

A Restricted No-Go Theorem for Exact, Coherent, and Isolated Quantum Learning Architectures

Mrittunjoy Guha Majumdar

`mrittunjoy@dl.amrita.edu`

Amrita Vishwa Vidyapeetham

Short Report

Keywords: Quantum Learning, No-Go Theorems, Quantum Computing, Quantum Artificial Intelligence

Posted Date: October 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10359300/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A Restricted No-Go Theorem for Exact, Coherent, and Isolated Quantum Learning Architectures

Mrittunjoy Guha Majumdar^{1,2,3*}

^{1*}School of Natural Sciences and Engineering, National Institute of Advanced Studies, Mathikere, Bengaluru, 560054, Karnataka, India.

²School of Artificial Intelligence, Amrita Vishwa Vidyapeetham, Mata Amritanandamayi Marg, Faridabad, 121002, Haryana, India.

³Department of Physics, Amrita Vishwa Vidyapeetham, Amritapuri, Clappana P. O., Kollam, 690525, Kerala, India.

Corresponding author(s). E-mail(s): mrirtunjoy@dl.amrita.edu;

Abstract

The straightforward translation of classical machine learning models into the quantum domain is obstructed by fundamental structural constraints of quantum mechanics. In this work, we rigorously examine the interplay between universality, coherent training, and isolation in quantum learning systems. We prove a restricted no-go theorem: within the setting of exact continuous unitary hypothesis classes, finite-dimensional closed quantum systems, and fully coherent parameter-update rules with no intermediate measurement or classical feedback, no architecture can simultaneously satisfy universality (U), coherent training (C), and isolation (I). This result follows from a synthesis of the no-programming theorem, linearity of quantum dynamics, and the necessity of measurement for performance evaluation. Crucially, we distinguish between (i) nonlinear functions applied to classical parameter values encoded in a computational basis, which are permitted, and (ii) deterministic nonlinear transformations on arbitrary unknown quantum states, which are forbidden. Our analysis shows that many naively quantized models, such as fully coherent versions of neural networks, support vector machines, and principal component analysis, either incur substantial state-preparation overheads or violate the structural rules of quantum theory. We conclude that viable quantum machine learning (QML) architectures must relax the isolation condition, thereby justifying the dominant hybrid classical-quantum paradigm not as a mere engineering compromise, but as a physical necessity imposed by the UCI constraints.

Keywords: Quantum Learning, No-Go Theorems, Quantum Computing, Quantum Artificial Intelligence

1 Introduction

Despite the growing interest in quantum machine learning as a means to achieve computational speedups in data-driven tasks, the compatibility of standard learning architectures with the operational constraints of quantum mechanics remains an open and foundational problem [1–4]. Quantum machine learning explores the integration of quantum information processing with learning theory, aiming to uncover models and algorithms that offer provable advantages over their classical counterparts [5, 6]. The prospect of quantum advantage in machine learning rests on the premise that quantum systems—by virtue of unitary dynamics, entanglement, and superposition—can represent and process information in exponentially larger Hilbert spaces, potentially enabling algorithmic speedups or improved generalization in certain learning tasks [7, 8]. Despite progress in variational algorithms, quantum-enhanced kernels, and quantum neural network proposals, a foundational question remains largely open: to what extent can classical learning architectures be meaningfully and coherently embedded into the quantum framework, give quantum no-go theorems? This question is nontrivial because the core assumptions underlying many classical learning models conflict directly with structural constraints imposed by quantum mechanics. Classical learning architectures routinely assume unrestricted access to training data, the ability to clone internal states, shared parameter updates across model copies, and local loss-based gradient descent [9–12]. However, quantum theory imposes specific restriction that may hinder a translation of various classical machine learning paradigms. Notably, the no-cloning theorem forbids the duplication of arbitrary quantum states [13–15], the no-broadcasting theorem excludes the consistent sharing of non-commuting quantum information with mixed states [16, 17], the no-programming theorem constrains the existence of universal programmable circuits for arbitrary unitaries [18], and the global linearity of quantum evolution rules out localized, nonlinear update rules ubiquitous in classical training protocols. The Kochen–Specker theorem prohibits non-contextual hidden variable assignments [19, 20], the no-deletion theorem forbids the perfect erasure of unknown quantum states [21], and the no-hiding theorem asserts that any quantum information lost from a subsystem under unitary evolution must reside entirely in the remainder of the global system, and not in correlations—highlighting a fundamental asymmetry with classical information, where data can be hidden purely in joint correlations [22].

In this work, we rigorously examine the implications of these no-go theorems for quantum machine learning architectures, focusing on models in which both training data and internal model representations are quantum states and learning proceeds via coherent or measurement-based quantum operations. We demonstrate that the no-cloning theorem imposes significant operational costs on architectures requiring adaptive, non-measurement access to quantum datasets across multiple stages of

computation unless the training states are mutually orthogonal. These costs manifest as $\mathcal{O}(N^2)$ state-preparation overheads for Gram-matrix estimation, rather than a strict impossibility. Similarly, the no-broadcasting theorem precludes coherent parameter sharing in quantum neural networks that rely on entangled weight registers or globally distributed internal representations, unless the parameter states commute. We further show that the unitarity of quantum evolution imposes global consistency constraints that fundamentally restrict the use of local error signals to drive learning in fully coherent, isolated quantum systems. This does not rule out hybrid approaches where classical backpropagation is performed on measurement outcomes or classical parameter registers. In particular, we establish that coherent quantum representations of model parameters cannot undergo loss-dependent updates via gradient descent without violating linearity, unless auxiliary classical controllers or irreversible measurement-induced dynamics are introduced. Our analysis culminates in a structural no-go result: there exists no general, fully coherent quantum learning architecture that simultaneously satisfies universality over a nontrivial hypothesis class, coherent training via state-dependent updates, and full isolation from measurement or classical control. This result follows from a synthesis of the no-programming theorem and bounds on programmable quantum gate arrays, which imply that programmable access to an arbitrary unitary hypothesis space necessitates either exponentially large quantum memory or external classical indexing. These constraints, taken together, reveal a fundamental mismatch between the assumptions underpinning many classical learning models and the operational restrictions of quantum theory. The consequences are twofold. First, they rule out a broad class of naïvely quantized learning models, particularly those modeled on classical neural networks with continuous parameter updates and shared internal states. Second, they point toward a new design space for quantum-native architectures that are explicitly constructed to respect the physical constraints of quantum information processing. These include models based on measurement-driven updates, quantum reservoirs, and hybrid classical-quantum training loops. By clarifying the architectural limitations imposed by quantum no-go theorems, this work lays the foundation for a principled theory of learnability in quantum systems and delineates the boundary between what is physically implementable and what remains classically inspired but quantum-inadmissible.

2 Preliminaries and Formal Definitions

We first establish a formal framework for quantum learning architectures.

Definition 1 (Quantum Learning Architecture). *A quantum learning architecture is a tuple*

$$(\mathcal{H}_D, \mathcal{H}_P, \mathcal{H}_A, \mathcal{U}_\Theta, \mathcal{T}, \mathcal{M})$$

where:

- $\mathcal{H}_D$ is the data register Hilbert space.
- $\mathcal{H}_P$ is the parameter register Hilbert space.
- $\mathcal{H}_A$ is an auxiliary workspace.

- $\mathcal{U}_\Theta = \{U_\theta \mid \theta \in \Theta\}$ is a hypothesis class of unitary transformations on $\mathcal{H}_D$.
- $\mathcal{T}$ is a training protocol, consisting of a sequence of completely positive trace-preserving (CPTP) maps acting on $\mathcal{H}_D \otimes \mathcal{H}_P \otimes \mathcal{H}_A$.
- $\mathcal{M}$ is a measurement strategy.

Table 1 Summary of quantum no-go theorems, their formal assumptions, forbidden operations, and implications for QML.

Theorem	Formal Assumption	Forbidden Operation	QML Affected	Primitive
No-Cloning	Unknown arbitrary pure states	$U \lvert \psi \rangle \lvert 0 \rangle = \lvert \psi \rangle \lvert \psi \rangle$	Data duplication for parallel processing; Gram matrix estimation requires fresh preparations	
No-Broadcasting	Unknown non-commuting mixed states	$\text{Tr}_2[\mathcal{E}(\rho)] = \text{Tr}_1[\mathcal{E}(\rho)] = \rho$	Coherent sharing of non-commuting parameter states	
No-Programming	Exact deterministic processors	$U(\lvert \psi \rangle \otimes \lvert \phi_f \rangle) = (U_f \lvert \psi \rangle) \otimes \lvert \phi_f \rangle$	Universal coherent hypothesis classes in finite dimensions	
No-Deletion	Unknown pure states	$D(\lvert \psi \rangle \otimes \lvert \psi \rangle) = \lvert \psi \rangle \otimes \lvert 0 \rangle$	No direct QML obstruction; tracing out ancillas is a valid CPTP operation	
No-Hiding	Unitary dynamics	Information loss $\Rightarrow$ transfer to system-environment correlations	(Privacy claims require explicit models; not a direct barrier)	
Kochen-Specker	Non-contextual hidden variables	Global valuation v for all observables	Context-independent feature extraction; measurement basis dependence	

Definition 2 (Universality (U)). *An architecture satisfies Universality if for every $U_\theta \in \mathcal{U}_\Theta$, there exists a parameter state $\lvert \theta \rangle \in \mathcal{H}_P$ such that for all input states $\lvert \psi \rangle \in \mathcal{H}_D$,*

$$\mathcal{U}(\lvert \psi \rangle \otimes \lvert \theta \rangle) = (U_\theta \lvert \psi \rangle) \otimes \lvert \theta \rangle,$$

where $\mathcal{U}$ is a fixed unitary or CPTP map acting on the joint registers. In this work, we consider exact universality over a continuous subset of $SU(d)$ with $d \geq 2$, unless explicitly stated otherwise.

Definition 3 (Coherent Training (C)). *An architecture satisfies Coherent Training if the parameter update is implemented by a CPTP map Λ acting jointly on the data, parameter, and auxiliary registers, with no intermediate measurement or classical*

feedback:

$$\Lambda(\rho_D \otimes |\theta\rangle\langle\theta| \otimes \rho_A) = \sigma_{DA},$$

such that $\text{Tr}_{D,A}[\sigma_{DA}] = |\theta'\rangle\langle\theta'|$, where θ' depends on the data and the previous parameter. The entire process is coherent (unitary if the total system is closed).

Definition 4 (Isolation (I)). *An architecture satisfies Isolation if the entire learning process occurs within a closed quantum system, with no classical information inflow or outflow, no measurement-based feedback, and no external classical control, except for the initial state preparation.*

Remark 1 (Nonlinearity clarification) We distinguish two distinct notions:

1. *Nonlinear functions of classically encoded parameter values* (e.g., ReLU applied to computational basis states). These can be implemented via reversible classical circuits and do not violate quantum linearity.
2. *Deterministic nonlinear transformations on arbitrary unknown quantum superposition states*. These are forbidden by the linearity of quantum dynamics. Our no-go results concern exclusively the latter.

3 Impact of Quantum Limitations on Classical Machine Learning Formalisms

Quantum mechanics imposes rigorous structural constraints on information processing that render classical machine learning architectures mathematically non-isomorphic to their quantum counterparts. Specifically, the Hilbert space formalism and the associated operator algebra are subject to a family of no-go theorems, no-cloning, no-broadcasting, Kochen–Specker contextuality, no-programming, no-deletion, and no-hiding, that fundamentally obstruct the algebraic primitives underlying classical ML pipelines. Formally, the no-cloning theorem prohibits the existence of a universal isometric map $U : \mathcal{H} \otimes \mathcal{H}_0 \rightarrow \mathcal{H} \otimes \mathcal{H}$ satisfying $U|\psi\rangle|0\rangle = |\psi\rangle|\psi\rangle$ for arbitrary $|\psi\rangle \in \mathcal{H}$, where $\mathcal{H}$ is a separable Hilbert space. The no-broadcasting theorem generalizes this constraint to mixed states, asserting that a completely positive trace-preserving (CPTP) map $\mathcal{E} : \mathcal{D}(\mathcal{H}) \rightarrow \mathcal{D}(\mathcal{H} \otimes \mathcal{H})$ satisfying $\text{Tr}_2[\mathcal{E}(\rho)] = \text{Tr}_1[\mathcal{E}(\rho)] = \rho$ for all ρ exists if and only if the set $\{\rho\}$ is pairwise commuting. The Kochen–Specker theorem precludes the existence of a global valuation function $v : \mathcal{O} \rightarrow \mathbb{R}$ such that $v(A) \in \sigma(A)$ and $v(f(A)) = f(v(A))$ for all self-adjoint operators $A \in \mathcal{O}$ and Borel functions f , over any context-independent subalgebra $\mathcal{O} \subset \mathcal{B}(\mathcal{H})$, invalidating the assumption of non-contextual feature extraction. The no-programming theorem mathematically asserts that for any fixed universal unitary U , there does not exist an orthogonal set of program states $\{|\phi_f\rangle\}$ such that $U(|\psi\rangle \otimes |\phi_f\rangle) = (U_f|\psi\rangle) \otimes |\phi_f\rangle$ holds universally for all unitaries U_f . Complementarily, the no-deletion theorem rules out the existence of any linear map D such that $D(|\psi\rangle \otimes |\psi\rangle) \rightarrow |\psi\rangle \otimes |0\rangle$ universally, and the no-hiding theorem asserts that loss of information from a subsystem corresponds to complete transfer into the system-environment correlations, as formalized by unitary invariance of global von Neumann entropy under CPTP maps.

These constraints translate into operational considerations, ranging from resource overheads to structural incompatibilities, across machine learning models when lifted into the quantum domain. We now examine these on a model-by-model basis, distinguishing between strict impossibilities and resource costs. Now, consider Support Vector Machines (SVMs): the classical kernel function $k(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle$ presupposes the ability to construct Gram matrices $K \in \mathbb{R}^{N \times N}$ via arbitrary data duplication. In the quantum regime, the analogous Gram matrix $K_{ij} = |\langle \psi_i | \psi_j \rangle|^2$ requires estimation of pairwise fidelities, but the no-cloning constraint enforces that for each $\langle \psi_i | \psi_j \rangle$, the states $|\psi_i\rangle$ and $|\psi_j\rangle$ must be destructively measured or sequentially prepared at an operational cost scaling as $\mathcal{O}(N^2)$. This is a resource limitation rather than a foundational impossibility, particularly when the data can be re-prepared from a known classical source or via a preparation oracle. Additionally, Kochen–Specker contextuality imposes that observables corresponding to feature projections $\{P_k\}$ do not generally commute, meaning that the measurement-induced realization of feature vectors is fundamentally basis-dependent. Principal Component Analysis (PCA) requires diagonalization of the covariance operator $C = \text{Tr}_{\text{env}}[\rho]$ where $\rho = \sum_{i=1}^N p_i |\psi_i\rangle \langle \psi_i|$ is the empirical density operator. The empirical density operator ρ can be prepared as a classical mixture by preparing each $|\psi_i\rangle$ with probability p_i . The no-broadcasting theorem does not prevent this preparation; it forbids a universal broadcasting map that would take ρ to two copies of ρ for arbitrary non-commuting states. Thus, quantum PCA is resource-intensive rather than fundamentally impossible. Spectral estimation via Quantum Phase Estimation (QPE) introduces additional ancilla registers that are typically traced out at the end of the protocol. This is a standard CPTP operation and is not prohibited by the no-deletion theorem, which concerns the impossibility of deleting arbitrary unknown pure states.

For distance-based models like k-Nearest Neighbors (k-NN), the classical distance function $d(x_i, x_j) = \|x_i - x_j\|$ translates quantumly into fidelity estimation via SWAP tests, with operational cost scaling $\mathcal{O}(MN)$ for M test and N training samples. Each SWAP test requires fresh state preparation due to the destructive nature of quantum projective measurement. The no-hiding theorem implies that fidelity measurement leaks nontrivial information about the training states into the measurement record, violating classical assumptions of privacy-preserving computation. In Random Forests, the no-programming theorem prohibits the construction of a universal circuit that encodes all decision tree classifiers within a common circuit template; formally, given a set of decision functions $\{f_t\}$, there does not exist a unitary U and orthogonal program states $\{|\phi_t\rangle\}$ such that $U |\psi\rangle |\phi_t\rangle = |f_t(\psi)\rangle |\phi_t\rangle$ holds for arbitrary f_t . Additionally, the no-cloning theorem invalidates classical bootstrap aggregation by forbidding simultaneous access to multiple instances of the training dataset. In compositional models like Neural Networks (NNs) and Generative Adversarial Networks (GANs), the constraint imposed by linearity of quantum mechanics is formally absolute. Classical nonlinear activation functions, such as ReLU or sigmoid, correspond to nonlinear maps over the vector space. However, quantum operations are strictly linear, and no completely positive trace-preserving (CPTP) map exists to implement

a nonlinear transformation on the amplitude vector without inducing measurement collapse. Consequently, quantum neural networks must implement nonlinearity via probabilistic post-selection on measurement outcomes, measurement probabilities, or classical post-processing. While this introduces overhead compared to classical backpropagation, a rigorous complexity analysis of this overhead is an open question and is not established in this work. In GANs, the discriminator’s measurement necessarily collapses the superposition state of the generator’s output, violating the classical requirement for differentiable adversarial feedback loops. The same holds for Variational Autoencoders (VAEs), where no-broadcasting prohibits the reuse of the encoded latent state for both reconstruction and gradient computation, while contextuality ensures that any intermediate sampling perturbs the latent representation in an irreducibly basis-dependent manner. As a result, the classical Evidence Lower Bound (ELBO) maximization objective lacks a quantum-consistent analogue.

These considerations suggest that quantum learning protocols face constraints beyond circuit depth or qubit count, including state re-preparation rates, measurement-induced collapse, and coherence-time requirements. However, whether these translate into provable complexity-theoretic separations remains an open question. The UCI theorem established in Section 4 provides a structural, rather than complexity-theoretic, limitation. This supports the view that QML is not merely an isomorphic port of classical ML into the quantum domain but constitutes a distinct computational paradigm bounded by the operator-theoretic and information-theoretic structure of quantum mechanics.

4 A No-Go Theorem for Coherent, Isolated Architectures

We now formalize the main result. The theorem is stated under four explicit assumptions:

- (A1) The Hilbert spaces are finite-dimensional.
- (A2) The hypothesis class contains a continuous subset of $SU(d)$ for $d \geq 2$ (e.g., all rotations about a fixed axis by arbitrary angles).
- (A3) The architecture is exact (no approximation error ϵ is tolerated).
- (A4) The architecture satisfies the definitions of Coherent Training (C) and Isolation (I) as given in Definitions 2 and 3.

Theorem 1 (Restricted UCI No-Go Theorem) *Under assumptions (A1)–(A4), no quantum learning architecture can simultaneously satisfy Universality (U), Coherent Training (C), and Isolation (I).*

Proof The proof proceeds via three lemmas, each showing that two of the three properties are incompatible under the given assumptions.

Lemma 1. *Under assumptions (A1)–(A2) and Isolation (I), Universality (U) cannot hold for a continuous hypothesis class if the parameter space is finite-dimensional.*

Proof Universality requires a fixed quantum operation $\mathcal{U}$ such that for any U_θ ,

$$\mathcal{U}(|\psi\rangle \otimes |\theta\rangle) = (U_\theta |\psi\rangle) \otimes |\theta\rangle.$$

This is precisely a programmable quantum processor. By the no-programming theorem [18], for two distinct unitaries $U_\theta \neq U_{\theta'}$, the corresponding program states must be orthogonal: $\langle \theta | \theta' \rangle = 0$. Since Θ contains a continuous subset, this implies an uncountably infinite set of pairwise orthogonal program states, which requires $\dim(\mathcal{H}_P) = \infty$. This contradicts (A1). Thus, (U) and (I) are incompatible for continuous hypothesis classes in finite dimensions. *Note:* This proof concerns exact universality. Approximate universality ($\epsilon > 0$) may relax this constraint and is outside the scope of this theorem. $\square$

Lemma 2. *Under assumptions (A3)–(A4) and Coherent Training (C), Isolation (I) is incompatible with generic state-dependent parameter updates.*

Proof Coherent training requires a CPTP map Λ that updates the parameter state: $|\theta\rangle \mapsto |f(\theta, \psi)\rangle$, where f depends on both the parameter and the input data state $|\psi\rangle$. Quantum dynamics is linear over density operators. For a superposition of parameter states $\sum_j c_j |\theta_j\rangle$, linearity forces

$$\Lambda \left(|\psi\rangle \langle \psi| \otimes \left(\sum_j c_j |\theta_j\rangle \right) \left(\sum_k c_k \langle \theta_k| \right) \otimes \rho_A \right) \quad (1)$$

$$= \sum_{j,k} c_j \bar{c}_k \Lambda(|\psi\rangle \langle \psi| \otimes |\theta_j\rangle \langle \theta_k| \otimes \rho_A). \quad (2)$$

For a deterministic, state-dependent update $|\theta\rangle \mapsto |f(\theta, \psi)\rangle$ to be implemented unitarily on *arbitrary unknown* superpositions, f must be linear in θ . However, generic learning updates (gradient descent, Bayesian updates) are nonlinear functions of θ . This does *not* forbid implementing nonlinear functions on computational-basis encoded parameters via reversible classical circuits; it forbids deterministic nonlinear transformations on arbitrary coherent superpositions of program states. Since (I) excludes measurement or classical readout that would collapse these superpositions, the architecture cannot implement generic coherent updates. Thus, (C) and (I) are incompatible for generic learning rules. $\square$

Lemma 3. *Under Universality (U) and Coherent Training (C), Isolation (I) is violated because evaluating performance requires measurement.*

Proof To update parameters coherently based on data, the architecture must have access to a performance metric (e.g., a loss function) $\mathcal{L}(U_\theta, \psi)$. In quantum mechanics, extracting information about the action of U_θ on ψ , such as the expectation value $\langle \psi | U_\theta^\dagger O U_\theta | \psi \rangle$, requires either a measurement (projective or weak) or classical post-processing of measurement statistics. Both operations break Isolation (I) by introducing classical information or irreversibility. A purely unitary, closed-system evolution cannot evaluate such performance-dependent quantities without ancillary interactions that effectively constitute a measurement. Hence, (U) and (C) together force a violation of (I). $\square$

Combining Lemmas 1–3, no architecture can satisfy all three properties (U), (C), and (I) simultaneously under assumptions (A1)–(A4). This completes the proof. $\square$

5 Discussion

The UCI theorem, as formulated, applies to a narrow but precisely defined class of architectures. Several important caveats and connections to the broader literature must be emphasized. First, the theorem does *not* rule out:

- Variational quantum algorithms (VQAs) where parameters are stored classically and updated via a classical optimizer (this relaxes (C) and (I)).
- Quantum kernel methods where the feature map is known and states can be repeatedly prepared from classical data.
- Restricted ansätze with finite, discrete hypothesis classes (these evade the infinite-dimensional orthogonality requirement).
- Approximate or probabilistic implementations, where a small error $\epsilon > 0$ or success probability $1 - \delta$ may allow finite-dimensional parameterizations via approximate compilation theorems.

Importantly, the theorem’s reliance on exactness (A3) means that it does not preclude architectures that achieve universality only approximately. Recent results on approximate programmable quantum gate arrays suggest that finite-dimensional program registers may suffice when a small error $\epsilon > 0$ is tolerated, at the cost of increased circuit depth or ancillary qubits. Quantifying this trade-off remains an important open direction.

Second, our results complement existing trainability literature. While barren plateaus concern the variance of gradients for specific ansätze and cost functions, the UCI theorem addresses a structural limitation: the impossibility of fully coherent, isolated universal learning. The two are distinct but mutually reinforcing: one limits *efficient* training, the other limits *coherent* training. Third, contextuality (Kochen-Specker) is often discussed as a resource for quantum advantage, rather than merely an obstacle. Our use of contextuality is limited to the observation that context-independent feature extraction lacks a quantum analogue; we do not claim contextuality is globally detrimental to QML. Finally, we stress that our theorem assumes exactness. In practice, NISQ devices operate with finite precision, and the infinite-dimensional obstruction from the no-programming theorem may be circumvented by approximate programmable gate arrays. Exploring the trade-off between approximation error, program state dimension, and learning performance remains an open direction.

Scope and Limitations of the Present Work. We emphasize the following limitations:

1. The UCI theorem assumes *exact* universality. Approximate universality ($\epsilon > 0$) may be achievable with finite-dimensional parameter registers, as shown by approximate compilation results.

2. The theorem applies to *arbitrary unknown* quantum data states. If the data are classical and encoded via a known fixed circuit, no-cloning does not constrain repeated preparation.
3. Continuous-variable (infinite-dimensional) quantum systems are not covered by the finite-dimensional assumption and may exhibit different trade-offs.
4. The theorem does not address statistical learning theory, generalization bounds, or sample complexity; it concerns only the structural implementability of coherent architectures.

While the theorem is stated for exact universality, recent quantitative results on approximate programmable quantum processors [23] imply that the required program dimension grows at least as $\Omega(\log(1/\epsilon))$ for approximation error ϵ . This suggests that even approximate universal coherent learning may incur significant resource overheads, though the precise trade-off with learning performance remains an open question.

6 Conclusion

In this work, we have rigorously examined the compatibility of classical machine learning primitives with the structural constraints of quantum mechanics. The central result, the restricted UCI no-go theorem, establishes that no finite-dimensional, fully coherent, and isolated quantum architecture can simultaneously achieve exact universality over a continuous hypothesis class, coherent parameter updates, and isolation from classical resources. The proof synthesizes three independent obstructions: the no-programming theorem (which forces infinite-dimensional program spaces for continuous universality), the linearity of quantum dynamics (which forbids deterministic nonlinear updates on arbitrary superpositions), and the necessity of measurement for performance evaluation. Crucially, we have clarified several common misconceptions. The no-broadcasting theorem does not prevent the preparation of empirical density operators; the no-deletion theorem does not prohibit tracing out ancillas; and quantum neural networks can achieve nonlinearity through measurement-based or hybrid strategies. The exponential complexity claims for backpropagation have been removed pending rigorous proof. Rather than signaling a failure of quantum machine learning, our results provide a formal foundation for the hybrid variational paradigm. By relaxing the isolation condition (I), storing parameters classically and employing measurement-based feedback, practical QML algorithms naturally circumvent the UCI obstructions. The theorem thus delineates the boundary between classically-inspired but quantum-inadmissible architectures and viable quantum-native designs that respect the physical laws of quantum information. Future work should explore approximate versions of the UCI theorem, the role of finite-sample training sets, and resource-theoretic formulations that quantify the classical communication overhead required to break isolation.

Statements and Declarations

Data Availability Statement

This manuscript is a theoretical study and does not involve the generation or analysis of any datasets. All results presented are derived from analytical derivations and do not rely on external data.

Author Contributions

MGM undertook conceptualization, formal analysis, investigation, methodology, project administration, validation, visualization, writing as well as review and editing.

Funding

No funding was received to assist with the preparation of this manuscript. No funding was received for conducting this study.

Competing Interests

The author declares no competing interests.

References

- [1] 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. *Nature Reviews Physics* **3**(9), 625–644 (2021)
- [2] Cerezo, M., Verdon, G., Huang, H.-Y., Cincio, L., Coles, P.J.: Challenges and opportunities in quantum machine learning. *Nature computational science* **2**(9), 567–576 (2022)
- [3] Banchi, L., Pereira, J., Pirandola, S.: Generalization in quantum machine learning: A quantum information standpoint. *PRX Quantum* **2**(4), 040321 (2021)
- [4] Huang, H.-Y., Kueng, R., Preskill, J.: Information-theoretic bounds on quantum advantage in machine learning. *Physical Review Letters* **126**(19), 190505 (2021)
- [5] Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., Lloyd, S.: Quantum machine learning. *Nature* **549**(7671), 195–202 (2017)
- [6] Alchieri, L., Badalotti, D., Bonardi, P., Bianco, S.: An introduction to quantum machine learning: from quantum logic to quantum deep learning. *Quantum Machine Intelligence* **3**(2), 28 (2021)
- [7] Schuld, M., Sinayskiy, I., Petruccione, F.: An introduction to quantum machine learning. *Contemporary Physics* **56**(2), 172–185 (2015)

- [8] Bai, Q., Hu, X.: Superposition-enhanced quantum neural network for multi-class image classification. *Chinese Journal of Physics* **89**, 378–389 (2024)
- [9] Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *nature* **323**(6088), 533–536 (1986)
- [10] LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. *Neural computation* **1**(4), 541–551 (1989)
- [11] Bottou, L., Curtis, F.E., Nocedal, J.: Optimization methods for large-scale machine learning. *SIAM review* **60**(2), 223–311 (2018)
- [12] Robbins, H., Monro, S.: A stochastic approximation method. *The annals of mathematical statistics*, 400–407 (1951)
- [13] Wootters, W.K., Zurek, W.H.: The no-cloning theorem. *Physics Today* **62**(2), 76–77 (2009)
- [14] Miyadera, T., Imai, H.: No-cloning theorem on quantum logics. *Journal of mathematical physics* **50**(10) (2009)
- [15] Bužek, V., Hillery, M.: Quantum copying: Beyond the no-cloning theorem. *Physical Review A* **54**(3), 1844 (1996)
- [16] Barnum, H., Caves, C.M., Fuchs, C.A., Jozsa, R., Schumacher, B.: Noncommuting mixed states cannot be broadcast. *Physical Review Letters* **76**(15), 2818 (1996)
- [17] Barnum, H., Barrett, J., Leifer, M., Wilce, A.: Generalized no-broadcasting theorem. *Physical review letters* **99**(24), 240501 (2007)
- [18] Nielsen, M.A., Chuang, I.L.: Programmable quantum gate arrays. *Physical Review Letters* **79**(2), 321 (1997)
- [19] Held, C.: Kochen—specker theorem. In: *Compendium of Quantum Physics*, pp. 331–335. Springer, ??? (2009)
- [20] Peres, A.: Generalized kochen-specker theorem. *Foundations of Physics* **26**(6), 807–812 (1996)
- [21] Kumar Pati, A., Braunstein, S.L.: Impossibility of deleting an unknown quantum state. *Nature* **404**(6774), 164–165 (2000)
- [22] Braunstein, S.L., Pati, A.K.: Quantum information cannot be completely hidden in correlations: Implications for the black-hole information paradox. *Physical review letters* **98**(8), 080502 (2007)
- [23] Kubicki, A.M., Palazuelos, C., Pérez-García, D.: Resource quantification for the

no-programing theorem. Physical review letters **122**(8), 080505 (2019)