MatchPoint Research Working Paper No. 2026/T23

Quantum Machine Learning for Alpha Generation: Hype versus Reality

An evidence-gated research and investment-decision framework for institutional allocators and family-office investment teams

Prepared by Matchpoint Partners
Author attribution pending CK approval
August 2026

ABSTRACT

Background. Quantum machine learning is frequently presented as a route to superior financial prediction, while investment alpha remains a low-signal, non-stationary and economically adversarial target. **Objective.** This paper tests what the current evidence can support for A1 international institutional allocators and A2 family-office CIOs and heads of alternatives. **Approach.** The analysis reviews 40 primary, authoritative and clearly labelled research sources available through 1 August 2026 and separates theorem, simulation, hardware experiment, historical backtest and production evidence. **Findings.** Published QML studies show task-specific promise, yet the reviewed evidence does not establish repeatable, net-of-cost production alpha or broad quantum advantage on investable classical market data. Data encoding, sampling, noise, trainability, baseline choice, leakage, multiple testing and regime change remain binding proof obligations. **Implications.** Institutions should run a classical-first benchmark tournament with frozen protocols, complete resource accounting, independent validation and zero investment authority until pre-approved gates pass. All worked inputs are unverified illustrative management assumptions; attributed Matchpoint or client revenue, cash cost reduction and loss reduction remain USD 0 until approved observed evidence exists.

Highlights

Proof burden. Alpha requires repeatable, leakage-free, net-of-cost out-of-sample performance and accountable attribution.

Evidence boundary. Simulation or a small hardware demonstration does not establish production quantum advantage.

Benchmark. The quantum candidate faces strong tuned classical models, simple economic baselines and quantum-inspired surrogates.

Architecture. Data, features, quantum execution, portfolio construction, costs, controls and approval remain separately auditable.

Decision. Capital authority stays with the authorised investment committee; the research lane has no trading authority.

JEL Classification: C45, C55, C58, G11, G12, G17, O33

Keywords: quantum machine learning, alpha generation, asset management, quantum kernels, variational quantum circuits, return prediction, classical benchmark, backtest overfitting, data leakage, model risk, institutional allocator, family office

Research paper only. This document is not investment, legal, regulatory, accounting, audit, tax, privacy, cybersecurity, employment, valuation or technology advice. All worked inputs, thresholds, times, rates and economics are unverified illustrative management assumptions.

1. INTRODUCTION

Quantum machine learning sits at the intersection of two domains that attract unusually strong expectations: quantum computing and investment prediction. The combination creates a demanding governance problem. A useful result must survive the technical burden of quantum execution, the statistical burden of financial forecasting, and the economic burden of implementation. It must also remain reproducible after the model specification, data period, transaction costs and reporting choices are frozen.

This paper addresses **Quantum Machine Learning for Alpha Generation: Hype versus Reality** for two Matchpoint Partners target audiences. **A1 International Institutional Allocators** are pensions, insurers, endowments and fund-of-funds evaluating or scaling UAE and GCC private-market exposure. **A2 Family-Office CIOs and Heads of Alternatives** are investment professionals at single-family offices, multi-family offices, private-wealth firms and external asset managers. Both groups may encounter managers, research teams or technology providers claiming that quantum models improve investment selection, timing, risk-adjusted returns or research productivity.

The governing question is precise: **what evidence would justify treating a QML output as investment-relevant alpha rather than an interesting research result?** The answer cannot be derived from qubit count, circuit novelty, model accuracy or one favourable backtest. Alpha is an economic residual after the investment mandate, benchmark, risk exposures, transaction costs, financing, capacity, taxes where relevant, data availability and implementation delay are accounted for. In a competitive market, the residual must also persist outside the research sample and across plausible regimes.

Quantum machine learning is a legitimate research field. Foundational work describes quantum models, feature maps, parameterised circuits and potential learning advantages [1-5]. Theoretical and empirical work also identifies material limitations: barren plateaus, noise-induced trainability problems, sampling cost, kernel concentration, classical learnability and dequantisation [6-15]. Finance-focused research reports promising results in selected forecasting and credit tasks, alongside studies where strong classical models remain superior [17-25]. These bodies of evidence support disciplined experimentation. They do not, in the sources reviewed here, establish broad production quantum advantage for investable alpha.

The paper therefore adopts an evidence ladder. A theorem establishes a result under stated assumptions. A simulator experiment establishes behaviour in an ideal or modelled environment. A hardware experiment adds device noise, compilation, queueing and finite sampling. A historical backtest adds data and market structure. A paper portfolio adds implementability, costs and risk. A production result adds controlled decision rights, real execution and observed attribution. Evidence does not move automatically between these levels.

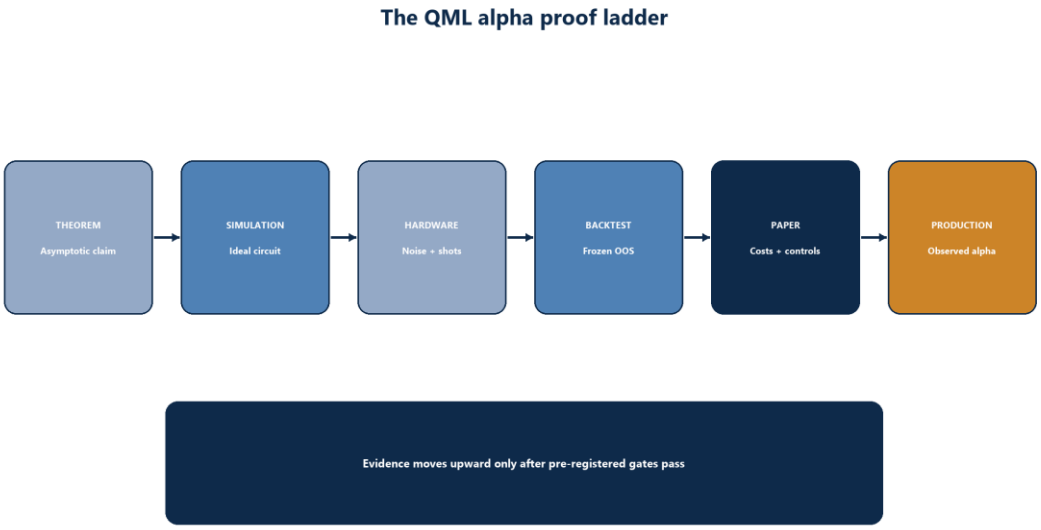

Figure 1. The QML alpha proof ladder

Source: Matchpoint evidence framework; research and regulatory context [1-25,31-40].

The contribution is a practical operating framework. It provides:

- a definition of alpha and quantum advantage that prevents category errors;
- an evidence map for quantum kernels, variational circuits and hybrid models;

- a classical-first benchmark tournament;
- a point-in-time data and leakage-control protocol;
- a complete resource and investment-cost ledger;
- an A1 and A2 decision-rights model;
- a hard-gated evaluation scorecard;
- unverified illustrative operating cases;
- a ninety-day research roadmap; and
- an evidence and claims register suitable for investment committees and manager due diligence.

The paper reviews 40 primary, authoritative and clearly labelled sources available through 1 August 2026. Preprints are identified as preprints. The analysis does not treat a claim of superiority as verified when the source does not provide production evidence, an investable comparison or sufficient information for independent replication.

2. SCOPE, DEFINITIONS AND EVIDENCE BOUNDARIES

2.1 What quantum machine learning means here

Quantum machine learning is used here for a model or pipeline in which a quantum computation contributes directly to a learning function such as feature construction, kernel estimation, parameterised prediction, sequence processing or probabilistic sampling [1-5]. A classical model inspired by quantum concepts is classified as **quantum-inspired**, not quantum-executed. A classical optimisation surrounding a quantum circuit is classified as a **hybrid quantum-classical** workflow.

Three families are most relevant to an alpha research programme:

1. **Quantum kernel methods.** Classical observations are encoded into quantum states or feature maps. Quantum circuits estimate similarities used by a classical kernel learner [2,3,11-13].
2. **Variational quantum circuits or quantum neural networks.** Parameterised circuits transform encoded inputs; a classical optimiser updates parameters using measured outputs [4-10,39,40].
3. **Quantum or hybrid temporal models.** A quantum circuit is embedded inside a recurrent, convolutional or multi-task architecture for time-series or cross-sectional prediction [19-22].

Quantum optimisation, Monte Carlo acceleration and quantum risk analysis are adjacent domains. They can affect portfolio construction or risk calculation without generating a predictive signal. This distinction matters because an optimisation speed-up is not evidence of superior expected-return forecasts, and a risk-estimation result is not alpha [17,18,37].

2.2 Alpha

For this paper, alpha is the residual return attributable to a defined investment process after the approved benchmark and risk model, costs and implementation are applied. At minimum, an alpha claim identifies:

- the target return and forecast horizon;
- the investable universe available at each decision time;
- the benchmark and risk adjustment;
- the signal formation time and execution time;
- turnover, spread, market impact, financing and borrow assumptions;
- portfolio constraints and capacity;
- gross and net performance;
- the number of models, features, hyperparameters and variants tested; and
- the out-of-sample and live or shadow period.

A classification AUC, forecast RMSE or directional accuracy can be useful intermediate evidence. It is not alpha. A statistically significant forecast can also fail economically after turnover, constraints and costs. The empirical asset-pricing literature demonstrates both the potential of machine learning and the unusually low signal-to-noise ratio, overfitting exposure and multiplicity problem in return prediction [26-30].

2.3 Quantum advantage

This paper reserves **quantum advantage** for a quantum-enabled method that improves a decision-relevant outcome over the best feasible classical alternative under a fair resource comparison. The comparison includes data preparation,

encoding, training, hyperparameter search, queueing, circuit execution, shots, mitigation, decoding, portfolio construction and validation. A smaller parameter count or circuit step count does not by itself establish lower elapsed time, energy, cost or error.

Four narrower labels are useful:

Label	Evidence established	Evidence not established
Representational result	A quantum model can express a stated function class	Trainability, data relevance or economic value
Prediction result	A model improves a stated metric on a stated dataset	Investable alpha or production advantage
Computational result	A complexity or resource result holds under assumptions	End-to-end wall-clock advantage on market data
Economic result	Net decision outcome improves under controlled implementation	Persistence outside the observed environment

The Power of Data study shows why computational hardness of a quantum process does not guarantee a prediction advantage when classical learners can use labelled data [11]. Quantum-kernel work shows that noise, finite measurements and kernel concentration can erase or trivialise an advantage [12,13]. Work on shadows and classical simulability shows that deployment or trainability choices can expose efficient classical surrogates [14,15]. The benchmark must therefore include classical approximations of the quantum model, not only familiar off-the-shelf classifiers.

2.4 Evidence classes

Each claim in this paper is classified by the strongest evidence it receives:

Class	Definition	Permitted wording
T	Theoretical result under explicit assumptions	proves under assumptions; derives; bounds
S	Numerical simulator result	simulated; numerical experiment
H	Quantum-hardware experiment	executed on named hardware; finite task
B	Historical market backtest	historical out-of-sample result; costs as stated
P	Paper or shadow portfolio	observed decisions without production authority
R	Production result with approved attribution	realised and independently attributed

The reviewed source set contains T, S, H and B evidence. It contains research claims that may inform paper or shadow portfolios. It does not provide approved Matchpoint or client R evidence. Attributed Matchpoint or client revenue, cash cost reduction and loss reduction therefore remain USD 0.

2.5 Current-state conclusion

As of 1 August 2026, the evidence reviewed supports a bounded conclusion: QML is credible as a research programme with task-specific theoretical and experimental promise. The evidence does not establish repeatable, net-of-cost production alpha or broad quantum advantage on investable classical market data. This conclusion is evidence-bounded and can change when stronger reproducible evidence appears.

3. THE A1 AND A2 DECISION PROBLEM

3.1 A1 international institutional allocators

An A1 allocator is unlikely to build a trading quantum stack solely because a manager uses QML. Its immediate decision is due diligence: whether a manager's claimed edge is understood, evidenced, governed and investable. The allocator requires access to enough information to distinguish a research narrative from a repeatable investment process.

A1 diligence questions include:

- Which component uses a quantum processor?
- What information was available when each decision was formed?
- Which classical baselines were tuned to a comparable standard?
- How many model variants were tried before the reported result?
- Did the result survive costs, capacity and portfolio constraints?
- Was the quantum component necessary, or can a surrogate reproduce it?

- Who validated the result independently of the research author?
- What happens when hardware access, calibration or vendor terms change?
- How is performance described in investor communications?

The A1 control objective is not to reproduce every circuit. It is to obtain a complete evidence chain from data and model through the portfolio and claim. Marketing statements require particular care. The SEC marketing rule addresses substantiation, fair and balanced treatment of risks and hypothetical performance [31]. The SEC's 2024 AI-washing cases show that claims about an investment adviser's use of advanced technology can become enforcement matters when representations are false or misleading [32]. A quantum claim should receive the same substantiation discipline.

3.2 A2 family-office CIOs and heads of alternatives

An A2 team can encounter QML in manager selection, direct public-market research, venture investment and technology procurement. The smaller team may have concentrated decision rights and limited specialist capacity. The operating model should separate scientific evaluation from capital authority.

A2 questions include:

- Is the proposed use case forecast generation, risk estimation or optimisation?
- Does the team have a decision-relevant bottleneck that the quantum component addresses?
- Can the result be challenged by an independent classical researcher?
- Are research data and cloud execution permitted under the family's confidentiality and residency requirements?
- How does the family preserve the evidence if a vendor, device or software version changes?
- Which committee can authorise a shadow portfolio, and which can authorise capital?

Family-office flexibility can accelerate experiments. The same flexibility can compress independent challenge. The paper therefore requires explicit ownership for data, research, validation, technology, risk and investment approval.

3.3 Shared decision perimeter

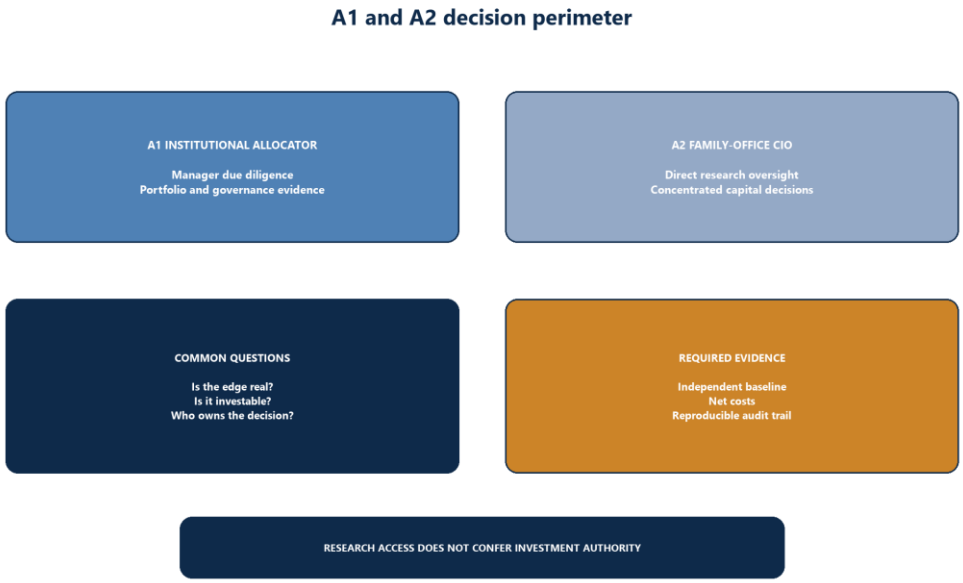

Figure 2. A1 and A2 decision perimeter

Source: Matchpoint ICP and decision-rights framework.

Decision	Research team	Independent validation	Risk / compliance	Investment committee
Define hypothesis	Proposes	Challenges	Reviews boundary	Approves material objective
Build data and models	Executes	Reproduces	Reviews data and vendor use	Receives evidence
Select reported result	Documents all trials	Tests selection process	Reviews claims	Cannot receive cherry-picked result
Run paper portfolio	Operates	Monitors	Sets controls	Approves scope

Decision	Research team	Independent validation	Risk / compliance	Investment committee
Allocate capital	No authority	Advises	Advises	Sole approval
Publish performance claim	Supplies evidence	Verifies	Approves wording	Approves where required

Research access does not confer trading authority. A favourable model result does not change that boundary.

4. HOW QML COULD CREATE VALUE

4.1 Representation

Quantum feature maps can place classical data into a high-dimensional Hilbert space and produce kernels that are difficult to estimate classically [2,3]. Parameterised circuits can express function classes shaped by the data encoding and circuit structure [4,5,39,40]. A useful investment hypothesis is that this representation may capture interactions that a constrained classical model misses.

The hypothesis requires an economic reason. Market data do not become suitable merely because they are high-dimensional. A valid research question identifies a mechanism: cross-asset interactions, conditional regime structure, non-linear feature combinations or a distributional pattern that relates to expected return. It then tests whether the QML representation improves the out-of-sample economic decision.

4.2 Sample efficiency

Some QML claims focus on performance with limited data. This may appear attractive in finance, where clean labelled regimes are scarce. Credit-scoring studies report results on small or imbalanced samples [23-25]. The interpretation must remain narrow. A small dataset can make flexible classical models unstable, while it can also make any comparison highly uncertain. Sample-efficiency claims require repeated splits, uncertainty intervals, stability tests and a baseline search budget that does not favour the proposed model.

4.3 Hybrid temporal processing

Finance studies have embedded circuits within recurrent, convolutional or multi-task architectures. The contextual QNN study reports improved multi-asset prediction against quantum single-task comparators [19]. A 2026 preprint benchmarks quantum and classical models across direction, simulated trading and volatility tasks, reporting gains in some assets and regimes while classical approaches lead elsewhere [20]. A DeFi benchmark reports classical ensembles outperforming the tested quantum models [21]. A cross-sectional QTCNN preprint reports a favourable out-of-sample long-short Sharpe comparison on the JPX dataset [22].

These results motivate replication. They do not form one consistent production finding. The studies differ in datasets, labels, architectures, comparators, cost treatment, simulation and hardware use. Each result stays within its source boundary.

4.4 Research productivity

Even without investment alpha, a controlled QML programme can create research options:

- a reusable point-in-time dataset;
- a stronger classical benchmark library;
- explicit experiment tracking;
- improved leakage and multiple-testing controls;
- vendor and hardware evaluation capability;
- staff understanding of quantum algorithms; and
- a documented basis for rejecting weak technology claims.

These are operating outputs. Their economic value should be measured through approved evidence such as accepted research throughput, elapsed time, avoided duplicate experiments and documented due-diligence quality. Productivity does not equal alpha.

5. WHY ALPHA IS AN UNUSUALLY HARD TARGET

5.1 Low signal and adaptive markets

Expected returns are difficult to measure because realised returns contain substantial unforecastable news. Gu, Kelly and Xiu show that classical machine-learning methods can improve risk-premium prediction in a large US equity study, while

also emphasising overfit controls and the low signal-to-noise environment [26]. This creates a high classical bar for QML. The relevant question is whether the quantum candidate adds stable economic information beyond well-tuned linear, tree, kernel and neural baselines.

Markets also adapt. McLean and Pontiff find that returns associated with published predictors decline out of sample and further after publication [28]. A model can therefore reproduce a historical relation that weakens after discovery or deployment. A useful quantum representation does not remove this economic decay.

5.2 Multiple testing

A QML research programme contains many degrees of freedom: universe, label, horizon, scaling, dimensionality reduction, encoding, circuit depth, entanglement, observable, optimiser, learning rate, shots, mitigation, seed, portfolio rule and cost model. If only the best result is reported, performance can be a selection artefact.

Harvey, Liu and Zhu show that the factor-discovery environment requires a higher statistical hurdle than a conventional isolated test [27]. White's Reality Check addresses data snooping across model alternatives [29]. The Deflated Sharpe Ratio corrects for selection bias, non-normal returns and the number of trials [30]. These methods do not certify a strategy automatically; they establish the type of multiplicity record that QML research must preserve.

5.3 Leakage

Financial leakage is any use of information that would not have been available to the real decision at the relevant time. Common paths include:

- standardising on the full sample;
- selecting features using test-period outcomes;
- using revised fundamentals before their release;
- including later index constituents in earlier universes;
- using same-period close information for an assumed close execution;
- tuning hyperparameters on the final test set;
- applying mitigation or seed selection after observing test results; and
- choosing a reported regime because it is favourable.

Quantum-specific preprocessing creates additional risk. Dimensionality reduction may use the full sample before a low-dimensional vector is encoded. A classical feature extractor may contain most of the predictive work while the final circuit receives credit. The data clock must extend through every classical and quantum step.

5.4 Costs and capacity

A small improvement in prediction can disappear in the portfolio. The test must include turnover, spread, market impact, financing, stock borrow, latency, stale quotes and capacity. If a QML model changes rankings only at the margin, the economic result depends on how those changes interact with position sizing and costs.

The resource ledger also includes research cost. Quantum execution can require repeated shots, calibration-aware compilation, error mitigation and cloud access. A fair comparison reports wall-clock time and monetary cost from raw input to accepted portfolio, not only circuit depth or the number of trainable parameters.

6. TECHNICAL REALITY: DATA, TRAINABILITY, NOISE AND CLASSICAL SURROGATES

6.1 Data encoding

Classical market data must be converted into operations on qubits. Angle encoding can be relatively direct but often requires one or more circuit elements per feature. Amplitude encoding can represent a larger vector compactly in a state description, while state preparation can impose material cost. Re-uploading data can increase expressivity by repeating encodings through the circuit [5,39].

The encoding is part of the model. It determines accessible frequencies and interactions [39]. A comparison that gives the QML model a heavily engineered input while giving the classical baseline raw inputs is not fair. Both lanes receive the same approved information, and the resource ledger attributes classical feature engineering explicitly.

6.2 Trainability

Random or overly expressive parameterised circuits can exhibit gradients that vanish exponentially with system size, known as barren plateaus [6]. Cost-function design affects this behaviour [7]. Noise can create barren plateaus as circuit depth grows [8]. Gradient-free optimisers do not remove the resource problem when cost differences become

exponentially suppressed [9]. Shallow models can avoid a simple barren plateau and still face landscapes dominated by poor local minima or statistical-query limits [10].

These results do not show that every QML model is untrainable. They require the research team to report gradient diagnostics, optimisation stability, initialisation, seeds, shot budgets, circuit depth and scaling. A small successful circuit is evidence for that circuit and task. Scaling is a separate claim.

6.3 Noise and finite sampling

Quantum outputs are estimated from measurements. Finite shots create sampling error. Device noise, drift, gate errors, readout errors and compilation choices add variation. Kernel estimation can require many circuit evaluations as the number of observations grows. Wang et al. show that noise, dataset size and measurement count can affect the utility of quantum kernels [12]. Thanasilp et al. show conditions under which quantum-kernel values concentrate, making informative differences expensive to resolve [13].

Every reported metric therefore needs an uncertainty decomposition:

- market-sample uncertainty;
- train/test split uncertainty;
- model-seed uncertainty;
- finite-shot uncertainty;
- device and calibration uncertainty; and
- portfolio-estimation uncertainty.

A single deterministic-looking Sharpe ratio hides these layers.

6.4 Classical learnability and dequantisation

QML advantage claims face a moving classical frontier. Classical learners can use labelled data to approximate functions that appear computationally difficult when considered as standalone quantum computations [11]. Shadow models can move quantum work into training and permit classical deployment [14]. Research on barren-plateau-free structures indicates that some structures enabling trainability may also enable classical or quantum-enhanced classical simulation [15].

The benchmark tournament includes:

- the best known classical task model;
- a parameter-matched model;
- a compute-budget-matched model;
- a kernel approximation or random-feature model;
- a tensor-network or other quantum-inspired surrogate where applicable; and
- an ablation that removes the quantum component while retaining preprocessing.

Failure to beat one weak neural network does not establish quantum advantage.

6.5 Hardware evidence

Preskill's NISQ framing remains important: near-term processors contain meaningful capabilities and material noise, without full fault tolerance [16]. Finance reviews identify potential algorithms and resource challenges [17,18]. Finance QML studies using hardware or simulations should be read at their demonstrated scale. A hardware run establishes that a defined circuit executed and produced a result. It does not establish scalable production economics unless the full pipeline and comparator are measured.

The current research architecture should remain portable across simulators, vendors and devices. Vendor dependence, queueing and calibration are recorded as model inputs, not hidden operational details.

7. EVIDENCE REVIEW: WHAT THE FINANCE STUDIES SHOW

7.1 Positive task-specific evidence

The contextual QNN study applies a multi-task quantum architecture to stock-price distribution prediction for several large US equities and reports improvement over quantum single-task models [19]. The comparator boundary matters: the paper supports a claim about its architecture and experiment; it does not independently establish net production alpha against the full classical frontier.

The 2026 quantum-versus-classical benchmark preprint standardises selected data splits, features and metrics across directional, simulated trading and volatility tasks. It reports QML gains in some asset-task combinations and classical leadership in others [20]. This is useful because it treats superiority as conditional. Its preprint status, task choices and simulated investment layer remain part of the evidence label.

The QTCNN preprint reports an out-of-sample long-short Sharpe improvement on a JPX cross-sectional dataset [22]. That is closer to an investment metric than forecast accuracy alone. Independent reproduction should test universe formation, delistings, corporate actions, timing, costs, turnover, capacity, hyperparameter search and sensitivity to the stated period before the result enters allocator due diligence.

Credit-scoring studies report promising hybrid results in limited-data settings, including simulations and a small hardware experiment [23-25]. Credit classification is economically important and can share modelling features with investment underwriting. It is not the same target as public-market alpha. Evidence can inform architecture and evaluation without being relabelled as return prediction.

7.2 Negative and mixed evidence

The DeFi yield benchmark compares classical ensembles, deep-learning and quantum models on a stated Curve Finance dataset. The reported classical ensemble results exceed the quantum models, whose directional accuracy is below 50 percent in that experiment [21]. This finding does not disprove QML. It demonstrates why strong classical baselines and publication of unfavourable quantum results are necessary.

Theoretical QML literature adds further boundaries. Quantum kernels can lose advantage with noise, measurement limits or uninformative concentration [12,13]. Trainability can fail through barren plateaus or traps [6-10]. Classical models can learn from data or exploit structure in ways that narrow the proposed separation [11,14,15]. These are design constraints and falsification tests.

7.3 Evidence matrix

Source group	Strongest class	Supports	Does not establish
Foundational QML [1-5,39,40]	T/S/H	Model classes, representations and bounded experiments	Investable finance alpha
Trainability and noise [6-10,12,13]	T/S	Failure modes and resource bounds	Failure of every QML design
Classical learnability [11,14,15]	T/S	Need for surrogates and careful advantage claims	Absence of all quantum advantage
Finance reviews [17,18,37]	Review / official analysis	Use-case map and resource considerations	Production QML alpha
Forecast studies [19-22]	S/B; some preprints	Task-specific comparative results	Broad, persistent net alpha
Credit studies [23-25]	S/H; preprints	Limited-data classification hypotheses	Public-market alpha
Empirical finance [26-30]	B/T	Strong baselines, decay, multiplicity and backtest controls	QML-specific result
Governance [31-36]	Official	Claims, validation and risk-management expectations	Investment performance

The evidence matrix prevents both overclaiming and premature dismissal. QML earns a place in a research tournament. Capital remains behind a higher gate.

8. THE CLASSICAL-FIRST BENCHMARK TOURNAMENT

8.1 The tournament rule

The research team defines the tournament before running the final test. Every candidate receives the same point-in-time data, forecast horizon, training and validation windows, economic constraints and cost model. Hyperparameter search budgets are recorded. The final holdout is opened once for the approved comparison.

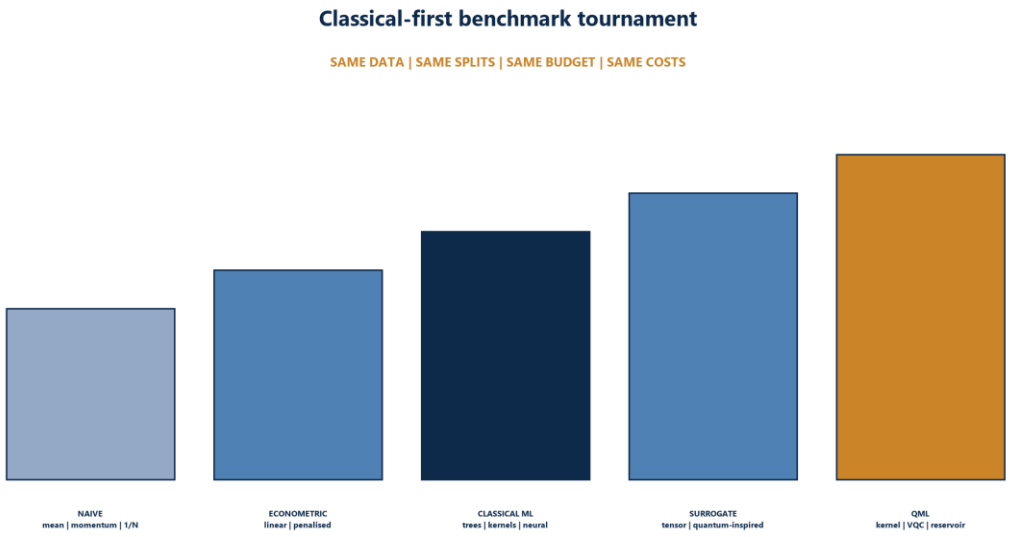

Figure 4. Classical-first benchmark tournament

Source: Matchpoint benchmark framework; empirical-finance and QML context [6-30].

Five lanes are required:

1. **Naive and economic baselines.** Historical mean, no-change forecast, momentum or reversal rule where relevant, equal weight, approved benchmark and simple linear factor model.
2. **Econometric baselines.** Ordinary and penalised regressions, dimension reduction and forecast combinations appropriate to the target.
3. **Classical machine learning.** Tree ensembles, classical kernels and neural networks. Gu, Kelly and Xiu provide evidence that trees and neural networks can be strong return-prediction comparators [26].
4. **Quantum-inspired or surrogate models.** Classical approximations that preserve the QML feature map or structural idea where feasible [11,14,15].
5. **Quantum-executed candidates.** Quantum kernel, variational circuit, reservoir or hybrid temporal model.

The winner is selected on the pre-approved decision objective, not the metric that happens to favour it. For an alpha problem, primary outcomes may include net information ratio, net Sharpe, drawdown, tail loss, turnover, capacity and stability. Prediction metrics remain diagnostic.

8.2 Equal information

All lanes receive the same permissible information. If the QML lane uses principal components, autoencoder features or supervised feature selection, the classical lanes receive the same approved transformed inputs or the transformation is treated as a separate model. This identifies whether value comes from preprocessing, the circuit or the portfolio layer.

The data register records:

Field	Required record
Dataset	Owner, licence, location, version and hash
Availability	Earliest permissible decision timestamp
Universe	Point-in-time eligibility and exclusions
Label	Formula, horizon, lag and revision policy
Transform	Fit sample, parameters and version
Missing data	Rule and indicator treatment
Corporate actions	Source and effective timestamp
Split	Train, validation, test and embargo periods

8.3 Equal search

The experiment ledger records every run, including failed and abandoned specifications. A fair comparison cannot spend hundreds of searches on the proposed QML architecture and one default run on a classical comparator. Search equivalence can be defined through maximum trials, wall-clock budget, researcher hours or monetary compute budget. The chosen rule is frozen before the final test.

Multiplicity controls use the full trial count. Harvey, Liu and Zhu, White, and Bailey and Lopez de Prado provide relevant frameworks for higher discovery thresholds, data-snooping adjustment and deflated performance assessment [27,29,30]. No single statistic substitutes for transparent disclosure of the research path.

8.4 Equal resources

The resource ledger captures:

- CPU, GPU and QPU time;
- simulator time;
- queue and compilation time;
- number of circuits and shots;
- calibration and mitigation runs;
- data preparation and feature-engineering time;
- hyperparameter search;
- software, cloud and vendor cost;
- failure and rerun rate; and
- inference cost at the intended production frequency.

A complexity result may still be strategically important when current wall-clock economics are unfavourable. The report labels the result accurately.

8.5 Equal investment implementation

Each predictive output enters the same portfolio constructor unless the experiment specifically tests portfolio construction. Position limits, sector and factor constraints, risk target, turnover penalty, rebalance rule and cost model stay constant. This prevents a stronger portfolio rule from being misattributed to a stronger predictor.

The portfolio record includes gross and net performance, exposure, drawdown, turnover, capacity, concentration and performance by regime. Any use of leverage or shorting is explicit. Results are shown against the approved benchmark and the strongest classical candidate.

9. CONTROLLED RESEARCH ARCHITECTURE

9.1 Data clock and feature layer

The data clock is the foundation. Each observation carries an event time, an availability time and a processing time. The research dataset is reconstructed as it could have appeared to the decision maker. Restatements and revised data are retained as separate versions.

Feature transformations are fitted only on the permitted training sample. Dimensionality reduction, scaling, winsorisation and imputation parameters are versioned. A QML input vector is traceable back to approved source observations. The architecture blocks the run when a critical timestamp, licence or lineage field is absent.

9.2 Quantum execution lane

The quantum lane records:

- provider and backend;
- software and library versions;
- device identifier and calibration timestamp;
- circuit source and compiled circuit hash;
- qubit mapping;
- gate counts and depth;
- shots by circuit;

- mitigation method and settings;
- random seeds;
- optimiser and stopping rule;
- failure, retry and queue events; and
- raw measurements and decoded outputs.

The record permits a simulator rerun, a same-device rerun where available, and a cross-device or surrogate comparison. The model cannot rely on an undocumented interactive notebook state.

9.3 Portfolio and control layer

The portfolio layer receives a versioned forecast with uncertainty. Deterministic rules apply universe eligibility, constraints, costs and risk limits. The research result cannot bypass these rules. A validator reproduces the signal and portfolio from immutable artefacts.

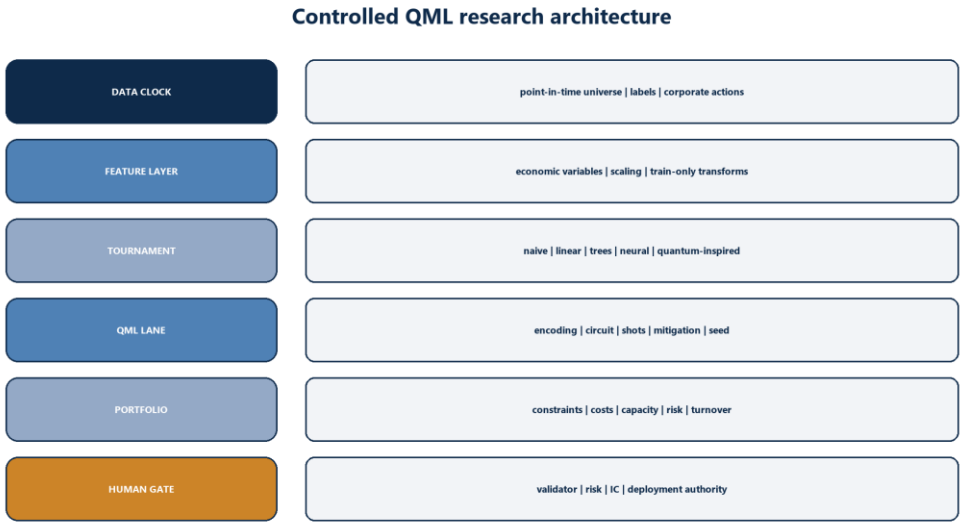

Figure 3. Controlled QML research architecture

Source: Matchpoint control architecture; technical and governance context [1-25,31-40].

Layer	Hard gate	Owner
Data clock	Point-in-time and licence checks pass	Data owner
Features	Train-only fit and lineage pass	Research owner
QML execution	Circuit, shots, device and seeds complete	Quantum research owner
Tournament	Approved baselines and search budgets complete	Independent validator
Portfolio	Costs, constraints and capacity complete	Portfolio construction owner
Claims	Wording matches evidence class	Compliance / legal owner
Capital	Committee approval recorded	Investment committee

9.4 Model-risk governance

The Federal Reserve, OCC and FDIC revised model-risk guidance in April 2026, superseding SR 11-7 for covered banking organisations and emphasising risk-based model governance [33]. The NIST AI RMF organises risk management around Govern, Map, Measure and Manage [34]. ESMA states that management bodies retain responsibility when AI is used in investment services and identifies data, opacity, overreliance, privacy and security risks [35]. IOSCO's capital-markets work identifies model, data, outsourcing, concentration, accountability and human-interaction risks [36].

These sources apply in different legal and supervisory contexts. Qualified legal and compliance advisers determine applicability. The common operating lesson is clear: advanced technology does not remove ownership, validation, monitoring or documentation.

9.5 Third-party and concentration risk

Cloud quantum access can concentrate critical research capability in a small number of providers. The vendor register covers subcontractors, data location, intellectual property, service levels, incident notification, audit rights, exit, model portability and continuity. The research design retains enough information to reproduce the economic conclusion on a simulator or alternative backend where practical.

10. EVALUATION SCORECARD AND HARD GATES

10.1 Scorecard design

Average performance can conceal a critical failure. The scorecard therefore combines hard gates with comparative metrics.

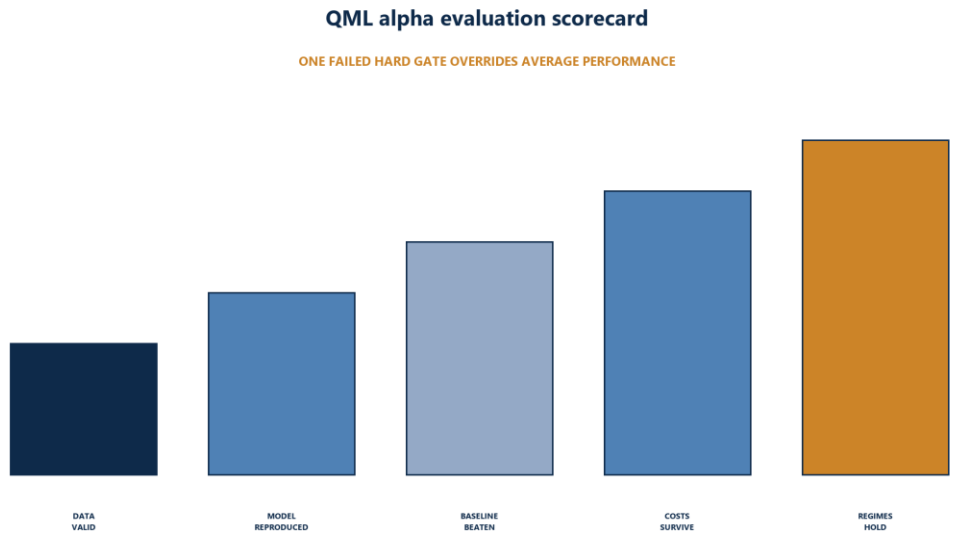

Figure 5. QML alpha evaluation scorecard
Source: Matchpoint evaluation framework.

Dimension	Hard gate	Comparative metric
Data integrity	No material point-in-time or lineage failure	Coverage, staleness, missingness
Reproducibility	Independent rerun within approved tolerance	Output dispersion across seeds and devices
Baseline fairness	All approved baselines and equal search complete	Relative forecast and portfolio performance
Statistical validity	Multiplicity and uncertainty disclosed	Adjusted significance, Deflated Sharpe Ratio
Economic validity	Net result survives approved costs	Net Sharpe, information ratio, drawdown
Capacity	Strategy remains feasible at stated capital	Impact, turnover and concentration
Regime stability	No unexplained dependence on one selected regime	Rolling and regime results
Model risk	Inventory, owner, validation and monitoring complete	Exceptions and remediation
Claims	Wording matches evidence class	Substantiation pack completeness

A failed data-integrity, reproducibility, baseline or claims gate blocks promotion regardless of the average score.

10.2 Statistical protocol

The protocol uses nested or walk-forward validation appropriate to time-ordered data. Hyperparameters are selected without viewing the final test. Embargo or purge windows address overlap where labels span time. Uncertainty is estimated across time splits, model seeds and quantum execution variability.

The research report includes:

- all tested model families and material variants;

- the selection rule;
- the final untouched holdout;
- confidence or uncertainty intervals;
- multiple-testing adjustment;
- performance before and after costs;
- sensitivity to the cost and delay assumptions; and
- failure cases.

10.3 Falsification tests

A strong QML claim should survive attempts to remove its advantage:

- shuffle labels within valid time blocks;
- replace quantum features with matched random features;
- approximate the kernel classically;
- remove the quantum component and retain preprocessing;
- vary shots and noise;
- use alternative point-in-time universes;
- delay execution;
- raise transaction costs;
- test adjacent regimes; and
- rerun on an alternative seed or backend.

The purpose is to locate the source of the result. If the advantage disappears when a minor operational assumption changes, that sensitivity is decision-relevant evidence.

10.4 Monitoring after a paper-portfolio gate

A paper or shadow portfolio preserves the frozen model while observing new data. The team monitors forecast distribution, exposure, turnover, costs, model disagreement, data drift, circuit and device drift, exceptions and performance attribution. Any material model change creates a new version and restarts the relevant validation gate.

The shadow period should be defined by information content rather than a convenient calendar promise. A strategy with monthly decisions needs enough independent observations and regimes to support the committee's stated confidence. The ninety-day roadmap below can reach a research gate; it does not prove investment alpha in ninety days.

11. UNVERIFIED ILLUSTRATIVE OPERATING CASES

All inputs, thresholds, hours, rates, costs and outcomes in this section are **unverified illustrative management assumptions**. They are not observed Matchpoint or client results. They do not establish revenue, cost reduction, loss reduction or investment performance.

11.1 Case A: A1 manager-due-diligence challenge

An A1 allocator receives a manager presentation claiming that a hybrid QML model improves cross-sectional equity selection. The allocator uses the evidence framework before considering the claim in a manager scorecard.

Illustrative input	Unverified assumption
Reported historical period	8 years
Reported assets	500 liquid equities
Reported quantum candidates	12
Undisclosed variants identified in diligence	96
Reported gross Sharpe	1.40
Reproduced net Sharpe under allocator costs	0.38
Best reproduced classical baseline net Sharpe	0.44
Approved capital attributable to claim	USD 0

The allocator does not conclude that QML has no value. The disclosed evidence does not support the manager's superiority claim under the allocator's cost and trial-count assumptions. The issue is recorded as a claims and evidence gap. Any commercial decision follows the allocator's full due-diligence process.

11.2 Case B: A2 controlled research pilot

An A2 family-office investment team wants to explore whether a quantum kernel improves a monthly regime classifier used as one input to asset-allocation discussion.

Illustrative input	Unverified assumption
Approved features	8 macro and market variables
Quantum lane	8-qubit kernel experiment
Classical lanes	Logistic, SVM, boosted trees, shallow neural network
Final holdout	24 monthly observations
Primary metric	Balanced accuracy
Secondary economic test	Paper allocation with costs
Trading authority	None

The final holdout is too small to support a strong economic claim. The team treats the pilot as architecture and process learning. It records reproducibility, research time, cost, vendor dependence and model disagreement. No capital decision is attributed to the QML output.

11.3 Case C: productivity measurement

The research team compares the elapsed time needed to reproduce a manager's technical appendix before and after a reusable experiment ledger and benchmark harness are introduced.

Illustrative measure	Baseline	Pilot	Evidence status
Research hours per reproducible comparison	40	24	Unverified scenario
Independent rerun pass rate	50%	85%	Unverified scenario
Missing trial-count disclosures	6	2	Unverified scenario
Attributed cash cost reduction	USD 0	USD 0	No approved evidence

The scenario measures a potential operating benefit of disciplined QML research. It does not claim investment alpha or realised financial benefit.

12. BUSINESS CASE AND VALUE ATTRIBUTION

12.1 Value bridge

The business case progresses through evidence states:

1. **Idea.** A research hypothesis has no recognised economic value.
2. **Reproduced.** An independent team reproduces the technical result.
3. **Paper.** The result survives investability and cost checks in a frozen historical or paper setting.
4. **Shadow.** The model operates on new data without capital authority.
5. **Attributed.** Approved evidence connects an observed decision or operating outcome to the intervention.

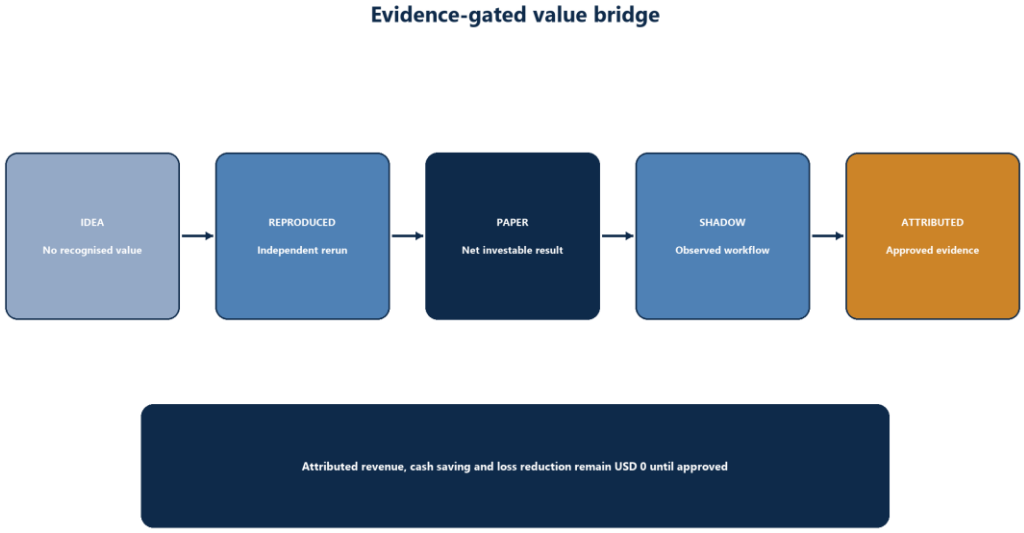

Figure 6. Evidence-gated value bridge

Source: Matchpoint business-case framework; all scenario values require approved observed evidence.

Research productivity can be measured separately from investment performance. Suitable operating measures include accepted experiments, elapsed time to reproduce, validation exceptions, evidence-pack completeness and avoided duplicate work. Investment measures include net active return, risk, drawdown, turnover, capacity and attribution. The two sets should not be combined into one unsupported return-on-investment number.

12.2 Illustrative economics formula

For a research-process use case:

Illustrative annual gross capacity value = accepted hours released x approved loaded cost per hour x approved realisation factor.

Each input must be observed and approved. Released hours that are not removed from spend or used for approved additional work do not automatically become cash savings. For an investment use case:

Illustrative net economic contribution = realised portfolio outcome minus approved benchmark outcome minus implementation costs, adjusted for risk and attribution.

This calculation requires investment-committee and finance approval. It must not be reverse-engineered from a favourable backtest.

12.3 Current attribution status

No approved observed Matchpoint or client evidence was supplied for T23 revenue, cash cost reduction, loss reduction or investment alpha. The current attributed values are:

Category	Attributed value
Matchpoint revenue	USD 0
Client revenue	USD 0
Matchpoint cash cost reduction	USD 0
Client cash cost reduction	USD 0
Loss reduction	USD 0
Investment alpha	USD 0

These zero values are evidence boundaries. They are not forecasts.

13. NINETY-DAY RESEARCH ROADMAP

13.1 Days 0-15: question and authority

- Name the A1 diligence or A2 research decision.
- Define alpha, benchmark, horizon and costs.
- Record research, validation, risk and committee owners.
- Confirm data, vendor, confidentiality and publication boundaries.
- Freeze evidence classes and prohibited claims.

Exit gate: signed research charter and no trading authority.

13.2 Days 16-30: data clock

- Build the point-in-time universe and label.
- Version sources, transformations and release lags.
- Establish train, validation, test and embargo periods.
- Run leakage tests.
- Freeze the final holdout.

Exit gate: independent data review passes.

13.3 Days 31-45: classical baselines

- Run naive, economic, econometric and classical ML lanes.
- Record equal search budgets and all trials.
- Implement the portfolio, cost and capacity layer.
- Define primary and secondary metrics.

Exit gate: benchmark pack is accepted before QML result selection.

13.4 Days 46-60: QML lane

- Implement encoding, circuit, measurements and optimiser.
- Record simulator and hardware settings.
- Measure shots, noise, queueing, failures and cost.
- Run quantum-inspired or surrogate comparators.
- Preserve every trial.

Exit gate: technical artefacts are complete and reproducible.

13.5 Days 61-75: independent validation

- Reproduce the chosen and rejected results.
- Apply multiplicity and uncertainty analysis.
- Run falsification, cost and regime sensitivity.
- Challenge the need for the quantum component.
- Draft evidence-bounded claims.

Exit gate: validator signs the evidence class and exceptions.

13.6 Days 76-90: committee gate

- Present the full tournament, including unfavourable results.
- Decide whether to stop, redesign or enter a restricted shadow period.
- Approve monitoring, incident and change controls.
- Approve external wording separately.
- Record capital authority as none unless a later investment gate is met.

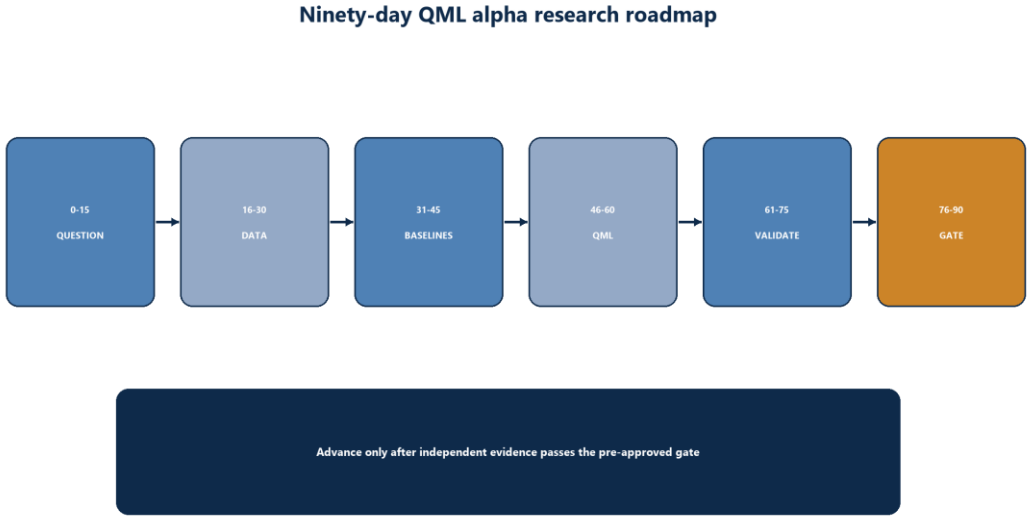

Figure 7. Ninety-day QML alpha research roadmap

Source: Matchpoint implementation framework.

The ninety-day result is a governed research decision. It is not a promise of alpha or deployment.

14. CLAIMS REGISTER

Claim	Evidence status	Permitted conclusion
QML provides expressive feature maps and model classes	Supported by T/S/H research [1-5,39,40]	Legitimate research family
Some QML tasks show advantages over selected comparators	Supported in bounded studies [11,19,20,22-25]	Replication hypothesis
Some tested quantum models underperform classical models	Supported in bounded benchmark [21]	Strong baselines are necessary
Trainability, noise and sampling can be binding	Supported by T/S research [6-10,12,13]	Scaling and resource tests are required
Classical learners and surrogates can narrow advantage	Supported by T/S research [11,14,15]	Include dequantisation and ablation tests
Machine learning can improve return prediction	Supported in large classical study [26]	Strong classical frontier exists
Published signals and selected backtests can decay or overfit	Supported by empirical and statistical research [27-30]	Trial records and fresh evidence are required
The reviewed sources prove production QML alpha	Not established	Prohibited claim
T23 proves Matchpoint or client financial benefit	Not established	Attributed value remains USD 0

15. LIMITATIONS AND CONCLUSION

15.1 Limitations

The quantum and finance literatures evolve quickly. Several recent finance comparisons are preprints [20-25], and their findings require peer review or independent reproduction. Published studies use different datasets, labels, architectures and comparators, limiting aggregation. Hardware capability and classical simulation methods continue to change. The paper does not perform a new empirical QML experiment and does not assess a named manager, vendor or portfolio.

Regulatory sources apply to particular jurisdictions, firms and activities [31-36]. This paper provides a governance framework and not legal advice. A1 and A2 institutions must obtain qualified advice for their facts.

Alpha is strategy-specific. A failure to establish broad production advantage does not establish impossibility. A favourable task result does not establish investable superiority. Both boundaries are material.

15.2 Conclusion

Quantum machine learning deserves disciplined research attention. It offers distinctive representations, kernels and hybrid architectures, and finance studies provide task-specific positive, negative and mixed results. The reviewed evidence through 1 August 2026 does not establish repeatable, net-of-cost production alpha or broad quantum advantage on investable classical market data.

For A1 allocators, the immediate value is a stronger due-diligence standard for manager and technology claims. For A2 family-office teams, the immediate value is a controlled research lane that preserves independent challenge and committee authority. In both cases, the correct operating model is classical-first, evidence-gated and fully attributable.

The decision rule is concise: **advance a QML alpha claim only when the point-in-time data, equal-information tournament, complete resource ledger, investable portfolio, independent reproduction, multiplicity control, regime evidence and accountable approval all pass.** Until then, the output is research evidence. Attributed Matchpoint or client revenue, cash cost reduction, loss reduction and investment alpha remain USD 0 without approved observed evidence.

REFERENCES

[1] Biamonte, J. et al. (2017). Quantum machine learning. Nature, 549, 195-202. https://doi.org/10.1038/nature23474
[2] Havlicek, V. et al. (2019). Supervised learning with quantum-enhanced feature spaces. Nature, 567, 209-212. https://doi.org/10.1038/s41586-019-0980-2
[3] Schuld, M. and Killoran, N. (2019). Quantum machine learning in feature Hilbert spaces. Physical Review Letters, 122, 040504. https://doi.org/10.1103/PhysRevLett.122.040504
[4] Benedetti, M. et al. (2019). Parameterized quantum circuits as machine learning models. Quantum Science and Technology, 4, 043001. https://doi.org/10.1088/2058-9565/ab4eb5
[5] Perez-Salinas, A. et al. (2020). Data re-uploading for a universal quantum classifier. Quantum, 4, 226. https://doi.org/10.22331/q-2020-02-06-226
[6] McClean, J. R. et al. (2018). Barren plateaus in quantum neural network training landscapes. Nature Communications, 9, 4812. https://doi.org/10.1038/s41467-018-07090-4
[7] Cerezo, M. et al. (2021). Cost function dependent barren plateaus in shallow parametrized quantum circuits. Nature Communications, 12, 1791. https://doi.org/10.1038/s41467-021-21728-w
[8] Wang, S. et al. (2021). Noise-induced barren plateaus in variational quantum algorithms. Nature Communications, 12, 6961. https://doi.org/10.1038/s41467-021-27045-6
[9] Arrasmith, A. et al. (2021). Effect of barren plateaus on gradient-free optimization. Quantum, 5, 558. https://doi.org/10.22331/q-2021-10-05-558
[10] Anschuetz, E. R. and Kiani, B. T. (2022). Quantum variational algorithms are swamped with traps. Nature Communications, 13, 7760. https://doi.org/10.1038/s41467-022-35364-5
[11] Huang, H.-Y. et al. (2021). Power of data in quantum machine learning. Nature Communications, 12, 2631. https://doi.org/10.1038/s41467-021-22539-9
[12] Wang, X. et al. (2021). Towards understanding the power of quantum kernels in the NISQ era. Quantum, 5, 531. https://doi.org/10.22331/q-2021-08-30-531
[13] Thanasilp, S. et al. (2024). Exponential concentration in quantum kernel methods. Nature Communications, 15, 5200. https://doi.org/10.1038/s41467-024-49287-w
[14] Abbe, E. et al. (2024). Shadows of quantum machine learning. Nature Communications, 15, 5676. https://doi.org/10.1038/s41467-024-49877-8
[15] Cerezo, M., Larocca, M. and Holmes, Z. (2025). Does provable absence of barren plateaus imply classical simulability? Nature Communications, 16, 7907. https://doi.org/10.1038/s41467-025-63099-6
[16] Preskill, J. (2018). Quantum Computing in the NISQ era and beyond. Quantum, 2, 79. https://doi.org/10.22331/q-2018-08-06-79
[17] Egger, D. J. et al. (2020). Quantum Computing for Finance: State-of-the-Art and Future Prospects. IEEE Transactions on Quantum Engineering, 1, 3101724. https://doi.org/10.1109/TQE.2020.3030314
[18] Herman, D. et al. (2023). Quantum computing for finance. Nature Reviews Physics, 5, 450-465. https://doi.org/10.1038/s42254-023-00603-1
[19] Mourya, S., Leipold, H. and Adhikari, B. (2026). Contextual quantum neural networks for stock price prediction. Scientific Reports, 16, 4454. https://doi.org/10.1038/s41598-025-34413-5
[20] Ahmad, R. et al. (2026). Quantum vs. Classical Machine Learning: A Benchmark Study for Financial Prediction. arXiv preprint arXiv:2601.03802. https://arxiv.org/abs/2601.03802
[21] Chen, C.-S. and Tsai, A. H.-W. (2025). Benchmarking Classical and Quantum Models for DeFi Yield Prediction on Curve Finance. arXiv preprint arXiv:2508.02685. https://arxiv.org/abs/2508.02685
[22] Chen, C.-S. et al. (2025). Quantum Temporal Convolutional Neural Networks for Cross-Sectional Equity Return Prediction: A Comparative Benchmark Study. arXiv preprint arXiv:2512.06630. https://arxiv.org/abs/2512.06630
[23] Schetakis, N. et al. (2023). Quantum Machine Learning for Credit Scoring. arXiv preprint arXiv:2308.03575. https://arxiv.org/abs/2308.03575
[24] Mancilla, J. et al. (2024). Empowering Credit Scoring Systems with Quantum-Enhanced Machine Learning. arXiv preprint arXiv:2404.00015. https://arxiv.org/abs/2404.00015
[25] Wang, Z.-A. et al. (2025). Hybrid Quantum-Classical Neural Networks for Few-Shot Credit Risk Assessment. arXiv preprint arXiv:2509.13818. https://arxiv.org/abs/2509.13818
[26] Gu, S., Kelly, B. and Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. The Review of Financial Studies, 33(5), 2223-2273. https://doi.org/10.1093/rfs/hhaa009
[27] Harvey, C. R., Liu, Y. and Zhu, H. (2016). ... and the Cross-Section of Expected Returns. The Review of Financial Studies, 29(1), 5-68. https://doi.org/10.1093/rfs/hhv059
[28] McLean, R. D. and Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? The Journal of Finance, 71(1), 5-32. https://doi.org/10.1111/jofi.12365
[29] White, H. (2000). A Reality Check for Data Snooping. Econometrica, 68(5), 1097-1126. https://doi.org/10.1111/1468-0262.00152
[30] Bailey, D. H. and Lopez de Prado, M. (2014). The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting and Non-Normality. The Journal of Portfolio Management, 40(5), 94-107. https://doi.org/10.2139/ssrn.2460551
[31] U.S. Securities and Exchange Commission (2020). Investment Adviser Marketing; Final Rule, Release No. IA-5653. https://www.sec.gov/files/rules/final/2020/ia-5653.pdf
[32] U.S. Securities and Exchange Commission (2024). SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence, Press Release 2024-36. https://www.sec.gov/newsroom/press-releases/2024-36
[33] Board of Governors of the Federal Reserve System, Office of the Comptroller of the Currency and Federal Deposit Insurance Corporation (2026). Revised Guidance on Model Risk Management, SR 26-2. https://www.federalreserve.gov/supervisionreg/srletters/SR2602.htm
[34] National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1. https://doi.org/10.6028/NIST.AI.100-1
[35] European Securities and Markets Authority (2024). Public Statement on the use of Artificial Intelligence in the provision of retail investment services, ESMA35-335435667-5924. https://www.esma.europa.eu/sites/default/files/2024-05/ESMA35-335435667-5924__Public_Statement_on_AI_and_investment_services.pdf
[36] International Organization of Securities Commissions (2025). Artificial Intelligence in Capital Markets: Use Cases, Risks, and Challenges, Board/2025/017. https://www.iosco.org/library/pubdocs/pdf/IOSCOPD788.pdf
[37] Auer, R. et al. (2024). Quantum computing and the financial system: opportunities and risks. BIS Papers No. 149. Bank for International Settlements. https://www.bis.org/publ/bppdf/bispap149.htm
[38] IBM Research and Qiskit contributors (2021). Qiskit Machine Learning: A Software Package for Quantum Machine Learning. arXiv preprint arXiv:2109.01584. https://arxiv.org/abs/2109.01584
[39] Schuld, M., Sweke, R. and Meyer, J. J. (2021). Effect of data encoding on the expressive power of variational quantum-machine-learning models. Physical Review A, 103, 032430. https://doi.org/10.1103/PhysRevA.103.032430
[40] Abbas, A. et al. (2021). The power of quantum neural networks. Nature Computational Science, 1, 403-409. https://doi.org/10.1038/s43588-021-00084-1

APPENDIX A. SCENARIO ASSUMPTIONS REGISTER

Assumption	Status	Approval required before use
Historical period, universe and decision frequency	Unverified illustrative management assumption	Research and investment owners
Model and hyperparameter counts	Unverified illustrative management assumption	Research owner and validator
Compute, QPU, software and vendor cost	Unverified illustrative management assumption	Technology and finance owners
Spread, impact, borrow and financing cost	Unverified illustrative management assumption	Trading and finance owners
Capacity and liquidity	Unverified illustrative management assumption	Portfolio and risk owners
Released research hours	Unverified illustrative management assumption	Operations and finance owners
Loaded cost and realisation factor	Unverified illustrative management assumption	Finance owner
Performance and economic contribution	Unverified illustrative management assumption	Investment committee and finance owner

APPENDIX B. MINIMUM EXPERIMENT RECORD

Category	Required fields
Hypothesis	Economic mechanism, target, horizon, benchmark, falsification rule
Data	Source, licence, timestamps, universe, label, transformations, hashes
Split	Train, validation, test, embargo and final holdout
Classical models	Architecture, search space, trials, seeds, compute and results
QML model	Encoding, circuit, observable, optimiser, depth, shots and seeds
Hardware	Provider, backend, calibration, compilation, queue and failure record
Portfolio	Construction, constraints, costs, capacity and benchmark
Statistics	Trial count, uncertainty, multiplicity adjustment and sensitivity
Validation	Independent code, data and result reproduction
Governance	Owners, approvals, exceptions, monitoring and change record
Claims	Exact wording, evidence class, audience, approver and date

APPENDIX C. GLOSSARY

Alpha. Residual investment return after the approved benchmark, risk model, costs and implementation are applied.
Barren plateau. A training landscape in which gradients become too small to estimate efficiently as a model scales under stated conditions.
Classical surrogate. A classical model that approximates or reproduces a quantum model's relevant input-output behaviour.
Data leakage. Use of information or fitted transformations unavailable to the real decision at the relevant time.
Dequantisation. Development of a classical algorithm or approximation that narrows or removes a proposed quantum computational advantage for a defined task.
NISQ. Noisy intermediate-scale quantum; devices with meaningful qubit counts and operations, material noise and no complete fault tolerance.
Quantum advantage. A decision-relevant improvement over the best feasible classical alternative under a fair end-to-end resource comparison.
Quantum kernel. A similarity function estimated from quantum-state or circuit overlaps and used by a kernel learning method.
Quantum neural network. A learning model using parameterised quantum circuits, commonly within a hybrid classical optimisation loop.
Shots. Repeated quantum circuit measurements used to estimate output probabilities or expectation values.
Shadow portfolio. A controlled portfolio process using new information without production capital authority.