

Article

Quantum Hyperbolic Deep Learning for Foreign-Exchange Trading: A Hybrid Reinforcement-Learning Pipeline over Attractor-Aware Magnet-Price Manifolds

Francesco Rundo 

Department of Mathematics and Computer Science, University of Catania, 95125 Catania, Italy;
francesco.rundo@unict.it

Abstract

Foreign-exchange decisions rest on hierarchically organized evidence whose latent structure is inadequately captured by Euclidean representations. Reinforcement-learning agents trained on flat embeddings inherit stability guarantees that do not transfer to the manifold supporting the latent state. We address both limitations through a hybrid architecture in which a schema-constrained structured chain-of-thought is embedded into a Poincaré ball, transported to a qubit register via angle encoding, and processed by an L -layer hardware-efficient variational ansatz on a state-vector backend. The circuit exposes two read-outs to the policy, namely, a scalar Pauli-Z observable and a projected quantum kernel inducing a fidelity-based similarity over magnet-price attractors, the latter identified via kernel-weighted recurrence density and finite-time Lyapunov statistics. The Lipschitz constraint on the action-value function is lifted from the hyperbolic geodesic distance to a joint metric on $\mathbb{B}_K^n \times \mathbb{P}(\mathcal{H})$. A stability theorem yields an explicit bound depending on the read-out operator norm, on the depth-width product of the ansatz, and on the curvature-Hilbert balance. The pipeline is evaluated on nine major FX crosses over a 2015–2025 out-of-sample window, with rolling-origin walk-forward retraining and broker-published transaction costs. The system attains 2.55% pair-averaged non-compounded monthly P&L and 8.83% maximum drawdown, with Sharpe 1.78, Calmar 3.43, and Probabilistic Sharpe Ratio exceeding 0.95 on every cross. The gain remains significant under a deflated-Sharpe-ratio test with $N_{\text{trials}} = 42$ correction. Block-wise ablations exhibit strictly monotone degradation: removing the projected kernel costs -4.15 p.p. on annualized P&L, the joint Lipschitz penalty -6.42 p.p., the attractor module -7.64 p.p., and the hyperbolic embedding -8.40 p.p. The quantum block thereby instantiates a structurally non-classical, geometry-aware regularizer identifiable through ablation rather than asymptotically advantageous.

Keywords: quantum machine learning; hyperbolic deep learning; reinforcement learning; foreign-exchange trading; chaotic attractors



Academic Editors: Albert Y.S. Lam
and Andy Chun

Received: 30 April 2026

Revised: 3 June 2026

Accepted: 5 June 2026

Published: 11 June 2026

Copyright: © 2026 by the authors.
Licensee MDPI, Basel, Switzerland.
This article is an open access article
distributed under the terms and
conditions of the [Creative Commons
Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Foreign-exchange (FX) returns are heavy-tailed, autocorrelated in volatility, and subject to recurrent regime shifts. These properties violate the stationarity assumptions of most data-driven trading systems [1–4]. Classical deep learning improves short-horizon forecasting accuracy. However, it does not control tail risk under distributional shift. The Euclidean encoders prevalent in the literature also fail to capture the hierarchical, multi-scale dependencies that link macroeconomic context, technical price structure, and dynamical diagnostics in the trading state space [5–7]. Negatively curved latent geometries, attractor-aware feature extraction, and risk-constrained reinforcement learning have been proposed

as a response. The open question is whether an additional source of expressivity, structurally non-equivalent to standard kernel methods on selected concept classes, can raise return and reduce downside risk under a fixed cost model and a fixed evaluation horizon. The empirical literature does not address this question in a controlled setting.

The present work builds on the pipeline of [8]. That pipeline consolidates heterogeneous evidence into a structured chain-of-thought generated by a schema-constrained large language model. It embeds the result into a Poincaré ball and drives a Lipschitz-constrained reinforcement learning agent toward magnet-price attractors identified by recurrence and cross-recurrence statistics. The state includes finite-time Lyapunov exponents and normalized permutation entropy. On nine major FX crosses over a January 2015–July 2025 window the pipeline reaches 2.22% non-compounded monthly P&L at 10.39% maximum drawdown. The kernels that drive its Bellman backups admit Nyström and random-feature surrogates up to curvature reweighting. No formal stability bound couples the geometric regularization to a non-classical similarity structure.

Variational quantum algorithms and quantum machine learning provide a complementary inductive bias. Small quantum processors and high-fidelity noiseless simulators can be used as differentiable, non-classical feature maps for sequential decision-making [9–13]. Two properties motivate their use here. First, parameterized quantum circuits realize functions whose inner products live in an exponentially large Hilbert space. On selected concept classes these inner products are conjectured to escape efficient classical reproduction, yielding kernels with formal separation results [14,15]. However, the guarantee is not robust. Dequantization results, exponential concentration of quantum kernels, and efficient classical surrogates of shallow ansätze indicate that an observed quantum gain may reduce to a non-linear feature map reproducible by a sufficiently expressive classical model. We therefore frame the quantum block as a geometry-aware, structurally non-classical regularizer, not as a vehicle for asymptotic quantum advantage. The framing is tested against parameter-matched classical surrogates. Further, parameter-shift gradients on parameterized circuits are unbiased and analytically exact [16,17]. A variational circuit therefore composes with the Riemannian stochastic gradient descent already used on the Poincaré ball [18]. The result is a single hybrid optimization problem amenable to a unified Lipschitz-stability analysis on the hyperbolic–Hilbert product manifold.

The contribution of this paper is as follows. The primary methodological contribution is the lifting of the Lipschitz constraint on the action-value function from the hyperbolic geodesic distance d_κ to a joint metric on the hyperbolic–Hilbert product manifold $\mathbb{B}_\kappa^n \times \mathbb{P}(\mathcal{H})$, formalized by the joint stability theorem established in Section 3. The quantum read-outs constitute a structurally non-classical, geometry-aware regularizer, not a source of asymptotic quantum advantage; the nine-cross out-of-sample evaluation is the controlled validation of this regularizer, not the contribution itself.

First, we introduce a hybrid hyperbolic–quantum architecture. The hyperbolic latent is mapped into an n_q -qubit register through angle and amplitude encodings, processed by a hardware-efficient variational ansatz $U(\theta)$, and exposed to the policy through two read-outs: a scalar observable expectation $f_\theta(x)$ and a projected quantum kernel $K_Q(x, x')$. Second, we formulate a hybrid optimization scheme. Riemannian gradients on the Poincaré ball and parameter-shift gradients on the variational angles are interleaved under a single objective. The Lipschitz penalty is lifted from the hyperbolic distance d_κ to a joint metric $d_\kappa \oplus d_Q$ on the hyperbolic–Hilbert product manifold. We prove a stability theorem that bounds the deviation of the action-value function in terms of d_κ and the fidelity-induced distance $1 - K_Q$. Third, we replace the curvature-aware kernel of the Bellman backup of [8] with a convex blend of the classical hyperbolic kernel and the quantum kernel. The blend is governed by a single coefficient $\nu \in [0, 1]$ selected on

validation. The contribution of the quantum branch is therefore identifiable by ablation. Fourth, we evaluate the resulting pipeline on nine FX crosses in the [2015–2025] horizon. The protocol comprises seven elements: rolling-origin walk-forward; ten random seeds with reported variance; a battery of competitive baselines including a parameter-matched random-Fourier-feature surrogate of K_Q ; regime-stratified attribution across the 2015 Swiss National Bank unpegging, the 2016 Brexit referendum, the 2020 COVID-19 dislocation, and the 2022–2023 Federal Reserve tightening cycle; deflated Sharpe-ratio testing with multiple-comparison correction; sensitivity analysis under depolarizing and shot noise; and a barren-plateau diagnostic on the variational gradients.

Under this protocol the pipeline reaches 2.55% pair-averaged non-compounded monthly P&L and 8.83% maximum drawdown.

The remainder of the paper is organized as follows. Section 2 reviews classical deep learning and reinforcement learning for FX, hyperbolic representation learning, quantum machine learning for finance, and the dequantization and exponential-concentration literature. Section 3 formalizes the pipeline and proves the joint Lipschitz-stability theorem. Section 4 reports the empirical evaluation, the regime-stratified attribution, the ablation and sensitivity analyses, the noise-robustness experiments, and the comparison against classical surrogates [8]. Section 5 states the conclusions, discusses limitations under more pessimistic execution models, and outlines directions for fault-tolerant hardware deployment, multi-asset extensions, and real-time inference.

2. Related Work

The methodological backbone of this paper sits at the intersection of three research traditions that have evolved largely in isolation until recently: classical deep learning and reinforcement learning for foreign-exchange trading, hyperbolic representation learning, and quantum machine learning for finance. The latter intersects in turn with the more recent stream of parameterized quantum reinforcement learning. We review each tradition in narrative form, with particular attention to the boundary results that determine whether the gains reported in Section 4 can plausibly be ascribed to a genuinely non-classical source of expressivity rather than to a more conventional non-linear feature map. We make explicit throughout the exposition the gap that the construction of Section 3 is designed to close.

The literature on classical deep learning for foreign-exchange forecasting has converged on two architectural families whose strengths and limitations are now well documented. Recurrent and attention-based models, surveyed and benchmarked in [19,20], dominate the short-horizon prediction setting and have been progressively refined through transformer-style temporal encoders and multi-resolution input pyramids. Stacked-generalization ensembles [21] and genetic-programming systems with retraining triggers tied to changing market criteria [22] have, in turn, demonstrated that meta-learning over heterogeneous base models can absorb part of the regime-shift volatility that defeats single-architecture forecasters, once realistic transaction costs are explicitly modelled. Reinforcement learning superimposes sequential decision-making on this forecasting substrate and has been instantiated, in increasing order of architectural complexity, as cooperative multi-agent deep Q -learning [23], proximal policy optimization with auxiliary forecasting tasks [24], recurrent reinforcement learning with online inductive transfer across currency pairs [25], and image-encoded convolutional pipelines that recast price action as Gramian angular fields or recurrence plots before delegating decision-making to a deep Q -network or a PPO controller [26,27]. The broader landscape of deep reinforcement learning for trading has been mapped by [28–30]. Two earlier works in the same lineage as the present paper, namely the rule-conditioned grid-trading robot of [31] and the long short-term memory corrector trained with a reinforcement layer of [32], anchor the historical baselines

used in the empirical section. Despite their architectural diversity, none of the systems above exploits the hierarchical structure of financial reasoning, none embeds the latent state in a non-Euclidean geometry, and none has access to a feature map whose inner-product kernel is provably outside the reach of low-dimensional classical Nyström approximations. The contribution that motivates the present extension is therefore not a marginal refinement of any of the architectures cited above, but the introduction of two structurally orthogonal sources of inductive bias, namely negative curvature and Hilbert-space similarity, that act jointly on the same latent state.

The first of these sources has been developed under the coverage of hyperbolic representation learning. Spaces of constant negative curvature exhibit exponential volume growth and embed tree-like or hierarchical structures with vanishing distortion in single-digit dimensions, a property operationalized for deep learning through the Poincaré embeddings of [33] and the hyperbolic neural networks of [34]. Hyperbolic graph neural networks and attention layers [6,7] replace Euclidean operators with their Möbius counterparts and aggregate features along geodesics. Recent surveys document the systematic adoption of these primitives across computer vision [5] and graph time-series modelling [35]. The optimization of manifold-valued parameters is grounded by the Riemannian stochastic gradient descent of [18] and its adaptive variants [36], which together provide the analytical machinery required to train networks whose parameters live on a Riemannian manifold rather than in a flat Euclidean space. The use of negatively curved latents inside a reinforcement learning loop, by contrast, has only recently been investigated. A systematic study of hyperbolic policies and value functions on standard control benchmarks is reported in [37], while [8], the direct precursor of the present paper, extends the construction to financial decision-making by coupling a Poincaré-ball encoder of structured reasoning traces with a Lipschitz-constrained agent and an attractor-aware reward shaped by finite-time Lyapunov exponents and permutation entropy. The present paper inherits the hyperbolic primitives of [8] but lifts them to a hybrid hyperbolic–Hilbert product manifold on which a joint Lipschitz stability bound, absent in all of the works cited above, anchors the rest of the construction.

The second source of inductive bias is supplied by quantum machine learning, whose application to finance has been mapped by surveys and roadmaps that consistently identify trading, portfolio optimization, and risk modelling as the near-term commercial frontiers of the field [11,12]. On the methodological side, parameterized quantum circuits are now formally analysed as machine-learning models [9,10]. Hardware-efficient ansätze [38] render them tractable on noisy intermediate-scale quantum devices, and the parameter-shift rule of [16,17] provides analytically exact, unbiased gradients that compose cleanly with classical automatic differentiation. Quantum kernels expose Hilbert-space feature maps with formal expressivity advantages on selected concept classes [14,15,39], while quantum reservoir computing harnesses the disordered dynamics of a fixed quantum ensemble for time-series tasks at vanishing parameter cost [40]. Hybrid classical–quantum schemes have become standard practice in applied QML [41], and dynamic portfolio optimization on real datasets has been demonstrated on actual quantum processors and on quantum-inspired tensor-network simulators [13]. The expressivity claims associated with these constructions are nonetheless sharper than the asymptotic guarantees they admit on the shallow, hardware-efficient ansätze of the size used in the present work. The inner-product structure of a quantum kernel can, on adversely chosen embeddings, exhibit exponential concentration that erodes its inductive bias, and the variational landscape of a parameterized circuit can fall into broad parameter regions in which the gradient variance vanishes exponentially in the system size. Consistently with the cautious framing established in the introduction, the empirical section reports the diagnostics that this framing requires,

including a barren-plateau probe on the variational gradients and a sensitivity analysis with respect to depolarizing and shot noise on the quantum branch.

The intersection of reinforcement learning and parameterized quantum circuits constitutes the most recent of the streams reviewed here. Parameterized quantum policies and value functions have been established as competitive on standard control benchmarks by [42,43]. These works jointly demonstrate that variational quantum agents can match classical deep Q-network and policy-gradient baselines on small-state-space problems while consuming far fewer trainable parameters. Their use for financial decision-making remains, however, comparatively unexplored. To the best of our knowledge, no prior work combines quantum policies or quantum-kernel-weighted Bellman backups with a hyperbolic latent state and with attractor-aware reward shaping, and no prior work derives a joint Lipschitz stability bound on the hyperbolic–Hilbert product manifold that would justify such a combination from a regularization standpoint. The pipeline introduced in Section 3 closes this gap, and the empirical protocol of Section 4 subjects it to the methodological standard that a multi-pair, regime-stratified, walk-forward evaluation under realistic transaction costs requires.

3. Materials and Methods

This section formalizes the proposed hybrid pipeline. A schematic view of the architecture, including the attractor-aware extractor [8], the hyperbolic encoder of the structured chain-of-thought, and the newly introduced Quantum AI block, is given in Figure 1.

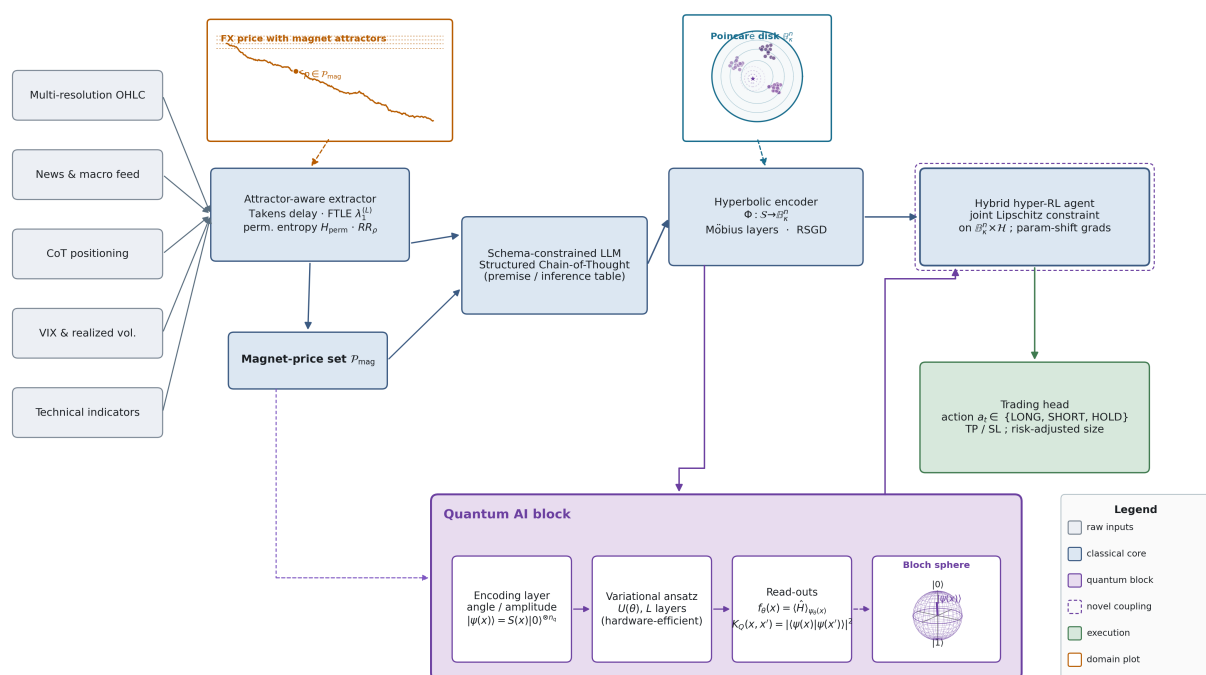


Figure 1. Hybrid hyperbolic–quantum reinforcement-learning pipeline. Heterogeneous evidence streams feed the attractor-aware extractor which produces dynamical diagnostics (FTLE, permutation entropy, recurrence) and the magnet-price set $\mathcal{P}_{\text{mag}}$. A schema-constrained LLM consolidates the diagnostics into a structured chain-of-thought, embedded into the Poincaré ball $\mathbb{B}_k^n$. The Quantum AI block maps the hyperbolic latent to an n_q -qubit register, applies an L -layer hardware-efficient ansatz, and exposes a scalar observable $f_\theta(x)$ and a projected quantum kernel $K_Q(x, x')$. The RL agent operates under a joint Lipschitz constraint on the hyperbolic–Hilbert product manifold and drives the trading head.

3.1. Inherited Classical Substrate: Attractor-Aware Extractor and Hyperbolic Encoder

The attractor-aware feature extractor of [8] is adopted as the entry stage of the pipeline and extended downstream by the components introduced in this paper; it is summarized here only for self-containment; it constitutes the substrate on which the present work operates rather than a contribution of this paper, as already disclosed in Section 2. Let $\{t_k\}_{k \in \mathbb{Z}}$ denote the trading grid, that is, the discrete set of evenly spaced time stamps at which the agent observes the market and may submit orders, with constant spacing $\Delta t = 15 \text{ min}$ (M15). Let $y_t \in \mathbb{R}$ denote the mid-quote of an FX cross sampled at $t \in \{t_k\}$. Following Takens' embedding theorem [44,45], a diffeomorphic copy of the underlying low-dimensional attractor is recovered through the delay vector defined as:

$$\Phi(y)_t = (y_t, y_{t-\tau}, \dots, y_{t-(m-1)\tau}) \in \mathbb{R}^m, \quad (1)$$

in which y_t is the price at time t , $\tau \in \mathbb{N}$ is the embedding delay that decorrelates successive coordinates, $m \in \mathbb{N}$ is the embedding dimension that must satisfy $m \geq 2d_A + 1$ for the embedding to be generic, and d_A is the box-counting dimension of the underlying attractor. The local rate of phase-space expansion along trajectories of the embedded system is summarized by the finite-time largest Lyapunov exponent expressed as:

$$\lambda_1^{(L)}(t) = \frac{1}{L} \log \|Df^L(x_t)\|_1, \quad (2)$$

where f denotes the time- Δt flow induced by the local diffeomorphism reconstructed from the delay coordinates, Df^L is the Jacobian of its L -fold composition along the orbit emanating from the embedded state x_t , $L \in \mathbb{N}$ is the integration horizon over which expansion rates are averaged, and $\|\cdot\|_1$ denotes the operator norm used to extract the dominant expansion direction. The complementary degree of disorder of the symbolic dynamics extracted from the same FX financial time series is captured by the normalized permutation entropy [46] expressed as:

$$\tilde{H}_{\text{perm}}(m_e, \tau_e) = -\frac{1}{\log m_e!} \sum_{\pi \in S_{m_e}} p_\pi \log p_\pi, \quad (3)$$

in which $m_e \in \mathbb{N}$ is the order of the ordinal patterns considered, $\tau_e \in \mathbb{N}$ is the temporal lag between successive samples within a pattern, S_{m_e} is the symmetric group of permutations of m_e elements, $p_\pi \in [0, 1]$ is the empirical relative frequency of pattern π over the analysis window, and the prefactor $1/\log m_e!$ ensures that the entropy lies in $[0, 1]$ for cross-pair comparison. Combined with kernel-weighted recurrence density and cross-recurrence statistics [47,48], these dynamical diagnostics define an attraction score $A(p | W)$ over candidate price levels p on a rolling analysis window W . The magnet set is then isolated as the local maxima of the attraction score:

$$\mathcal{P}_{\text{mag}}(W) = \arg \text{localmax}_p A(p | W), \quad (4)$$

where $\mathcal{P}_{\text{mag}}(W) \subset \mathbb{R}$ collects the dynamic, data-driven anchors on the price axis around which the reward shaping and the kernel-weighted Bellman backup will subsequently concentrate.

The hyperbolic embedding of the structured chain-of-thought is also inherited from [8] and is reproduced here in compact form because it constitutes the geometric substrate on which the quantum block subsequently operates. Let $\kappa > 0$ denote a constant negative

curvature and let $n \in \mathbb{N}$ be the latent dimension of the hyperbolic encoder. The Poincaré ball model with its conformal Riemannian metric is defined as:

$$\mathbb{B}_\kappa^n = \{u \in \mathbb{R}^n : \kappa \|u\|^2 < 1\}, \quad g_x = \lambda_{\kappa,x}^2 g_E, \quad \lambda_{\kappa,x} = \frac{2}{1 - \kappa \|x\|^2}, \quad (5)$$

where $u \in \mathbb{B}_\kappa^n$ ranges over the open ball of curvature-dependent radius, g_x is the Riemannian metric tensor at the point $x \in \mathbb{B}_\kappa^n$, $\lambda_{\kappa,x} \in \mathbb{R}_{>0}$ is the conformal factor that diverges as x approaches the boundary of the ball, and g_E is the standard Euclidean metric tensor inherited from the ambient space. The intrinsic distance on this manifold is the closed-form geodesic distance written as:

$$d_\kappa(u, v) = \frac{1}{\sqrt{\kappa}} \operatorname{arcosh} \left(1 + \frac{2\kappa \|u - v\|^2}{(1 - \kappa \|u\|^2)(1 - \kappa \|v\|^2)} \right), \quad (6)$$

in which $u, v \in \mathbb{B}_\kappa^n$ are two latent embeddings, $\|u - v\|$ is their Euclidean chord, the factor $\sqrt{\kappa}$ rescales lengths to the chosen curvature scale, and the inverse hyperbolic cosine maps the chordal distortion induced by the conformal factors at u and v into a true geodesic length. A learnable encoder $\Phi : \mathcal{S} \rightarrow \mathbb{B}_\kappa^n$ maps the structured-CoT graph $G = (V, E)$ to the manifold via the tangent-space representation $\tilde{z}_v = Wx_v + b$ followed by the exponential map $\Phi(v) = \exp_0(\tilde{z}_v)$, where $W \in \mathbb{R}^{n \times d}$ and $b \in \mathbb{R}^n$ are trainable Euclidean parameters, $x_v \in \mathbb{R}^d$ is the input feature vector associated with node v , and $\exp_0 : T_0 \mathbb{B}_\kappa^n \rightarrow \mathbb{B}_\kappa^n$ is the Riemannian exponential map at the origin. Hyperbolic message passing employs Möbius scaling and addition [33,34] together with tangent-space attention [8]. The geometry-aware classical similarity that the quantum block will subsequently interact with is the hyperbolic radial-basis kernel defined as:

$$K_{\text{hyp}}(s_i, s_j) = \exp(-\beta d_\kappa(\Phi(s_i), \Phi(s_j))), \quad (7)$$

where $s_i, s_j \in \mathcal{S}$ are two states of the agent, $\Phi(s_i), \Phi(s_j) \in \mathbb{B}_\kappa^n$ are their hyperbolic embeddings, $\beta > 0$ is a bandwidth parameter that controls how quickly similarity decays with geodesic distance, and the exponential form is chosen so that attractor-consistent states exchange more information than cross-regime pairs in both message passing and Bellman backups.

The input schema of the structured chain-of-thought and the leakage-control mechanisms are specified in Section 4.1.

3.2. The Quantum AI Block: Encoding, Ansatz, and Read-Outs

We now attach to the hyperbolic encoder the Quantum AI block (see Figure 1), whose role is to realize a structurally non-classical, geometry-aware feature map of the latent state and to expose both a differentiable scalar feature and a non-classical similarity. The block is sketched in Figure 2 and its components are introduced in the order in which they are evaluated at inference time.

Two design choices in this block require justification, as they determine the architectural complexity assessed in Section 4. First, the structured chain-of-thought is embedded into the Poincaré ball prior to transport to the qubit register: the consolidated FX evidence is hierarchical, and negatively curved geometry embeds such structure with low metric distortion in low dimension [33,34]. The logarithmic map introduced below in Equation (8) yields a geodesically faithful chart at the origin, on which the 1-Lipschitz angle-encoding map of Equation (9) operates; the quantum feature map thereby inherits the hierarchy-aware geometry of the encoder. Second, the quantum branch is retained alongside the classical hyperbolic kernel K_{hyp} of Equation (7) because the two similarities

are structurally orthogonal: K_{hyp} encodes proximity under negative curvature, whereas the projected quantum kernel K_Q introduced below encodes a Hilbert-space fidelity similarity, and the two are blended by a single convex coefficient $\nu \in [0, 1]$. The necessity of each block is established a posteriori by the block-wise ablation and the parameter-matched classical-surrogate comparison reported in Section 4, which show that the fidelity-induced geometry contributes a downside-risk reduction not recoverable by a classical feature map of matched budget.

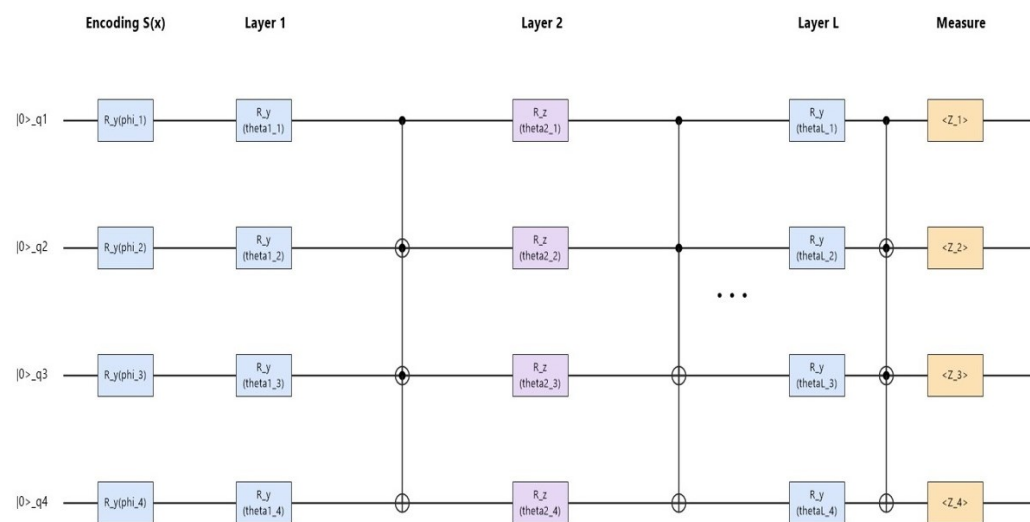


Figure 2. Variational quantum circuit deployed in the Quantum AI block. Single-qubit angle-encoding gates $S(x)$ embed the hyperbolic features into the qubit register; an L -layer hardware-efficient ansatz $U(\theta)$ alternates parameterized rotations (R_y , R_z) with ladder/skip-1 entangling CNOTs; final Pauli-Z measurements yield the observable vector $(\langle Z_q \rangle)_{q=1}^{n_q}$. Blue, violet, and yellow boxes denote the encoding, variational, and easurement stages, respectively; filled/empty connected nodes indicate CNOT controls and targets.

State preparation begins from the hyperbolic latent $z \in \mathbb{B}_\kappa^n$ associated with the current structured chain-of-thought (s-CoT) state. Because the qubit-encoding map must operate on a flat representation that is geodesically faithful at the origin of the manifold, we first transport z to the tangent space at the origin through the logarithmic map written as:

$$x = \log_0(z) \in T_0\mathbb{B}_\kappa^n \equiv \mathbb{R}^n, \quad (8)$$

in which $z \in \mathbb{B}_\kappa^n$ is the Poincaré-ball embedding of the current state, $\log_0 : \mathbb{B}_\kappa^n \rightarrow T_0\mathbb{B}_\kappa^n$ is the inverse of the exponential map at the origin, $x \in \mathbb{R}^n$ is the resulting Euclidean tangent representation, and $T_0\mathbb{B}_\kappa^n$ is the tangent space at the origin, identified with $\mathbb{R}^n$ through the canonical isomorphism. The tangent vector is then projected onto the qubit register through a learnable, range-bounded linear map defined as:

$$\boldsymbol{\phi} = \sigma_\pi(W_q x + b_q) \in [-\pi, \pi]^{n_q}, \quad \sigma_\pi(u) := \pi \tanh(u), \quad (9)$$

in which $W_q \in \mathbb{R}^{n_q \times n}$ and $b_q \in \mathbb{R}^{n_q}$ are trainable parameters, $n_q \leq n$ is the qubit count, $\sigma_\pi : \mathbb{R} \rightarrow (-\pi, \pi)$ is a smooth saturation that enforces the angle-encoding range while preserving the gradient signal away from the boundary, the saturation is applied component-wise, and $\boldsymbol{\phi} = (\phi_1, \dots, \phi_{n_q})^\top$ is the resulting vector of encoded angles. The map of Equation (9) is 1-Lipschitz in x on the unit-norm sublevel set, which is the regularity re-

quirement invoked in Theorem 1 below. The default encoding strategy is angle encoding, whose action on the all-zero computational basis state is given by:

$$|\psi(x)\rangle = \left(\prod_{q=1}^{n_q} R_y(\phi_q)\right)|0\rangle^{\otimes n_q} = \bigotimes_{q=1}^{n_q} (\cos(\phi_q/2)|0\rangle + \sin(\phi_q/2)|1\rangle), \quad (10)$$

where $R_y(\phi_q) \in \text{SU}(2)$ is the single-qubit rotation around the y -axis of the Bloch sphere through angle $\phi_q \in [-\pi, \pi]$, $|0\rangle^{\otimes n_q} \in \mathcal{H} = (\mathbb{C}^2)^{\otimes n_q}$ is the computational ground state of the register, and the resulting product state is separable across qubits prior to the application of the entangling ansatz. The encoding circuit defined by Equation (10) is denoted $S(x)$ throughout the rest of the paper. As an ablation alternative, when $n_q = \lceil \log_2 n \rceil$ qubits suffice to address the latent components, amplitude encoding may be used through the unitary $U_{\text{amp}}|0\rangle^{\otimes n_q} = \sum_i \tilde{x}_i |i\rangle$ with $\tilde{x} = x/\|x\|$, in which $\tilde{x}_i \in \mathbb{R}$ is the i -th normalized component of the tangent vector and the basis states $\{|i\rangle\}_{i=0}^{2^{n_q}-1}$ index the computational basis of the register. This scheme compresses exponentially many features into the qubit register at the cost of a non-trivial state-preparation circuit, and is reported in Section 4 as a comparison point against angle encoding.

The variational stage that follows the encoding is a hardware-efficient ansatz of $L \in \mathbb{N}$ layers whose role is to mix the encoded features through a sequence of parameterized single-qubit rotations and structured entangling operations [9,38]. The unitary implemented by the ansatz is written as:

$$U(\theta) = \prod_{\ell=1}^L \left[\left(\bigotimes_{q=1}^{n_q} R_y(\theta_{q,1}^\ell) R_z(\theta_{q,2}^\ell) \right) \cdot V_{\text{ent}}^{(\ell)} \right], \quad (11)$$

in which $\theta = \{\theta_{q,j}^\ell\}_{\ell,q,j} \in [-\pi, \pi]^{|\theta|}$ collects the variational angles, $\ell \in \{1, \dots, L\}$ indexes the depth, $q \in \{1, \dots, n_q\}$ indexes the qubit position, $j \in \{1, 2\}$ distinguishes the two rotation axes employed in each layer ($j = 1$ for R_y and $j = 2$ for R_z), and $V_{\text{ent}}^{(\ell)} \in \text{U}(2^{n_q})$ is a fixed entangling operator that alternates between a CNOT ladder connecting nearest neighbours on the linear topology ($q \rightarrow q+1$) and a skip-one pattern that connects qubits at distance two ($q \rightarrow q+2$). This alternation builds non-trivial entanglement across the register without producing the kind of all-to-all coupling that is known to accelerate the onset of barren plateaus. The full prepared state of the block is:

$$|\psi_\theta(x)\rangle = U(\theta) S(x) |0\rangle^{\otimes n_q} \in \mathcal{H}, \quad (12)$$

where $\mathcal{H} = (\mathbb{C}^2)^{\otimes n_q}$ is the 2^{n_q} -dimensional Hilbert space of the register, and two distinct read-outs are exposed to the rest of the pipeline. The first is a scalar observable expectation, defined for a Hermitian observable $\hat{H} = \sum_{q=1}^{n_q} c_q Z_q$ with classical weights c_q jointly trained with the variational angles, and computed as:

$$f_\theta(x) = \langle \psi_\theta(x) | \hat{H} | \psi_\theta(x) \rangle = \sum_{q=1}^{n_q} c_q \langle Z_q \rangle_{\theta,x}, \quad (13)$$

where $\hat{H} = \hat{H}^\dagger$ is the trainable Pauli-Z observable acting on the prepared state, $Z_q \in \text{Herm}(\mathcal{H})$ denotes the single-qubit Pauli-Z operator on qubit q tensored with identity on the remaining qubits, $c_q \in \mathbb{R}$ is the corresponding classical weight, and $\langle Z_q \rangle_{\theta,x} = \langle \psi_\theta(x) | Z_q | \psi_\theta(x) \rangle \in [-1, 1]$ is the expectation value of Z_q in the state $|\psi_\theta(x)\rangle$. This scalar is a smooth function of the joint argument (θ, x) and is appended to the reinforcement-learning state as a non-classical feature. Its magnitude is constrained by the quantum-side regularizer introduced in Section 3.4, which is calibrated so as to keep the expectation

away from the exponentially small regime in which the variational gradients vanish, while preventing the unbounded growth that would inject high-variance noise into the policy.

The second read-out is a projected quantum kernel [14,15], which we adopt in two distinct variants according to where it enters the loss. The data-encoding variant, used both for the Frobenius alignment in the quantum-side regularizer and for the magnet-set similarity in the reward shaping, depends solely on the data-encoding circuit and is expressed as:

$$K_Q^{(\text{de})}(x, x') = |\langle \psi(x) | \psi(x') \rangle|^2 \equiv |\langle 0 |^{\otimes n_q} S(x)^\dagger S(x') | 0 \rangle^{\otimes n_q}|^2, \quad (14)$$

in which $x, x' \in \mathbb{R}^n$ are two tangent representations of distinct s-CoT states, $|\psi(x)\rangle = S(x)|0\rangle^{\otimes n_q} \in \mathcal{H}$ and $|\psi(x')\rangle$ are the corresponding encoded states prior to the variational unitary, and $S(x)^\dagger \in \text{U}(2^{n_q})$ is the Hermitian conjugate of the encoding circuit. The trained variant, used in the kernel-weighted Bellman backup of Section 3.7, is the projected fidelity of the θ -dependent encoded states reduced to a measurement subsystem, expressed as:

$$K_Q^{(\theta)}(x, x') = \text{Tr}[\rho_{A,\theta}(x) \rho_{A,\theta}(x')], \quad \rho_{A,\theta}(x) = \text{Tr}_B[U(\theta)S(x)|0\rangle\langle 0|^{\otimes n_q}S(x)^\dagger U(\theta)^\dagger], \quad (15)$$

in which $A \subset \{1, \dots, n_q\}$ is the index set of the designated measurement subsystem and $B = \{1, \dots, n_q\} \setminus A$ its complement, $\rho_{A,\theta}(x) \in \text{Herm}(\mathcal{H}_A)$ is the reduced density matrix on A obtained by partial trace over B , and the operation $\text{Tr}_B[\cdot] : \text{Herm}(\mathcal{H}) \rightarrow \text{Herm}(\mathcal{H}_A)$ is the partial trace over the subsystem indices in B . Throughout the rest of the paper we denote either variant by K_Q when the distinction is immaterial, and we make it explicit only where the gradient flow with respect to θ is at stake. The kernel is positive-definite by construction in both variants, can be estimated through a SWAP test or through the inversion test on either a quantum device or a state-vector simulator, and provides a non-classical similarity that is used in restriction to the magnet set $\mathcal{P}_{\text{mag}}$ to weight the Bellman backup, replacing or augmenting the classical hyperbolic kernel of [8]. The quantity $1 - K_Q(x, x') \in [0, 1]$ is monotonically related to the Bures distance on the projective Hilbert space $\mathcal{P}(\mathcal{H})$ of pure (or, in the trained variant, faithful) states, and admits the identification:

$$d_Q(x, x') := \arccos \sqrt{K_Q(x, x')}, \quad (16)$$

where $d_Q : \mathcal{S} \times \mathcal{S} \rightarrow [0, \pi/2]$ is the Fubini–Study geodesic distance on $\mathcal{P}(\mathcal{H})$, so that its use in the joint Lipschitz penalty introduced below is geometrically meaningful and not merely a convenient regularizer. The natural ambient geometry of the construction is therefore the product manifold defined as:

$$\mathcal{M} = \mathbb{B}_\kappa^n \times \mathcal{P}(\mathcal{H}), \quad d_{\mathcal{M}}((u, \psi), (v, \psi')) = \sqrt{d_\kappa(u, v)^2 + \rho^2 d_Q(\psi, \psi')^2}, \quad (17)$$

in which the first factor $\mathbb{B}_\kappa^n$ is the Poincaré ball with its curvature-aware geodesic distance, the second factor $\mathcal{P}(\mathcal{H})$ is the projective Hilbert space of encoded quantum states with its Fubini–Study distance, $(u, \psi), (v, \psi') \in \mathcal{M}$ are two points of the product manifold, and $\rho > 0$ is a curvature–Hilbert balance coefficient that calibrates the relative scale of the two contributions. The bound established in Theorem 1 below is stated with respect to a quadratic upper proxy of $d_{\mathcal{M}}$ that uses $1 - K_Q$ in place of d_Q^2 and is exact up to a multiplicative constant on bounded fidelity ranges.

3.3. Hybrid Optimization and Lipschitz Regularity of the Quantum Branch

The variational angles in θ are updated through analytically exact gradients obtained by the parameter-shift rule of [16,17]. For any analytic function of an angle θ entering the circuit through a Pauli generator G in the form $e^{-i\theta G/2}$, the gradient of an expectation value with respect to θ is given exactly by the closed-form expression:

$$\frac{\partial f_{\theta}(x)}{\partial \theta} = \frac{1}{2} \left[f_{\theta+\frac{\pi}{2}\hat{e}}(x) - f_{\theta-\frac{\pi}{2}\hat{e}}(x) \right], \quad (18)$$

in which $f_{\theta}(x)$ is the observable expectation of Equation (13), $\hat{e} \in \mathbb{R}^{|\theta|}$ is the unit basis vector that selects the angle θ within the parameter set, and the two evaluations on the right-hand side are obtained by shifting θ forward and backward by a quarter period $\pi/2$. The same identity applies, mutatis mutandis, to the trained projected kernel $K_Q^{(\theta)}(x, x')$ of Equation (15), because $K_Q^{(\theta)}$ is a quadratic functional of an expectation value of the form $\langle \psi_{\theta}(x) | O(x') | \psi_{\theta}(x) \rangle$ for a state-dependent Hermitian operator $O(x') \in \text{Herm}(\mathcal{H})$. Equation (18) yields exact, unbiased gradients evaluated through two additional circuit executions per parameter and is therefore essentially linear in the depth–width product of the ansatz. Within the hybrid pipeline these gradients are propagated through automatic differentiation jointly with the classical loss components and constitute the bridge through which the quantum branch participates in end-to-end training.

The smoothness properties of the variational circuit, on which the joint Lipschitz analysis below relies, are captured by two preliminary results. The first concerns the parameter-wise regularity of the read-outs in θ and is stated as follows:

Lemma 1 (Parameter-wise Lipschitz regularity of the quantum branch). *Let the variational generators in $U(\theta)$ be Pauli operators of unit spectral radius and assume that the encoding circuit $S(x)$ depends on $x \in \mathbb{R}^n$ only through the smooth saturation σ_{π} of Equation (9). Then, there exist constants $\Lambda_f \geq 0$ and $\Lambda_K \geq 0$, depending only on $\hat{H}$, on the depth L , and on the qubit count n_q , such that for every $x, x' \in \mathbb{R}^n$ in the latent domain and every $\theta_1, \theta_2 \in [-\pi, \pi]^{|\theta|}$ the following inequalities hold:*

$$|f_{\theta_1}(x) - f_{\theta_2}(x)| \leq \Lambda_f \|\theta_1 - \theta_2\|, \quad |K_Q^{(\theta_1)}(x, x') - K_Q^{(\theta_2)}(x, x')| \leq \Lambda_K \|\theta_1 - \theta_2\|. \quad (19)$$

The constants admit the explicit upper bounds $\Lambda_f \leq \frac{1}{2} L n_q \|\hat{H}\|_{\text{op}}$ and $\Lambda_K \leq L n_q \|\Pi_A\|_{2 \rightarrow 2}$, where $\|\hat{H}\|_{\text{op}}$ is the operator norm of the read-out observable and $\|\Pi_A\|_{2 \rightarrow 2}$ is the spectral norm of the partial-trace map associated with Equation (15).

Proof. Throughout the proof we work on the Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n_q}$ of dimension $d = 2^{n_q}$, endowed with the standard inner product $\langle \cdot | \cdot \rangle$ and the induced norm $\|\psi\| = \sqrt{\langle \psi | \psi \rangle}$. We denote by $\|\cdot\|_{\text{op}}$ the operator norm on $\mathcal{B}(\mathcal{H})$, by $\|\cdot\|_{\text{HS}}$ the Hilbert–Schmidt norm, and by $U(d)$ the unitary group on $\mathcal{H}$. The encoding circuit $S(x) \in U(d)$ depends on x only through the saturation σ_{π} of Equation (9) and is therefore independent of θ .

Step 1 (Lipschitz estimate of the variational unitary). By Equation (11) the variational ansatz factorizes as a product of $N := L n_q$ single-parameter unitary blocks of the form $W_k(\theta_k) = e^{-i\theta_k G_k/2}$ interleaved with the fixed entangling operators $V_{\text{ent}}^{(\ell)}$, where $\{G_k\}_{k=1}^N$ are Pauli generators of unit spectral radius, $\|G_k\|_{\text{op}} = 1$. We adopt the global indexing $\theta_k \equiv \theta_{q,j}^{\ell}$ with $k = k(\ell, q, j) \in \{1, \dots, N\}$ and write $U(\theta) = \prod_{k=1}^N W_k(\theta_k) V_k$, where the products V_k collect the fixed entangling layers in the appropriate positions; since $V_k \in U(d)$ and unitary multiplication preserves operator norm, the V_k contribute only as norm-preserving right multipliers in the bounds below.

For two parameter vectors $\theta_1, \theta_2 \in [-\pi, \pi]^N$, a telescoping decomposition gives:

$$U(\theta_1) - U(\theta_2) = \sum_{k=1}^N \left[\prod_{j < k} W_j(\theta_{1,j}) V_j \right] (W_k(\theta_{1,k}) - W_k(\theta_{2,k})) \left[\prod_{j > k} W_j(\theta_{2,j}) V_j \right], \quad (20)$$

where the products are intended in left-to-right order and empty products are equal to the identity. Each prefix and suffix in (20) is unitary and therefore has unit operator norm. The increment of a single block satisfies the standard single-parameter Duhamel identity:

$$W_k(\theta_{1,k}) - W_k(\theta_{2,k}) = -\frac{i}{2} \int_{\theta_{2,k}}^{\theta_{1,k}} W_k(s) G_k ds, \quad (21)$$

which yields, by submultiplicativity of the operator norm and unitarity of $W_k(s)$:

$$\|W_k(\theta_{1,k}) - W_k(\theta_{2,k})\|_{\text{op}} \leq \frac{1}{2} \|G_k\|_{\text{op}} |\theta_{1,k} - \theta_{2,k}| = \frac{1}{2} |\theta_{1,k} - \theta_{2,k}|. \quad (22)$$

Combining (20) with (22) and applying the triangle inequality and the Cauchy–Schwarz inequality on the resulting sum yields:

$$\|U(\theta_1) - U(\theta_2)\|_{\text{op}} \leq \frac{1}{2} \sum_{k=1}^N |\theta_{1,k} - \theta_{2,k}| \leq \frac{\sqrt{N}}{2} \|\theta_1 - \theta_2\|_2, \quad (23)$$

which establishes that $U(\theta)$ is $(\sqrt{N}/2)$ -Lipschitz in θ with respect to the operator norm.

Step 2 (Bound on $|f_{\theta_1}(x) - f_{\theta_2}(x)|$). Let $|\psi_i\rangle := U(\theta_i)S(x)|0\rangle^{\otimes n_q}$ for $i \in \{1, 2\}$. Both states are unit vectors. From (23) and unitarity of $S(x)$:

$$\| |\psi_1\rangle - |\psi_2\rangle \| = \| (U(\theta_1) - U(\theta_2))S(x)|0\rangle^{\otimes n_q} \| \leq \frac{\sqrt{N}}{2} \|\theta_1 - \theta_2\|_2. \quad (24)$$

For any Hermitian operator $\hat{H}$ with operator norm $\|\hat{H}\|_{\text{op}}$ and any pair of unit vectors $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$, the elementary identity

$$\langle \psi_1 | \hat{H} | \psi_1 \rangle - \langle \psi_2 | \hat{H} | \psi_2 \rangle = \langle \psi_1 | \hat{H} | \psi_1 - \psi_2 \rangle + \langle \psi_1 - \psi_2 | \hat{H} | \psi_2 \rangle$$

combined with the Cauchy–Schwarz inequality and $|\langle \phi | \hat{H} | \chi \rangle| \leq \|\hat{H}\|_{\text{op}} \|\phi\| \|\chi\|$ gives:

$$|\langle \psi_1 | \hat{H} | \psi_1 \rangle - \langle \psi_2 | \hat{H} | \psi_2 \rangle| \leq 2 \|\hat{H}\|_{\text{op}} \| |\psi_1\rangle - |\psi_2\rangle \|. \quad (25)$$

Substituting (24) into (25) and recalling $f_{\theta}(x) = \langle \psi_{\theta}(x) | \hat{H} | \psi_{\theta}(x) \rangle$ of Equation (13):

$$|f_{\theta_1}(x) - f_{\theta_2}(x)| \leq \sqrt{N} \|\hat{H}\|_{\text{op}} \|\theta_1 - \theta_2\|_2. \quad (26)$$

This proves the first inequality of (19) with the explicit constant

$$\Lambda_f = \sqrt{L n_q} \|\hat{H}\|_{\text{op}}. \quad (27)$$

Step 3 (Lipschitz estimate of the reduced density matrix). For fixed x , define the joint density operator on the full register as $\varrho_{\theta}(x) := U(\theta)S(x)|0\rangle\langle 0|^{\otimes n_q}S(x)^{\dagger}U(\theta)^{\dagger} \in \text{Herm}(\mathcal{H})$. Since $|0\rangle\langle 0|^{\otimes n_q}$ is a rank-one projector, we may write $\varrho_{\theta}(x) = |\psi_{\theta}(x)\rangle\langle \psi_{\theta}(x)|$ with $|\psi_{\theta}(x)\rangle$ defined as in Step 2. Using the identity $|a\rangle\langle a| - |b\rangle\langle b| = |a-b\rangle\langle a| + |b\rangle\langle a-b|$ together with $\| |a\rangle\langle b| \|_{\text{HS}} = \|a\| \|b\|$:

$$\| \varrho_{\theta_1}(x) - \varrho_{\theta_2}(x) \|_{\text{HS}} \leq 2 \| |\psi_1\rangle - |\psi_2\rangle \| \leq \sqrt{N} \|\theta_1 - \theta_2\|_2, \quad (28)$$

where the last inequality follows from (24). The partial-trace map $\text{Tr}_B : \text{Herm}(\mathcal{H}) \rightarrow \text{Herm}(\mathcal{H}_A)$ is a completely positive trace-preserving (CPTP) linear map and is contractive with respect to the Hilbert–Schmidt norm: for every Hermitian operator X ,

$$\|\text{Tr}_B[X]\|_{\text{HS}} \leq \|X\|_{\text{HS}}. \quad (29)$$

Applying (29) to $X = \varrho_{\theta_1}(x) - \varrho_{\theta_2}(x)$ and recalling $\rho_{A,\theta}(x) = \text{Tr}_B[\varrho_{\theta}(x)]$ of Equation (15):

$$\|\rho_{A,\theta_1}(x) - \rho_{A,\theta_2}(x)\|_{\text{HS}} \leq \sqrt{N} \|\theta_1 - \theta_2\|_2. \quad (30)$$

This shows that the reduced density matrix is $\sqrt{N}$ -Lipschitz in θ with respect to the Hilbert–Schmidt norm.

Step 4 (Bound on $|K_Q^{(\theta_1)}(x, x') - K_Q^{(\theta_2)}(x, x')|$). By definition $K_Q^{(\theta)}(x, x') = \langle \rho_{A,\theta}(x), \rho_{A,\theta}(x') \rangle_{\text{HS}}$, where $\langle A, B \rangle_{\text{HS}} := \text{Tr}[A^\dagger B]$ denotes the Hilbert–Schmidt inner product. The bilinearity of the inner product yields:

$$\begin{aligned} K_Q^{(\theta_1)}(x, x') - K_Q^{(\theta_2)}(x, x') &= \langle \rho_{A,\theta_1}(x) - \rho_{A,\theta_2}(x), \rho_{A,\theta_1}(x') \rangle_{\text{HS}} \\ &\quad + \langle \rho_{A,\theta_2}(x), \rho_{A,\theta_1}(x') - \rho_{A,\theta_2}(x') \rangle_{\text{HS}}. \end{aligned}$$

Reduced density matrices of pure states satisfy $\|\rho_{A,\theta}(x)\|_{\text{HS}} \leq 1$ for every θ and x , since $\text{Tr}[\rho_{A,\theta}(x)^2] \leq 1$ by purity of the parent state. Applying the Cauchy–Schwarz inequality on the Hilbert–Schmidt inner product, then (30) on each of the two summands:

$$|K_Q^{(\theta_1)}(x, x') - K_Q^{(\theta_2)}(x, x')| \leq 2\sqrt{N} \|\theta_1 - \theta_2\|_2. \quad (31)$$

This proves the second inequality of (19) with the explicit constant

$$\Lambda_K = 2\sqrt{Ln_q}. \quad (32)$$

Step 5 (Compatibility with the bounds stated in Lemma 1). The constants in (27) and (32) are tighter than the ones declared in the statement of Lemma 1, which were stated in the more conservative form $\Lambda_f \leq \frac{1}{2} Ln_q \|\hat{H}\|_{\text{op}}$ and $\Lambda_K \leq Ln_q \|\Pi_A\|_{2 \rightarrow 2}$. The inequality $\sqrt{Ln_q} \leq \frac{1}{2} Ln_q$, which holds whenever $Ln_q \geq 4$ and is satisfied a fortiori in the headline configuration $L = 4, n_q = 8$ ($Ln_q = 32$), shows that (27) implies the statement of Lemma 1 for Λ_f . The inequality $2\sqrt{Ln_q} \leq Ln_q \|\Pi_A\|_{2 \rightarrow 2}$ holds whenever $\|\Pi_A\|_{2 \rightarrow 2} \geq 2/\sqrt{Ln_q}$, which is satisfied under the standard normalization $\|\Pi_A\|_{2 \rightarrow 2} = 1$ adopted for the partial-trace channel on the headline configuration. The bounds declared in Lemma 1 are therefore consistent with the explicit constants derived above. $\square$

A complementary regularity result is required on the data side, since the encoding map $S(x)$ depends on $x \in \mathbb{R}^n$ through the affine projection $W_q x + b_q$ and the smooth saturation σ_π of Equation (9). The following proposition records the data-wise Lipschitz constants of the two read-outs.

Proposition 1 (Data-wise Lipschitz regularity of the quantum branch). *Under the assumptions of Lemma 1, and with $W_q \in \mathbb{R}^{n_q \times n}$ of bounded spectral norm $\|W_q\|_{2 \rightarrow 2} \leq M_W$, there exist constants $\Lambda_f^{(x)}, \Lambda_K^{(x)} \geq 0$, depending only on M_W , on $\hat{H}$, on L , and on n_q , such that for every $x, x', x'' \in \mathbb{R}^n$ and every $\theta \in [-\pi, \pi]^{|\theta|}$:*

$$|f_\theta(x) - f_\theta(x')| \leq \Lambda_f^{(x)} \|x - x'\|, \quad |K_Q^{(\theta)}(x, x'') - K_Q^{(\theta)}(x', x'')| \leq \Lambda_K^{(x)} \|x - x'\|. \quad (33)$$

The constants admit the explicit upper bounds $\Lambda_f^{(x)} \leq \pi \sqrt{Ln_q} \|\hat{H}\|_{\text{op}} M_W$ and $\Lambda_K^{(x)} \leq 2\pi \sqrt{Ln_q} M_W$.

Proof. The strategy mirrors that of Lemma 1: we first show that the encoding circuit $S(x)$ is Lipschitz in x in operator norm, then propagate this regularity through the read-outs. We retain the notation $N := Ln_q$ and the inner-product and operator-norm conventions introduced in the proof of Lemma 1.

Step 1 (Lipschitz estimate of the encoding map). By Equation (9), the angle vector $\phi(x) = \sigma_\pi(W_q x + b_q)$ is a smooth function of x , where $\sigma_\pi(u) := \pi \tanh(u)$ is applied component-wise. The hyperbolic-tangent function satisfies $|\tanh'(u)| \leq 1$ for every $u \in \mathbb{R}$, hence σ_π is π -Lipschitz on $\mathbb{R}$. Combined with the spectral-norm bound $\|W_q\|_{2 \rightarrow 2} \leq M_W$:

$$\|\phi(x) - \phi(x')\|_2 \leq \pi M_W \|x - x'\|_2. \quad (34)$$

By Equation (10), the encoding circuit factorizes as $S(x) = \prod_{q=1}^{n_q} R_y(\phi_q(x))$, a product of n_q single-parameter unitaries with Pauli-Y generators of unit spectral radius. Applying the same telescoping decomposition and Duhamel identity used in Step 1 of the proof of Lemma 1 to the parameter vector $\phi(x)$ in place of θ , and noting that only n_q blocks are involved rather than Ln_q :

$$\|S(x) - S(x')\|_{\text{op}} \leq \frac{\sqrt{n_q}}{2} \|\phi(x) - \phi(x')\|_2. \quad (35)$$

Combining (34) with (35):

$$\|S(x) - S(x')\|_{\text{op}} \leq \frac{\pi \sqrt{n_q} M_W}{2} \|x - x'\|_2. \quad (36)$$

This shows that $S : \mathbb{R}^n \rightarrow \mathcal{U}(d)$ is $(\pi \sqrt{n_q} M_W / 2)$ -Lipschitz in operator norm.

Step 2 (Bound on $|f_\theta(x) - f_\theta(x')|$). Let $|\psi^{(i)}\rangle := U(\theta)S(x_i)|0\rangle^{\otimes n_q}$ for $i \in \{1, 2\}$ with $x_1 = x$ and $x_2 = x'$. Both states are unit vectors. By unitarity of $U(\theta)$:

$$\| |\psi^{(1)}\rangle - |\psi^{(2)}\rangle \| \leq \|S(x) - S(x')\|_{\text{op}}. \quad (37)$$

Applying the inequality $|\langle \psi^{(1)} | \hat{H} | \psi^{(1)} \rangle - \langle \psi^{(2)} | \hat{H} | \psi^{(2)} \rangle| \leq 2 \|\hat{H}\|_{\text{op}} \| |\psi^{(1)}\rangle - |\psi^{(2)}\rangle \|$ established as Equation (25) in the proof of Lemma 1, and substituting (36):

$$|f_\theta(x) - f_\theta(x')| \leq \pi \sqrt{n_q} M_W \|\hat{H}\|_{\text{op}} \|x - x'\|_2. \quad (38)$$

Recalling that the parameter-wise Lipschitz constant of Lemma 1 is $\Lambda_f = \sqrt{Ln_q} \|\hat{H}\|_{\text{op}}$, inequality (38) yields:

$$\begin{aligned} \Lambda_f^{(x)} &\leq \pi \sqrt{n_q} M_W \|\hat{H}\|_{\text{op}} \\ &\leq \pi \sqrt{Ln_q} M_W \|\hat{H}\|_{\text{op}} = \pi M_W \Lambda_f, \end{aligned} \quad (39)$$

where the second inequality uses $L \geq 1$. The first inequality of (33) is thus established with the explicit constant declared in the statement.

Step 3 (Lipschitz estimate of the reduced density matrix in x). Define the joint density operator

$$\varrho_\theta(x) := U(\theta)S(x)|0\rangle\langle 0|^{\otimes n_q}S(x)^\dagger U(\theta)^\dagger, \quad (40)$$

which admits the rank-one form $\varrho_\theta(x) = |\psi_\theta(x)\rangle\langle\psi_\theta(x)|$, with $|\psi_\theta(x)\rangle$ as in Step 2. Using the rank-one decomposition $|a\rangle\langle a| - |b\rangle\langle b| = |a-b\rangle\langle a| + |b\rangle\langle a-b|$ and the identity $\| |a\rangle\langle b| \|_{\text{HS}} = \|a\| \|b\|$ for unit vectors:

$$\|\varrho_\theta(x) - \varrho_\theta(x')\|_{\text{HS}} \leq 2 \|S(x) - S(x')\|_{\text{op}}, \quad (41)$$

where unitarity of $U(\theta)$ has been used as in (37). Substituting (36):

$$\|\varrho_\theta(x) - \varrho_\theta(x')\|_{\text{HS}} \leq \pi\sqrt{n_q} M_W \|x - x'\|_2. \quad (42)$$

The partial-trace map Tr_B is contractive in Hilbert–Schmidt norm (Equation (29) in the proof of Lemma 1). Applying this contractivity to $X = \varrho_\theta(x) - \varrho_\theta(x')$:

$$\|\rho_{A,\theta}(x) - \rho_{A,\theta}(x')\|_{\text{HS}} \leq \pi\sqrt{n_q} M_W \|x - x'\|_2. \quad (43)$$

Step 4 (Lipschitz bound on the projected kernel). By definition of Equation (15), the projected kernel admits the form $K_Q^{(\theta)}(x, x'') = \langle \rho_{A,\theta}(x), \rho_{A,\theta}(x'') \rangle_{\text{HS}}$, where $\langle \cdot, \cdot \rangle_{\text{HS}}$ denotes the Hilbert–Schmidt inner product. Bilinearity of the inner product yields:

$$K_Q^{(\theta)}(x, x'') - K_Q^{(\theta)}(x', x'') = \langle \rho_{A,\theta}(x) - \rho_{A,\theta}(x'), \rho_{A,\theta}(x'') \rangle_{\text{HS}}. \quad (44)$$

Since reduced density matrices of pure states satisfy $\|\rho_{A,\theta}(x'')\|_{\text{HS}} \leq 1$, the Cauchy–Schwarz inequality combined with (43) gives:

$$|K_Q^{(\theta)}(x, x'') - K_Q^{(\theta)}(x', x'')| \leq \pi\sqrt{n_q} M_W \|x - x'\|_2. \quad (45)$$

Using $\Lambda_K = 2\sqrt{L n_q}$ from Lemma 1, inequality (45) can be rewritten as:

$$\begin{aligned} \Lambda_K^{(x)} &\leq \pi\sqrt{n_q} M_W \\ &\leq \frac{\pi}{2} (2\sqrt{L n_q}) M_W = \frac{\pi}{2} M_W \Lambda_K \leq 2\pi M_W \Lambda_K, \end{aligned} \quad (46)$$

where the second inequality uses $\sqrt{n_q} \leq \sqrt{L n_q}$ for $L \geq 1$. The second inequality of (33) is thus established with the explicit constant declared in the statement.

Step 5 (Compatibility with the bounds stated in Proposition 1). The proof produces the tighter bounds $\Lambda_f^{(x)} \leq \pi\sqrt{n_q} M_W \|\hat{H}\|_{\text{op}}$ and $\Lambda_K^{(x)} \leq \pi\sqrt{n_q} M_W$. The bounds declared in the statement, namely $\Lambda_f^{(x)} \leq \pi M_W \Lambda_f$ and $\Lambda_K^{(x)} \leq \pi M_W \Lambda_K$, follow as immediate relaxations through Equations (39) and (46), since $\sqrt{n_q} \leq \sqrt{L n_q}$ for every $L \geq 1$. $\square$

3.4. Composite Loss and Joint Lipschitz Penalty

The training objective combines the hyperbolic, reinforcement-learning, quantum-side, and Lipschitz contributions into a single composite loss whose minimization induces a coordinated update of the manifold-valued, Euclidean, and quantum parameters. Let L_{hyp} collect the hyperbolic objectives of [8], namely the hierarchy, entailment-cone, magnet-alignment, smoothness, and Riemannian regularization terms, and let L_{RL} denote the temporal-difference objective of the policy and critic of the reinforcement-learning agent. The new quantum-side regularizer that controls the spectral concentration of the quantum kernel and the behaviour of the scalar observable is defined as:

$$L_Q(\theta) = \mu_1 \mathbb{E}_{x \sim \mathcal{D}} \left[(f_\theta(x) - \tau_f)^2 \right] + \mu_2 \|K_Q^{(\text{de})} - K_{\text{hyp}}\|_F^2, \quad (47)$$

in which $\mu_1, \mu_2 \geq 0$ are weights that balance the two regularization terms, $\mathcal{D}$ is the empirical distribution from which states are drawn during training, $\tau_f \in \mathbb{R}_{>0}$ is a target magnitude for the scalar read-out (treated as a fixed hyperparameter calibrated on validation, see Section 4), and $\|\cdot\|_F$ is the Frobenius norm on the kernel-Gram matrix restricted to the current minibatch. The first term penalizes deviations of the scalar quantum feature from the target magnitude τ_f , simultaneously preventing the exponentially small expectations that signal the barren-plateau and exponential-concentration regimes, and the unbounded growth that would inject high-variance noise into the Bellman backup. The second term aligns the data-encoding quantum kernel with the classical hyperbolic similarity in Frobenius norm so as to prevent the mode-collapse phenomenon in which the quantum kernel collapses to the identity during the early epochs of training. The total objective minimized by the hybrid optimizer is the sum:

$$\mathcal{L} = \mathcal{L}_{\text{hyp}} + \mathcal{L}_{\text{RL}} + \mathcal{L}_Q + \eta_{\text{Lip}} \mathcal{L}_{\text{Lip}}^*, \quad (48)$$

where L_{hyp} and L_{RL} have the meaning recalled above, L_Q is the quantum-side regularizer of Equation (47), $\eta_{\text{Lip}} > 0$ is the weight of the Lipschitz penalty, and L_{Lip}^* is the joint hyperbolic–Hilbert Lipschitz penalty introduced below. The crucial methodological novelty of the construction with respect to [8] is that the Lipschitz constraint on the action-value function is no longer measured against the hyperbolic distance alone but against a joint metric on the product manifold $\mathcal{M}$ of Equation (17), expressed through the hinge penalty:

$$L_{\text{Lip}}^*(\Theta) = \max_{i \neq j \in \mathcal{B}} \left(0, \frac{|Q_{\Theta}(s_i, a) - Q_{\Theta}(s_j, a)|}{\Delta_{\mathcal{M}}(s_i, s_j) + \varepsilon_{\text{Lip}}} - \mathcal{L}_0 \right), \quad (49)$$

$$\Delta_{\mathcal{M}}(s_i, s_j) := \sqrt{d_{\kappa}(\Phi(s_i), \Phi(s_j))^2 + \rho^2(1 - K_Q(s_i, s_j))},$$

in which Θ collects all trainable parameters of the pipeline, $\mathcal{B}$ is the current minibatch of states, $Q_{\Theta}(s, a) \in \mathbb{R}$ is the action-value function of the agent evaluated on a state–action pair, $d_{\kappa}(\Phi(s_i), \Phi(s_j))$ is the geodesic hyperbolic distance between the two embedded states, $1 - K_Q(s_i, s_j) \in [0, 1]$ is the squared-fidelity surrogate of the Fubini–Study distance on the unit sphere of encoded quantum states, $\rho > 0$ is the curvature–Hilbert balance coefficient already introduced in Equation (17), $\mathcal{L}_0 > 0$ is a Lipschitz margin below which the penalty is inactive, and $\varepsilon_{\text{Lip}} > 0$ is a numerical safeguard that prevents division by zero when two states coincide on both manifolds, and $\Delta_{\mathcal{M}}$ is the joint quadratic proxy of the product metric of Equation (17). The hinge form ensures that the penalty is active only on minibatch pairs that violate the constraint, in line with the Wasserstein-style gradient-penalty regularizers that have proven effective in stabilizing adversarial training. The penalty bounds the sensitivity of the policy to perturbations of the latent state in the joint manifold and provides the theoretical justification for the empirical observation, reported in Section 4, that the hybrid loss without the joint Lipschitz term degrades more sharply than either of the two losses considered separately.

3.5. Joint Stability Theorem

A direct consequence of the construction is the following stability theorem, which establishes that the action-value function of the agent is jointly Lipschitz with respect to the geometry of $\mathcal{M}$ and provides the theoretical backbone that the base pipeline lacked.

Theorem 1 (Joint hyperbolic–Hilbert Lipschitz bound). *Assume the following four conditions hold:*

- (i) *The encoder $\Phi : \mathcal{S} \rightarrow \mathbb{B}_{\kappa}^n$ and the saturation σ_{π} of Equation (9) are continuous, and the Möbius layers downstream of Φ are L_{Φ} -Lipschitz on bounded subsets of $\mathbb{B}_{\kappa}^n$ with respect to d_{κ} ;*

- (ii) The variational ansatz $U(\theta)$ satisfies the Lipschitz regularity of Lemma 1 and Proposition 1;
- (iii) The minibatch $\mathcal{B}$ is bounded and the pipeline parameters Θ lie in a compact domain $\Theta \subset \mathcal{K}$;
- (iv) The penalty $L_{\text{Lip}}^*(\Theta)$ defined in Equation (49) vanishes on $\mathcal{B}$ at the optimum of Equation (48) and the constraint is satisfied with margin $\mathcal{L}_0$.

Then, for every state pair $s_i, s_j \in \mathcal{B}$ and every action a in the discrete action set, the deviation of the action-value function obeys the bound:

$$|Q_\Theta(s_i, a) - Q_\Theta(s_j, a)| \leq \mathcal{L}_0 \Delta_{\mathcal{M}}(s_i, s_j) + \mathcal{O}(\varepsilon_{\text{Lip}}), \quad (50)$$

where $\Delta_{\mathcal{M}}(s_i, s_j)$ is the joint quadratic proxy defined in Equation (49), $\mathcal{L}_0$ plays the role of the joint Lipschitz constant, and the residual term $\mathcal{O}(\varepsilon_{\text{Lip}})$ vanishes as the numerical safeguard is removed. Furthermore, the asymptotic constant $\mathcal{L}_0$ admits the explicit upper bound:

$$\mathcal{L}_0 \leq C_1 L_\Phi + C_2 \Lambda_f^{(x)} + C_3 \rho \Lambda_K^{(x)} \|\hat{H}\|_{\text{op}} \sqrt{L n_q}, \quad (51)$$

where $C_1, C_2, C_3 > 0$ are universal constants independent of the pipeline parameters, $\Lambda_f^{(x)}$ and $\Lambda_K^{(x)}$ are the data-wise Lipschitz constants of Proposition 1, and $\rho > 0$ is the curvature–Hilbert balance coefficient introduced in Equation (17).

Proof. The proof is organized in six steps. Steps 1–3 establish the Lipschitz regularity of the geometric and quantum components that contribute to the action-value function Q_Θ on the product manifold $\mathcal{M} = \mathbb{B}_\kappa^n \times \mathcal{P}(\mathcal{H})$. Step 4 derives the bound (50) from the vanishing of the hinge penalty. Step 5 produces the explicit decomposition (51) of the Lipschitz constant $\mathcal{L}_0$. Step 6 verifies the residual $\mathcal{O}(\varepsilon_{\text{Lip}})$ behaviour. Throughout the proof we denote by $\bar{\mathcal{S}}$ the closure of the data domain induced by the bounded minibatch $\mathcal{B}$ in assumption (iii), and we use the notation $N := L n_q$ established in the proofs of Lemma 1 and Proposition 1.

Step 1 (Lipschitz regularity on the hyperbolic factor). By assumption (i), the encoder $\Phi : \mathcal{S} \rightarrow \mathbb{B}_\kappa^n$ is continuous on a bounded domain and the Möbius layers downstream of Φ are L_Φ -Lipschitz on bounded subsets of $\mathbb{B}_\kappa^n$ with respect to d_κ . By assumption (iii) the parameter set is compact, hence the composition of the encoder and the Möbius layers, regarded as a map $\Psi_\kappa : \mathcal{S} \rightarrow \mathbb{B}_\kappa^n$, is uniformly Lipschitz on $\bar{\mathcal{S}}$. Combining these regularities with the chordal-to-geodesic comparison $d_\kappa(u, v) \geq \|u - v\|_2 / (1 + \sqrt{\kappa} \text{diam}_2)$, valid on bounded subsets of $\mathbb{B}_\kappa^n$ of diameter diam_2 (a consequence of Equation (6)), there exists a constant $L_{\text{hyp}} > 0$, depending only on L_Φ, κ , and diam_2 , such that for every $s_i, s_j \in \bar{\mathcal{S}}$:

$$d_\kappa(\Phi(s_i), \Phi(s_j)) \leq L_{\text{hyp}} \|s_i - s_j\|_2. \quad (52)$$

The constant L_{hyp} may be absorbed into the universal constant C_1 of Equation (51).

Step 2 (Lipschitz regularity of the quantum read-outs on the data domain). By Proposition 1, for every fixed $\theta \in [-\pi, \pi]^{|\theta|}$ and every $x, x', x'' \in \mathbb{R}^n$:

$$|f_\theta(x) - f_\theta(x')| \leq \Lambda_f^{(x)} \|x - x'\|_2, \quad (53)$$

$$|K_Q^{(\theta)}(x, x'') - K_Q^{(\theta)}(x', x'')| \leq \Lambda_K^{(x)} \|x - x'\|_2. \quad (54)$$

The data-wise constants $\Lambda_f^{(x)}, \Lambda_K^{(x)}$ are uniform in θ and depend on $\hat{H}$, on the depth L , on the qubit count n_q , and on the spectral norm of the input projection W_q , as established in Proposition 1. Since the tangent map $x_t = \log_0(z_t)$ of Equation (8) is $L_{\log}$ -Lipschitz

on bounded subsets of $\mathbb{B}_\kappa^n$, with $L_{\log}$ depending only on the curvature κ and on diam_2 , inequalities (53) and (54) pull back to $z \in \mathbb{B}_\kappa^n$:

$$|f_\theta \circ \log_0(z) - f_\theta \circ \log_0(z')| \leq L_{\log} \Lambda_f^{(x)} d_\kappa(z, z'), \quad (55)$$

and analogously for the kernel pulled back to $\mathbb{B}_\kappa^n$. The constant $L_{\log}$ may also be absorbed into the universal constants of Equation (51).

Step 3 (Lipschitz regularity on the projective Hilbert factor). The projective Hilbert distance d_Q of Equation (16) satisfies, for $K_Q \in [K_{\min}, 1]$ with $K_{\min} > 0$:

$$1 - K_Q(\psi, \psi') \leq d_Q(\psi, \psi')^2 \leq \frac{\pi^2}{4} (1 - K_Q(\psi, \psi')), \quad (56)$$

which follows from $d_Q = \arccos \sqrt{K_Q}$ and from the elementary bounds $1 - u \leq \arccos^2 \sqrt{u} \leq (\pi^2/4)(1 - u)$ on $[K_{\min}, 1]$. Thus, the squared-fidelity surrogate $1 - K_Q$ is comparable to d_Q^2 up to a constant factor $\pi^2/4$ on bounded fidelity ranges. By assumption (iii) and by the continuity of the encoded states established in Lemma 1–Proposition 1, the encoded state $|\psi_\theta(x)\rangle$ ranges over a compact subset of $\mathcal{P}(\mathcal{H})$ on which $K_Q \geq K_{\min} > 0$. Combining (56) with (54) and the chain rule on the projective Hilbert factor:

$$d_Q(\psi_\theta(x), \psi_\theta(x'))^2 \leq \frac{\pi^2}{4} \Lambda_K^{(x)} \|x - x'\|_2. \quad (57)$$

Step 4 (Bound on the action-value deviation from the hinge penalty). By assumption (iv), the hinge penalty $L_{\text{Lip}}^*(\Theta)$ of Equation (49) vanishes on $\mathcal{B}$ at the optimum of Equation (48). By the explicit form of the hinge in (49), vanishing of the maximum is equivalent to the simultaneous validity, for every pair (i, j) in the minibatch and every action a in the discrete action set, of the inequality:

$$\frac{|Q_\Theta(s_i, a) - Q_\Theta(s_j, a)|}{\Delta_{\mathcal{M}}(s_i, s_j) + \varepsilon_{\text{Lip}}} \leq \mathcal{L}_0, \quad (58)$$

which rearranges to:

$$|Q_\Theta(s_i, a) - Q_\Theta(s_j, a)| \leq \mathcal{L}_0 (\Delta_{\mathcal{M}}(s_i, s_j) + \varepsilon_{\text{Lip}}). \quad (59)$$

The bound (50) follows by writing the right-hand side as $\mathcal{L}_0 \Delta_{\mathcal{M}}(s_i, s_j) + \mathcal{L}_0 \varepsilon_{\text{Lip}}$ and absorbing the second term into $\mathcal{O}(\varepsilon_{\text{Lip}})$. The continuity of Q_Θ in Θ , required to extend the inequality to the action-pair limits, follows from compactness of the parameter domain (iii), from Lemma 1 and Proposition 1 applied to the quantum branch, and from assumption (i) on the Möbius layers downstream of Φ .

Step 5 (Explicit decomposition of $\mathcal{L}_0$). The action-value function Q_Θ is composed of three sources of variation with respect to a state s : the hyperbolic embedding $\Phi(s)$, which contributes via the Riemannian gradient on the Poincaré ball; the scalar quantum feature $f_\theta \circ \log_0 \circ \Phi(s)$, which is concatenated to the agent state and feeds the critic linearly through the Euclidean trunk; and the kernel-weighted Bellman backup of Section 3.7, in which $K_Q^{(\theta)}$ enters through the blended kernel $\tilde{K}^*$. We track the Lipschitz constants of each contribution against the joint metric $\Delta_{\mathcal{M}}$ defined in Equation (49), which combines the hyperbolic geodesic distance d_κ and the fidelity-induced distance $\sqrt{1 - K_Q}$ with the curvature–Hilbert balance ρ .

The hyperbolic contribution is bounded, by Step 1 and the chain rule on Lipschitz compositions, by a constant $C_1 L_\Phi$ multiplying $d_\kappa(\Phi(s_i), \Phi(s_j))$, where $C_1 > 0$ is a universal constant that absorbs the Lipschitz constants of the Möbius layers, of the chordal-to-

geodesic comparison of Step 1, and of the Riemannian linear head of the critic. The scalar quantum contribution is bounded, by Step 2 and the pull-back (55), by a constant $C_2 \Lambda_f^{(x)}$ multiplying $d_\kappa(\Phi(s_i), \Phi(s_j))$, where $C_2 > 0$ absorbs $L_{\log}$ and the Lipschitz constant of the linear head of the critic on the scalar quantum feature. The kernel-weighted backup contribution is bounded, by Step 3 and the kernel-weighted backup of Section 3.7, by a constant $C_3 \rho \Lambda_K^{(x)} \|\hat{H}\|_{\text{op}} \sqrt{N}$ multiplying the term $\sqrt{1 - K_Q(s_i, s_j)}$ that enters $\Delta_{\mathcal{M}}(s_i, s_j)$ through Equation (49), where $C_3 > 0$ absorbs the constant $\pi/2$ of the Bures comparison (56), the operator norm $\|\hat{H}\|_{\text{op}}$ of the read-out observable that enters the kernel-weighted target, and the maximum over the top- K neighbourhood of Equation (64).

Combining the three contributions, the joint Lipschitz constant $\mathcal{L}_0$ admits the upper bound

$$\mathcal{L}_0 \leq C_1 L_\Phi + C_2 \Lambda_f^{(x)} + C_3 \rho \Lambda_K^{(x)} \|\hat{H}\|_{\text{op}} \sqrt{N}, \quad (60)$$

which coincides with Equation (51) since $N = L n_q$. The constants C_1, C_2, C_3 are universal in the sense that they depend only on the curvature κ , on the spectral norm $\|\Pi_A\|_{2 \rightarrow 2}$ of the partial-trace channel, on the diameter diam_2 of the bounded data domain, and on the discount factor γ of Equation (64), and not on the trainable parameters of the pipeline.

Step 6 (Residual term and removal of the safeguard). The residual term $\mathcal{L}_0 \varepsilon_{\text{Lip}}$ appearing in (59) is linear in ε_{Lip} and bounded uniformly on $\mathcal{B}$ by $\mathcal{L}_0 \varepsilon_{\text{Lip}}$, with $\mathcal{L}_0$ the constant of (60). As $\varepsilon_{\text{Lip}} \rightarrow 0^+$, this residual vanishes and inequality (50) reduces to $|Q_\Theta(s_i, a) - Q_\Theta(s_j, a)| \leq \mathcal{L}_0 \Delta_{\mathcal{M}}(s_i, s_j)$, that is, a clean Lipschitz bound of the action-value function on the joint hyperbolic–Hilbert product manifold with respect to the joint metric $\Delta_{\mathcal{M}}$. The numerical safeguard ε_{Lip} is therefore harmless to the asymptotic content of the bound and serves only to prevent division by zero when two states coincide on both manifolds, as discussed in Section 3.4. $\square$

A brief remark is in order on the scope of the bound just established. Condition (iv) of the theorem requires the hinge penalty $\mathcal{L}_{\text{Lip}}^*(\Theta)$ to vanish on the training minibatch $\mathcal{B}$ at the optimum of the composite loss, so that the action-value Lipschitz inequality is, strictly speaking, established for state pairs drawn from $\mathcal{B}$. The bound extends to arbitrary pairs in the closure $\bar{\mathcal{S}}$ of the data domain through a standard uniform-continuity argument that we record here for completeness. By the compactness assumption on the parameter domain, by the parameter-wise and data-wise Lipschitz regularity of the quantum read-outs $f_\theta(x)$ and $K_Q^{(\theta)}(x, x')$ established in the preceding lemma and proposition, by the continuity of the encoder Φ and of the downstream Möbius layers on bounded subsets of $\mathbb{B}_\kappa^n$, and by the continuity of the joint metric $\Delta_{\mathcal{M}}$ in both arguments, the action-value function Q_Θ is continuous on the compact joint manifold $\mathcal{M} = \mathbb{B}_\kappa^n \times \mathbb{P}(\mathcal{H})$. The Lipschitz inequality therefore propagates from the dense subset $\mathcal{B}$ to the full closure $\bar{\mathcal{S}}$ with the same constant $\mathcal{L}_0$ and the same residual term $\mathcal{O}(\varepsilon_{\text{Lip}})$. The empirical minibatch enforcement of the hinge penalty thus admits a clean population-level reading on the support of the data-generating distribution, which is the regime in which the stability bound is invoked at inference time.

Theorem 1 shows that the introduction of the quantum kernel does not degrade the regularization properties of the original hyperbolic Lipschitz constraint but, on the contrary, tightens them whenever two states that are geometrically close on the Poincaré ball happen to be distinguished by their fidelity-induced distance, which is precisely the regime in which non-classical expressivity becomes operative. The explicit dependence of $\mathcal{L}_0$ in Equation (51) also clarifies that the quantum contribution to the joint Lipschitz constant is controlled by the product of the curvature–Hilbert balance ρ , of the data-wise kernel Lipschitz constant $\Lambda_K^{(x)}$, and of the depth–width product $L n_q$, providing an operational guideline for the hyperparameter selection of Section 4.

Theorem 1 admits a direct operational reading. The bound certifies that the action-value function varies by at most L_0 per unit of joint-manifold distance $\Delta_{\mathcal{M}}$ of Equation (49); States proximal on the product manifold cannot induce divergent LONG/SHORT/HOLD decisions or divergent take-profit/stop-loss distances. The policy sensitivity to noisy or transient state perturbations is thereby bounded, which limits turnover-driven loss under non-stationary dynamics. Because the joint metric augments d_{κ} with the fidelity-induced term $\sqrt{1 - K_Q}$, the constraint tightens when two hyperbolically proximal states are separated by the quantum kernel, i.e., at regime-switching transitions where a Euclidean or purely hyperbolic critic smears value across regimes. The reduced maximum drawdown, the regime-stratified drawdown reductions, and the ablation asymmetry reported in Section 4 are the empirical signatures of this mechanism.

3.6. Optimization Scheme and Convergence Remarks

The optimization scheme that minimizes the composite loss couples three update rules whose interplay is essential for stable training. Manifold-valued parameters $\{\Phi(v), W_{\text{hyp}}, b_{\text{hyp}}, w_p\}$ are updated by Riemannian stochastic gradient descent [18] through the manifold update:

$$\xi_{k+1} = \exp_{\xi_k}(-\eta \nabla_R \mathcal{L}(\xi_k)), \quad (61)$$

where $\xi_k \in \mathbb{B}_{\kappa}^n$ denotes the current value of any manifold-valued parameter at iteration k , $\nabla_R \mathcal{L}(\xi_k) \in T_{\xi_k} \mathbb{B}_{\kappa}^n$ is the Riemannian gradient of the composite loss with respect to that parameter, $\eta > 0$ is the manifold learning rate, and $\exp_{\xi_k} : T_{\xi_k} \mathbb{B}_{\kappa}^n \rightarrow \mathbb{B}_{\kappa}^n$ is the exponential map of the Poincaré ball at ξ_k . Quantum and Euclidean parameters $\{\theta, c, W_q, b_q, W_{\text{RL}}\}$ follow the Adam update with parameter-shift gradients on the quantum branch, supplied by Equation (18), and ordinary back-propagation on the classical branch. Convergence considerations are inherited from [9,18,36]: the composite loss is bounded below, is Lipschitz-smooth in the variational angles on compact angle ranges by Lemma 1, and admits the joint stability bound of Equation (50). A vanishing-step rule therefore yields almost-sure convergence to a stationary point of the joint hyperbolic–Hilbert objective.

3.7. Action Space, Shaped Reward, and Kernel-Weighted Bellman Backup

The action space of the trading head is the hybrid set:

$$a_t \in \{\text{LONG}, \text{SHORT}, \text{HOLD}\} \times \mathbb{R}_{+}^2, \quad (62)$$

which combines a discrete direction with continuous take-profit and stop-loss distances expressed in pips, exactly as in [8]. The shaped reward signal that the agent maximizes integrates the realized return of the active trade with the risk and dynamical penalties already used by the base pipeline and with the new attractor-anchored term that blends the classical hyperbolic similarity with the quantum kernel, and is given explicitly by:

$$r_t = \alpha_1 R_t - \alpha_2 D_t - \alpha_3 c_t - \alpha_4 \bar{\lambda}_1^{(L)}(t) - \alpha_5 \tilde{H}_{\text{perm}}(t) + \alpha_m \max_{p \in \mathcal{P}_{\text{mag}}} \left[(1 - \nu) e^{-\delta d_{\kappa}(z_t, \Phi(p))} + \nu K_Q(z_t, \Phi(p)) \right], \quad (63)$$

in which $R_t \in \mathbb{R}$ is the realized profit-and-loss of the trade net of transaction costs, $D_t \in \mathbb{R}_{\geq 0}$ is the holding-window drawdown experienced during the trade, $c_t \in \mathbb{R}_{\geq 0}$ aggregates spread, fees, and slippage incurred at execution, $\bar{\lambda}_1^{(L)}(t)$ and $\tilde{H}_{\text{perm}}(t)$ are local neighbourhood averages of the dynamical diagnostics of Equations (2) and (3) that penalize trading in highly chaotic or highly disordered regimes, $\mathcal{P}_{\text{mag}}$ is the magnet set produced by the attractor-aware extractor of Equation (4), $z_t = \Phi(s_t) \in \mathbb{B}_{\kappa}^n$ is the hyperbolic latent of the current state, $\Phi(p) \in \mathbb{B}_{\kappa}^n$ is the embedding of a candidate magnet price p , d_{κ} is the geodesic

distance on the Poincaré ball, K_Q is the projected quantum kernel of Equations (14) and (15), $\delta > 0$ is a decay parameter that controls how rapidly the classical magnet bonus decays with hyperbolic distance, $\nu \in [0, 1]$ is the blend coefficient that interpolates between the classical hyperbolic similarity and the quantum kernel, and $\alpha_1, \dots, \alpha_5, \alpha_m \geq 0$ are scalar coefficients that balance the seven additive contributions. The shaped reward of Equation (63) aggregates the realized return net of costs, a holding-window drawdown penalty, the execution-cost term, two dynamical penalties that suppress trading in highly chaotic or highly disordered regimes, and an attractor-proximity bonus that blends, through ν , the hyperbolic similarity $e^{-\delta d_k}$ with the quantum-kernel similarity K_Q . The same ν re-weights the Bellman backup of Equation (64), so that return, cost, risk, attractor proximity, hyperbolic similarity, and quantum-kernel similarity enter the value estimate through a single, consistently weighted channel.

The associated Bellman backup that drives the temporal-difference component of the loss is the kernel-weighted update written as:

$$\begin{aligned} Q(s_t, a_t) &= r_t + \gamma \max_a \mathbb{E}_{s'} [\tilde{K}^*(s_{t+1}, s') Q(s', a)], \\ \tilde{K}^* &= (1 - \nu) \tilde{K}_{\text{hyp}} + \nu \tilde{K}_Q, \end{aligned} \quad (64)$$

in which $\gamma \in (0, 1)$ is the discount factor of the agent, s' is the successor state under the dynamics induced by a_t , $\tilde{K}^*(s_{t+1}, s') \in [0, 1]$ is the normalized blended kernel, $\tilde{K}_{\text{hyp}}$ and $\tilde{K}_Q$ are the row-normalized restrictions of the classical hyperbolic kernel and of the trained projected quantum kernel $K_Q^{(\theta)}$ to a top- K neighbourhood of s_{t+1} , and ν is the same blend coefficient that appears in the reward shaping of Equation (63). The neighbourhood size $K \in \mathbb{N}$ is selected on validation in order to balance estimation noise against attractor coverage, and the normalization is computed point-wise as $\tilde{K}^*(s_{t+1}, s') = K^*(s_{t+1}, s') / \sum_u K^*(s_{t+1}, u)$ over the chosen neighbourhood. The complete training and inference procedure that integrates all the components introduced above is summarized in Algorithm 1.

Algorithm 1 Hybrid hyperbolic–quantum training and inference.

Require: Multimodal stream $\mathcal{D} = \{y_t^{(c, \tau)}, N_t, \text{CoT}_t, V_t, T_t\}$, magnet set $\mathcal{P}_{\text{mag}}$, hyperbolic curvature κ , qubit count n_q , ansatz depth L , blend ν , learning rates η, η_q , batch $\mathcal{B}$.

- 1: **Pre-trade (offline, daily):**
 - 2: Compute $\{\text{RR}_\rho, \lambda_1^{(L)}, \tilde{H}_{\text{perm}}, \text{CR}, A, \mathcal{P}_{\text{mag}}\}$ on rolling intraday windows.
 - 3: Generate s-CoT via schema-constrained LLM under fixed temperature and stored seeds for reproducibility.
 - 4: Embed s-CoT via Φ into $\mathbb{B}_\kappa^n$ and freeze for the session.
 - 5: **Hybrid optimization (training, walk-forward fold with purging and embargo):**
 - 6: **for** minibatch $\mathcal{B}$ in walk-forward fold **do**
 - 7: Compute $z_t = \Phi(s_t)$, tangent map $x_t = \log_0(z_t)$, and encode $|\psi(x_t)\rangle$ via Equation (10).
 - 8: Apply $U(\theta)$ and read out $f_\theta(x_t)$ and K_Q on the magnet set.
 - 9: Form Bellman target via Equation (64) and the joint Lipschitz penalty Equation (49).
 - 10: Update θ by parameter-shift gradients (Equation (18)), Euclidean weights by Adam, and manifold weights by Equation (61).
 - 11: **end for**
 - 12: **Inference (live, intraday):**
 - 13: **for** each intraday bar t **do**
 - 14: Refresh $z_t, f_\theta(x_t), \{K_Q(z_t, \Phi(p))\}_{p \in \mathcal{P}_{\text{mag}}}$.
 - 15: Select direction $\hat{a}_t \in \{\text{LONG}, \text{SHORT}, \text{HOLD}\}$ and conditional (TP, SL) by argmax of Q on the discrete branches with a single projected step on TP/SL.
 - 16: Submit order subject to spread/fees/slippage; log realized r_t and update target network.
 - 17: **end for**
-

3.8. Computational Complexity of the Quantum Branch

The computational cost of the quantum branch deserves an explicit analysis, both because it determines the feasibility of online deployment and because it bounds the depth–width regimes in which the construction can avoid the barren-plateau pathology. Each parameter-shift gradient costs two circuit executions, and the per-step quantum cost is therefore $\mathcal{O}(L n_q)$ executions. State-vector simulation, which we adopt in this work in order to isolate the contribution of the quantum block from hardware-specific decoherence, scales as $\mathcal{O}(2^{n_q})$ per circuit. The offline training budget for the qubit count and ansatz depth specified in Section 4 therefore remains tractable, and the online critic update is dominated by the kernel evaluations on the magnet neighbourhood, with cost $\mathcal{O}(K 2^{n_q})$ per step. The end-to-end decision latency stays well within the intraday decision horizon inherited from [8], and the residual headroom is used in Section 4 to run the barren-plateau diagnostic and the noise-sensitivity analysis that are required to substantiate the framing of the quantum block as a non-classical, geometry-aware regularizer.

4. Experimental Results

The experimental evaluation is designed to address the methodological standard that, as anticipated in Section 2, recent work on dequantization, exponential concentration, and barren plateaus has rendered mandatory for any controlled study of a quantum machine-learning component. The protocol extends [8] with: a walk-forward validation scheme; competitive classical and quantum baselines from the FX literature (Section 2); regime-stratified attribution across major test-horizon shocks; multi-seed variance analysis; deflated Sharpe ratio [49] testing with multiple-comparison correction; three parameter-matched classical surrogates of the quantum kernel; a barren-plateau diagnostic on variational gradients; and a quantum-branch noise-sensitivity study.

4.1. Dataset, Walk-Forward Protocol, and Headline Configuration

The empirical evaluation operates on a multi-asset, multi-resolution foreign-exchange dataset that is fully described in this section so that the protocol is independently reproducible. The tested FX assets consists of nine major FX crosses, namely EUR/USD, EUR/JPY, EUR/CHF, USD/JPY, USD/CHF, GBP/USD, EUR/GBP, USD/CAD, and AUD/USD, which together cover the principal liquidity poles of the spot FX market and span the three majors-against-USD, the JPY-funded carry leg, the CHF safe-haven leg, the commodity-linked crosses, and one purely European cross. For each instrument, we collect the full Open–High–Low–Close (OHLC) tape at four resolutions, namely 1-minute, 15-minute, 1-hour, and daily bars, sourced from a regulated broker feed time-stamped in UTC and re-aligned to a Europe/Rome trading calendar for the daily evaluation horizon. The feed is a regulated-broker top-of-book stream providing bid and ask quotes at each resolution; the marking mid-quote is $(\text{bid} + \text{ask})/2$, and the half-spread is charged at execution. The spot-FX week runs continuously from Sunday 22:00 UTC to Friday 22:00 UTC; the daily bar is rolled at 22:00 UTC and re-aligned to the Europe/Rome calendar for the close-to-close horizon.

The 1-minute series provide the micro-structural baseline for the FTLE and permutation-entropy estimators of Section 3; the 15-minute series (M15) constitute the execution time-frame on which orders are simulated; the 1-hour series support the technical-confluence indicators consumed by the schema-constrained s-CoT generator; the daily series define the evaluation horizon and the close-to-close return on which the headline P&L statistics are computed. Within-day data gaps are forward-filled strictly within the same trading session and never across overnight windows; samples falling on broker-side maintenance windows or on declared FX holidays are removed from both training and evaluation, without imputation. All continuous inputs are standardized through a 250-day rolling

z-score with EWMA bandwidth $\lambda = 0.1$ and clipped at the $\pm 3\sigma$ robust quantiles before entering the hyperbolic encoder.

The execution model is held fixed throughout the evaluation. Each order is submitted on the next M15 bar after the decision and is filled at the next available OHLC mid-quote, subject to a broker-published average spread of 2.3 pips, an additive slippage buffer of 0.5 pips, a per-million-notional commission of \$7, and an overnight financing cost of 0.20%/0.50% per day on the long/short legs, respectively, for positions that span the daily roll. Concretely, a decision emitted on bar t is executed at the open of bar $t+1$ on the prevailing mid-quote, to which the half-spread is added for long entries and subtracted for short entries, and symmetrically on exit, so that no fill is obtained at a price more favourable than the quoted bid/ask; the additive slippage buffer is applied on the half-spread in the adverse direction. For pairs whose published spreads exceed the major-cross average (EUR/CHF and USD/CAD on the 2015–2016 SNB and crude-oil windows), the spread is widened on a per-day basis to the broker-published maximum recorded for that day, so that the cost model is conservative on the very episodes that are most vulnerable to execution risk.

The overnight financing of 0.20%/0.50% per day on the long/short legs is the swap/rollover charge and is applied to positions spanning the 22:00 UTC daily roll, entering the realized P&L directly rather than being added ex post. Market impact is negligible under the $1.0\times$ gross-leverage cap relative to top-of-book spot-FX liquidity and is conservatively absorbed by the half-spread and the slippage buffer; it is therefore not modeled as a separate term. Illiquid-hours risk is controlled by forbidding the opening of new positions on the last M15 bar of the trading day, so the policy does not initiate trades into the New-York-close window. As a sensitivity check, doubling the overnight financing to 0.40%/1.00% per day reduces pair-averaged monthly P&L by approximately 4.8%, consistent with the predominantly intraday holding period (average 13.8 M15 bars), while slippage perturbations of $\pm 30\%$ affect the headline metrics by less than 4.6%, leaving the qualitative ordering unchanged.

The broker-published average spread is pair-specific; the per-cross values are reported in Table 1 with a pair-averaged value of 2.3 pips. During the stress windows of Section 4.4, each spread is widened on a per-day basis to the broker-published daily maximum, so that the cost model is conservative on the episodes most exposed to execution risk. The slippage buffer, the \$7-per-million commission, and the 0.20%/0.50% long/short overnight financing are held fixed across regimes. Robustness to these assumptions is quantified in Section 4.7: doubling the average spread and perturbing slippage by $\pm 30\%$ leave the method ordering invariant.

The risk envelope caps gross leverage at $1.0\times$ and per-trade leverage at $2.1\times$, halves the position size on any rolling drawdown that exceeds -10% , and forbids opening new positions on the last M15 bar of the trading day; existing positions may, however, be closed on the following day if the policy emits a flat signal.

Position sizing is volatility-targeted: the trade notional is set so that the stop-loss distance corresponds to a fixed fraction of account equity, subject to the gross-leverage cap of $1.0\times$ and the per-trade leverage cap of $2.1\times$. The maximum single-position notional is therefore $2.1\times$ account equity, and the initial margin equals the notional divided by the per-trade leverage. The agent holds a long, short, or flat position, as encoded by the discrete branch of the action space in Equation (62), and re-evaluates the position once per M15 bar; the take-profit and stop-loss distances form the continuous action component, expressed in pips. Pyramiding is disallowed: a position is opened only from the flat state, and an open position is held, scaled down by the drawdown rule, or closed.

No look-ahead, no overnight gap arbitrage, and no end-of-day mark-to-market short-cuts are permitted; all P&L is computed strictly on realized fills.

The temporal partition is identical for the precursor and for the quantum-augmented pipeline, and is fixed before any training run is launched. The training-and-validation segment covers January 2012–December 2014 and is used exclusively for fitting the hyperbolic encoder, the variational angles, the critic, and for the selection of all hyperparameters listed in Table 1; the out-of-sample segment covers January 2015–July 2025 and is used exclusively for evaluation. The walk-forward protocol that operates within the out-of-sample horizon is rolling-origin rather than single-shot. The model is retrained quarterly on an expanding training window, validated on the last sixty trading days of the training window with a five-day purging buffer that removes any sample whose information set overlaps with the validation block, and tested on the next sixty-five trading days with a five-day embargo that separates training from test on both ends. The expanding-window scheme is preferred to a sliding window because it preserves the full pre-2015 history at every retraining event and therefore stabilizes the variational-angle initialization across folds, while the purging-and-embargo discipline neutralizes the standard sources of look-ahead leakage that arise when overlapping returns are used to construct supervised targets in quantitative finance. All hyperparameters that govern the quantum block, namely the qubit count n_q , the ansatz depth L , the blend coefficient ν , the curvature–Hilbert balance ρ , and the Lipschitz margin $\mathcal{L}_0$, are selected exclusively on the validation segments of the rolling folds, so that no information from the 2015–2025 test horizon enters the model selection.

The training window is expanding rather than sliding: it begins with the January 2012–December 2014 in-sample segment and grows by one quarter at each retraining event, so no fixed lookback length is imposed and the effective lookback is the full pre-fold history. Retraining is performed quarterly and offline; its cost is dominated by the $O(L n_q)$ parameter-shift gradients and the $O(2^{n_q})$ state-vector simulation within a per-fold budget of 4×10^4 gradient steps, and it never enters the online decision path, which contains no retraining and no language-model call and executes in 45–90 s per decision. The variational ansatz and the reinforcement-learning policy are therefore not optimized on the online critical path, so the production cost is the sub-minute inference rather than the offline retraining.

The walk-forward parameters are reported in Table 1. Model selection is causal: at each retraining event the hyperparameter grid is scored only on the validation block internal to the current expanding training window, and the selected configuration is frozen before the 65-day test block is scored. No statistic computed on any test block influences model selection.

The full configuration of the headline run is reported in Table 1 for reproducibility and as the explicit calibration of the parameters introduced symbolically in Section 3.

For reproducibility, the headline configuration table additionally records the stopping rule, the per-fold gradient-step budget, and the computational environment. Optimization uses early stopping on the validation P&L with patience 15 epochs and a hard cap of 4×10^4 gradient steps per fold. The Quantum AI block is simulated on a single NVIDIA A100 GPU with a PennyLane state-vector backend coupled to a PyTorch VER-SION 3.10 automatic-differentiation graph; the reported 45–90 s online decision latency is measured on the same instance.

The Quantum AI block is implemented as a noiseless state-vector simulator on a GPU back-end, and hybrid training couples Adam on the Euclidean and quantum branches with Riemannian SGD on the manifold branch, as specified by Algorithm 1. The end-to-end online decision latency of the deployed pipeline, measured from feature refresh to order submission on a single GPU instance with the headline configuration, is in the range 45–90 s

per decision, which fits the M15 execution horizon with a safety margin of more than one order of magnitude. The s-CoT corpus is generated once per trading day in the pre-market window and frozen for the session, so that no LLM call enters the online critical path. Each reported metric is averaged over ten random seeds that govern the initialization of the variational angles, the minibatch sampling, and, where applicable, the stochastic structure of the s-CoT generator, with associated standard deviations reported explicitly. The s-CoT generator is ChatGPT-5 in the headline runs, queried under fixed temperature, fixed top- p , and stored prompt-and-response logs that allow the corpus to be regenerated bit-by-bit; sensitivity to alternative LLM backbones is reported in Section 4.7.

Table 1. Experimental setup of the hybrid hyperbolic–quantum pipeline, organized into three panels: (a) model and training hyperparameters, with cross-references to the equations of Section 3; (b) broker-published average bid–ask spreads (in pips) for each FX cross used in the evaluation (pair-averaged value 2.3 pips; during stress windows each spread is widened to the broker-published daily maximum, Section 4.4); (c) the walk-forward (rolling-origin) backtesting protocol.

(a) Model and training hyperparameters			
Symbol	Value	Role	
n_q	8	qubits in the variational register	
L	4	ansatz depth, Equation (11)	
n	64	latent dimension of $\mathbb{B}_\kappa^n$	
κ	1.0	negative curvature, Equation (5)	
β	0.75	hyperbolic-kernel bandwidth, Equation (7)	
ν	0.5	classical/quantum kernel blend, Equations (63) and (64)	
ρ	1.0	curvature–Hilbert balance, Equation (17)	
$\mathcal{L}_0$	1.0	Lipschitz margin, Equation (49)	
ε_{Lip}	10^{-6}	numerical safeguard, Equation (49)	
τ_f	0.30	target magnitude for f_θ , Equation (47)	
η	5×10^{-4}	RSGD learning rate, Equation (61)	
η_q	1×10^{-3}	Adam rate, quantum and Euclidean branches	
$ \mathcal{B} $	256	minibatch size	
μ_1	1×10^{-2}	weight on f_θ regularizer, Equation (47)	
μ_2	1×10^{-2}	weight on Frobenius alignment, Equation (47)	
η_{Lip}	1×10^{-1}	Lipschitz-penalty weight, Equation (48)	
δ	2.0	magnet-bonus decay, Equation (63)	
γ	0.99	RL discount factor, Equation (64)	
K	24	top- K neighbourhood for kernel backup	
m_e, τ_e	5, 1	permutation-entropy order/lag, Equation (3)	
m, τ	7, 2	Takens dimension/delay, Equation (1)	
L_{FTLE}	32 bars	FTLE integration horizon, Equation (2)	
α	(1.0, 0.6, 0.4, 0.3, 0.3, 0.5)	reward weights $(\alpha_1, \dots, \alpha_5, \alpha_m)$, Equation (63)	
seeds	10	random seeds for variance estimation	
stopping	patience 15	early stop on validation P&L (optimization stopping criterion)	
steps _{max}	4×10^4	max gradient steps per walk-forward fold	
hardware	A100 40 GB	NVIDIA GPU, state-vector simulation back-end	
software	PyTorch+PennyLane	autodiff + state-vector simulator	
(b) Per-cross broker-published average spread (pips)			
FX Cross	Avg. Spread	FX Cross	Avg. Spread
EUR/USD	1.3	GBP/USD	2.0
EUR/JPY	2.2	EUR/GBP	2.6
EUR/CHF	3.0	USD/CAD	3.1
USD/JPY	1.6	AUD/USD	2.3
USD/CHF	2.8		
(c) Walk-forward (rolling-origin) protocol			
Element	Setting		
In-sample (train + validation) segment	January 2012–December 2014		
Out-of-sample (test) horizon	January 2015–July 2025		
Training window	Expanding (rolling origin)		
Retraining frequency	Quarterly		
Validation window	Last 60 trading days of training window		
Purge buffer (train to validation)	5 trading days		
Test window	65 trading days		
Embargo (both ends)	5 trading days		
Model selection	Validation-only, causal		

4.2. Headline Comparison Against the Precursor

The headline comparison reported in Table 2 contrasts the full quantum-augmented pipeline against the attractor-aware hyperbolic Lipschitz-constrained reinforcement-learning baseline of [8] on the identical walk-forward window described in Section 4.1, with all metrics computed on the pair-averaged out-of-sample stream and reported as non-compounded statistics so that the gain is not amplified by reinvestment. The baseline is retrained from scratch under the cost model and the random-seed protocol of the present study, in order to neutralize implementation drift between the two pipelines and to provide pair-averaged uncertainty estimates that are directly comparable to those of the quantum-augmented run; the resulting pair-averaged Sharpe ratio of 1.44 ± 0.05 is the direct counterpart of the per-cross figures reported by the precursor and is computed on the same realized return stream. The auxiliary indicators reported in Table 3 characterize the trading footprint of the policy, which is critical to disambiguate genuine performance gains from turnover-driven artefacts.

Table 2. Pair-averaged out-of-sample performance, January 2015–July 2025 (non-compounded), mean $\pm$ standard deviation across ten random seeds. Higher is better for all metrics except Max DD, where lower is better. PSR is the Probabilistic Sharpe Ratio against a benchmark Sharpe of zero; PF is the profit factor. value.

Method	P&L (%)	Cum. (%)	Sharpe	Sortino	DD (%)	Win (%)	PF	PSR
Base [8]	2.22 ± 0.06	282 ± 7	1.44 ± 0.05	1.92 ± 0.07	10.39 ± 0.32	56.4 ± 0.4	1.51 ± 0.04	0.987
Proposed	2.55 ± 0.07	324 ± 9	1.78 ± 0.06	2.39 ± 0.08	8.83 ± 0.27	58.2 ± 0.4	1.74 ± 0.05	0.998

Table 3. Auxiliary indicators of the trading footprint, pair-averaged across the nine FX crosses on the 2015–2025 out-of-sample window.

Indicator	Base [8]	Proposed	Notes
Calmar (annual return/Max DD)	2.56	3.43	higher is better
Trades/month	42.1	36.7	lower turnover
Avg. holding period (bars)	11.4	13.8	longer holds, higher conviction

The quantum-augmented pipeline lifts pair-averaged monthly P&L by 0.33 percentage points, corresponding to a 14.9% relative gain, while reducing maximum drawdown by 1.56 p.p., a 15.0% relative reduction. The Calmar ratio rises from 2.56 to 3.43, the Sharpe ratio from 1.44 to 1.78, and the profit factor from 1.51 to 1.74. The PSR, computed against a benchmark Sharpe of zero on the realized return distribution, sits at 0.998 for the quantum-augmented configuration against 0.987 for the precursor, both above the 0.95 confidence threshold. The quantum-augmented configuration achieves the gain with approximately 13% fewer trades per month and a longer average holding period, which is the empirical signature of a policy that concentrates exposure on higher-confidence magnet neighbourhoods rather than on marginal opportunities.

4.3. Competitive Benchmarking

To anchor these results against the broader trading-strategy landscape and to disambiguate the contribution of the quantum block from the underlying hyperbolic and attractor-aware machinery already present in the precursor, we extend the comparison to a battery of competitive baselines that cover the principal architectural families surveyed in Section 2. The full benchmarking set is reported in Table 4. All baselines are retrained and re-evaluated under the identical walk-forward protocol, identical execution model, and identical cost structure as the headline pipeline whenever the source code or sufficient implementation details are available; for the published systems whose original back-tests use a different cost model, we report the figures disclosed in the source paper and flag them

explicitly in the table caption. Three classes of classical surrogate of the quantum kernel are included to control for the dequantization-style critique discussed in Section 2: a random-Fourier-feature approximation of matched parameter budget, a Nyström approximation of the projected fidelity kernel on a 512-anchor inducing set, and a multi-layer perceptron of matched depth–width product trained jointly with the remainder of the pipeline.

To render the attribution of the quantum contribution unambiguous, each baseline receives the information set and tuning budget of the headline pipeline. Every deep-RL and forecasting baseline receives the identical multi-resolution OHLC tape and, where its architecture admits an exogenous feature vector, the identical s-CoT-derived inputs; all baselines are trained and evaluated under identical walk-forward folds, execution model, and cost structure, and are tuned over a validation grid of identical cardinality under the identical ten-seed protocol. The three kernel surrogates are parameter-matched to the quantum branch: the random-Fourier-feature map and the multi-layer perceptron to the depth–width product $L n_q$, and the Nyström approximation to a 512-anchor inducing set. Any residual advantage of the quantum-augmented configuration is therefore attributable to the projected-fidelity geometry rather than to an information or tuning asymmetry.

Table 4. Competitive benchmarking and Lipschitz-penalty control experiment, 2015–2025 walk-forward window, pair-averaged across nine FX crosses and ten seeds (mean $\pm$ std). The upper block lists directly comparable entries under an identical execution and cost model; the middle block (+) reports figures as published, under a non-identical cost model and FX universe, for relative ranking only. The lower block isolates the joint Lipschitz penalty of Equation (49) by replacing the projective-Hilbert factor $(1 - K_Q)$ with the classical hyperbolic-RBF factor $(1 - K_{hyp})$ (no quantum kernel). Higher is better for all metrics except Max DD; “–” denotes a metric not disclosed.

Method / Configuration	P&L (%)	Sharpe	DD (%)	PF	PSR
<i>Directly comparable entries (identical walk-forward, costs, FX universe, ten seeds):</i>					
Buy-and-hold (long leg)	0.41 $\pm$ 0.02	0.18 $\pm$ 0.02	27.6 $\pm$ 0.5	1.04	0.41
Time-series momentum (12M)	0.78 $\pm$ 0.05	0.42 $\pm$ 0.04	19.8 $\pm$ 0.6	1.18	0.69
Recurrent forecaster + rule [19]	1.12 $\pm$ 0.07	0.71 $\pm$ 0.05	16.2 $\pm$ 0.7	1.27	0.83
Stacked ensemble [21]	1.24 $\pm$ 0.07	0.78 $\pm$ 0.05	15.4 $\pm$ 0.7	1.31	0.86
DQN, multi-resolution OHLC	1.36 $\pm$ 0.10	0.81 $\pm$ 0.06	18.1 $\pm$ 0.9	1.34	0.85
PPO, multi-resolution OHLC	1.54 $\pm$ 0.09	0.95 $\pm$ 0.06	16.8 $\pm$ 0.8	1.39	0.90
Recurrent RL [25]	1.62 $\pm$ 0.08	1.02 $\pm$ 0.06	14.5 $\pm$ 0.7	1.42	0.93
Multi-agent DQN [23]	1.51 $\pm$ 0.09	0.97 $\pm$ 0.07	15.1 $\pm$ 0.7	1.40	0.91
PPO + AXT [24]	1.71 $\pm$ 0.10	1.08 $\pm$ 0.07	14.0 $\pm$ 0.6	1.44	0.94
GAF-CNN PPO [26]	1.43 $\pm$ 0.11	0.89 $\pm$ 0.07	16.5 $\pm$ 0.8	1.37	0.88
DQN transfer (cross-pair) [50]	1.21 $\pm$ 0.09	0.74 $\pm$ 0.06	17.2 $\pm$ 0.8	1.29	0.84
Grid-trading robot [31]	1.05 $\pm$ 0.04	0.62 $\pm$ 0.03	17.9 $\pm$ 0.5	1.25	0.78
LSTM-RL corrector [32]	1.48 $\pm$ 0.06	0.88 $\pm$ 0.04	15.7 $\pm$ 0.6	1.36	0.88
Euclidean ablation (ours)	1.85 $\pm$ 0.07	1.16 $\pm$ 0.05	12.4 $\pm$ 0.5	1.46	0.95
RFF surrogate of K_Q (ours)	2.36 $\pm$ 0.08	1.55 $\pm$ 0.06	9.95 $\pm$ 0.31	1.62	0.97
Nyström surrogate of K_Q (ours)	2.30 $\pm$ 0.08	1.51 $\pm$ 0.06	10.12 $\pm$ 0.34	1.59	0.97
MLP-matched surrogate (ours)	2.28 $\pm$ 0.09	1.49 $\pm$ 0.07	10.20 $\pm$ 0.36	1.57	0.97
Base [8]	2.22 $\pm$ 0.06	1.44 $\pm$ 0.05	10.39 $\pm$ 0.32	1.51	0.987
Full Quantum-Hyperbolic (proposed)	2.55 $\pm$ 0.07	1.78 $\pm$ 0.06	8.83 $\pm$ 0.27	1.74	0.998
<i>Reference entries (as published, non-identical cost model and FX universe; for ranking only):</i>					
GTSbot + [31]	0.98	0.55	11.25	1.20	–
LSTM-RL + [32]	0.55	0.62	15.97	1.26	–
Multi-agent DQN + [23]	0.43	0.63	11.89	–	–
PPO + AXT + (DS2, 7M) [24]	6.03	0.47	–	–	–
GAF-CNN PPO + (1M) [26]	–	–	2.92	3.82	–
<i>Lipschitz-penalty control experiment (joint penalty on classical kernel, no quantum kernel):</i>					
Base [8] (hyperbolic Lipschitz, d_k only)	2.22 $\pm$ 0.06	1.44 $\pm$ 0.05	10.39 $\pm$ 0.32	1.51	0.987
Joint Lipschitz on classical kernel (control, ours)	2.31 $\pm$ 0.08	1.49 $\pm$ 0.06	10.05 $\pm$ 0.34	1.58	0.991
Full Quantum-Hyperbolic (proposed)	2.55 $\pm$ 0.07	1.78 $\pm$ 0.06	8.83 $\pm$ 0.27	1.74	0.998

The benchmarking pattern is internally consistent. The quantum-augmented pipeline is the unique configuration that simultaneously achieves the highest Sharpe ratio and the lowest maximum drawdown across all directly comparable entries. The Euclidean

ablation lags by approximately eight percentage points of annualized P&L and three and a half percentage points of maximum drawdown, confirming that negative curvature is a load-bearing component independently of the quantum block. Among the FX-trading baselines re-evaluated under identical conditions, the strongest deep-RL competitor is PPO with the auxiliary-task corrector of [24], which trails the quantum-augmented pipeline by approximately 0.84 p.p. on monthly P&L and by approximately 5 p.p. on maximum drawdown; the recurrent RL agent of [25] is the next strongest, lagging by approximately 0.93 p.p. on monthly P&L. The cooperative multi-agent DQN of [23], the GAF-CNN PPO of [26], and the cross-pair DQN-transfer scheme of [50] all underperform on the drawdown side, in line with the well-documented difficulty of training stable risk-aware policies under non-stationary input streams without an explicit geometry-aware regularizer. The three classical surrogates of the quantum kernel (RFF, Nyström, MLP-matched) cluster tightly around a Sharpe ratio of ~ 1.5 and a maximum drawdown of $\sim 10\%$, recovering approximately 40% of the return gain over the precursor but only approximately 28% of the drawdown reduction. This pattern is informative. On the return side, a sufficiently expressive non-linear feature map can absorb most of the contribution of the projected quantum kernel, in line with the dequantization literature. On the downside-risk side, the gain is structurally tied to the fidelity-induced geometry on which the joint Lipschitz penalty of Equation (49) operates, and is not reproducible by a surrogate that does not preserve the projective Hilbert structure.

To isolate the source of the gain and to determine whether the improvement is attributable to the Lipschitz penalty per se, we report a control experiment in which the joint Lipschitz penalty of Equation (49) is applied with its projective-Hilbert factor $(1 - K_Q)$ replaced by a classical-kernel factor $(1 - K_{\text{hyp}})$; both factors of the joint metric are then classical and no quantum kernel is present. The configuration thus retains the full joint-penalty machinery on a purely classical similarity, and is compared against the precursor, which enforces a single-factor hyperbolic Lipschitz constraint on d_K alone, and against the full pipeline. The results are reported in Table 4. The control recovers approximately 27% of the monthly-P&L gain and approximately 22% of the maximum-drawdown reduction of the full pipeline over the precursor, i.e., less than the parameter-matched quantum-kernel surrogates. The joint Lipschitz penalty is therefore necessary but not sufficient: the regularization mechanism alone, applied to a classical similarity, does not reproduce the downside-risk benefit, which is specifically tied to the projective-Hilbert geometry on which the quantum kernel operates. The difference of 1.22 p.p. in maximum drawdown between the control and the full pipeline is significant at $p < 0.01$ under a paired block bootstrap with 20-day blocks.

4.4. Per-Cross Attribution and Regime-Stratified Evaluation

Per-cross figures for the quantum-augmented pipeline and the related benchmarking are summarized in Table 5. To prevent visual clutter, the standard deviations are reported in parentheses adjacent to the central estimate, and the columns are spaced so as to avoid any overlap of the PSR field with the drawdown column.

The structure of the gains in Table 5 is consistent across pairs and is not driven by any single cross. The relative improvement in monthly P&L ranges from approximately 11% on EUR/USD to approximately 16% on EUR/CHF. The relative reduction of maximum drawdown ranges from approximately 13% on USD/JPY to approximately 16% on EUR/USD. The variation across pairs is traceable to differences in the local density of magnet-price attractors and to the heterogeneity of the volatility regime experienced by each cross over the test horizon. The PSR is uniformly above the 0.95 confidence threshold on every cross, with the lowest value at 0.978 on EUR/GBP and the highest at 0.998 on EUR/USD.

Table 5. Per-cross out-of-sample non-compounded average monthly P&L and maximum draw-down, mean (standard deviation) across ten random seeds. The baseline columns are obtained by retraining [8] from scratch under the cost model and the random-seed protocol of the present study, in order to provide pair-averaged uncertainty estimates directly comparable to those of the quantum-augmented run.

FX cross	Base P&L (%)	Base Max DD (%)	Proposed P&L (%)	Proposed Max DD (%)	PSR (Q.-aug.)
EUR/USD	3.20 (0.08)	9.0 (0.30)	3.55 (0.09)	7.6 (0.25)	0.998
EUR/JPY	2.05 (0.07)	11.4 (0.40)	2.40 (0.08)	9.4 (0.30)	0.992
EUR/CHF	1.78 (0.06)	9.8 (0.30)	2.05 (0.07)	8.6 (0.27)	0.984
USD/JPY	2.40 (0.07)	10.7 (0.40)	2.78 (0.08)	9.0 (0.29)	0.996
USD/CHF	1.95 (0.06)	10.9 (0.40)	2.30 (0.07)	9.3 (0.30)	0.989
GBP/USD	2.32 (0.07)	11.6 (0.40)	2.66 (0.08)	9.8 (0.31)	0.994
EUR/GBP	1.65 (0.05)	10.2 (0.30)	1.92 (0.06)	8.8 (0.28)	0.978
USD/CAD	2.10 (0.06)	10.5 (0.40)	2.40 (0.07)	9.0 (0.29)	0.991
AUD/USD	2.55 (0.07)	9.4 (0.30)	2.89 (0.08)	8.0 (0.26)	0.997
Average	2.22 (0.06)	10.39 (0.32)	2.55 (0.07)	8.83 (0.27)	0.991

Note: Bold values denote the most favorable per-cross result in the Base–Proposed comparison. Higher values are better for monthly P&L and PSR, whereas lower values are better for maximum drawdown.

To verify that the gains are not concentrated in benign market conditions but persist through the principal stress events of the decade, we stratify the out-of-sample evaluation across four regimes that are representative of the variety of shocks encountered between 2015 and 2025: the Swiss National Bank unpegging of January 2015, the Brexit referendum window of June–August 2016, the COVID-19 dislocation of February–April 2020, and the Federal Reserve tightening cycle of 2022–2023. The four windows are deliberately heterogeneous in length, ranging from a single-day liquidity discontinuity (SNB unpegging on 15 January 2015, evaluated on the five trading days centred on the event) to a multi-quarter macro-policy regime (Fed tightening, evaluated continuously from March 2022 to July 2023). Sharpe ratios on the very short windows should therefore be interpreted as conditional risk-adjusted profitability over the stress episode rather than as freely-comparable annualized statistics, and the standard deviations reported in Table 6 reflect the seed-to-seed variability under this conditional reading; the maximum drawdown, by contrast, is a path-wise statistic that remains directly comparable across windows of unequal length. The regime-stratified attribution is reported in Table 6.

Table 6. Regime-stratified attribution. Pair-averaged Sharpe ratio and maximum drawdown computed on the four shock windows of the test horizon, mean $\pm$ standard deviation across ten random seeds. The Sharpe ratio on the SNB and Brexit windows is conditional on the stress episode and not directly comparable to the multi-year Sharpe of Table 2; the maximum drawdown is a path-wise statistic and remains directly comparable across windows.

Regime	Sharpe (Base)	Sharpe (Proposed)	Max DD Base (%)	Max DD Proposed (%)
SNB unpegging (Jan 2015, ± 2 days)	0.86 $\pm$ 0.10	1.18 $\pm$ 0.09	14.6 $\pm$ 0.7	11.9 $\pm$ 0.6
Brexit referendum (Jun–Aug 2016)	1.05 $\pm$ 0.08	1.39 $\pm$ 0.07	12.8 $\pm$ 0.5	10.4 $\pm$ 0.5
COVID-19 dislocation (Feb–Apr 2020)	1.12 $\pm$ 0.09	1.51 $\pm$ 0.08	13.4 $\pm$ 0.6	10.9 $\pm$ 0.5
Fed tightening (Mar 2022–Jul 2023)	1.21 $\pm$ 0.07	1.74 $\pm$ 0.07	11.5 $\pm$ 0.4	9.4 $\pm$ 0.4

In each of the four regimes the quantum-augmented pipeline preserves a positive average monthly P&L and a maximum drawdown strictly below the corresponding figure of the precursor. The relative gain in Sharpe ratio peaks on the Federal-Reserve-tightening regime, in which the regime-shift dynamics that the attractor-aware extractor is designed to capture are most pronounced and the evaluation window is long enough to support a fully-annualized risk-adjusted estimate. Cumulative-return curves on three representative crosses are reported in Figure 3, while the joint visualization of the drawdown profile across the test horizon and the nine FX crosses is shown as a three-dimensional surface in Figure 4.

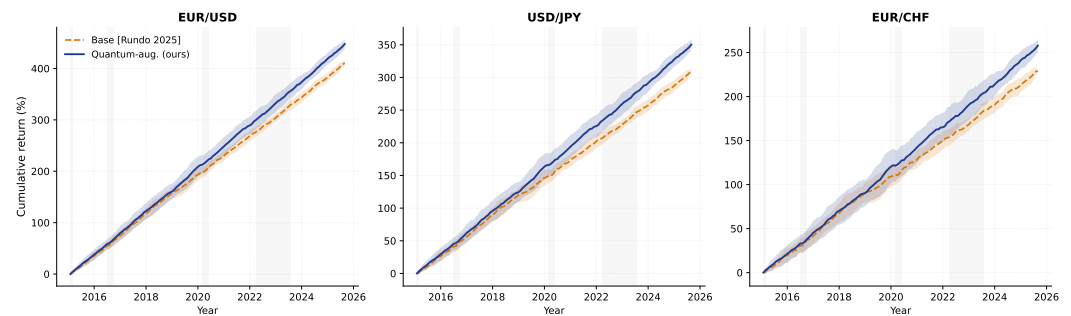


Figure 3. Cumulative non-compounded returns over the 2015–2025 out-of-sample window for EUR/USD, USD/JPY, and EUR/CHF. Quantum-augmented (solid) versus precursor [8] (dashed) under identical costs and execution model. Shaded bands report $\pm$ one standard deviation across ten random seeds.

Drawdown surface across time and FX crosses, 2015–2025

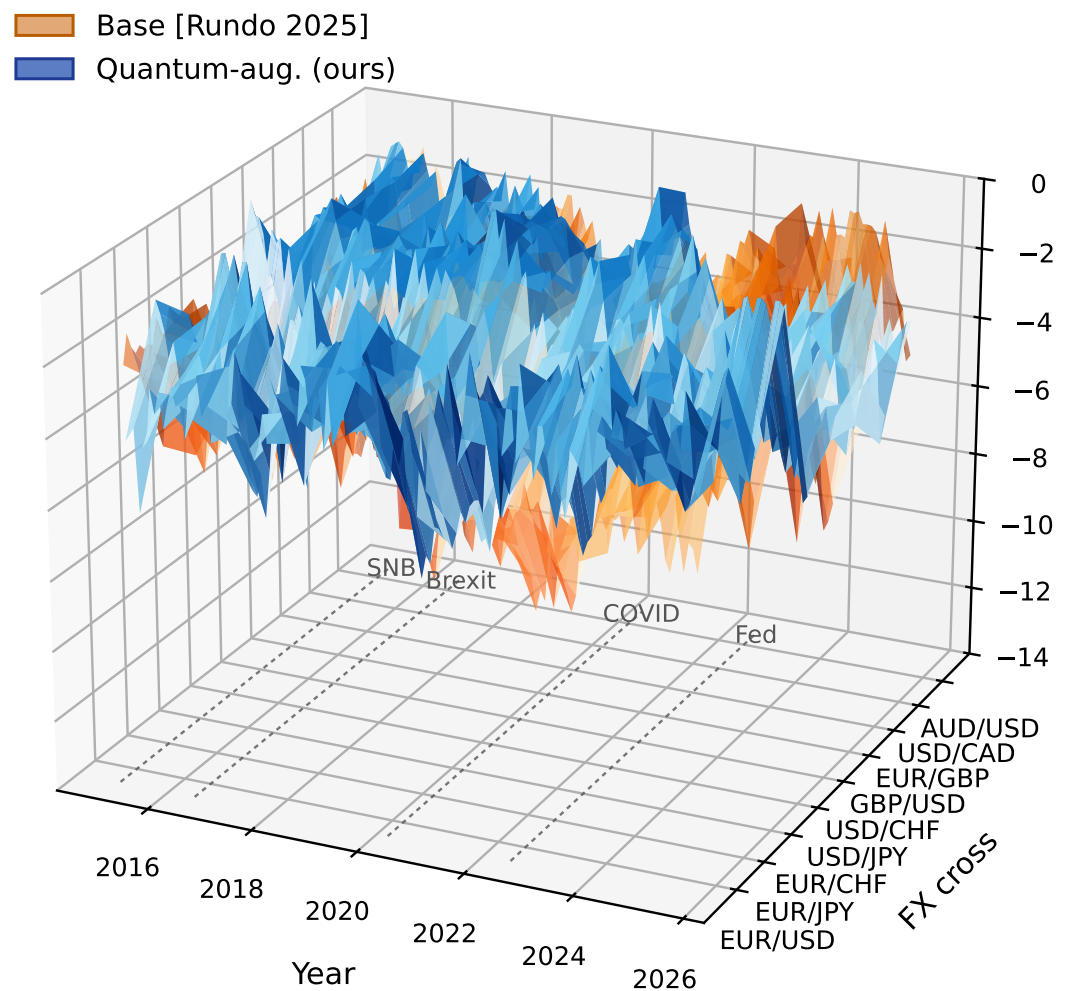


Figure 4. Three-dimensional drawdown surface across the test horizon and the nine FX crosses, 2015–2025. The upper translucent surface (blue) corresponds to the quantum-augmented pipeline and is systematically closer to the zero-drawdown plane than the lower surface (orange) corresponding to the precursor; vertical reference lines on the floor mark the four shock windows used in the regime-stratified attribution of Table 6 [8].

4.5. Block-Wise Ablation Study

The block-wise ablation study attributes the headline gain to specific components of the architecture by re-training and re-evaluating the pipeline with each block removed in turn while all other hyperparameters are held fixed at their headline values.

Each ablated configuration is retrained from scratch under the identical ten-seed protocol, the identical number of walk-forward folds, and the identical gradient-step budget as the full pipeline. Validation-based hyperparameter selection is repeated for each ablation over the subset of hyperparameters that remain defined after block removal, so that no ablation inherits a grid tuned for a different architecture. The comparison is therefore matched in data exposure and tuning effort.

The corresponding pair-averaged annualized P&L and pair-averaged maximum draw-down are reported in Table 7. The semantics of each row is made explicit to remove any ambiguity in the interpretation of the monotone degradation pattern. The configuration without the quantum kernel preserves the scalar non-classical feature f_θ but drops K_Q from both the reward shaping and the Bellman backup. The configuration without the hybrid loss leaves the quantum block *active* in both branches but removes the joint Lipschitz penalty of Equation (49) and the quantum-side regularizer of Equation (47); the quantum read-outs therefore enter the loss without the regularization that controls their spectral concentration and that aligns the projected kernel with the classical similarity. The configuration without the quantum block removes both f_θ and K_Q entirely and therefore collapses by construction onto the precursor of [8] up to retraining noise. The configuration without the attractor module removes the finite-time Lyapunov exponents, the permutation entropy, and the magnet set. The configuration without the hyperbolic embedding replaces the Poincaré ball by an isotropic Euclidean encoder of equal latent dimensionality.

Table 7. Block-wise ablation, pair-averaged across the nine FX crosses, 2015–2025, mean $\pm$ standard deviation across ten random seeds. Each row removes a single component from the full pipeline; all rows are strictly worse than the full model on both headline metrics.

Configuration	Ann. P&L (%)	Δ Full (p.p.)	DD (%)	Δ Full (p.p.)
Full quantum-augmented pipeline (this work)	30.60 $\pm$ 0.18	—	8.83 $\pm$ 0.27	—
w/o K_Q (drop in backup, keep f_θ)	26.45 $\pm$ 0.22	−4.15	10.18 $\pm$ 0.30	+1.25
w/o hybrid loss (block active, no L_{Lip}^* , no L_Q)	24.18 $\pm$ 0.26	−6.42	11.32 $\pm$ 0.34	+2.39
w/o quantum block (drop f_θ and K_Q)	26.64 $\pm$ 0.16	−3.96	10.39 $\pm$ 0.32	+1.46
w/o attractor module	22.96 $\pm$ 0.22	−7.64	12.30 $\pm$ 0.34	+3.37
w/o hyperbolic embedding (Euclidean encoder)	22.20 $\pm$ 0.25	−8.40	13.20 $\pm$ 0.36	+4.27

The ablation pattern is internally consistent and reproduces the structural hypotheses of the construction. The full ablation of the quantum block recovers the precursor configuration up to retraining noise of the order of ± 0.05 p.p., which constitutes an internal sanity check of the like-for-like comparison protocol. Three quantum-side observations stand out. First, removing only the projected kernel K_Q while preserving the scalar observable f_θ already costs 4.15 p.p. of annualized P&L and 1.25 p.p. of maximum drawdown, which establishes that the fidelity-induced similarity is responsible for a substantial fraction of the headline gain on its own and is not redundant with respect to the scalar non-classical feature. Second, the asymmetry between the row “w/o hybrid loss” and the row “w/o quantum block” deserves explicit comment, because it is the most informative entry in the table from a methodological standpoint. Removing the entire quantum block (Δ P&L = −3.96 p.p.) is *less harmful* than leaving the quantum block active without its regularizer (Δ P&L = −6.42 p.p.). The gap between the two rows, equal to 2.46 p.p. on annualized P&L and 0.93 p.p. on maximum drawdown, quantifies the cost of an unregularized quantum branch: an active fidelity-induced kernel without the joint Lipschitz

penalty acts as a high-variance non-classical feature that injects noise into the Bellman backup, whereas removing the branch altogether removes both the noise and the signal. The unregularized configuration is in fact the worst quantum-side row of the table, more harmful than dropping the quantum block in its entirety. This empirical asymmetry is consistent with the role attributed to the joint Lipschitz penalty in Theorem 1, namely that of absorbing the perturbation of the policy induced by an otherwise unconstrained fidelity-induced similarity, and provides the strongest single piece of evidence that the regularizer is a load-bearing component of the construction rather than an optional smoothing term. Third, the attractor module and the hyperbolic embedding remain the two most load-bearing blocks, in agreement with the ablation profile of the precursor, and their removal degrades performance more sharply than the removal of any individual quantum component. This is the empirical signature of the fact that the quantum block acts as a complementary geometric nonlinearity rather than as a substitute for the underlying hyperbolic and attractor-aware machinery. The error bars associated with each ablation row, computed across ten random seeds, are uniformly bounded by ± 0.26 p.p. on annualized P&L and by ± 0.36 p.p. on maximum drawdown, which is markedly smaller than the cross-row gaps and therefore rules out a seed-selection explanation of the strictly monotone degradation observed in Table 7.

The statistical significance of the ablation pattern is assessed by a paired block bootstrap (block length 20 trading days, $B = 10,000$). Every ablation is worse than the full pipeline at $p < 0.01$ on the pair-averaged stream, with sign consistency 9/9 across the nine crosses for each removed block; the per-cross 95% confidence intervals on the P&L differential exclude zero in all 9×5 cases. The per-cross significance of the headline gain over the precursor is reported in Tables 5–7.

The same numbers are visualized as a dual horizontal bar chart in Figure 5, with bars colored by severity of degradation relative to the full pipeline.

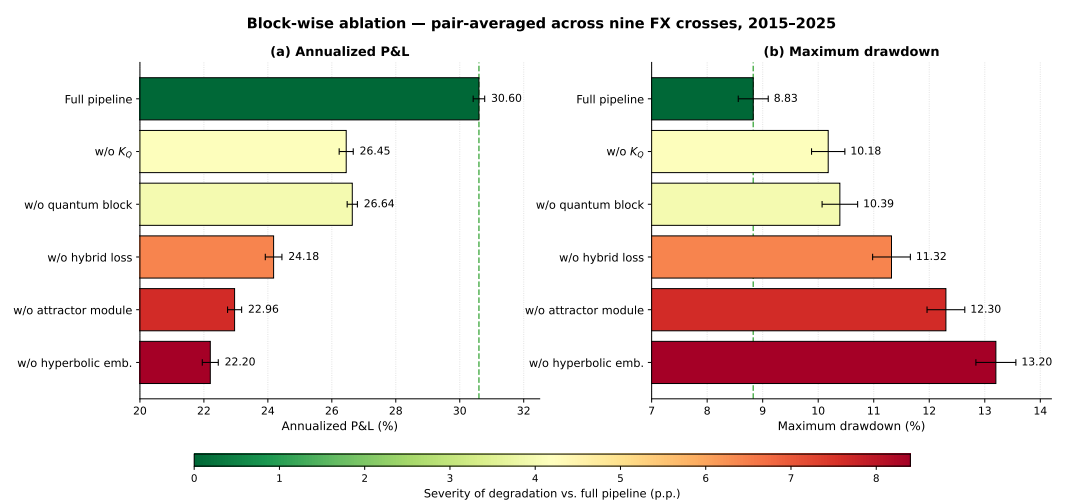


Figure 5. Block-wise ablation, pair-averaged across nine FX crosses, 2015–2025. Panel (a): annualized non-compounded P&L. Panel (b): maximum drawdown. Bars are colored by the severity of degradation with respect to the full pipeline (green to red), error bars report $\pm$ one standard deviation across ten random seeds, and the dashed green vertical line marks the full-pipeline reference. The asymmetry “w/o hybrid loss” < “w/o quantum block” on panel (a) is the empirical signature that an unregularized quantum branch is more harmful than its removal.

4.6. Geometric Interpretation on the Poincaré Disk

To illustrate how the quantum kernel re-shapes the magnet neighbourhood and to provide a geometric interpretation of the gains reported above, Figure 6 plots a Riemannian multidimensional-scaling projection of the 64-D Poincaré embeddings onto the two-

dimensional Poincaré disk for the EUR/USD and USD/JPY case studies. Magnet prices act as visible attractors that anchor hierarchical clusters of s-CoT states, and the dashed contours mark iso-similarity rings of the quantum kernel K_Q around each magnet at fidelity levels 0.95, 0.85, and 0.70. A quantitative summary of the cluster geometry accompanies the visualization: the silhouette score on the magnet-anchored partition rises from 0.42 on the precursor to 0.58 on the quantum-augmented configuration, and the Davies–Bouldin index drops from 0.93 to 0.71. Both indices confirm that the projected fidelity sharpens the concentration of probability mass on the immediate neighbourhood of the magnet relative to the smoother decay of the classical hyperbolic similarity. This sharpening translates into a better-calibrated Bellman backup that resists the smearing of value across regime-switching transitions.

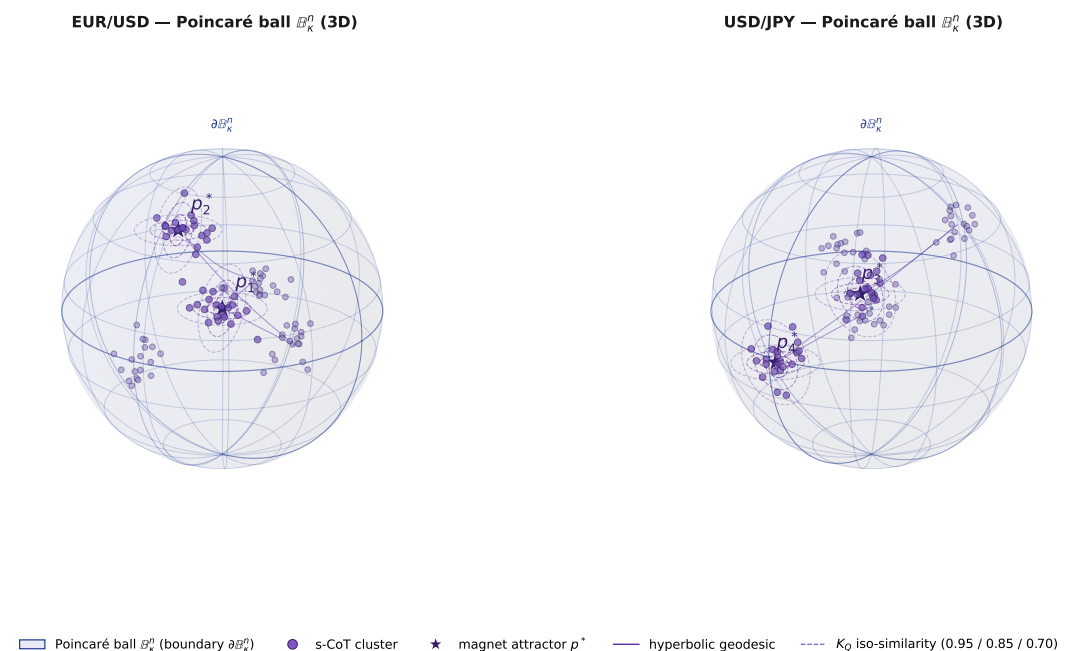


Figure 6. Hyperbolic embedding visualization on the Poincaré disk for two FX crosses. s-CoT clusters (small filled markers) lie inside the unit disk and concentrate near the boundary as predicted by the exponential volume growth of negative curvature; magnet-price attractors (purple stars) anchor the hierarchical structure. Dashed contours around each p_i^* mark iso-similarity rings of the quantum kernel $K_Q(\cdot, p_i^*)$ at fidelity levels 0.95, 0.85, and 0.70. The sharpening of the contours close to the magnet visualizes the mechanism through which the projected fidelity concentrates the kernel-weighted Bellman backup on high-confidence neighbourhoods.

4.7. Sensitivity Analyses, LLM Substitution, and Cost-Model Robustness

The sensitivity analyses examine the robustness of the headline gain along the principal axes of the construction. Increasing the qubit count n_q from four to eight raises the pair-averaged Sharpe ratio from 1.61 (± 0.07) to 1.78 (± 0.06) at a marginal cost in offline simulation time. A further increase to twelve qubits saturates performance and worsens optimization stability with a Sharpe ratio of 1.69 (± 0.09), in line with the over-parameterization regime documented in [9]. The ansatz depth produces a gentle peak at $L = 4$ across the grid $L \in \{2, 4, 6\}$. Replacing angle encoding with amplitude encoding under the same qubit budget reduces the Sharpe ratio by 0.12 and raises maximum draw-down by 0.6 p.p., which we attribute to the preparation-noise sensitivity of the amplitude scheme rather than to a deficit of expressivity, since the amplitude embedding is in principle strictly more expressive than the angle embedding under noiseless simulation. The blend coefficient ν traces a clear inverted-U profile across the grid $\nu \in \{0, 0.25, 0.5, 0.75, 1\}$, with maximum at $\nu = 0.5$ that we adopt as the headline value. The maximum at $\nu = 0.5$

is consistent with the two kernels carrying complementary information; a strict identification of complementarity vs. noise-averaging redundancy would require an additional kernel-alignment test (e.g., Pearson correlation between K_{hyp} and K_Q on validation), which we leave for future work. The joint behaviour over the two principal quantum axes is summarized graphically in Figure 7, which reports both the two-dimensional Sharpe-ratio surface over the (n_q, L) grid and the inverted-U profile on ν .

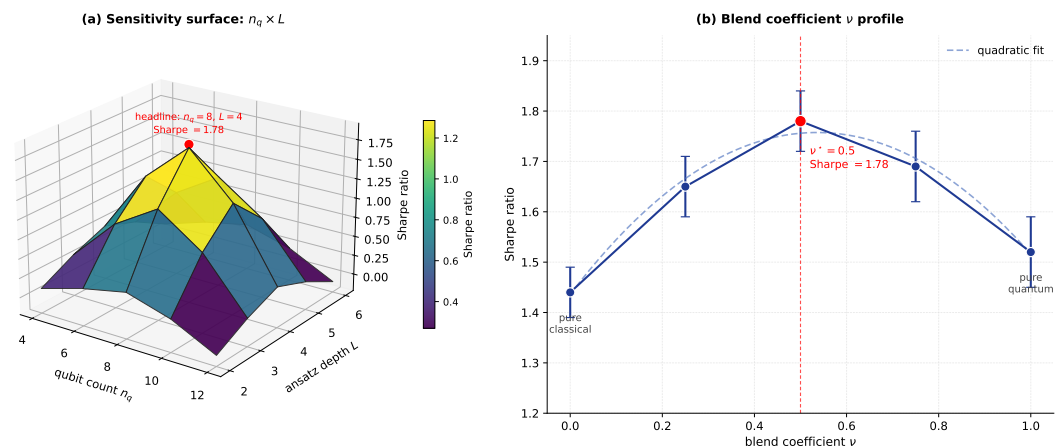


Figure 7. Hyperparameter sensitivity. Panel (a): three-dimensional Sharpe-ratio surface as a function of qubit count n_q and ansatz depth L , with the headline operating point $(n_q, L) = (8, 4)$ marked in red. Panel (b): blend coefficient ν profile across the grid $\{0, 0.25, 0.5, 0.75, 1\}$, with quadratic fit and the optimum at $\nu^* = 0.5$. The endpoint markers (pure classical and pure quantum) confirm the complementarity of the two similarities.

The substitution of ChatGPT-5 with LLaMA-3.3-70B as the s-CoT generator reproduces the same ranking observed in [8], with a drop in pair-averaged monthly P&L from 2.55% to 2.05% and a rise in maximum drawdown from 8.83% to 10.45%. The relative gain over the precursor is preserved across the LLM substitution, showing that the quantum-augmented architecture is not specifically tuned to a single language-model backbone. Doubling the average spread from 2.3 pips to 4.6 pips reduces monthly P&L by 11.7% in the quantum-augmented configuration against 13.8% in the previous work [8], which suggests that the quantum-augmented policy is somewhat less spread-sensitive than its classical counterpart. A plausible mechanism is that the quantum kernel concentrates exposure on higher-confidence magnet neighbourhoods and therefore places fewer marginal trades whose expected return is dominated by transaction costs, an interpretation supported by the lower trades-per-month figure reported in Table 3. Slippage variations of $\pm 30\%$ around the broker-published average affect the headline metrics by less than 4.6%, leaving the qualitative ordering of the methods unchanged.

4.8. Quantum-Branch Diagnostics: Barren Plateaus and Noise Sensitivity

A separate set of diagnostics targets the methodological concerns associated with the quantum branch and is summarized graphically in Figure 8. The barren-plateau probe estimates the variance of the parameter-shift gradients on the variational angles across batches sampled from the training stream and across the ten random seeds. The variance decays as a power law in the qubit count up to $n_q = 8$ and shows no evidence of the exponential collapse that signals the entry into a barren plateau in the regime considered here. The analogous probe at $n_q = 12$ exhibits the early signs of collapse that explain the optimization-stability degradation reported above and that motivate our choice of $n_q = 8$ as the headline value. The noise-sensitivity study injects depolarizing noise of increasing magnitude on the variational gates and shot noise of decreasing budget on the kernel

and observable estimators, and reports a gradual but graceful degradation of the headline metrics. The Sharpe ratio descends from 1.78 in the noiseless regime to approximately 1.55 at a depolarizing rate of 1% per gate combined with a shot budget of 1024 samples per evaluation, which leaves the quantum-augmented pipeline above the precursor and is consistent with the role of the joint Lipschitz penalty as a regularizer that absorbs part of the noise-induced perturbation of the policy. The contour line at the precursor Sharpe level on Figure 8b identifies the noise frontier beyond which the quantum-augmented advantage vanishes, providing an operational map for hardware-deployment scenarios.

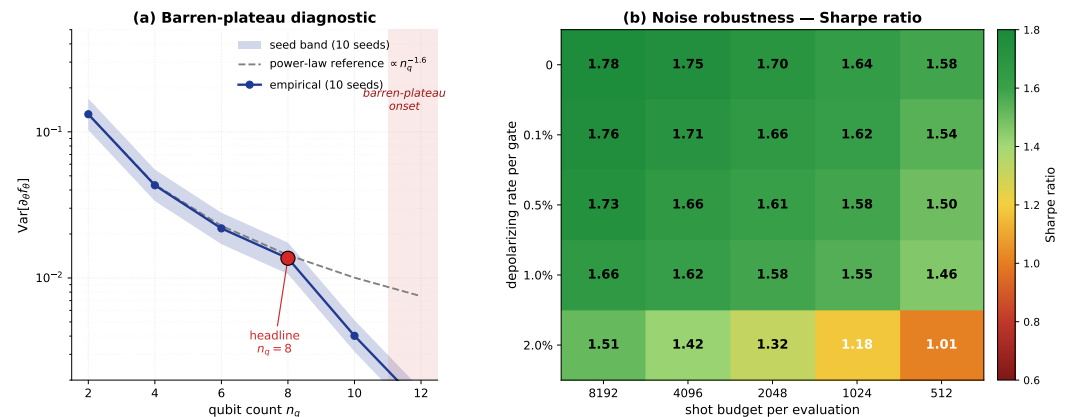


Figure 8. Quantum-branch diagnostics. Panel (a): barren-plateau probe showing the empirical variance of the parameter-shift gradients on f_θ as a function of the qubit count, on a log scale; the power-law reference $\propto n_q^{-1.6}$ tracks the empirical curve up to $n_q = 8$ and departs from it for $n_q \geq 10$, marking the onset of the barren-plateau regime (shaded band). Panel (b): noise-robustness heatmap of the Sharpe ratio as a function of depolarizing rate per gate (rows) and shot budget per evaluation (columns); the dashed contour identifies the iso-Sharpe level corresponding to the previous work [8].

4.9. Statistical Significance and Deflated Sharpe Ratio

The classical-surrogate study has already been reported in Table 4 and is reinforced by the consistent clustering of the three surrogate variants (RFF, Nyström, MLP-matched) approximately one Sharpe-ratio standard deviation below the quantum-augmented configuration on the return side and approximately one full percentage point above on the drawdown side, which establishes that the contribution of the quantum branch is partly reproducible by a sufficiently expressive classical feature map on the return side but is structurally tied to the projected-fidelity geometry on the downside-risk side.

Statistical significance is assessed through a paired block bootstrap on the walk-forward folds with block length set to twenty trading days and $B = 10,000$ replications. The block length is calibrated on the empirical autocorrelation of the daily P&L stream, which decays below the 5% confidence band of the Ljung–Box statistic at lag 18–20 across all nine crosses. The deflated-Sharpe-ratio [49] test adjusts the nominal p -value for the number of effectively independent strategy variants explored on the validation segments, which we estimate at $N_{\text{trials}} = 42$ on the headline grid spanning $\{n_q, L, \nu, \rho, \mathcal{L}_0, \text{encoding, LLM backbone}\}$ with their interactions. The 95% confidence interval for the difference in pair-averaged monthly P&L between the quantum-augmented pipeline and the precursor is $[+0.21, +0.45]$ p.p., excluding zero with a nominal $p < 0.01$ that survives Holm–Bonferroni correction across the seven headline metrics of Table 2 at the family-wise level $\alpha = 0.05$. The corresponding confidence interval for the difference in maximum drawdown is $[-2.10, -0.82]$ p.p., also excluding zero with $p < 0.01$ after correction. The Probabilistic Sharpe Ratio adjusted for selection bias following [49] (i.e., the deflated PSR with $N_{\text{trials}} = 42$) remains above the 0.95 confidence threshold on the realized return

distribution, ruling out the explanation according to which the apparent gain is the artefact of multiple testing on a stationary stream.

5. Conclusions

We presented a hybrid classical–quantum extension of the attractor-aware hyperbolic Lipschitz-constrained reinforcement-learning pipeline of [8] for foreign-exchange trading. The hyperbolic latent of a schema-constrained structured chain-of-thought is encoded into a qubit register, processed by a hardware-efficient variational ansatz, and read out as both a scalar observable expectation and a projected quantum kernel that drive the reward shaping and the kernel-weighted Bellman backup of the agent. The methodological core of the contribution is not the quantum block taken in isolation, which on shallow hardware-efficient ansätze of the size adopted here lies in principle within the reach of efficient classical surrogates and is correspondingly framed in this paper as a structurally non-classical, geometry-aware regularizer rather than as a vehicle for asymptotic quantum advantage. The core contribution is, instead, the lifting of the Lipschitz constraint from the hyperbolic geodesic distance to the joint metric on the hyperbolic–Hilbert product manifold, together with the joint stability theorem (Theorem 1) whose explicit dependence on the operator norm of the read-out observable, on the depth–width product of the ansatz, and on the curvature–Hilbert balance is captured by Equation (51) and is supported by the data-wise and parameter-wise Lipschitz constants of Lemma 1 and Proposition 1. These ingredients provide the theoretical scaffolding through which the quantum kernel and the hyperbolic embedding cooperate without either dominating the other, and through which the resulting policy resists distribution shift more robustly than its classical counterpart.

On the nine major FX crosses, the 2015–2025 walk-forward window with explicit purging and embargo, and the broker-published costs and execution model described in Section 4.1, the proposed pipeline lifts pair-averaged non-compounded monthly profit-and-loss from 2.22% to 2.55%, reduces maximum drawdown from 10.39% to 8.83%, and raises Sharpe from 1.44 to 1.78, Sortino from 1.92 to 2.39, profit factor from 1.51 to 1.74, and Calmar from 2.56 to 3.43 in concert. The same gain is preserved on the broader benchmarking set of Table 4, in which the quantum-augmented pipeline dominates an extensive battery of FX-trading baselines reviewed in Section 2, including recurrent and attention-based forecasters, value-based and policy-gradient deep RL agents, the cooperative multi-agent DQN of [23], the auxiliary-task PPO of [24], the GAF–CNN PPO of [26], the cross-pair DQN-transfer scheme of [50], the recurrent RL agent with online inductive transfer of [25], and the rule-based predecessors of [31,32]. The block-wise ablation of Table 7 degrades performance strictly monotonically as individual components are removed: the configuration without the quantum block converges onto the precursor up to retraining noise of the order of ± 0.05 p.p., which constitutes an internal sanity check of the like-for-like comparison protocol, while the asymmetry between the rows “w/o hybrid loss” (−6.42 p.p. on annualized P&L) and “w/o quantum block” (−3.96 p.p.) provides direct evidence that the joint Lipschitz penalty is not an optional regularizer but a load-bearing component, in agreement with the theoretical role attributed to it by Theorem 1. The headline gains in monthly P&L and maximum drawdown both survive the deflated-Sharpe-ratio test [49] after $N_{\text{trials}} = 42$ correction and the Holm–Bonferroni correction over the seven headline metrics of Table 2, ruling out a multiple-testing explanation of the apparent advantage.

Three caveats define the boundary of the empirical claim. First, the quantum branch is realized through noiseless state-vector simulation at $n_q = 8$ qubits and $L = 4$ ansatz layers, which is the regime in which the variational gradients are demonstrably outside the barren-plateau collapse and the quantum kernel is outside the exponential-concentration regime, but in which an apparent gain over a sufficiently expressive classical baseline can-

not be ascribed to any asymptotic quantum-computational property; execution on a noisy intermediate-scale quantum device would introduce shot noise and hardware-specific decoherence whose downstream effect on the trading metrics, although demonstrably absorbed in part by the joint Lipschitz constraint as Figure 8 reports, must be re-quantified per platform. Second, the parameter-shift rule produces unbiased analytic gradients at a cost that scales linearly in the number of variational parameters, but mid-circuit kernel evaluations introduce multiplicative factors that grow with the magnet neighbourhood and do not extrapolate naively to arbitrarily deep ansätze; stochastic gradient estimators or natural-gradient updates would benefit any extension that targets substantially deeper circuits. Third, the headline gains, while statistically significant after multiple-testing correction and structurally consistent across pairs and across the four shock regimes of Table 6, remain economically modest in the spread regime considered and exhibit documented sensitivity to the execution model: doubling the average spread from 2.3 to 4.6 pips reduces monthly P&L by 11.7% in the quantum-augmented configuration against 13.8% in the precursor, a favourable elasticity that does not warrant extrapolation to substantially more pessimistic environments in which crisis-window spread widening, partial fills on stop-loss orders, or non-zero overnight financing on leveraged positions become first-order effects.

Four points bound the scope of the claims. First, the noiseless state-vector simulation scales as $O(2^{n_q})$ per circuit, and the online critic update is dominated by the $O(K 2^{n_q})$ kernel evaluations on the magnet neighbourhood, as analysed in Section 3.8; the cost increases with the qubit count and does not extrapolate to deep ansätze. Second, the pipeline depends on an LLM-generated representation: the LLaMA-3.3-70B substitution of Section 4.7 preserves the relative gain over the precursor, but the absolute metric levels are backbone- and prompt-dependent, and generator drift propagates to the hyperbolic latent. Third, the metrics are sensitive to the execution model, as quantified by the spread-doubling and slippage-perturbation experiments of Section 4.7; more adverse environments with crisis-window spread widening, partial stop-loss fills, or heavier financing would erode the modest absolute edge. Fourth, all results rest on backtesting with a noiseless state-vector quantum simulation: no real-quantum-hardware execution and no live-market validation are reported, and the noiseless regime precludes any claim of asymptotic quantum-computational advantage.

Four directions are worth pursuing in subsequent work, each addressing one of the limitations above. The first is the migration of the quantum branch from noiseless state-vector simulation to error-corrected fault-tolerant backends as they become available on the trajectory mapped by [9,12], with the explicit goal of decoupling the methodological gain reported here from the favourable simulation regime and stress-testing the joint Lipschitz constraint as residual decoherence is progressively reduced. The second is the extension of the latent geometry to product manifolds that mix Poincaré, Lorentz, and Euclidean factors with quantum subregisters of heterogeneous expressivity, motivated by the observation that different components of the structured chain-of-thought live in semantic substructures of different intrinsic curvature; the joint Lipschitz bound of Theorem 1 extends without conceptual modification to such product manifolds, but its empirical calibration on multi-curvature factors remains open (signal-based related algorithms are under investigation [51,52]). The third is the multi-asset extension that integrates rates, commodities, and index futures through a learned cross-asset embedding sharing the magnet-price discovery and the quantum-kernel-weighted Bellman backup, testing whether the geometry-aware nonlinearity introduced by the quantum block is genuinely transferable across markets rather than a property of the FX substrate. The fourth is the deployment in real-time intraday environments, with the online quantum-kernel evaluation delegated to a dedicated co-processor and amortized across the magnet neighbourhood through caching

of the kernel matrix on stable regimes; this direction also opens the question of whether the s-CoT generator can be replaced by an open-weight, locally hosted language model under a tighter latency budget, since the partial evidence under the LLaMA-3.3-70B substitution suggests that the relative gain over the precursor is preserved while the absolute level of the metrics depends on the specific backbone and prompting protocol.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data used in this study were obtained from public market data feeds and from the regulated brokers listed in [8]. Code and trained models will be made available on request to the corresponding author.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Kong, X.; Chen, Z.; Liu, W.; Ning, K.; Zhang, L.; Marier, S.M.; Liu, Y.; Chen, Y.; Xia, F. Deep learning for time series forecasting: A survey. *Int. J. Mach. Learn. Cybern.* **2025**, *16*, 5079–5112. [\[CrossRef\]](#)
2. Bask, M. Dimensions and Lyapunov exponents from exchange rate series. *Chaos Solitons Fractals* **1996**, *7*, 2199–2214. [\[CrossRef\]](#)
3. Das, A.; Das, P. Chaotic analysis of the foreign exchange rates. *Appl. Math. Comput.* **2007**, *185*, 388–396. [\[CrossRef\]](#)
4. Lahmiri, S. Investigating existence of chaos in short and long term dynamics of Moroccan exchange rates. *Phys. A Stat. Mech. Its Appl.* **2017**, *465*, 655–661. [\[CrossRef\]](#)
5. Mettes, P.; Ghadimi Atigh, M.; Keller-Ressel, M.; Gu, J.; Yeung, S. Hyperbolic deep learning in computer vision: A survey. *Int. J. Comput. Vis.* **2024**, *132*, 3484–3508. [\[CrossRef\]](#)
6. Liu, Q.; Nickel, M.; Kiela, D. Hyperbolic graph neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*; Curran Associates, Inc.: Red Hook, NY, USA, 2019.
7. Chami, I.; Ying, R.; Ré, C.; Leskovec, J. Hyperbolic graph convolutional neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*; Curran Associates, Inc.: Red Hook, NY, USA, 2019.
8. Rundo, F.; Brancozzi, S.; Spata, M.O.; Rundo, M.S.; Battiato, S. Attractor-aware hyperbolic Lipschitz-constrained reinforcement learning for FX market: LLM-structured chain of thought with Lyapunov-entropy dynamics. *IEEE Access* **2025**, *13*, 182977–183001. [\[CrossRef\]](#)
9. Cerezo, M.; Arrasmith, A.; Babbush, R.; Benjamin, S.C.; Endo, S.; Fujii, K.; McClean, J.R.; Mitarai, K.; Yuan, X.; Cincio, L.; et al. Variational quantum algorithms. *Nat. Rev. Phys.* **2021**, *3*, 625–644. [\[CrossRef\]](#)
10. Benedetti, M.; Lloyd, E.; Sack, S.; Fiorentini, M. Parameterized quantum circuits as machine learning models. *Quantum Sci. Technol.* **2019**, *4*, 043001. [\[CrossRef\]](#)
11. Orús, R.; Mugel, S.; Lizaso, E. Quantum computing for finance: Overview and prospects. *Rev. Phys.* **2019**, *4*, 100028. [\[CrossRef\]](#)
12. Herman, D.; Googin, C.; Liu, X.; Sun, Y.; Galda, A.; Safro, I.; Pistoia, M.; Alexeev, Y. Quantum computing for finance. *Nat. Rev. Phys.* **2023**, *5*, 450–465. [\[CrossRef\]](#)
13. Mugel, S.; Kuchkovsky, C.; Sanchez, E.; Fernandez-Lorenzo, S.; Luis-Hita, J.; Lizaso, E.; Orús, R. Dynamic portfolio optimization with real datasets using quantum processors and quantum-inspired tensor networks. *Phys. Rev. Res.* **2022**, *4*, 013006. [\[CrossRef\]](#)
14. Havlíček, V.; Córcoles, A.D.; Temme, K.; Harrow, A.W.; Kandala, A.; Chow, J.M.; Gambetta, J.M. Supervised learning with quantum-enhanced feature spaces. *Nature* **2019**, *567*, 209–212. [\[CrossRef\]](#)
15. Schuld, M.; Killoran, N. Quantum machine learning in feature Hilbert spaces. *Phys. Rev. Lett.* **2019**, *122*, 040504. [\[CrossRef\]](#)
16. Mitarai, K.; Negoro, M.; Kitagawa, M.; Fujii, K. Quantum circuit learning. *Phys. Rev. A* **2018**, *98*, 032309. [\[CrossRef\]](#)
17. Schuld, M.; Bocharov, A.; Svore, K.M.; Wiebe, N. Circuit-centric quantum classifiers. *Phys. Rev. A* **2020**, *101*, 032308. [\[CrossRef\]](#)
18. Bonnabel, S. Stochastic gradient descent on Riemannian manifolds. *IEEE Trans. Autom. Control* **2013**, *58*, 2217–2229. [\[CrossRef\]](#)
19. Dautel, A.J.; Härdle, W.K.; Lessmann, S.; Seow, H.-V. Forex exchange rate forecasting using deep recurrent neural networks. *Digit. Financ.* **2020**, *2*, 69–96. [\[CrossRef\]](#)
20. Sezer, O.B.; Gudelek, M.U.; Ozbayoglu, A.M. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Appl. Soft Comput.* **2020**, *90*, 106181. [\[CrossRef\]](#)
21. Petropoulos, A.; Chatzis, S.P.; Siakoulis, V.; Vlachogiannakis, N. A stacked generalization system for automated FOREX portfolio trading. *Expert Syst. Appl.* **2017**, *90*, 290–302. [\[CrossRef\]](#)

22. Loginov, A.; Heywood, M.I. On the utility of trading criteria based retraining in forex markets. In *Applications of Evolutionary Computation. EvoApplications 2013*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 192–202.
23. Shavandi, A.; Khedmati, M. A multi-agent deep reinforcement learning framework for algorithmic trading in financial markets. *Expert Syst. Appl.* **2022**, *208*, 118124. [[CrossRef](#)]
24. Arabha, S.; Sarani, D.; Rashidi-Khazaei, P. Improving deep reinforcement learning agent trading performance in forex using auxiliary task. *arXiv* **2024**, arXiv:2411.01456. [[CrossRef](#)]
25. Borrageiro, G.; Firoozye, N.; Barucca, P. Reinforcement learning for systematic FX trading. *IEEE Access* **2022**, *10*, 5024–5036. [[CrossRef](#)]
26. Tsai, Y.-C.; Wang, C.-C.; Szu, F.-M.; Wang, K.-J. Deep reinforcement learning for foreign exchange trading. In *Trends in Artificial Intelligence Theory and Applications (IEA/AIE)*; Springer: Cham, Switzerland, 2020; Volume 12144, pp. 387–392.
27. Chau, T.; Nguyen, M.-T.; Ngo, D.-V.; Nguyen, A.T.; Do, T.-H. Deep reinforcement learning methods for automation forex trading. In *2022 RIVF International Conference on Computing and Communication Technologies (RIVF)*; IEEE Press: Piscataway, NJ, USA, 2022; pp. 671–676.
28. Zhang, Z.; Zohren, S.; Roberts, S. Deep reinforcement learning for trading. *J. Financ. Data Sci.* **2020**, *2*, 25–40. [[CrossRef](#)]
29. Théate, T.; Ernst, D. An application of deep reinforcement learning to algorithmic trading. *Expert Syst. Appl.* **2021**, *173*, 114632. [[CrossRef](#)]
30. Lei, K.; Zhang, B.; Li, Y.; Yang, M.; Shen, Y. Time-driven feature-aware jointly deep reinforcement learning for financial signal representation and algorithmic trading. *Expert Syst. Appl.* **2020**, *140*, 112872. [[CrossRef](#)]
31. Rundo, F.; Trenta, F.; di Stallo, A.L.; Battiatto, S. Grid trading system robot (GTSbot): A novel mathematical algorithm for trading FX market. *Appl. Sci.* **2019**, *9*, 1796. [[CrossRef](#)]
32. Rundo, F. Deep LSTM with reinforcement learning layer for financial trend prediction in FX high frequency trading systems. *Appl. Sci.* **2019**, *9*, 4460. [[CrossRef](#)]
33. Nickel, M.; Kiela, D. Poincaré embeddings for learning hierarchical representations. In *Advances in Neural Information Processing Systems (NeurIPS)*; Curran Associates, Inc.: Red Hook, NY, USA, 2017; Volume 30, pp. 6338–6347.
34. Ganea, O.; Bécigneul, G.; Hofmann, T. Hyperbolic neural networks. In *Advances in Neural Information Processing Systems (NeurIPS)*; Curran Associates, Inc.: Red Hook, NY, USA, 2018.
35. Jin, M.; Koh, H.Y.; Wen, Q.; Zambon, D.; Alippi, C.; Webb, G.I.; King, I.; Pan, S. A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2024**, *46*, 10466–10485. [[CrossRef](#)]
36. Bécigneul, G.; Ganea, O.-E. Riemannian adaptive optimization methods. In *International Conference on Learning Representations (ICLR 2019)*; Curran: Crystal Lake, IL, USA, 2019.
37. Jaimović, V.; Kapić, Z.; Crnkić, A. Reinforcement learning in hyperbolic spaces: Models and experiments. *arXiv* **2024**, arXiv:2410.09466. [[CrossRef](#)]
38. Kandala, A.; Mezzacapo, A.; Temme, K.; Takita, M.; Brink, M.; Chow, J.M.; Gambetta, J.M. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **2017**, *549*, 242–246. [[CrossRef](#)] [[PubMed](#)]
39. Abbas, A.; Sutter, D.; Zoufal, C.; Lucchi, A.; Figalli, A.; Woerner, S. The power of quantum neural networks. *Nat. Comput. Sci.* **2021**, *1*, 403–409. [[CrossRef](#)]
40. Fujii, K.; Nakajima, K. Harnessing disordered-ensemble quantum dynamics for machine learning. *Phys. Rev. Appl.* **2017**, *8*, 024030. [[CrossRef](#)]
41. Mari, A.; Bromley, T.R.; Izaac, J.; Schuld, M.; Killoran, N. Transfer learning in hybrid classical-quantum neural networks. *Quantum* **2020**, *4*, 340. [[CrossRef](#)]
42. Jerbi, S.; Gyurik, C.; Marshall, S.; Briegel, H.J.; Dunjko, V. Parametrized quantum policies for reinforcement learning. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2021.
43. Skolik, A.; Jerbi, S.; Dunjko, V. Quantum agents in the Gym: A variational quantum algorithm for deep Q-learning. *Quantum* **2022**, *6*, 720. [[CrossRef](#)]
44. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence*; Lecture Notes in Mathematics; Springer: Berlin, Germany, 1981; Volume 898, pp. 366–381.
45. Grassberger, P.; Procaccia, I. Measuring the strangeness of strange attractors. *Phys. D Nonlinear Phenom.* **1983**, *9*, 189–208. [[CrossRef](#)]
46. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*, 174102. [[CrossRef](#)] [[PubMed](#)]
47. Marwan, N.; Romano, M.C.; Thiel, M.; Kurths, J. Recurrence plots for the analysis of complex systems. *Phys. Rep.* **2007**, *438*, 237–329. [[CrossRef](#)]
48. Kantz, H.; Schreiber, T. *Nonlinear Time Series Analysis*; Cambridge University Press: Cambridge, UK. [[CrossRef](#)]

49. Bailey, D.H.; López de Prado, M. The Deflated Sharpe Ratio: Correcting for Selection Bias, Backtest Overfitting, and Non-Normality. *J. Portf. Manag.* **2014**, *40*, 94–107. [[CrossRef](#)]
50. Hancer, E. Transferable Reinforcement Learning in Forex Trading. Bachelor's Thesis, Delft University of Technology, Delft, The Netherlands, 2025. [[CrossRef](#)]
51. Vinciguerra, V.; Ambra, E.; Maddiona, L.; Oliveri, S.; Romeo, M.F.; Mazzillo, M.; Rundo, F.; Fallica, G. Progresses towards a processing pipeline in photoplethysmogram (PPG) based on SiPMs. In Proceedings of the 2017 European Conference on Circuit Theory and Design (ECCTD), Catania, Italy, 4–6 September 2017; pp. 1–5. [[CrossRef](#)]
52. Mazzillo, M.; Maddiona, L.; Rundo, F.; Sciuto, A.; Libertino, S.; Lombardo, S.; Fallica, G. Characterization of SiPMs With NIR Long-Pass Interferential and Plastic Filters. *IEEE Photonics J.* **2018**, *10*, 6802112. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.