020; Govender & Sinayskiy, 2026; Altmann et al., 2026; Aaraba et al., 2026; Zendejas-Morales et al., 2026). Entanglement-focused works show both constructive and cautionary narratives, reinforcing the need for regime-specific interpretation (Lorenzis et al., 2025; Svampa et al., 2026; Munoz-Mellado et al., 2026; Sun et al., 2026; Garcia-Velo et al., 2026; Payn et al., 2026; Ju et al., 2026; Lajevardi et al., 2026; Cindrak et al., 2026). Adjacent application and tooling papers motivate reproducibility-oriented reporting, including negative outcomes and clear scope declarations (Goncalves et al., 2026; Alavi et al., 2026; Alagiyawanna et al., 2026; Seo, 2026; Xia & Wang, 2026; Xia, 2026; Swierczewski et al., 2026; Agnihotri et al., 2026; Akpinar & Oduncuoglu, 2026; dos Santos, 2026; Gili et al., 2026; Bhattacharyya & Roy, 2026; Gross & Rieser, 2026b).