Quantum Deep Learning: A Comprehensive Review

Yanjun Ji,[1,*] Zhao-Yun Chen,[2,†] Marco Roth,[3,‡] David A. Kreplin,[4,§] Christian Schiffer,[5,6] Martin King,[7,8] Oliver Anton,[9,10,¶] M. Sahnawaz Alam,[11] Markus Krutzik,[9,10] Dennis Willsch,[12,13,**] Ludwig Mathey,[14,15,16,††] Frank K. Wilhelm,[1,17,‡‡] and Guo-Ping Guo[2,18,19,20,21,§§]

[1] *Institute for Quantum Computing Analytics (PGI-12), Forschungszentrum Jülich, 52425 Jülich, Germany*
[2] *Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei, Anhui, 230088, P. R. China*
[3] *Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), Nobelstrasse 12, Stuttgart 70569, Germany*
[4] *Heilbronn University of Applied Sciences, Max-Planck-Str. 39, 74081 Heilbronn, Germany*
[5] *Institute of Neuroscience and Medicine (INM-1), Research Centre Jülich, Jülich, Germany*
[6] *Helmholtz AI, Research Centre Jülich, Jülich, Germany*
[7] *Munich Center for Mathematical Philosophy, Ludwig Maximilian University of Munich, Munich, Germany*
[8] *Lichtenberg Group for History and Philosophy of Physics, University of Bonn, Bonn, Germany*
[9] *Institut für Physik and Center for the Science of Materials Berlin (CSMB) Adlershof, Humboldt-Universität zu Berlin, Newtonstrß 15, 12489 Berlin, Germany*
[10] *Ferdinand-Braun-Institut (FBH), Gustav-Kirchoff-Straße 4, 12489 Berlin, Germany*
[11] *Institute for Automation of Complex Power Systems, RWTH Aachen University, Aachen, Germany*
[12] *Jülich Supercomputing Centre, Forschungszentrum Jülich, 52425 Jülich, Germany*
[13] *Faculty of Medical Engineering and Technomathematics, University of Applied Sciences Aachen, 52428 Jülich, Germany*
[14] *Zentrum für Optische Quantentechnologien, Universität Hamburg, 22761 Hamburg, Germany*
[15] *Institut für Quantenphysik, Universität Hamburg, 22761 Hamburg, Germany*
[16] *The Hamburg Centre for Ultrafast Imaging, 22761 Hamburg, Germany*
[17] *Theoretical Physics, Saarland University, 66123 Saarbrücken, Germany*
[18] *Laboratory of Quantum Information, University of Science and Technology of China, Hefei, Anhui, 230026, P. R. China*
[19] *CAS Center for Excellence and Synergetic Innovation Center in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, P. R. China*
[20] *Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, P. R. China*
[21] *Origin Quantum Computing Technology (Hefei) Co., Ltd., Hefei, Anhui, 230026, P. R. China*

(Dated: March 10, 2026)

Quantum deep learning (QDL) explores the use of both quantum and quantum-inspired resources to determine when deep learning's core capabilities, such as expressivity, generalization, and scalability, can be enhanced based on specific resource constraints. Distinct from broader quantum machine learning, QDL emphasizes compositional depth at the pipeline level and the integration of quantum or quantum-inspired components within end-to-end workflows. This review provides an operational definition of QDL and introduces a taxonomy comprising four primary paradigms: hybrid quantum-classical models, quantum deep neural networks, quantum algorithms for deep learning primitives, and quantum-inspired classical algorithms. Theoretical principles are connected to advanced architectures, software toolchains, and experimental demonstrations across superconducting, trapped-ion, photonic, semiconductor spin, and neutral-atom systems, as well as quantum annealers. Claims of quantum advantage are critically assessed by distinguishing provable complexity-theoretic separations from empirical observations. The analysis characterizes trade-offs between model expressivity, trainability, and classical simulability, while systematically detailing the bottlenecks imposed by optimization landscapes, input-output access models, and hardware constraints. Applications are surveyed in domains encompassing image classification, natural language processing, scientific discovery, quantum data processing, and quantum optimal control, underscoring fair benchmarking against optimized classical counterparts and a comprehensive assessment of resource requirements. This review serves as a tutorial entry point for graduate students while guiding readers to specialized literature. It concludes with a verification-aware roadmap to transition QDL from near-term demonstrations to scalable and fault-tolerant implementations.

CONTENTS

* y.ji@fz-juelich.de
† chenzhaoyun@iai.ustc.edu.cn
‡ marco.roth@ipa.fraunhofer.de
§ david.kreplin@hs-heilbronn.de
¶ oliver.anton@physik.hu-berlin.de
** d.willsch@fz-juelich.de
†† ludwig.mathey@uni-hamburg.de
‡‡ f.wilhelm-mauch@fz-juelich.de
§§ gpguo@ustc.edu.cn

I. INTRODUCTION

The intersection of artificial intelligence (AI) and quantum computing (QC) has grown into a rapidly expanding interdisciplinary domain spanning quantum learning theory, algorithm design, and hardware-driven resource constraints (Acampora *et al.*, 2025; Alexeev *et al.*, 2025; Arunachalam and de Wolf, 2017; Biamonte *et al.*, 2017; Chang and Cerezo, 2025; Schuld and Petruccione, 2021). While deep learning (DL) provides the prevailing AI paradigm for representation learning and large-scale function approximation (Goodfellow *et al.*, 2016; LeCun *et al.*, 2015), QC has progressed from foundational theory (Nielsen and Chuang, 2010) to tangible experimental architectures (Arute *et al.*, 2019; Wu *et al.*, 2021b). At the same time, continued performance gains increasingly demand disproportionate increases in computational resources, energy, and data (Hoffmann *et al.*, 2022; Kaplan *et al.*, 2020; Patterson *et al.*, 2021), reflecting both the economics of large-scale training and slowing improvements in classical hardware scaling (Theis and Wong, 2017). This dynamic motivates the central question that frames this review: can quantum and quantum-inspired resources change the cost-accuracy scaling of learning primitives once data access, readout, and noise are specified? Against this backdrop of growing convergence, the present review uses the term quantum deep learning (QDL) in an operational sense. We distinguish it from AI for QC, an adjacent field where classical machine learning (ML) is used primarily for quantum-system design, control, compilation, or characterization (Alexeev *et al.*, 2025). Instead, our emphasis lies on learning primitives that explicitly exploit coherent quantum resources under realistic measurement and classical-control constraints.

A. The convergence into QDL

Classical DL originally drew on foundational neural concepts such as perceptrons (Rosenblatt, 1958), multilayer networks (Ivakhnenko and Lapa, 1965), associative memory (Hopfield, 1982), and energy-based probabilistic models (Ackley *et al.*, 1985). Subsequent progress relied on developments in gradient-based optimization and representational structure, including reverse-mode automatic differentiation (Linnainmaa, 1970), error backpropagation for multilayer networks (Rumelhart *et al.*, 1986), and convolutional architectures ranging from the neocognitron (Fukushima, 1980) to convolutional networks trained via backpropagation (LeCun *et al.*, 1989). The 2024 Nobel Prize in Physics recognized foundational contributions to artificial neural networks, awarding Hopfield and Hinton for discoveries enabling ML with neural networks (The Royal Swedish Academy of Sciences, 2024).

Modern DL's resurgence was enabled by large labeled datasets, graphics processing unit (GPU)-accelerated training, and architectural and optimization advances, marked by AlexNet's accuracy gains on the ILSVRC-2012 ImageNet benchmark (Deng *et al.*, 2009; Krizhevsky *et al.*, 2012) over hand-engineered pipelines, followed by architectural innovations including VGGNet (Simonyan and Zisserman, 2014), GoogLeNet (Szegedy *et al.*, 2015), and ResNet (He *et al.*, 2016). These breakthroughs rapidly expanded beyond computer vision, delivering major advances in natural language processing (NLP) via attention-based sequence modeling (Vaswani *et al.*, 2017), protein structure prediction (Jumper *et al.*, 2021; Tunyasuvunakool *et al.*, 2021), and materials discovery (Merchant *et al.*, 2023).

In the domain of QC, modern theory took shape in the 1980s with Benioff's Hamiltonian formulation of reversible Turing machines (Benioff, 1980), Manin's identification of complexity in simulating quantum dynamics (Manin, 1980), and Feynman's argument that simulating generic quantum mechanical systems was plausibly hard for classical computation in general, motivating devices operating on quantum principles as an alternative (Feynman, 1982, 1985). Deutsch further formalized the universal quantum computer, establishing a theoretical framework for QC (Deutsch, 1985). Subsequent algorithmic advances included Shor's polynomial-time quantum factoring and discrete logarithms (Shor, 1994, 1997), Grover's quadratic speedup for unstructured search (Grover, 1996, 1997), and Lloyd's efficient simulation of local quantum systems (Lloyd, 1996). Quantum error correction (QEC) (Shor, 1996; Steane, 1996) and threshold theorems (Aharonov and Ben-Or, 1997; Knill *et al.*, 1996, 1998; Preskill, 1997) clarified the path toward physical realization, establishing that scalable fault-tolerant computation is possible under suitable locality assumptions if physical error rates lie below an architecture- and noise-model-dependent threshold.

As a theoretical prelude to QDL, early quantum neural proposals drew heuristic correspondences between neural primitives and quantum state manipulation (Chrisley, 1995; Kak, 1995; Menneer and Narayanan, 1995; Purushothaman and Karayiannis, 1997). Related associative-memory constructions leveraged amplitude-amplification-style primitives (Grover, 1997; Trugenberger, 2001; Ventura and Martinez, 2000). This early literature revealed constraints absent in classical neural networks, particularly the limitations imposed by linear unitary evolution and the no-cloning theorem (Dieks, 1982; Wootters and Zurek, 1982). Consequently, effective nonlinearity was typically realized via measurement-conditioned classical postprocessing and feedback (Schuld and Petruccione, 2021), although coherent alternatives leveraging nonlinear media or many-body interactions remain a significant theoretical frontier.

Proposals oriented toward fault tolerance connected learning primitives to quantum subroutines, notably quantum linear-algebraic (QLA) routines (Harrow *et al.*, 2009) and quantum principal component analysis primitives (Lloyd *et al.*, 2014), typically under strong input-access assumptions. In the mid-2010s, Wiebe *et al.* (2015) framed "quantum deep learning" as accelerating deep Boltzmann machines under strong data-access assumptions. Related studies explored classical classification as a task suitable for quantum implementation via kernel estimation or linear-algebraic primitives (Rebentrost *et al.*, 2014). The advent of noisy intermediate-scale quantum (NISQ) hardware (Preskill, 2018) fundamentally recalibrated these expectations, as severe limitations in coherent circuit depth, gate fidelities, and sampling overhead strictly bounded practical algorithmic viability. Attention shifted to variational hybrid loops (Cerezo *et al.*, 2021a; McClean *et al.*, 2016; Peruzzo *et al.*, 2014) in which a parameterized quantum circuit (PQC) prepares a state, measurements estimate an objective, and a classical optimizer updates parameters in a feedback loop (Sim *et al.*, 2019; Tilly *et al.*, 2022), with hardware-efficient ansätze reducing compiled depth under native constraints (Kandala *et al.*, 2017). Subsequent architectural diversification introduced concepts such as hierarchy, locality, and compositionality (Cong *et al.*, 2019; Farhi and Neven, 2018), data re-uploading (Pérez-Salinas *et al.*, 2020), transfer learning (Mari *et al.*, 2020), and representative generative and structured models (Bausch, 2020; Dallaire-Demers and Killoran, 2018; Lloyd and Weedbrook, 2018; Verdon *et al.*, 2019b), alongside early experimental end-to-end training demonstrations (Beer *et al.*, 2020; Hu *et al.*, 2019).

However, trainability constraints emerged as circuit depth and complexity increased. These manifested primarily as barren plateau phenomena (McClean *et al.*, 2018), which motivate locality-aware costs and struc-

tured designs (Cerezo *et al.*, 2021b; Pesah *et al.*, 2021) (see Sec. III.B.3). In parallel, variational classifiers were theoretically reframed as kernel machines, with kernels defined by quantum feature maps (Schuld, 2021), e.g., the ZZ feature map (Havlíček *et al.*, 2019), which encodes data via entangling ZZ interactions between qubits. Complexity-theoretic arguments identify regimes where the induced feature states are conjectured to be costly to simulate classically, while information-theoretic analyses delimit when any end-to-end advantage can survive finite data, finite shots, and noise (Huang *et al.*, 2021a,b). Furthermore, dequantization and quantum-inspired classical algorithms emphasize that apparent separations can disappear when classical baselines are granted comparable access models (Tang, 2019, 2021). This synthesis motivates the current focus on task-specific quantum utility under explicit resource contracts (Chang and Cerezo, 2025; Preskill, 2025). A central unresolved issue in the field is to characterize and control the three-way tradeoff among expressivity, trainability, and classical simulatability. Architectures that are sufficiently expressive to represent useful hypothesis classes often become difficult to optimize due to vanishing-gradient phenomena. Conversely, architectures engineered for stable optimization are typically shallow or structured enough to admit strong classical simulation and competitive classical baseline performance.

Recent breakthroughs in quantum hardware are signaling a gradual transition from NISQ heuristics toward the fault-tolerant regime assumed by early QLA proposals. Experiments have successfully demonstrated encoded logical qubits and repeated error-correction cycles across multiple architectures, notably in trapped ions (Paetznick *et al.*, 2024), neutral atoms (Bluvstein *et al.*, 2024), and superconducting platforms (Acharya *et al.*, 2025). Concurrently, the maturation of differentiable software toolchains (Bergholm *et al.*, 2022; Bian *et al.*, 2023; Broughton *et al.*, 2021; Dou *et al.*, 2022; Kreplin *et al.*, 2025) has transformed QDL from theoretical speculation into an operational reality, enabling end-to-end optimization across quantum-classical boundaries at small and intermediate scales. As physical deployments scale, navigating the associated trade-offs increasingly hinges on rigorous benchmarking against strong classical baselines. This requires leveraging high-performance classical simulators (De Raedt *et al.*, 2019) and adhering to best-practice protocols for fair, resource-matched evaluation (Bowles *et al.*, 2024). However, rapid interdisciplinary progress has also produced a deeply fragmented literature, with results scattered across physics, computer science, and ML venues using incompatible terminology and evaluation methodology. These coupled tensions motivate the central theme of this review: identifying regimes in which end-to-end quantum utility survives realistic access, noise, and benchmarking constraints, and articulating the corresponding resource contracts under which such claims are meaningful. To unify this fragmented landscape, we establish a rigorous terminology and structural taxonomy.

B. Definition and taxonomy

To formalize our scope, we adopt the following operational definition.

Definition 1 (QDL).—QDL refers to compositional learning architectures and workflows exhibiting hierarchical depth at the level of the overall learning pipeline, in which at least one quantum or quantum-inspired component is functional and architecturally central. Concretely, the learning objective or update rule depends on this component's outputs, and its inclusion nontrivially determines the hypothesis class and/or estimator family under an explicitly stated access and resource model.

This definition distinguishes QDL from broader QML by requiring compositional depth at the pipeline level and by making access models and resource assumptions part of the problem specification. Consequently, QDL includes both hardware-executed and quantum-inspired classical realizations of quantum components. However, it explicitly excludes AI-for-QC studies as well as standalone shallow quantum feature maps, kernels, or single-circuit predictors, provided they are used solely as end-to-end models rather than embedded components of a hierarchically deep pipeline. Examples of QDL under this framework include: (i) a PQC head on a frozen or trainable classical backbone; (ii) a quantum sampling subroutine used only at deployment within a deep generative pipeline; and (iii) dequantization baselines embedded in a matched-access evaluation. We adopt this convention to align "deep" with hierarchical composition and to keep access models and end-to-end costing explicit.

Complementary perspectives on learning quantum systems and AI for QC are reviewed in (Gebhart *et al.*, 2023) and (Alexeev *et al.*, 2025), respectively. For foundational overviews we recommend Nielsen and Chuang (2010) and Goodfellow *et al.* (2016). The broader QML landscape is well documented, spanning canonical surveys and perspectives (Biamonte *et al.*, 2017; Cerezo *et al.*, 2022; Dunjko and Briegel, 2018; Lamichhane and Rawat, 2025; Rodríguez-Díaz *et al.*, 2025; Schuld and Petruccione, 2021; Schuld *et al.*, 2015), reviews (Chen *et al.*, 2024b; Du *et al.*, 2025; Klusch *et al.*, 2024; Mishra *et al.*, 2020; Moreau *et al.*, 2025; Peral-García *et al.*, 2024; Wang and Liu, 2024; Zeguendry *et al.*, 2023), pedagogical introductions (Alchieri *et al.*, 2021; Khan and Robles-Kelly, 2020), and domain-focused analyses and roadmaps (Acampora *et al.*, 2025; Jia *et al.*, 2019; Kharsa *et al.*, 2023; Orka *et al.*, 2025; Tian *et al.*, 2023; Ullah and Garcia-Zapirain, 2024). In contrast, our aim is to synthesize the QDL-specific lineage, architectural archetypes, implementation stack, and comparative-evaluation disci-

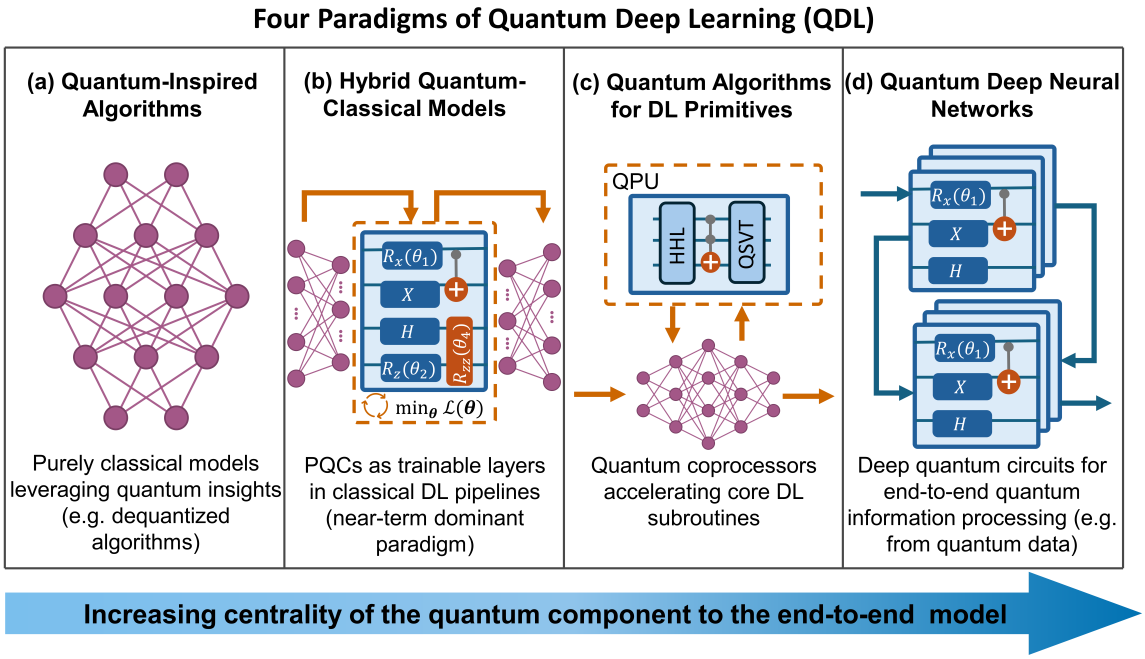

Figure 1 Four paradigms of quantum deep learning (QDL), ordered left to right by increasing centrality of the quantum component to the end-to-end model. (a) Quantum-inspired algorithms: purely classical models in which the structure or training algorithms are motivated by quantum information theory. Quantum content is expressed through the model class, not through hardware-executed quantum states. (b) Hybrid quantum-classical models: a parameterized quantum circuit (PQC) with trainable gate parameters $\boldsymbol{\theta}$ is embedded as a differentiable module within a classical pipeline. The cycle denotes the classical outer optimization loop that updates $\boldsymbol{\theta}$ from measurement outcomes by minimizing a loss function $\mathcal{L}(\boldsymbol{\theta})$. (c) Quantum algorithms for deep learning (DL) primitives: the quantum processing unit (QPU) acts as a specialized coprocessor to accelerate core subroutines, e.g., linear-system solvers such as HHL and the quantum singular value transformation (QSVT), within an overarching classical model under explicitly stated access and resource assumptions. (d) Quantum deep neural networks: deep, hierarchical quantum circuits designed for end-to-end quantum information processing, e.g., feature extraction directly from quantum data. $R_x(\theta_i)$ and $R_z(\theta_i)$: single-qubit rotation gates; X: Pauli-X gate; H: Hadamard gate.

pline.

Figure 1 summarizes four recurring paradigms we use throughout the review, conceptually ordered along the indicated axis by increasing centrality of the quantum component in the end-to-end model. At the classical end of the spectrum (Fig. 1(a)) lie quantum-inspired algorithms that import quantum-motivated structure or arise from dequantization (Tang, 2019, 2021). These algorithms use tools inspired by quantum information theory to build powerful classical models that can serve as strong, structure-exploiting baselines. In this regime, "quantum" content is expressed through the model class, linear-algebraic primitives, or sampling and optimization structure, not through hardware-executed quantum states. Figure 1(b) shifts from inspiration to explicit quantum components embedded in otherwise classical pipelines. The schematic emphasizes the placement of a PQC as a trainable module between classical feature processing stages and the hybrid feedback loop required for end-to-end optimization. This paradigm is dominant in the NISQ era because it amortizes quantum resources into repeated circuit evaluations while delegating representation learning and scaling to classical deep networks. The third paradigm (Fig. 1(c)) isolates a distinct, and often conflated, use case: quantum processors as coprocessors for specific DL primitives. The key scientific constraint, highlighted directly in the panel, is that any claimed speedup or advantage is inseparable from an explicit access and resource model. The same high-level primitive, such as Harrow-Hassidim-Lloyd (HHL) algorithm (Harrow *et al.*, 2009) and quantum singular value transformation (QSVT) (Gilyén *et al.*, 2019), can map to qualitatively different end-to-end costs depending on data access, state preparation, and measurement constraints. Finally, Fig. 1(d) represents the quantum-native extreme, in which depth and hierarchy reside primarily in the quantum circuit itself: layered unitary blocks process quantum data inputs and yield quantum outputs and/or measurements. This setting naturally interfaces with learning tasks defined on quantum states or quantum processes, where the quantum circuit is not merely a subroutine but the principal representational backbone.

C. Contributions and organization

This review provides a synthesis of QDL, offering both a tutorial-style, textbook-quality entry point and a comprehensive guide to the specialized literature on the topic. It serves a dual audience: providing an accessible roadmap for graduate students and researchers entering the field, while delivering an authoritative, structured evaluation for experts in QC or AI. Specifically, the core contributions of this review are to (i) weave the interdisciplinary literature into a coherent narrative that clarifies how QDL emerged and matured; (ii) introduce a precise operational definition of QDL and a unifying taxonomy that categorizes its four primary architectural paradigms; (iii) connect abstract learning principles to physical implementation by adopting a full-stack perspective spanning models, software toolchains, and hardware platforms; (iv) critically assess claims of quantum advantage by distinguishing between provable theoretical separations and heuristic empirical evidence, culminating in a rigorous protocol for fair, task-specific benchmarking; and (v) outline a forward-looking roadmap across near-term and fault-tolerant regimes. These elements address the central questions driving the field: in which domains can quantum models provide a genuine, defensible learning advantage, and what are the fundamental resource constraints to realizing that potential?

The remainder of the review is organized as follows. Section II establishes the foundational concepts of QC and classical DL, formally defining the framework of computational symbiosis that evaluates hybrid systems under explicit resource contracts. Building on this foundation, Sec. III constructs the theoretical blueprint of QDL architectures. It systematically surveys data encoding strategies, hybrid quantum-classical models, quantum deep neural networks, and quantum algorithms for classical deep learning primitives. It also examines quantum-inspired classical algorithms as essential baselines, details the prerequisite optimization techniques, and critically evaluates the theoretical limits of trainability and quantum advantage. Section IV grounds these abstract models in physical reality. We introduce a five-layer QDL implementation stack and survey the leading quantum hardware platforms that form the bedrock of experimental execution. Following a review of the software frameworks that bridge algorithms and hardware, this section synthesizes milestone experimental demonstrations and critically analyzes the formidable scaling barriers imposed by hardware noise, connectivity constraints, and measurement bottlenecks.

In Sec. V, we survey the application domains of QDL, distinguishing between classical data inputs (e.g., image classification and NLP) and coherent quantum data processing. We conclude this section by introducing a rigorous, four-pillar comparative protocol for evaluating empirical QDL performance against matched classical baselines. Finally, Sec. VI synthesizes our key findings and presents a strategic research roadmap. We highlight the three coupled tensions, including the expressivity-trainability trade-off, the trainability-simulability constraint, and the quantum-classical interface bottleneck, that fundamentally govern the field's trajectory. The review concludes by outlining actionable milestones to navigate these challenges, charting the transition of QDL from near-term heuristics to early fault-tolerant systems and, ultimately, application-scale fault-tolerant quantum intelligence.

II. FOUNDATIONS AND PRELIMINARIES

This section reviews the quantum mechanical postulates and computational models that govern quantum information processing, alongside the mathematical frameworks that define classical DL. Building upon these independent foundations, we introduce the concept of computational symbiosis (Sec. II.C), a rigorous preliminary framework that formalizes how quantum and classical resources interact, exchange data, and jointly optimize objective functions under realistic physical constraints.

A. Quantum computing basics

1. Quantum states and resources for learning

The basic unit of quantum information is a quantum bit, or qubit, represented in Dirac notation by a normalized vector in a two-dimensional complex Hilbert space, expanded in the computational basis $|0\rangle$ and $|1\rangle$ as (Dirac, 1930; Nielsen and Chuang, 2010):

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1}$$

where $\alpha, \beta \in \mathbb{C}$ satisfy $|\alpha|^2 + |\beta|^2 = 1$. Pure single-qubit states admit a geometric parametrization on the Bloch sphere up to an overall physically irrelevant global phase (Bloch, 1946; Nielsen and Chuang, 2010):

$$|\psi\rangle = \cos(\vartheta/2)|0\rangle + e^{i\varphi}\sin(\vartheta/2)|1\rangle, \tag{2}$$

with polar angle $\vartheta \in [0, \pi]$ and azimuthal angle $\varphi \in [0, 2\pi)$.

For n qubits, the Hilbert space dimension scales as 2^n and the tensor-product structure admits nonseparable (entangled) states in which correlations cannot be reproduced by product states (Amico *et al.*, 2008; Eisert, 2021; Horodecki *et al.*, 2009; Jozsa and Linden, 2003). Entanglement is analyzed as a circuit-template property alongside ansatz expressibility (see Sec. II.C) (Sim *et al.*, 2019), but it can also degrade trainability in regimes where typical states approach volume-law entanglement, yielding entanglement-induced barren plateaus (Ortiz Marrero *et al.*, 2021; Patti *et al.*, 2021). Decoherence, arising from coupling to uncontrolled environmental degrees

of freedom, drives pure states toward mixed states and suppresses phase coherences. Key timescales include the energy-relaxation time T_1 and the dephasing time T_2, with $T_2^{-1} = (2T_1)^{-1} + T_\phi^{-1}$ in common Markovian models, where T_ϕ is the pure-dephasing time.

2. Models of computation

Quantum computation is commonly formulated either in a discrete-time gate model or in continuous-time Hamiltonian models, including adiabatic quantum computation (AQC) (Albash and Lidar, 2018) and quantum annealing, typically formulated as an open-system heuristic for optimization at finite temperature (Kadowaki and Nishimori, 1998; Rajak *et al.*, 2022). In closed-system AQC, the system starts in the ground state of an initial Hamiltonian H_{init} and evolves toward a problem Hamiltonian H_{prob}. Adiabatic theorems ensure approximate tracking to the ground state of H_{prob} (Jansen *et al.*, 2007) under sufficiently slow schedules set by the minimum spectral gap and the schedule's smoothness. In the gate-based model, dynamics are expressed as a sequence of unitaries acting on a qubit register (Barenco *et al.*, 1995; Deutsch, 1985; DiVincenzo, 1995). Continuous-time evolution can be compiled into gate sequences via product-formula methods (Childs *et al.*, 2021; Faehrmann *et al.*, 2022; Suzuki, 1976a,b, 1985), while qubitization and quantum signal processing can achieve improved or optimal asymptotic query scaling for Hamiltonian simulation (Low and Chuang, 2019), providing a bridge between Hamiltonian time evolution and gate-based implementations. Moreover, a rigorous polynomial equivalence between AQC and the gate-based model can be proven (Aharonov *et al.*, 2004).

Given this computational equivalence, and because modern QDL architectures are overwhelmingly constructed using PQCs, the remainder of this review focuses primarily on the discrete-time gate model. In this paradigm, arbitrary unitary evolutions are decomposed into sequences of elementary operations. Specifically, a universal gate set comprising arbitrary single-qubit unitaries and a fixed two-qubit entangling gate can approximate any n-qubit unitary to arbitrary precision (Barenco *et al.*, 1995; DiVincenzo, 1995; Nielsen and Chuang, 2010). Within these gate sets, a fundamental complexity-theoretic distinction exists between Clifford and non-Clifford operations. The n-qubit Clifford group, generated, for example, by Hadamard, phase, and CNOT gates, maps Pauli operators to Pauli operators under conjugation. This property underlies the stabilizer formalism, meaning that a stabilizer state is the simultaneous $+1$ eigenstate of an Abelian group of commuting Pauli operators. Consequently, by the Gottesman-Knill theorem, Clifford circuits combined with Pauli measurements can be efficiently simulated classically (Aaronson and Gottesman, 2004; Gottesman, 1998). By contrast, non-Clifford resources, such as the T gate, are strictly required for universal computation. In fault-tolerant architectures, these non-Clifford gates dominate computational overhead due to the necessity of magic state distillation for error-correcting codes where such gates are non-transversal (Bravyi and Kitaev, 2005).

3. Parameterized quantum circuits as trainable models

Near-term QDL architectures predominantly employ continuously parameterized gates to enable gradient-based optimization. Standard single-qubit rotations $R_\alpha(\theta) = e^{-i\theta\sigma_\alpha/2}$ for $\alpha \in \{x, y, z\}$ combined with two-qubit entangling operations such as the controlled-NOT (CNOT)

$$\text{CNOT} = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X, \tag{3}$$

or parameterized $R_{ZZ}(\theta)$ constitute the elementary vocabulary from which PQC ansätze are assembled. A PQC serves as the foundational hypothesis class for variational and hybrid QDL, specifying a parameter-dependent unitary evolution characterized by tunable parameters $\boldsymbol{\theta}$ and L sequential ansatz layers (Benedetti *et al.*, 2019; Cerezo *et al.*, 2021a; McClean *et al.*, 2016). The effective hypothesis class is determined by the induced input-output map once the data-encoding (state preparation), measurement observables, and any classical postprocessing are specified. The specific structural decomposition of this unitary imposes an inductive bias that governs both the expressibility and trainability of the model (Abbas *et al.*, 2021; Sim *et al.*, 2019), a critical distinction that is formalized in Sec. II.C. Broadly, PQC architectures fall into two broad families: structure-driven constructions that encode specific problem symmetries or geometric priors (Cong *et al.*, 2019; Larocca *et al.*, 2022b), and hardware-efficient ansätze designed to minimize compilation overhead by aligning strictly with native gate sets and physical qubit connectivity (Kandala *et al.*, 2017). A standard hardware-efficient ansatz alternates single-qubit parameterized rotations with topology-respecting entangling layers:

$$U_{\text{HE}}(\boldsymbol{\theta}) = \prod_{l=1}^{L} \left[\left(\bigotimes_{i=1}^{n} U_i(\theta_{l,i}) \right) U_{\text{ent}}^{(l)} \right], \tag{4}$$

where n is the number of qubits, $U_i(\theta_{l,i})$ denotes a parameterized single-qubit gate on qubit i in layer l, and $U_{\text{ent}}^{(l)}$ is a fixed or parameterized entangling layer composed of native two-qubit interactions respecting the hardware topology. To balance expressivity with hardware constraints, the design of these circuits can also be partially automated using quantum architecture search or structure-learning methods, which dynamically optimize

circuit connectivity and depth alongside the continuous parameter (Du *et al.*, 2022; Ostaszewski *et al.*, 2021; Rapp *et al.*, 2025).

Once the architecture is defined, the PQC is embedded within a hybrid quantum-classical optimization loop. In a standard supervised-learning framing, the computational pipeline composes a data-dependent state preparation map $\rho_{\text{in}}(x)$, where x denotes classical input data, the parameterized unitary evolution $U(\boldsymbol{\theta})$, and a final measurement step. Crucially, the direct output of the quantum device is a discrete measurement sample (bitstrings) drawn from the final state's probability distribution. These finite-shot statistics are subsequently processed on a classical host to estimate empirical quantities of interest, such as sample probabilities or expectation values. A common scalar feature used for model evaluation is the expectation value of a designated Hermitian observable O:

$$f(x;\boldsymbol{\theta}) = \mathrm{Tr}\big[\, O \, \big(U(\boldsymbol{\theta})\, \rho_{\text{in}}(x)\, U^{\dagger}(\boldsymbol{\theta})\big)\big]\,, \tag{5}$$

Figure 2 makes this hybrid loop explicit. A central practical consequence of this explicit feedback path is that wall-clock training time is heavily constrained by the execute-measure-communicate cycle. Optimization bottlenecks are typically dominated by the massive shot budgets required to suppress estimator variance, compounded by the quantum-classical latency inherent to the loop (Lubinski *et al.*, 2022; Menickelly *et al.*, 2023). These stringent operational constraints strongly motivate the use of shallow circuits, noise-aware compilation and execution strategies (Ji *et al.*, 2025a; Montanez-Barrera *et al.*, 2025) that maximize the effective accuracy per circuit evaluation (Endo *et al.*, 2021).

4. Quantum measurement and classical interface

Measurement is the quantum-classical interface, mapping a quantum state to classical outcomes and, except for quantum-nondemolition settings, disturbs the post-measurement state (Nielsen and Chuang, 2010). For an observable $O = \sum_i \lambda_i M_i$ and a quantum state ρ, projective measurement yields outcome λ_i with probability $p(\lambda_i) = \mathrm{Tr}(M_i\rho)$ and collapses the state to $M_i\rho M_i/p(\lambda_i)$, where M_i are orthogonal projectors (Nielsen and Chuang, 2010; Von Neumann, 1955). More general measurements are described by positive operator-valued measures (POVMs) (Holevo, 1973; Nielsen and Chuang, 2010; Peres, 2002) specified by effects $\{E_m\}$ with $E_m \geq 0$, and $\sum_m E_m = I$, yielding

$$p(m) = \mathrm{Tr}(E_m\rho). \tag{6}$$

In hybrid algorithms, measured observables are typically decomposed into sums of Pauli strings:

$$O = \sum_k c_k P_k, \tag{7}$$

where $P_k = \bigotimes_{j=1}^{n} \sigma_{k_j}^{(j)}$ with $\sigma_{k_j}^{(j)} \in \{I, X, Y, Z\}$ and $c_k \in \mathbb{R}$. Here, I is the identity and X, Y, Z are the Pauli matrices. Because measurement outcomes are stochastic and noncommuting terms require different measurement bases, objectives such as $\langle O\rangle = \mathrm{Tr}(O\rho)$ are estimated from repeated state preparations and measurements, by grouping commuting terms. For independent single-copy measurements and bounded-outcome estimators, achieving additive precision ϵ at failure probability at most δ requires a number of shots that scales as $\mathcal{O}(\log(1/\delta)/\epsilon^2)$ for a single bounded observable. For sums of many Pauli terms, the shot cost additionally depends on coefficient weights and the chosen grouping strategy, and can become the dominant runtime bottleneck (Chen *et al.*, 2024c; Gonthier *et al.*, 2022).

Advanced schemes target this sampling bottleneck. Randomized measurement schemes (classical shadows) framework (Aaronson, 2018; Huang *et al.*, 2020) constructs a compact classical record ("shadow") from randomized single-copy measurements. To estimate C expectation values $\{\mathrm{tr}(O_i\rho)\}_{i=1}^{C}$ to additive error ϵ with failure probability at most δ, it suffices to take

$$N = \mathcal{O}\left(\frac{\log(C/\delta)}{\epsilon^2} \max_{i\leq C} \|O_i\|_{\text{shadow}}^2\right), \tag{8}$$

where $\|\cdot\|_{\text{shadow}}$ depends on the chosen measurement ensemble (Huang *et al.*, 2020). For random single-qubit Pauli measurements, $\|O_i\|_{\text{shadow}}^2$ scales at most exponentially with locality k. In particular, for an observable acting nontrivially on at most k qubits $\|O_i\|_{\text{shadow}}^2 \leq 4^k\|O_i\|_{\infty}^2$ (Huang *et al.*, 2020), implying

$$N = \mathcal{O}\left(\frac{4^k \log(C/\delta)}{\epsilon^2}\right) \tag{9}$$

at fixed k and bounded operator norm. Classical shadows also serve as a canonical "measure-first" protocol (Gyurik *et al.*, 2025) in quantum-data applications. In this framework, quantum states are first converted into classical representations via a fixed measurement strategy before being processed by a downstream classical learning algorithm.

5. Near-term and fault-tolerant quantum computing

As introduced in Sec. I.A, the current landscape is dominated by the NISQ era (Preskill, 2018). This regime is characterized by architectures containing 10^2–10^3 physical qubits in leading platforms with imperfect operations restricting reliable execution to relatively shallow circuits (Chen *et al.*, 2023a; Kim *et al.*, 2023b; Leymann and Barzen, 2020). Within QDL, these hardware constraints are particularly severe because most trainable models are instantiated as PQCs (Cerezo *et al.*, 2021a) and their performance degrades markedly under

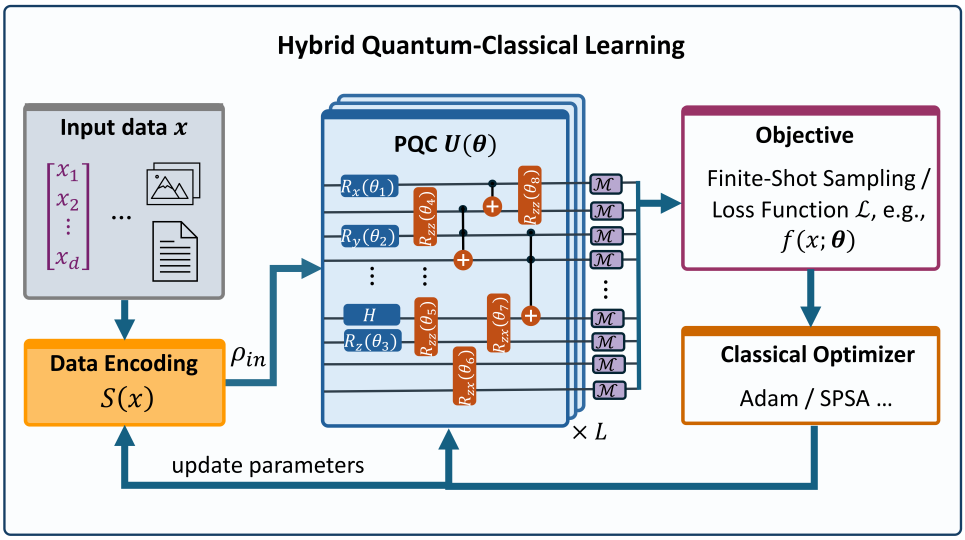

Figure 2 The hybrid quantum-classical learning loop. Classical data x are encoded into an input state $\rho_{\text{in}}(x)$ via the data-encoding unitary $S(x)$. A parameterized quantum circuit (PQC) $U(\boldsymbol{\theta})$ is executed, measurements of an observable yield finite-shot statistics used to estimate a loss function (e.g., expectation value $f(x;\boldsymbol{\theta})$), and a classical optimizer updates $\boldsymbol{\theta}$ by minimizing an objective built from these estimates. The loop repeats until convergence to an optimal parameter set $\boldsymbol{\theta}^*$.

hardware noise (Buonaiuto *et al.*, 2024; Skolik *et al.*, 2023b). Furthermore, compilation overheads required to map algorithmic ansätze onto restricted physical topologies further increase circuit depth, compounding accumulated errors (Buonaiuto *et al.*, 2024; Zhang *et al.*, 2026b). Consequently, the maximum feasible model depth is determined by a strict bottleneck between algorithmic connectivity demands and physical hardware fidelities.

To combat these effects without full fault tolerance, quantum error mitigation (QEM) techniques, such as zero-noise extrapolation and probabilistic error cancellation, attempt to estimate ideal expectation values from noisy circuit executions by trading sampling complexity for improved accuracy (Cai *et al.*, 2023). However, this sampling overhead generally scales exponentially with circuit depth and noise rates. For QDL, these steep mitigation costs severely constrain the utility of QEM for deep or highly entangled networks, heavily incentivizing architectures that are shallow and natively hardware-aligned.

Overcoming these fundamental depth limits requires a transition to fault-tolerant quantum computing (FTQC). In FTQC, logical information is redundantly encoded across multiple physical qubits and protected via repeated syndrome extraction and correction (Gottesman, 2009; Shor, 1995; Steane, 1996; Terhal, 2015). Below a critical physical error threshold, topological schemes such as the surface code can suppress logical error rates exponentially with code distance (Fowler *et al.*, 2012). However, this protection incurs massive resource overheads, typically demanding 10^2–10^4 physical qubits per logical qubit depending on the target logical error and ancilla requirements. Furthermore, non-Clifford resources, most notably T-gate distillation factories, often dominate the total spatial footprint and runtime of the architecture (Fowler *et al.*, 2012; Gidney *et al.*, 2025), while the classical control stack demands high-throughput, low-latency decoders to keep pace with rapid syndrome cycles (Bausch *et al.*, 2024). To bridge the gap between NISQ and FTQC, researchers are also exploring "almost fault-tolerant" or early-fault-tolerant regimes, which selectively reduce distillation overheads by leaving certain non-Clifford resources less strictly protected (Kang *et al.*, 2025).

Contextualizing this trajectory, Eisert and Preskill (2025) outline four transitional gaps the field needs to cross to reach practical utility: (i) from error mitigation to active error detection and correction, (ii) from rudimentary error correction to scalable fault tolerance, (iii) from early heuristics to mature, verifiable algorithms, and (iv) from exploratory simulators to credible advantage in quantum simulation. For the specific domain of QDL, gaps (i) and (ii) strictly bound the feasible circuit depth, and therefore the representational capacity, of any quantum model under a fixed resource contract. Meanwhile, gaps (iii) and (iv) underscore the imperative for resource-efficient validation and rigorous benchmarking against matched classical baselines. Ultimately, achieving end-to-end practical performance in QDL requires deep codesign across the entire stack, optimizing algorithms, compilation strategies, and QEC architectures in tandem (Babbush *et al.*, 2025).

B. Classical DL basics

1. Model classes and energy-based formalisms

DL studies high-capacity parametric hypothesis classes, most commonly multilayer neural networks, trained by variants of empirical risk minimization on differentiable objectives that enable efficient gradient-based optimization (Goodfellow *et al.*, 2016; LeCun *et al.*, 2015). Formally, a learning problem specifies a parametric family $f_\theta : \mathcal{X} \to \mathcal{Y}$, an objective, and a data-generating distribution or empirical sample over inputs and targets. Universal approximation theorems (Cybenko, 1989; Hornik *et al.*, 1989; Lu *et al.*, 2017) establish approximation capacity at sufficient width or depth under standard regularity assumptions, but do not address optimization, data requirements, or generalization. Practical performance is governed by trainability, sample efficiency, and architectural constraints that encode inductive bias.

Energy-based models (EBMs) are particularly relevant to QDL because they share structural features with quantum many-body systems and provide a natural interface for quantum sampling primitives (Amin *et al.*, 2018). These models specify a scalar energy function $E_\theta(z)$ over configurations z, such as observations, labels, and latent variables. In probabilistic EBMs, this energy induces an unnormalized density with explicit normalization by a partition function (Ackley *et al.*, 1985; Hopfield, 1982; LeCun *et al.*, 2006). For binary configurations $z \in \{0,1\}^n$ this induces the Gibbs distribution

$$p_\theta(z) = \frac{1}{Z_\theta} \exp(-\beta E_\theta(z)), \tag{10}$$

where $Z_\theta = \sum_z \exp(-\beta E_\theta(z))$ is the partition function and $\beta > 0$ is an inverse temperature parameter.

Hopfield and Boltzmann-type models provide a useful historical and conceptual bridge between classical DL and statistical-mechanics formalisms. The original Hopfield network can be cast as deterministic dynamics that perform energy minimization, yielding a prototypical associative-memory mechanism (Hopfield, 1982). Boltzmann machines generalize this picture by introducing stochasticity and making expectation estimation under the induced Gibbs distribution a central computational primitive, which is precisely where likelihood training becomes expensive (Ackley *et al.*, 1985; Hinton *et al.*, 2006). Restricted Boltzmann machines (RBMs) and related EBMs leverage tractable conditional structure to enable approximate learning via gradient surrogates (Hinton, 2002; Hinton *et al.*, 2006; LeCun *et al.*, 2006). Modern Hopfield formulations have connected content-addressable retrieval to attention-style readout rules, providing a concrete link between associative-memory energy landscapes and contemporary sequence modeling (Ramsauer *et al.*, 2020; Vaswani *et al.*, 2017).

2. Training dynamics and theoretical regimes

Deep networks are trained by stochastic optimization of nonconvex objectives using gradients computed by reverse-mode differentiation through composed operations (Goodfellow *et al.*, 2016). In large-scale settings, gradients are estimated on minibatches, and adaptive methods such as Adam (Kingma and Ba, 2014) modify updates using running moment estimates. Modern practice frequently operates in an overparameterized regime, with parameter count comparable to or exceeding samples in supervised settings. In such settings, many low-loss (and often interpolating) solutions can coexist, and the particular solution reached can depend strongly on implicit regularization arising from optimization dynamics, initialization, and algorithmic choices (Advani *et al.*, 2020; Bahri *et al.*, 2020).

Width scaling provides analytically tractable reference regimes. In the infinite-width limit, under appropriate initialization and scaling, training can enter the "lazy" or neural tangent kernel (NTK) regime, where the network is well approximated by its linearization around initialization. For a model $f(x;\theta) \in \mathbb{R}^m$, where m is the output dimension, let $J_\theta f(x) \in \mathbb{R}^{m\times p}$ denote the Jacobian with respect to parameters θ, with p the number of parameters. The NTK is

$$\Theta(x, x'; \theta) = J_\theta f(x)\, J_\theta f(x')^\top. \tag{11}$$

In the strict NTK limit, Θ converges to a deterministic kernel that remains effectively constant during training(Jacot *et al.*, 2018), so gradient descent is well approximated by kernel regression dynamics. Chizat *et al.* (2019) clarify when this approximation is appropriate and emphasize that substantial representation change (feature learning) requires departing from the strictly kernelized regime. At finite widths and standard learning rates, networks can enter a "rich" regime in which internal representations evolve substantially and the effective kernel becomes time-dependent and data-adaptive (Bordelon *et al.*, 2025; Mei *et al.*, 2018).

3. Symmetry and inductive bias

In high-capacity settings, empirical risk minimization alone does not uniquely specify a solution: many functions can interpolate a finite training set. Consequently, achieving true *generalization*, i.e., the ability of a model to maintain predictably low expected error on novel, unseen data drawn from the underlying distribution, rather than merely memorizing the training examples, cannot rely on fitting the data alone. Because the learning algorithm inherently interpolates behavior between sparsely observed data points, successful generalization depends strictly on *inductive bias*: the structural constraints or algorithmic preferences that favor particular solutions

within the hypothesis space. Architectures encode inductive bias by restricting hypothesis classes through symmetry, locality, and connectivity constraints. Formally, when data live on domains with group actions, inductive biases can be implemented by designing networks to be invariant or equivariant under those actions (Bronstein *et al.*, 2021). Convolutional neural networks, for example, encode a bias toward local, translation-equivariant features via weight sharing and local receptive fields (LeCun *et al.*, 1989). Group-equivariant generalizations extend this principle to broader transformation groups by enforcing equivariance at the level of the hypothesis class (Cohen and Welling, 2016).

For structured non-Euclidean data, graph neural networks implement permutation-equivariant message passing by aggregating neighbor information through permutation-invariant pooling operations (Battaglia *et al.*, 2018; Gilmer *et al.*, 2017; Kipf and Welling, 2017). For sequence and language modeling, the dominant architectural prior shifts from local receptive fields to global context aggregation. Transformer architectures implement this via self-attention, which computes data-dependent pairwise interactions across tokens in a sequence (Vaswani *et al.*, 2017). This structural and group-theoretic framing provides the essential lens for evaluating hybrid quantum-classical models. Establishing a genuine quantum advantage requires explicitly articulating how the quantum module alters the underlying inductive bias. Consequently, validating these theoretical separations demands rigorous empirical benchmarking against classical baselines equipped with equivalent geometric symmetries.

4. Generative modeling and reinforcement learning

Beyond supervised prediction, deep networks are widely used for generative modeling and sequential decision-making, two settings in which probabilistic inference and sampling play a central role and are often discussed as potential targets for quantum acceleration. Generative modeling requires sampling from complex distributions and, in many formulations, estimating expectations or score functions. As noted in Sec. II.B.1, likelihood-based training of EBMs requires expectations under the model distribution and therefore relies on approximate inference. A broad class of modern generative methods avoids explicit energy normalization. Denoising diffusion probabilistic models and related score-based methods reparameterize learning in terms of non-equilibrium stochastic dynamics, learning to reverse a prescribed forward noising process by estimating the score function $\nabla_x \log p(x)$ or an equivalent denoising objective (Ho *et al.*, 2020; Song *et al.*, 2021).

Reinforcement learning (RL) uses deep networks as function approximators to solve sequential decision problems, commonly formalized as a Markov decision process (Mnih *et al.*, 2015; Puterman, 1994; Sutton and Barto, 2018). In this framework, an agent interacts with its environment by observing a state $s \in \mathcal{S}$ and selecting an action (control) $c \in \mathcal{C}$. The objective of RL is to learn a parameterized stochastic policy $\pi_\theta(c|s)$, where θ denotes the trainable parameters (e.g., network weights), that maximizes the expected cumulative reward over time, typically when the environment dynamics are unknown and learning relies on sampled trajectories rather than differentiating through the environment. Because the environment's transition dynamics cannot be analytically differentiated, policy-gradient methods (Sutton *et al.*, 1999; Williams, 1992) instead estimate policy gradients from sampled trajectories using log-derivative identities. This sampling-centric structure motivates connections to quantum RL and quantum optimal control (QOC) (Dong *et al.*, 2008; Dunjko *et al.*, 2016), where measurement outcomes provide intrinsic stochastic samples that can be used to estimate gradients or returns under explicit access models.

5. Scaling laws and emergent regimes

Classical DL provides a moving baseline because performance often improves approximately with resources over measured ranges. In large-scale language modeling and related settings, empirical scaling laws describe approximate power-law relationships between loss or downstream performance and resources such as model size, dataset size, and training compute over broad ranges (Hoffmann *et al.*, 2022; Kaplan *et al.*, 2020). Analyses aim to explain these scalings by distinguishing noise-limited from resolution-limited regimes and relating them to target-function structure and data geometry (Bahri *et al.*, 2024). In contemporary practice, self-supervised pretraining at scale underpins transfer across tasks, notably in foundation-model settings (Bommasani *et al.*, 2022; Chen *et al.*, 2020b).

Scaling can also induce qualitative regime transitions. When capacity increases at fixed data, the model may cross an interpolation threshold where training error first becomes approximately zero. Double descent refers to the resulting nonmonotonic test error curve as a function of capacity or effective degrees of freedom: error decreases in the underparameterized regime (effective degrees of freedom below the number of training examples), peaks near the interpolation threshold (where training error first becomes approximately zero), and decreases again in the overparameterized regime (effective degrees of freedom above the number of training examples) as optimization selects interpolating solutions with improved generalization (Belkin *et al.*, 2019; Nakkiran *et al.*, 2021). Grokking highlights a distinct axis: at fixed architecture and data, continued optimization, often with regulariza-

tion, can yield delayed generalization, where a model fits the training set early yet only abruptly transitions to a generalizing solution late in training (Power *et al.*, 2022). Phase-transition analogies provide a compact language for organizing regime diagrams (Ziyin and Ueda, 2023), but the mathematical structure of these transitions differs from equilibrium statistical mechanics and does not imply universality in the sense of critical exponents or renormalization-group scaling.

C. Computational symbiosis

Hybrid QDL architectures compose quantum and classical modules that are codesigned and jointly optimized under a shared objective (Beer *et al.*, 2020; Cong *et al.*, 2019; Grant *et al.*, 2018; Long *et al.*, 2025; Skolik *et al.*, 2021). To rigorously evaluate these systems, we operate hybrid learning as minimizing an expected loss over a task distribution $\mathcal{D}$ with a fixed evaluation protocol and primary metric, subject to a resource contract $\mathcal{R}$ and an access-and-readout model $\mathcal{A}$. We denote this triple of specifications collectively as $(\mathcal{D}, \mathcal{R}, \mathcal{A})$ and use it to parametrize learning systems under comparison. This framework allows us to formally define the target operational regime of QDL.

Definition 2 (Computational symbiosis).—A learning system exhibits computational symbiosis if (i) it composes at least one quantum module and one classical module coupled through a shared objective; (ii) functional roles are partitioned via explicit interface variables; and (iii) the system is defined and evaluated relative to a specified task distribution $\mathcal{D}$, including the evaluation protocol and primary metric, together with a jointly specified pair $(\mathcal{R}, \mathcal{A})$ consisting of a resource contract $\mathcal{R}$ that fixes budget ceilings on end-to-end resources and an access-and-readout model $\mathcal{A}$ that fixes allowed data-access and measurement permissions.

Computational symbiosis is a systems-level comparative property. It is not an intrinsic property of the quantum circuit alone but of the full hybrid system evaluated under the stated $(\mathcal{D}, \mathcal{R}, \mathcal{A})$. The access-and-readout model $\mathcal{A}$ specifies allowed data-access and query operations together with measurement permissions, whereas the resource contract $\mathcal{R}$ specifies ceilings on end-to-end variables, including qubit count, circuit depth, total circuit executions, wall-clock time, hyperparameter tuning budget, classical memory, and error mitigation overhead, subject to a target statistical precision. In deployed settings, abstract ceilings map to operational budgets such as compiled circuit depth D_{eff}, time per circuit execution t_{shot}, iteration latency T_{iter}, and a stability window τ_{stable} over which device behavior is approximately stationary (see Sec. IV.E.1).

Symbiosis is evidenced on $\mathcal{D}$ when, under fixed $(\mathcal{R}, \mathcal{A})$, the hybrid yields a statistically significant improvement in a pre-specified primary metric, or matches the metric while reducing at least one ceiling in $\mathcal{R}$, compared against the strongest admissible classical baseline. Crucially, this requires that the quantum module implement a clearly specified encoding-ansatz-readout protocol in which the signal remains learnable within D_{eff} and τ_{stable}, allowing any metric gain or budget reduction to be confidently credited to the quantum component.

Achieving computational symbiosis in practice requires navigating the interplay among model expressivity, trainability, and classical simulability under finite measurement cost and bounded resources. It is crucial to distinguish *ansatz expressibility*, a circuit-ensemble notion quantifying how broadly a PQC explores state space (Sim *et al.*, 2019), from *model expressivity* or *expressive power*, which is the task-level representational capacity of the induced hypothesis class (Abbas *et al.*, 2021; Wu *et al.*, 2021c). Highly expressive unstructured circuits are prone to gradient concentration, i.e., the barren-plateau phenomenon (Sec. III.B.3), creating a fundamental tension between representational power and optimization feasibility. This tension directly impacts trainability, which denotes the physical feasibility of extracting an optimization signal to a target precision within the operational budgets of $\mathcal{R}$ (Cerezo *et al.*, 2021b; Holmes *et al.*, 2022; McClean *et al.*, 2018). Increased compiled depth further suppresses the signal-to-noise ratio by exposing the system to gate and decoherence errors, raising the shot budget required to estimate gradients to a fixed precision (Arrasmith *et al.*, 2021; Chinzei *et al.*, 2025). Shifting representational capacity into the classical component can improve robustness to finite-shot noise, but it reallocates cost to classical compute and memory, both of which are counted against $\mathcal{R}$.

Addressing trainability, however, often exacerbates the *trainability-simulability tension*. Provable trainability is often achieved by imposing structural restrictions on the quantum circuit that can also enable efficient classical evaluation of the same objective. In several analyses, the objective becomes classically efficient to evaluate after a one-time quantum data-acquisition stage (Cerezo *et al.*, 2025). Consequently, trainability certificates alone do not establish classical intractability unless compared against classical algorithms granted identical $(\mathcal{R}, \mathcal{A})$ assumptions. Recent analyses emphasize that many existing trainability guarantees apply only to restricted circuit families or regimes that also admit efficient classical simulation (Cerezo *et al.*, 2025). While this implication is not known to be universal, it underscores the need to strictly specify the architecture class when claiming "trainable yet beyond-classical" behavior.

Computational symbiosis is a comparative systems property demonstrable at current scales, whereas establishing formal *quantum advantage* is a significantly stronger claim that requires scalable asymptotic separation relative to an explicit classical comparator (Lanes

et al., 2025).

Definition 3 (QDL advantage).—A QDL advantage claim requires specifying: (i) a task and data-generating model defining $\mathcal{D}$; (ii) an access-and-readout model $\mathcal{A}$; (iii) the classical comparator class restricted to the same access-and-readout assumptions and evaluated under the same resource contract $\mathcal{R}$; (iv) the error metric, target accuracy ϵ, and success probability $1-\delta$; (v) an end-to-end resource contract $\mathcal{R}$; and (vi) a validation or verification procedure credible at the target scale under $\mathcal{A}$ together with an explicit separation statement identifying which ceilings in $\mathcal{R}$ are improved at fixed (ϵ, δ) relative to the comparator.

Throughout this review, we reserve the term "QDL advantage" for statements intended to satisfy Definition 3. Otherwise, we describe empirically observed improvements as computational symbiosis or task- and budget-conditional performance gains.

III. QDL ARCHITECTURES

This section surveys the core architectures and algorithmic building blocks that define QDL, concluding with a critical evaluation of the theoretical advantages and inherent limitations.

A. Data encoding and preparation

Classical data, such as real vectors or matrices, require encoding into quantum states for processing on quantum hardware (Lloyd *et al.*, 2020; Schuld *et al.*, 2021). The encoding scheme determines the model's inductive bias, the geometry of the induced feature map (Larocca *et al.*, 2022b; Nguyen *et al.*, 2024), and the dominant state-preparation and readout costs. Encodings exhibit dataset- and implementation-dependent accuracy-runtime tradeoffs. Benchmark analyses indicate strong task-, noise-, and implementation-dependence, with no encoding consistently dominating across studied regimes (Bowles *et al.*, 2024; Rath and Date, 2024).

1. Basis and amplitude encodings

Basis encoding discretizes the input data into classical bit strings and then transfers these bits one by one into computational basis states in the quantum system (Schuld and Petruccione, 2018). Given a classical data point with an N-bit binary representation $x = (x_1, x_2, \ldots, x_N)$ with $x_j \in \{0, 1\}$, the basis-encoded state is

$$|x\rangle = |x_1, x_2, \ldots, x_N\rangle \ . \tag{12}$$

Initialization of $|x\rangle$ from $|0\rangle^{\otimes N}$ only requires at most N Pauli-X operations. Furthermore, because these encoded states are classical bit strings forming an orthogonal computational basis, they permit copying via basis-preserving unitaries, such as a CNOT fan-out operation ($\mathrm{CNOT}|x_j\rangle|0\rangle = |x_j\rangle|x_j\rangle$), without violating the no-cloning theorem, which only strictly prohibits the universal copying of unknown, arbitrary superpositions. Despite these initialization advantages, basis encoding suffers from severe spatial bottlenecks when applied to data that is not originally binary. Because each real-valued feature is discretized into multiple bits at a fixed numerical precision, the encoding requires linear qubit overhead. For example, if a continuous variable requires a τ-bit binary representation to achieve the target precision, an n-qubit register can encode at most $\lfloor n/\tau \rfloor$ scalar features, making this approach prohibitively expensive for high-dimensional classical deep learning tasks.

Amplitude encoding is a common conceptual primitive in QLA-based approaches to QML (Lloyd *et al.*, 2013, 2014; Rebentrost *et al.*, 2014; Wiebe *et al.*, 2012). In this scheme, a d-dimensional data vector $\boldsymbol{x} = (x_0, x_1, \ldots, x_{d-1})$ is normalized and encoded as amplitudes of an $n = \lceil \log_2 d \rceil$ qubit state

$$|\psi_{\mathrm{amp}}(\boldsymbol{x})\rangle \ = \ \sum_{j=0}^{2^n-1} \frac{x_j}{\|\boldsymbol{x}\|_2} \, |j\rangle \, . \tag{13}$$

Although qubit count scales logarithmically with data dimension, this advantage is offset by state-preparation cost in the worst case: preparing a generic n-qubit amplitude-encoded state can require resources exponential in n, and even achieving a circuit depth of $\Theta(n)$, where the number of sequential gate layers scales linearly with the number of qubits, preparation may demand an exponential number of ancillas unless additional structure is exploited (Zhang *et al.*, 2022). Moreover, extracting global quantities can require exponentially many shots in n under standard access models (Aaronson, 2015; Zhang and Yuan, 2024).

2. Data-access models and QRAM

Speedup claims in amplitude-encoding and block-encoding pipelines are conditional on the input-access model, which determines how classical data is loaded into quantum states or unitaries, and the output model, which determines what classical information is extracted. When inputs are provided by explicit classical descriptions, constructing the state-preparation and access primitives can itself dominate the cost. Zhang and Yuan (2024) prove near-optimal Clifford+T circuit complexity bounds for constructing typical quantum access primitives from explicit classical descriptions and show that for unstructured matrices this cost can scale near linearly with matrix dimension in worst-case settings. Consequently, end-to-end resources can be set by data access

unless stronger access assumptions or exploitable structure are present.

Many proposals therefore assume quantum random access memory (QRAM) (Aaronson, 2015; Giovannetti *et al.*, 2008) to enable coherent superposition queries to a classical dataset. However, QRAM feasibility and overhead are highly architecture-dependent. For standard circuit-model QRAM, architectural analyses show that a superposition query can activate $\mathcal{O}(N_{\text{mem}})$ routers, where N_{mem} denotes the number of stored addresses. Because this circuit cost scales at least proportionally to the number of stored bits, it threatens to erase putative asymptotic advantages at large N_{mem} (Jaques and Rattew, 2025). Furthermore, causality-based bounds constrain how the QRAM size N_{mem} can scale for fixed clock-cycle time and spatial layout (Wang *et al.*, 2024b). Surprisingly, hardware noise is not necessarily the decisive obstruction in these limits; for instance, Hann *et al.* (2021) proved that the tree-structured bucket brigade QRAM can retain favorable (poly)logarithmic-in-N_{mem} query-infidelity scaling even under arbitrary local error channels, distinguishing it from naive fanout designs.

Instead, the primary obstruction for large-scale QRAM is often the fault-tolerant gate compilation cost. Because QRAM routers typically require heavily controlled operations (e.g., Toffoli gates), they accumulate massive non-Clifford depth. As established in Sec. II.A.2, non-Clifford gates require highly resource-intensive magic state distillation, which frequently dictates the dominant spatial footprint and temporal bottleneck of the entire FTQC architecture. Consequently, reducing non-Clifford depth is an imperative optimization target for the practical viability of data-loading oracles. Nevertheless, recent polynomial-encoding circuit constructions have been developed specifically to reduce the non-Clifford depth of QRAM lookups (Mukhopadhyay, 2025). Parallel efforts seek to mitigate these access bottlenecks by bypassing exact QRAM entirely, utilizing approximate amplitude encoding to trade strict fidelity for dramatically shallower circuits (Nakaji *et al.*, 2022), alongside alternative state-preparation protocols, such as repetitive amplitude encoding and structure-aware function-loading schemes (Balewski *et al.*, 2024; Gonzalez-Conde *et al.*, 2024; Li *et al.*, 2025b; Rosenkranz *et al.*, 2025).

3. Dynamic encoding and the Fourier perspective

Dynamic or Hamiltonian encoding embeds data via the unitary time evolution of Hermitian operators, fundamentally defining the model's inductive bias through the algebraic structure of the map (Gil-Fuster *et al.*, 2024a; Schuld *et al.*, 2021). Multivariate inputs $x \in \mathbb{R}^d$ are encoded via factorized unitaries

$$S(x) = \prod_{j=1}^{d} e^{-ih_j x_j t}, \tag{14}$$

where h_j are local generators and $t \in \mathbb{R}$ is a fixed or trainable scale parameter. Common choices include single-qubit Pauli operators $h_j \in \{X, Y, Z\}$, sometimes termed angle encoding (Schuld and Petruccione, 2021). To induce correlations in the feature map, the encoding can include multi-qubit interaction terms, as in Ising-type feature maps (Havlíček *et al.*, 2019).

This encoding scheme admits a useful analytical framework. Unlike classical bias, which arises from architecture and regularization (Goodfellow *et al.*, 2016), QDL bias is strongly shaped by the spectral properties of h_j and the circuit topology. This framework admits a Fourier series representation of the model, where the accessible frequency spectrum is determined by the generators' eigenvalue gaps and the number and placement of data-dependent layers, and the chosen measurement (Jaderberg *et al.*, 2024; Schuld *et al.*, 2021; Shin *et al.*, 2023). In kernelized formulations the same bias is captured by the kernel eigenspectrum (Kübler *et al.*, 2021; Thanasilp *et al.*, 2024). Repeating the encoding layers interleaved with trainable unitaries, a strategy known as data reuploading, expands the accessible trigonometric feature family and, in theory, can yield universal approximation guarantees for continuous functions on compact domains given sufficiently many reuploading layers and parameters (Pérez-Salinas *et al.*, 2020), but in practice, the resulting extended circuits frequently encounter severe trainability bottlenecks. Concentration phenomena can render naive fidelity kernels uninformative, necessitating metrics that quantify when quantum decision geometries diverge meaningfully from classical baselines (Huang *et al.*, 2021a; Thanasilp *et al.*, 2024). To mitigate this, one may impose structural constraints: utilizing equivariant constructions that enforce group structure to align the Hilbert space geometry with task symmetries (Larocca *et al.*, 2022b; Meyer *et al.*, 2023; Nguyen *et al.*, 2024).

Beyond trainability, increasing the complexity of the quantum encoding directly impacts model generalization. Specifically, encoding-dependent generalization bounds tend to worsen as the accessible frequency spectrum expands, unless explicit structural constraints or regularization techniques are applied (Caro *et al.*, 2021). Fortunately, the training dynamics themselves can provide a form of implicit regularization: under standard gradient-based optimization, these models frequently exhibit a *spectral bias*, preferentially learning low-frequency components before fitting high-frequency noise (Barthe and Pérez-Salinas, 2024). Alternatively, if high expressivity is required without incurring the generalization and trainability penalties of deep quantum circuits, practition-

ers can enhance the model's complexity entirely off-chip by composing a shallow quantum encoding with classical nonlinear preprocessing of the input (Gil Vidal and Theis, 2020).

4. Emerging encoding paradigms

Beyond fixed maps, neural quantum embedding schemes parameterize the encoding circuit and train it to improve separability in the induced quantum feature space before downstream classification layers (Hur *et al.*, 2024; Liu *et al.*, 2025b). Alternatively, reservoir computing exploits the natural high-dimensional dynamics of quantum many-body systems for temporal feature injection (Fujii and Nakajima, 2017; Hayashi *et al.*, 2025; Mujal *et al.*, 2021), while continuous-variable and photonic platforms encode data directly in field quadratures, yielding high-dimensional feature spaces without qubit discretization (Anai *et al.*, 2024; Killoran *et al.*, 2019a; Maier *et al.*, 2025).

B. Hybrid quantum-classical models

Building on the foundational hybrid loops introduced in Sec. II.A.3, hybrid quantum-classical models integrate parameterized quantum subroutines into classical computational graphs, delegating data processing, optimization, and hardware orchestration to classical components (Benedetti *et al.*, 2019; Cerezo *et al.*, 2021a; McClean *et al.*, 2016; Mitarai *et al.*, 2018; Peruzzo *et al.*, 2014). Architecturally, it is useful to partition these hybrid systems based on their training dynamics. "Boundary" families utilize the quantum device strictly as a fixed feature extractor, restricting all gradient updates to the classical readout layers. Prominent examples of this approach include fixed-feature quantum kernel methods (Havlíček *et al.*, 2019; Schuld *et al.*, 2020b; Schuld and Petruccione, 2018), quantum reservoir computing (Fujii and Nakajima, 2017; Mujal *et al.*, 2021), and fixed-feature quanvolution (Henderson *et al.*, 2020a). In contrast, core variational models establish an end-to-end feedback loop that actively updates the quantum circuit's internal parameters during training. However, this boundary naturally dissolves when quantum modules are embedded within hierarchically deep pipelines. For instance, when the parameters of a quantum feature map or quantum kernel are jointly optimized in-loop alongside classical weights, these trainable variants cross the threshold, merging into the core regime of fully variational hybrid QDL (Alvarez-Estevez, 2025; Gentinetta *et al.*, 2023; Rodriguez-Grasa *et al.*, 2025).

1. Architectural paradigms for hybrid models

Hybrid quantum-classical architectures can be organized most cleanly by *where* the quantum module appears in the computational graph, *which* resources dominate end-to-end cost, e.g., circuit-evaluation budgets versus classical preprocessing, and *what* inductive bias is enforced by the circuit topology, embedding map, and readout. Figure 3 provides a schematic map of these design choices across four recurring paradigms at the level of computation graphs and estimator interfaces, while Table I lists representative instantiations and key references.

(i) placement and composition.—The most direct approach treats the PQC as a shallow and trainable quantum readout or classifier integrated as a compositional stage, often a terminal head, within a classical pipeline (Farhi and Neven, 2018; Schuld *et al.*, 2020a, 2014). Figure 3(a) fixes the corresponding computation graphs, emphasizing where measurement-derived features reenter the classical model in interleaved variants. In transfer learning, a pretrained classical network serves as a frozen or lightly fine-tuned feature extractor, while a quantum head is trained on a low-dimensional vector (Mari *et al.*, 2020). Domain-driven variants further emphasize latency and deployment constraints (Chittem and Kumar Pradhan, 2025). Recent proposals explore non-sequential layered hybrid architectures with parallel feature exchange, rather than a single pass from classical features to a quantum head (Guo *et al.*, 2025; Kabir *et al.*, 2025).

(ii) Hierarchical and geometric structure.—Hierarchical architectures encode symmetries and locality into circuit topology and parameter tying, imposing inductive biases that can improve trainability and sample efficiency when the enforced symmetry and locality match the task (Bian *et al.*, 2024; Grant *et al.*, 2018; Meyer *et al.*, 2023; Schatzki *et al.*, 2024). Figure 3(b) groups these approaches by how structure is enforced. Quantum convolutional neural networks (QCNNs) (Cong *et al.*, 2019) are structured PQC architectures that mirror the locality and hierarchy of classical CNNs for quantum data. Architecturally, a QCNN alternates between quantum convolutional layers, which apply parameterized, local entangling unitaries across neighboring qubits, and quantum pooling layers, which systematically reduce the system's dimensionality by measuring or discarding a subset of the qubits. By instantiating this hierarchical pooling for quantum inputs, QCNNs naturally align with the physical locality of quantum many-body systems and have been rigorously shown to avoid the barren-plateau phenomenon (Pesah *et al.*, 2021), while in some locally structured settings their induced decision rules can also admit efficient classical simulation (Sec. III.E.1). Representative extensions include higher-dimensional inputs (Sander *et al.*, 2025), application-driven simulation studies (Chen

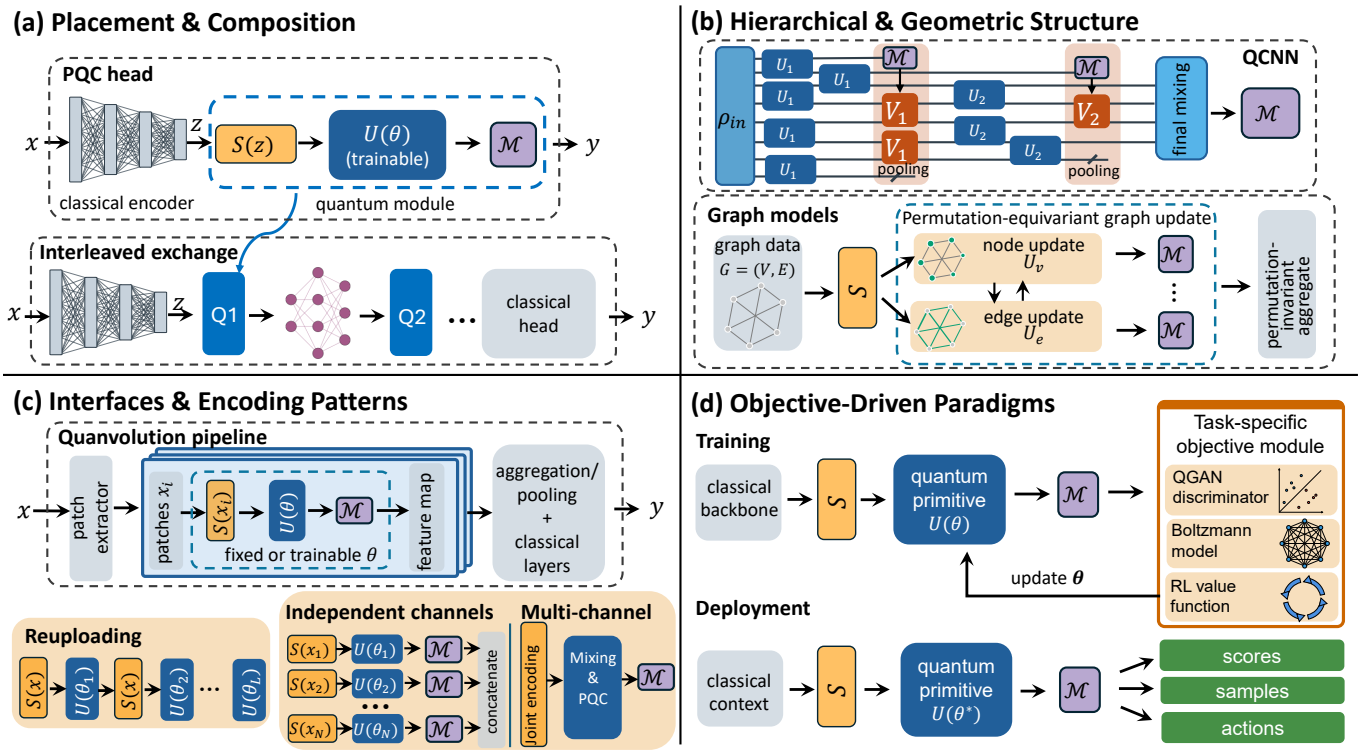

Figure 3 Taxonomy of hybrid quantum-classical architectures across four paradigms. (a) Placement and composition. A classical encoder compresses the input x to a latent code z, which is embedded via $S(z)$; this encoding of z rather than the raw data x directly alters the induced hypothesis class. An interleaved-exchange variant (lower) routes x through alternating quantum modules $Q_1, Q_2, \ldots$ and classical layers where each Q_i subsumes the full encode-process-measure pipeline $S(\cdot) \to U(\boldsymbol{\theta}_i) \to \mathcal{M}$ of the upper subpanel. (b) Hierarchical and geometric structure. Quantum convolutional neural network (QCNN, upper): alternating trainable-unitary convolutional and measurement-based pooling layers act on ρ_{in}, followed by a fully connected final mixing layer; Graph models (lower): node and edge unitaries U_v, U_e implement permutation-equivariant message passing on $G{=}(V, E)$. (c) Interfaces and encoding patterns. Quanvolution (upper): local PQCs process image patches x_i and output a spatial feature map for classical aggregation. Data reuploading (lower left): interleaving $S(x)$ with layers $U(\boldsymbol{\theta}_l)$ implements a Fourier-feature expansion. Independent and multi-channel (lower right): independent channels concatenate per-channel measurement outcomes classically, whereas in multi-channel, joint encoding applies coherent premeasurement mixing, enabling coherent cross-channel mixing in the quantum feature map prior to measurement. (d) Objective-driven paradigms. Training loop (upper): a classical backbone, quantum primitive $U(\boldsymbol{\theta})$, and task objective jointly update $\boldsymbol{\theta}$; instantiations include quantum generative adversarial networks (QGAN) discriminators, Boltzmann models, and reinforcement learning (RL) value functions. Deployment loop (lower): $U(\boldsymbol{\theta}^*)$ operates at trained and fixed parameters under a sampling bottleneck on quantum devices, producing scores, samples, or actions.

et al., 2022c), and alternative-platform realizations (Monbroussou *et al.*, 2025). Variants that modify the effective receptive field or adaptivity (Chen, 2022), and noise-aware architecture optimization further explore robustness-resource trade-offs relevant for deployment (Lee *et al.*, 2025a; MacCormack *et al.*, 2022). Quantum blocks in classical CNN pipelines formalize convolutional hybrids in which local PQC modules act as trainable feature extractors coupled to classical nonlinearities and pooling (Dongfen *et al.*, 2025; Liang *et al.*, 2021; Long *et al.*, 2025). Comparative studies emphasize strong sensitivity to qubit and layer choices and therefore the need for controlled, matched benchmarking (Zaman *et al.*, 2025).

Mirroring the classical group-equivariant networks of Sec. II.B.3, equivariant QNNs construct channels or layers that commute with a target group action, enabling symmetry-preserving hypothesis classes with provable guarantees (Nguyen *et al.*, 2024; Schatzki *et al.*, 2024). Resource-efficient equivariant QCNN constructions can reduce measurement overhead via circuit parallelization (Chinzei *et al.*, 2024). The quantum analog of classical GNNs are quantum graph neural networks (QGNNs), with graph-interaction ansätze (Lu *et al.*, 2025; Verdon *et al.*, 2019b) and explicitly permutation-equivariant circuits via S_n-equivariant constructions or parameter tying (Bradshaw *et al.*, 2025; Liu *et al.*, 2025d; Skolik *et al.*, 2023a). Recent work also explores gate-level geometric constraints such as "horizontal" gates on homogeneous spaces (Wiersema *et al.*, 2025). Symmetry-informed en-

Table I Representative taxonomy of hybrid quantum-classical models. Architectures are classified by the computational role of the quantum circuit and the primary structural constraint that shapes the hypothesis class and associated estimator family. Figure 3 depicts canonical *core motifs* for each paradigm; the table lists representative instantiations and closely related variants. Abbreviations: PQC, parameterized quantum circuit; QCNN, quantum convolutional neural network; QNN, quantum neural network; RL, reinforcement learning; QGAN, quantum generative adversarial network.

Paradigm	Quantum-module role	Inductive bias	Key refs.
Placement and composition (graph location)			
PQC head	terminal readout / classifier	encoding-induced hypothesis restriction; restricted readout	(Farhi and Neven, 2018; Schuld *et al.*, 2020a)
Transfer hybrid	quantum head on frozen classical features	representation bottleneck; low-dimensional embedding	(Chittem and Kumar Pradhan, 2025; Mari *et al.*, 2020)
Interleaved exchange	multiple quantum blocks separated by classical processing	non-sequential composition; measurement-mediated feature exchange across modules	(Guo *et al.*, 2025; Kabir *et al.*, 2025; Liang *et al.*, 2021; Long *et al.*, 2025)
Hierarchical and geometric structure (symmetry + locality priors)			
QCNN	hierarchical pooling	multiscale locality via pooling; parameter tying	(Cong *et al.*, 2019; Herrmann *et al.*, 2022)
Equivariant QNN	equivariant channels / symmetry-preserving layers	commutation with group action; parameter tying / twirling constructions	(Nguyen *et al.*, 2024; Schatzki *et al.*, 2024)
Graph models	graph-structured PQCs / graph learning blocks	permutation-equivariant graph embeddings; graph-structured interactions	(Skolik *et al.*, 2023a; Verdon *et al.*, 2019b)
Interfaces and encoding patterns (encode-measure motifs)			
Data reuploading	repeated embedding; trainable layers	Fourier-feature periodic function class; universal approximation	(Pérez-Salinas *et al.*, 2020; Schuld *et al.*, 2021)
Quanvolution	patchwise feature maps	local receptive fields; fixed random features or trainable filters	(Dou *et al.*, 2023; Henderson *et al.*, 2020a)
Channel-aware kernels	multi-channel circuit kernels	cross-channel mixing in circuit design	(Smaldone *et al.*, 2023)
Objective-driven paradigms (estimator and readout regime)			
Transformer substitution	attention-like mixing block	structured mixing in classical transformer (no asymptotic claim)	(Chen and Kuo, 2025a; Chen *et al.*, 2025)
Generative / unsupervised	generator/encoder modules	sampling-based or energy-based objectives	(Benedetti *et al.*, 2017; Lloyd and Weedbrook, 2018; Shen *et al.*, 2025; Zoufal *et al.*, 2021)
RL hybrids	policy/value module	stochastic policy/value via circuit + readout in RL loop	(Jerbi *et al.*, 2021b; Skolik *et al.*, 2022)

codings and architectures reduce hypothesis complexity and improve generalization and trainability (Bian *et al.*, 2024; Larocca *et al.*, 2022b; Meyer *et al.*, 2023).

(iii) Classical-quantum data interfaces and encoding patterns.—In hybrid visual or spatial models, the classical-quantum interface itself often serves as the primary inductive bias. As shown in Fig. 3(c), quanvolutional architectures achieve this by encoding small, localized input patches into a quantum register, processing them via a local circuit, and measuring the output to produce a spatial feature map for subsequent classical layers (Henderson *et al.*, 2020a). In the original proposal these patch circuits and quanvolution layers are fixed, meaning only the downstream classical postprocessing is trained. However, subsequent variants treat the patch kernel as a fully trainable PQC, effectively learning specialized local quantum filters (Dou *et al.*, 2023). Recent work shows that making quanvolution layers trainable can substantially improve performance while raising nontrivial gradient-access challenges in deeper layers (Kashif and Shafique, 2025).

Figure 3(c) also contrasts temporal data re-uploading with spatial multi-channel mixing. From an architectural perspective, the classical data-access frequency defines the effective depth of the interface. Rather than encoding the classical patch only once, data reuploading repeatedly injects the input data interleaved with trainable unitaries (Pérez-Salinas *et al.*, 2020) (see Sec. III.A.3. Beyond depth, the interface also handles the structural complexity of the input, which is particularly challenging for multi-channel data. Naive encoding strategies face a strict physical bottleneck: processing each channel in

an independent quantum circuit fails to capture critical inter-channel correlations prior to measurement, whereas classically concatenating all channels into a single, massive input vector causes the required qubit count to scale prohibitively. To resolve this, modern quantum convolutional kernels employ coherent pre-measurement mixing. As illustrated in Fig. 3(c), this approach injects distinct channels into separate partitions of the qubit register and applies hardware-adaptable entangling operations to coherently mix the channel information before measurement (Smaldone *et al.*, 2023). This ensures cross-channel correlations are captured at the quantum level without requiring exponential qubit overhead, yielding significant empirical improvements over single-channel QCNN baselines.

(iv) Objective-driven paradigms.—While fundamentally relying on the architectural primitives described above, objective-driven hybrids are organized most cleanly by their readout regime and deployment loop. The choice of task, such as expectation estimation versus sampling, fixes the estimator family and the dominant circuit-evaluation cost. Transformer hybrids, for example, substitute a PQC for primitives in classical pipelines, with parameters updated via classical optimization from finite-shot measurement estimates (Chen and Kuo, 2025a; Chen *et al.*, 2025). Here, the quantum circuit serves as a structured mixing or attention-like block on classically computed token embeddings encoded into quantum states. Across generative and unsupervised objectives, practical cost is typically controlled by the number of circuit evaluations required for readout and gradients rather than by circuit depth alone. Depth-dependent expressivity-trainability trade-offs for implicit deep generative circuits are discussed in Sec. III.C. Quantum generative adversarial networks (QGANs) train PQC generators to match target distributions via samples (Dallaire-Demers and Killoran, 2018; Hu *et al.*, 2019; Leyton-Ortega *et al.*, 2021; Lloyd and Weedbrook, 2018; Zoufal *et al.*, 2019). Expectation-value-sampler models use prescribed observable families for classical data, with tunable observables reducing shot costs (Shen *et al.*, 2025), while self-supervised hybrids employ variational quantum encoders in contrastive pipelines adapted to finite-shot measurement statistics (Jaderberg *et al.*, 2022). One proposal is to train certain quantum generative models classically when the training objective is classically tractable, while using the QPU primarily for deployment-time sampling, thereby shifting quantum cost to regimes where sampling is the bottleneck (Kurkin *et al.*, 2025; Recio-Armengol *et al.*, 2025a); Fig. 3(d) summarizes this training and deployment loop.

Alternative formulations include density and mixed-state QNN models that prepare mixtures of trainable unitaries (density QNNs) to improve expressivity–trainability trade-offs, especially on quantum hardware (Coyle *et al.*, 2025), and Bayesian QDL frameworks aimed at providing principled predictive uncertainty estimates in quantum-enhanced models (Zhao *et al.*, 2019). Diffusion variants use mixed-states, replacing scrambling unitaries with depolarizing channels and learning denoising via parameterized circuits and projective measurements (Kwun *et al.*, 2025). Energy-based approaches model Gibbs distributions with hybrid training (Adachi and Henderson, 2015; Amin *et al.*, 2018; Benedetti *et al.*, 2017; Melko *et al.*, 2019; Shingu *et al.*, 2021; Zoufal *et al.*, 2021), including continuous-variable (Bangar *et al.*, 2025), and sample-efficient (Coopmans and Benedetti, 2024) variants. Related thermodynamic tasks use variational quantum thermalizers for approximate thermal-state preparation (Verdon *et al.*, 2019a), while quantum autoencoders compress quantum states (Romero *et al.*, 2017).

In hybrid quantum RL, PQCs represent stochastic policies or value functions, trained via gradients including policy-gradient methods (Chen *et al.*, 2020a; Jerbi *et al.*, 2021a,b; Lockwood and Si, 2020) and deep Q-learning variants (Skolik *et al.*, 2022), studied numerically on benchmark tasks. Hybrid RL can be employed as a design mechanism for hybrid pipelines under hardware constraints by jointly learning a latent observation representation tailored to the QPU architecture (Nagy *et al.*, 2025). Hybrid tensor network (TN) and variational circuit constructions have also been used to scale to higher-dimensional inputs and larger environments, reporting parameter savings and competitive performance in specific benchmarks (Chen *et al.*, 2022b).

2. Differentiation and gradient access

Gradient-based training estimates gradients $\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})$ from repeated circuit evaluations and measurements. Unlike classical automatic differentiation (Baydin *et al.*, 2018) via single backward passes through deterministic primitives, hardware PQCs expose objective value only through finite-shot measurement statistics, with gradient access mediated by circuit evaluations in which cost and statistical accuracy are set jointly by the differentiation rule, observable count, and shot budget (Mitarai *et al.*, 2018; Schuld *et al.*, 2019; Sweke *et al.*, 2020).

The cornerstone of hardware-native gradient estimation is the parameter-shift rule, rooted in quantum optimal control techniques (Li *et al.*, 2017), and subsequently adapted for PQCs (Mitarai *et al.*, 2018; Schuld *et al.*, 2019). For gate $U(\mu) = e^{-i\mu G_\mu}$ with generator G_μ having two distinct eigenvalues $\pm r$ and $G_\mu^2 = r^2\mathbb{I}$, such as qubit rotations with $G_\mu = \sigma_k/2$ for a Pauli σ_k, the derivative of an expectation value $\langle O\rangle$ with respect to μ can be obtained from two shifted circuit evaluations:

$$\partial_\mu\langle O\rangle = r\left(\langle O\rangle_{\mu+s} - \langle O\rangle_{\mu-s}\right), \qquad s = \frac{\pi}{4r}. \tag{15}$$

Here, $\langle O\rangle_{\mu\pm s}$ denotes the expectation value at shifted

μ, while holding all other parameters fixed. On hardware, each $\langle O\rangle_{\mu\pm s}$ is replaced by a finite-shot estimator, rendering the resulting gradient estimator stochastic even when the underlying identity is exact. Additionally, exact parameter-shift rules for linear-optical photonic gates in Fock space have been derived for Fock-state inputs (Pappalardo *et al.*, 2025). For multi-eigenvalue generators, such as Hamiltonian coefficients, generalized parameter-shift rules express derivatives as linear combination of expectations (Kyriienko and Elfving, 2021; Wierichs *et al.*, 2022). For generators with complex spectra, stochastic parameter-shift rules (Banchi and Crooks, 2021) circumvent the proliferation of shift terms by sampling offsets from a fixed distribution to produce an unbiased gradient estimator.

Beyond access to $\nabla\mathcal{L}$, geometry-aware approaches require access to an information metric on the variational manifold. Quantum natural gradient methods (Dell'Anna *et al.*, 2025; Minervini *et al.*, 2025; Sohail *et al.*, 2025; Stokes *et al.*, 2020; Wierichs *et al.*, 2020) use a quantum information metric to precondition parameter updates for steepest descent directions on the variational manifold. For a parameterized state $|\psi(\boldsymbol{\theta})\rangle$, preconditioned gradients are

$$\tilde{\nabla}\mathcal{L} = g^{-1}\nabla\mathcal{L}, \tag{16}$$

where g is the regularized quantum Fisher information metric tensor estimated from circuit data. Estimating a dense g requires $\mathcal{O}(m^2)$ metric elements for m parameters, motivating structured approximations and stochastic estimators. To bypass explicit metric evaluation, quasi-Newton natural-gradient variants, such as the quantum Broyden adaptive natural gradient (Fitzek *et al.*, 2024), approximate the preconditioner from iterative secant updates. Alternatively, randomized natural-gradient variants estimate a usable preconditioner from randomized measurements with substantially reduced overhead compared to dense estimation (Kolotouros and Wallden, 2024). Beyond ideal unitary circuits, these natural-gradient methods have also been generalized to noisy and nonunitary regimes, natively integrating metric estimation and regularization into the oracle model (Koczor and Benjamin, 2022).

3. Resource-aware codesign principles

Hybrid QDL design is a codesign problem in which architectural capacity needs to be balanced with trainability and measurement cost under hardware constraints. A central limitation is the concentration of measure. In the standard measurement-based setting without the ability to store and reuse quantum states coherently, i.e., without quantum memory, each partial derivative is typically estimated from fresh state preparations via repeated circuit executions and measurements. Achieving backpropagation-like scaling where gradient computation cost scales linearly with depth in fully coherent gradient-propagation schemes proposed to date (Abbas *et al.*, 2023) generally requires access to multiple copies of intermediate quantum states, which is not available in standard measurement-based training. Complementarily, structured circuit families can reduce the circuit-count overhead for gradient estimation (Bowles *et al.*, 2025). Moreover, if the number of trainable parameters is very large, parameter-shift-based gradient evaluation can become measurement-dominated, where shot accumulation time exceeds all other costs, even before accounting for shot noise.

Deep unstructured random ansätze approximate unitary 2-designs (Brandão *et al.*, 2016; Harrow and Low, 2009), inducing the barren-plateau phenomenon where gradient variance is exponentially suppressed with system size for broad global cost functions (Cerezo *et al.*, 2021b; Holmes *et al.*, 2022; Larocca *et al.*, 2025; McClean *et al.*, 2018; Pesah *et al.*, 2021; Wang *et al.*, 2021b). Hence expressivity (Wu *et al.*, 2021c) needs to be balanced against trainability: increasing representational capacity can worsen gradient concentration (Du *et al.*, 2020; Friedrich and Maziero, 2023), leading to tradeoffs between expressivity and gradient-measurement efficiency in deep PQCs (Chinzei *et al.*, 2025). Recent results provide tight upper and lower bounds on loss and gradient concentration for broad classes of parameterized circuits and arbitrary observables, and show that these quantities can be estimated efficiently, furnishing diagnostics and sufficient conditions that can rule out barren plateaus in structured settings (Letcher *et al.*, 2024). These considerations clarify that gradient access needs to be specified jointly with circuit structure, observable choice, and shot budgets when assessing the practicality of gradient-based training.

Task-informed initialization, drawing on problem-specific priors (Liu *et al.*, 2023b), and staged depth growth, incrementally increasing circuit layers (Grant *et al.*, 2019; Skolik *et al.*, 2021) sustain training within locally optimizable regions. Matching entangling patterns to device connectivity and native gates minimizes compilation overhead (Kandala *et al.*, 2017), while tailored strategies further reduce routing costs (Ji *et al.*, 2025a, 2024). Moreover, QEM proves effective only when its sampling overhead yields a net gain in gradient signal-to-noise ratio (Cai *et al.*, 2023; Endo *et al.*, 2021; Kandala *et al.*, 2017; Murali *et al.*, 2019). Codesigning observable selection, grouping, estimation strategies, and shot allocation with the objective and architecture minimizes total circuit executions while satisfying accuracy thresholds (Gresch and Kliesch, 2025; Hadfield *et al.*, 2022; Huang *et al.*, 2020; Huggins *et al.*, 2021; Verteletskyi *et al.*, 2020). Additionally, shot noise provides regularization for optimization dynamics in specific regimes (Liu *et al.*, 2025c). Finally, as capacity variations often produce non-

monotonic generalization and mismatched inductive biases can dominate outcomes, symmetry-aligned encodings and ansätze, along with capacity diagnostics, enable precise calibration of data and shot budgets (Caro *et al.*, 2022; Gil-Fuster *et al.*, 2024a; Meyer *et al.*, 2023).

C. Quantum deep neural networks

Quantum deep neural networks (QDNNs) are parameterized quantum models that scale representation depth by using design variables such as alternating data-encoding and trainable layers or hierarchical circuits. This section focuses on depth-driven theory, including expressivity, scaling limits, and training dynamics, as well as motifs that enable stable training in deep regimes.

Motivated by the success of classical deep neural networks, QDNNs (Abbas *et al.*, 2021; Beer *et al.*, 2020; Jia *et al.*, 2019; Pérez-Salinas *et al.*, 2020) have been studied since the early development of variational quantum algorithms (VQAs) (McClean *et al.*, 2016; Peruzzo *et al.*, 2014). Deep regimes can expand the accessible function classes and frequency spectra via repeated embeddings and layers (Schuld *et al.*, 2021; Shin *et al.*, 2023). Encoding-dependent generalization bounds complement this expressivity perspective (Caro *et al.*, 2021). Theoretical analysis has adapted classical neural network theory to the quantum setting. Recent work shows that certain ensembles of QDNNs converge to Gaussian processes in appropriate large-dimension or large-width limits for specific observable choices (García-Martín *et al.*, 2025; Girardi and De Palma, 2025; Melchor Hernandez *et al.*, 2025). In these regimes, gradient descent can admit a kernel description analogous to the classical NTK framework, motivating quantum NTKs that capture aspects of training dynamics and representation learning when training remains close to initialization (Incudini *et al.*, 2023; Liu *et al.*, 2022a; Nakaji *et al.*, 2023; Tang and Yan, 2022). Related analyses identify a quantum lazy-training regime in which gradient descent remains close to initialization and learning is effectively kernelized (Abedi *et al.*, 2023). These results separate circuit-controlled expressivity from kernel-controlled optimization and provide depth-scaling diagnostics when ensemble and measurement assumptions hold (Anschuetz, 2025; Melchor Hernandez *et al.*, 2025). Additionally, overparameterization can improve optimization and generalization in studied regimes, analogous to classical DL phenomena and counter to worst-case intuitions (Larocca *et al.*, 2023; Peters and Schuld, 2023). These overparameterization benefits have an analogue in quantum kernel methods: related double-descent behavior in quantum kernel regression reinforces that generalization is nonmonotone in model capacity (Kempkes *et al.*, 2025).

QDNNs naturally extend to generative pipelines. Architectural families and benchmarking protocols are surveyed in Sec. III.B.1. Here we focus on depth-dependent expressivity and trainability phenomena in generative settings. Implicit circuit Born machines provide a canonical example in which deeper circuits act as richer implicit generative families learned from measurement samples (Liu and Wang, 2018). Deep-circuit generative priors for inverse problems provide an application-driven instantiation of this depth-as-capacity viewpoint (Xiao *et al.*, 2024). For implicit quantum generative models implemented by deep circuits, depth increases the class of representable distributions, but it can simultaneously exacerbate loss concentration and gradient suppression depending on the chosen divergence and how it is estimated from samples (Du *et al.*, 2020). Rudolph *et al.* (2024) formalize trainability barriers arising in quantum generative modeling from the interplay between sample-based divergence estimation and circuit depth, showing that commonly used divergences can induce new barren-plateau mechanisms, while trainable low-body losses can fail to distinguish higher-order correlations, creating a depth-dependent tension between fidelity and trainability.

To mitigate deep-regime pathologies, the field has increasingly adopted structural inductive biases inspired by classical success stories. Residual and skip connections, which are central to classical deep residual learning (He *et al.*, 2016), have quantum analogues that improve gradient propagation and/or enrich effective representations in numerical and theoretical studies (Kashif and Al-Kuwari, 2024; Wen *et al.*, 2024). Reduced-domain parameter initialization scales the initial parameter range inversely with the square root of depth and yields depth-explicit trainability guarantees in the stated settings (Wang *et al.*, 2024c). Recent work explicitly targets deep quanvolutional stacks and analyzes architectural mechanisms for stabilizing training as depth grows (Kashif and Shafique, 2025). Beyond purely unitary ansätze, dynamic circuits with intermediate measurements and classical feedforward have been proposed as depth-scalable architectures that can avoid barren-plateau mechanisms in the stated models while retaining expressivity (Deshpande *et al.*, 2024). Similarly, permutation-equivariant QNN constructions admit setting-specific guarantees linking symmetry constraints to improved trainability proxies (Schatzki *et al.*, 2024). Beyond feedforward depth, recurrent QNN formulations have been proposed for sequence learning (Bausch, 2020), and recent work explores architectures designed for variable-size inputs (Vargas-Calderón, 2025). A central open problem is to identify scalable, reliably trainable QDNN families that remain beyond known efficient classical simulation under the stated access, noise, and observable model, outside restricted or warm-started regimes (Meyer *et al.*, 2025; Mhiri *et al.*, 2025; Puig *et al.*, 2025).

D. Quantum algorithms for deep neural networks

Quantum algorithms for deep neural networks (DNNs) aim to accelerate the training and inference of their *classical* counterparts under explicit input and data-access models. Distinct from PQC-based approaches that define new trainable hypothesis classes, these algorithms seek to reproduce the input-output mapping of a specific, predefined classical network, such as a CNN (Krizhevsky *et al.*, 2012), ResNet (He *et al.*, 2016), or Transformer (Vaswani *et al.*, 2017), with a theoretical speedup. Many proposals build on QLA primitives, e.g., HHL (Harrow *et al.*, 2009) and QSVT (Gilyén *et al.*, 2019; Martyn *et al.*, 2021) to accelerate linear-algebraic subroutines under FTQC assumptions (Preskill, 1997). End-to-end complexity is conditional on the access-and-readout model $\mathcal{A}$, particularly QRAM availability and its substantial architectural overhead as analyzed in Sec. III.A.2. The resource assessment there applies directly to the algorithms surveyed below.

1. Nonlinearity and the measure-and-reprepare paradigm

Applying an element-wise nonlinearity to a classical vector encoded via amplitude encoding, while remaining within the same amplitude-encoding model, cannot in general be achieved exactly by a deterministic unitary alone since unitaries act linearly, so implementing a layer-wise classical nonlinearity typically requires non-unitary primitives such as measurement, post-selection, or probabilistic subroutines. A foundational work by Kerenidis *et al.* (2019) introduced a quantum algorithm for implementing and training a CNN via a "quantum CNN" construction that explicitly incorporates nonlinearities and pooling. Their approach combines quantum routines for the linear parts of the network with a measurement-mediated mechanism to implement nonlinear activation and pooling steps. To handle the output, they introduced a novel tomography protocol with ℓ_∞-norm guarantees with query complexity $\mathcal{O}(\log N_{\text{out}}/\epsilon^2)$ calls to the state-preparation unitary and associated measurements, where N_{out} is the dimension of the output vector and ϵ is the maximum allowable error per component.

The example above illustrates the *measure-and-reprepare* paradigm. To apply the nonlinearity, a measurement is performed at the end of each layer, the classical result is processed, and a new quantum state is prepared for the next layer. The need for intermediate measurement and repreparation introduces an explicit per-layer overhead, including precision-dependent sampling and tomography costs, which can dominate at large depth L and undermine coherent end-to-end scaling. Related transformer-inference proposals based on QLA similarly assume a pretrained classical model. When a classical output vector is required, they obtain it by measurement or tomography, and for multilayer settings this entails iterating the subroutine over tokens and layers (Guo *et al.*, 2024b).

2. The coherent approach

An alternative to intermediate measurements is to implement the nonlinearity coherently, i.e., without layer-wise readout. A coherent approach approximates the activation via a polynomial and applies it using QSVT. Guo *et al.* (2024a) formalize nonlinear amplitude transformation in an oracle model, constructing block-encodings from state-preparation unitaries to apply polynomial transforms coherently. Subsequent refinements, such as randomized QSVT (Wang *et al.*, 2025c), aim to reduce block-encoding overheads and ancilla complexity, trading them for randomized sampling and different polynomial-degree dependences. Rattew *et al.* (2025) propose a fully coherent multilayer inference construction analyzing multiple quantum data-access regimes for ResNet-like architectures. Their analysis highlights that skip connections can help stabilize the propagation of vector norms across depth, improving the stability of the coherent inference procedure.

While coherent approaches can yield polylogarithmic-in-N dependence under strong input oracles for the encoded vector dimension N, the overall cost is governed by state-preparation complexity and the polynomial degree required to approximate nonlinearities to target error, and by conditioning parameters in the underlying QLA primitives, which typically scales exponentially with depth to maintain constant end-to-end precision. This highlights a central trade-off: end-to-end coherence is achieved at the cost of algorithm- and access-model-dependent depth overhead, which may become prohibitive at large depth under stringent precision guarantees. Consequently, such algorithms are most compelling when the required depth and approximation degree remain moderate. This suggests that a scalable, fully coherent quantum algorithm for general deep networks with polynomial dependence on depth under realistic access models remains an open problem.

3. Component-wise acceleration for modern architectures

Given the limitations of full-network coherent simulation discussed above, particularly the depth-exponential overhead, a more pragmatic direction has been to focus on accelerating individual, computationally expensive components of modern architectures. This modular approach accepts the current limitations while still offering potential speedups for critical subroutines. For instance, the attention mechanism, a central computational bottleneck in Transformers, has been a major tar-

get. As detailed in a recent survey by Zhang and Zhao (2025), approaches here span both PQC-based paradigms and QLA-based paradigms, with complexity–resource trade-offs and data-access assumptions playing a decisive role. Cherrat *et al.* (2024) propose quantum vision-transformer variants in which patch vectors are loaded via amplitude-encoding style data loaders and processed by quantum orthogonal layers, yielding quantum attention mechanisms that (under their access model) can offer asymptotic runtime and parameter-count advantages relative to classical attention scaling in the number and dimension of patches. Gao *et al.* (2023) present a fast quantum algorithm specifically for computing a sparse attention matrix using Grover-search–based techniques, achieving a polynomial speedup for that sparse-attention setting rather than a generic QLA reduction of dense $QK^\top$ and Softmax. Similarly, the paradigm has been extended beyond discriminative models. Xiao *et al.* (2025) adapted the DeepONet architecture, designed for solving partial differential equations, to the quantum domain. They proposed a quantum DeepONet formulation aimed at reducing the evaluation complexity (e.g., in the input dimension). Meanwhile, Wang *et al.* (2025d) developed a quantum algorithm for diffusion models (Sohl-Dickstein *et al.*, 2015), focusing on accelerating the iterative denoising step by using quantum differential equation solvers and QLA to apply the denoising function.

4. Alternative paradigms and challenges

An entirely different path is to bypass the direct simulation of the network's forward pass. Zlokapa *et al.* (2021) give a quantum algorithm for approximately training overparameterized classical networks in a linearized NTK-type regime by reducing the training update to structured linear-system inversion. Any end-to-end speedup over classical gradient-based training additionally requires efficient quantum state preparation and readout for the relevant data distribution. Related NTK or quantum NTK-based constructions have been developed for structured settings, for example, by Tang and Yan (2022) for graph data. In the infinite-width limit where training a DNN is equivalent to a kernel method problem under lazy-training scaling assumptions, this approach elegantly sidesteps the layer-by-layer complexity, though it is restricted to the regime where the NTK theory holds.

The practical relevance of asymptotic speedups hinges on the full access and readout contract. In particular, many QLA and QSVT-based proposals implicitly assume QRAM-style superposition access and/or efficient state preparation, which can dominate end-to-end cost (see Sec. III.A.2). Moreover, QDL faces a moving target problem (Gundlach *et al.*, 2025), i.e., speedup claims that ignore access costs or benchmark only against weak baselines can erode under stronger classical competitors (see Sec. III.G.4). As emphasized in a unifying perspective by Liu *et al.* (2024), provable quantum advantage often requires explicit access models and end-to-end accounting of state preparation, measurement, and classical post-processing, and typically hinges on problem structure that keeps key algorithmic parameters controlled. The immense overheads hidden within the polylogarithmic factors of QLA and QSVT, combined with the stringent demands of fault-tolerance and the monumental challenge of realizing QRAM, mean that practical advantage remains distant. Nevertheless, recent work (Rattew *et al.*, 2025) has begun to provide more realistic cost models by analyzing multilayer performance under weakened data-access assumptions, showing that the speedup can degrade from exponential to polynomial when QRAM-like loading is not available, demonstrating a critical step towards assessing viability.

E. Quantum-inspired classical algorithms

Quantum-inspired classical algorithms leverage quantum information theoretic concepts to design methods executable entirely on conventional hardware. Representative families include (i) dequantization frameworks, which replicate purported quantum speedups when granted the same oracle access and sampling assumptions as the quantum algorithm (Aaronson, 2015; Chia *et al.*, 2022; Kerenidis and Prakash, 2016; Tang, 2019, 2022); (ii) TN architectures, which exploit entanglement structures as inductive biases for compression and simulability (Cichocki *et al.*, 2016; Orús, 2014; Schollwöck, 2011; Stoudenmire and Schwab, 2016); and (iii) quantum-inspired network architectures, which incorporate motifs such as complex-valued representations and unitary layers into classical pipelines (Scardapane *et al.*, 2020; Trabelsi *et al.*, 2017). These approaches constitute an essential baseline family for QDL within the access-model framework of Sec. II.C.

1. Dequantization

Early QML proposals suggested that coherent access to data could accelerate linear-algebraic primitives underlying learning, including matrix inversion (Harrow *et al.*, 2009; Rebentrost *et al.*, 2014) and eigenvalue estimation (Lloyd *et al.*, 2014). Dequantization makes the input model explicit. As detailed in Sec. III.A.2, several exponential speedups rely on QRAM-style data oracles. Classical algorithms with comparable sample and query (SQ) access can achieve similar complexity for several low-rank linear-algebraic primitives and approximate output goals, typically up to polynomial factors in rank, condition number, and inverse precision, shifting the bur-

den of proof to end-to-end, access-model-matched comparisons (Aaronson, 2015; Chia *et al.*, 2022; Mande and Shao, 2025; Tang, 2022).

A canonical example is Tang (2019), who constructed a quantum-inspired classical analogue of the recommendation system of Kerenidis and Prakash (2016). Under length-squared sampling access and a low-rank promise (rank k for an $m \times n$ matrix), the classical runtime becomes poly(k) and polylog(mn) up to additional polynomial dependence on accuracy- and conditioning-related parameters, eliminating the apparent exponential separation under the same oracle assumptions (Chia *et al.*, 2022; Tang, 2019, 2021). Subsequent work generalized this perspective beyond recommendation-style tasks to additional SQ-access linear-algebraic primitives and to dequantizations of certain QSVT-related transformations under structural promises, emphasizing that precision and conditioning can dominate the true complexity (Chia *et al.*, 2022; Gharibian and Le Gall, 2022). These frameworks can also be made robust to approximate SQ access, with complexity acquiring explicit polynomial dependence on the SQ-approximation error parameters (Le Gall, 2025).

Crucially, the principle that strict structural constraints enable efficient classical emulation extends beyond data access models to the QNN architectures themselves. For instance, recent work demonstrates that for certain unitary QCNN architectures and locally structured datasets, the induced decision rule can be efficiently classically emulated given access to local measurement data (Bermejo *et al.*, 2024). More broadly, increasing evidence shows that highly constrained ansätze, such as those restricted to small dynamical Lie algebras (Goh *et al.*, 2025) or built with strong symmetry-constrained constructions (Anschuetz *et al.*, 2023), inherently admit efficient classical simulation or have low effective dimension. Ultimately, just as low-rank data promises allow classical algorithms to match quantum recommendation systems, these rigid architectural constraints often undermine the necessity of a quantum processor.

However, dequantization has principled limits. Using communication-complexity methods, Mande and Shao (2025) derive lower bounds for SQ-access quantum-inspired algorithms on several matrix problems, implying regimes where at least quadratic quantum-classical separations persist under the stated access model and accuracy parameters. Moreover, Cotler *et al.* (2021) show that SQ access can be strictly stronger than receiving copies of the corresponding quantum state, so SQ-based dequantization baselines may be inappropriate for native quantum data settings where the learner's interface is restricted to state preparation and measurement outcomes. In practice, these dequantized analogues can incur large polynomial overheads in conditioning and accuracy parameters, limiting competitiveness outside favorable regimes (Arrazola *et al.*, 2020).

2. Tensor network architectures

Tensor networks provide a concrete bridge between quantum many-body structure and classical DL by representing high-order tensors through networks of low-rank cores (Berezutskii *et al.*, 2025; Cohen *et al.*, 2016; Orús, 2014; Rieser *et al.*, 2023). In QDL, TNs serve two roles: (i) as standalone models, where the network geometry and bond dimensions impose an entanglement-motivated inductive bias on classical data (Stoudenmire and Schwab, 2016), and (ii) as tensorized parameterizations that replace dense weight tensors by factored TN forms to reduce memory and parameter counts while preserving expressive structure (Novikov *et al.*, 2015; Wu *et al.*, 2023). The mechanism is controlled by two coupled quantities: the bond dimensions, which set the representational capacity, and the contraction complexity, which sets the computational cost (Orús, 2014). In quantum physics, TN efficiency is motivated by low-entanglement structure, formalized in area-law behavior for broad families of ground states (Eisert *et al.*, 2010). In ML, an analogous motivation is that many natural signals, such as images and time series, exhibit structured, predominantly local correlations (Ruderman and Bialek, 1994; Simoncelli and Olshausen, 2001), so low-bond TN priors can act as controlled low-rank constraints rather than generic overparameterization. When utilizing role (ii), parameter and compute costs are explicitly governed by a rank and bond hyperparameter (Cichocki *et al.*, 2016), which typically yields parameter efficiency and an implicit regularization effect by limiting the effective correlation capacity that the layer can represent. However, benefits are regime dependent. TN constraints can be helpful when the task admits low-rank structure, but they do not provide uniform improvements in worst-case learnability and generalization guarantees without additional alignment between data structure and network geometry (Jahromi and Orús, 2024; Wu *et al.*, 2025a). Figure 4 summarizes four fundamental TN architectures, which we detail below.

The matrix product state (MPS) (Bengua *et al.*, 2015; Cirac *et al.*, 2021; Perez-Garcia *et al.*, 2007; Schollwöck, 2011; Verstraete *et al.*, 2008, 2004), also known in ML as the tensor train decomposition (Grasedyck *et al.*, 2013; Oseledets, 2011), factorizes an order-N tensor into a 1D chain with bond dimension χ, yielding $O(Nd\chi^2)$ parameters and typical contraction costs scaling polynomially in χ. MPS can serve as tractable generative models, such as Born-machine variants, when the dominant dependencies are effectively one-dimensional (Han *et al.*, 2018; Wall *et al.*, 2021). As a learning model, the 1D geometry biases toward ordered structure, but capturing long-range correlations typically requires the bond dimension χ to grow rapidly with problem size (Harvey *et al.*, 2025; Novikov *et al.*, 2015; Stoudenmire and Schwab, 2016). Recent analysis further shows that ob-

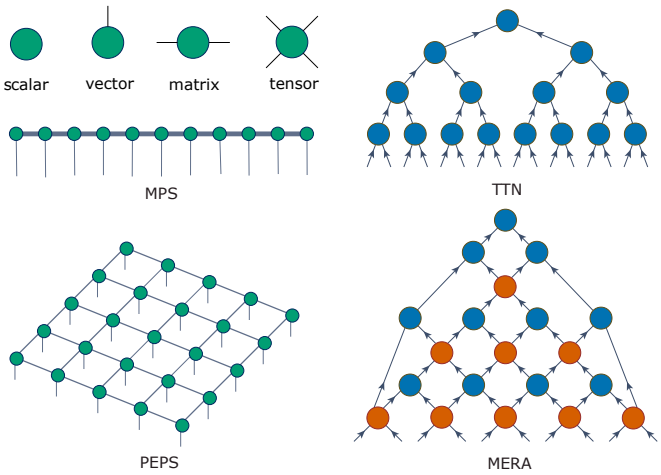

Figure 4 Fundamental tensor network (TN) architectures. Shown are matrix product state (MPS; also known as tensor train), projected entangled pair state (PEPS), tree tensor network (TTN), and the multiscale entanglement renormalization ansatz (MERA). The top row illustrates increasing tensor order (scalar, vector, matrix, tensor) as a visual legend for node types. Adapted from Berezutskii *et al.* (2025).

jectives defined by global observables can exhibit vanishing derivatives, whereas local-observable objectives can remain trainable, and identifies regimes admitting efficient classical evaluation (Basheer *et al.*, 2025). Relatedly, variational schemes for preparing short-range MPS on quantum hardware have been developed to reduce circuit depth (Jaderberg *et al.*, 2026).

The projected entangled pair state (PEPS) (Cirac *et al.*, 2021; Stoudenmire, 2018; Verstraete and Cirac, 2004; Verstraete *et al.*, 2008) generalize MPS to 2D lattices, matching spatial locality and motivating applications to image-like data. PEPS can be viewed as locality-respecting graphical models over 2D neighborhoods, loosely reminiscent of convolutional inductive biases but without the defining CNN mechanisms such as weight sharing and explicit translation equivariance. They have also been investigated for supervised learning and for modeling spatially structured datasets (Cheng *et al.*, 2021). Their central limitation is computational. In the worst case, exact contraction of 2D PEPS is #P-hard (Schuch *et al.*, 2007), so practical use relies on approximate contraction schemes whose cost can dominate training and inference.

Tree tensor networks (TTNs) are a hierarchical generalization of MPS, organizing tensors in a tree-like structure (Shi *et al.*, 2006; Stoudenmire, 2018). This multiscale architecture has been used successfully as a hierarchical classifier or a feature extractor analogous to pooling layers in a classical CNN (Huggins *et al.*, 2019). TTNs are efficiently contractible and offer a compromise between the limited 1D expressivity of MPS and the costly contraction of PEPS. They can capture nonlocal correlations along the tree's branches, imposing an inductive bias well-suited for data with a natural compositional or hierarchical structure. The main limitation of TTNs is their rigidity: the imposed tree structure may not align with the translational or rotational symmetries of many natural datasets.

The multiscale entanglement renormalization ansatz (MERA) (Evenbly and Vidal, 2009; Vidal, 2007) augments hierarchical structure with explicit disentanglers and isometries to model correlations across scales, originally targeting critical states with logarithmic corrections to area laws. As a classical architecture it can encode scale structure but incurs higher contraction costs. As a quantum inspiration, its connectivity patterns motivate QCNN-style ansätze and associated trainability analyses (Cong *et al.*, 2019; Pesah *et al.*, 2021).

Selecting a TN architecture requires balancing the data's correlation structure with the TN's inherent geometry. The bond dimension χ is a primary hyperparameter governing expressivity for a fixed TN parameterization, with larger values capturing more complex correlations at the cost of elevated training and inference demands for the classical model. Training and model selection remain practical bottlenecks. TN learning can be performed via sweep-style local updates or via gradient-based optimization through differentiable contractions. Both are sensitive to conditioning, gauge choices (redundant parameterizations of the same TN), and approximation error, so reliable performance typically requires careful regularization and topology and rank selection (Evenbly, 2018; Liao *et al.*, 2019; Stoudenmire and Schwab, 2016).

Constructive mappings relate several neural-network families to TN representations, providing an explicit classical lens on correlation capacity via entanglement-inspired measures (Chen *et al.*, 2018; Glasser *et al.*, 2018; Levine *et al.*, 2019). For RBMs, the induced TN structure is controlled by the hidden-visible interface. Local (short-range) RBMs satisfy area-law entanglement bounds, whereas dense RBMs can realize substantially stronger scaling (Chen *et al.*, 2018; Glasser *et al.*, 2018). These results have a direct benchmarking implication for QDL on classical data. Arguments based on entanglement-like expressivity need to be tested against classical models whose induced TN capacity matches the same correlation regime. These correspondences are primarily representational, not algorithmic. A TN representation does not, in general, imply efficient contraction or efficient inference. Depth-separation results further delineate regimes where low-bond TN surrogates are inadequate, showing that depth and nonlocal connectivity can compactly represent correlation families, including families generated by polynomial-size quantum circuits, that shallow constructions cannot (Deng *et al.*, 2017; Gao and Duan, 2017).

3. Quantum-inspired network architectures

Beyond TN-based compression, which exploits quantum entanglement structure, a distinct family of fully classical baselines imports mathematical motifs from quantum theory into neural architectures without invoking coherent quantum resources. They test whether purported "quantum" gains on classical data are better explained by classical inductive biases that can be implemented and optimized on conventional hardware.

Complex-valued and hypercomplex parameterizations introduce phase-like degrees of freedom and structured coupling between channels. Complex linear layers can be represented as real block-structured maps on doubled channels, but the resulting constraints and nonlinearities can still change optimization and inductive bias relative to generic real-valued networks (Bassey *et al.*, 2021; Trabelsi *et al.*, 2017). Quaternion and hypercomplex networks extend this idea by tying groups of channels through algebraic products, often improving parameter efficiency and sometimes empirical performance in settings with correlated feature structure (Gaudet and Maida, 2018; Parcollet *et al.*, 2020, 2018). Representative instantiations include quantum-measurement motivated complex-valued models for conversational emotion recognition (Li *et al.*, 2021a), entanglement-inspired embedding constructions for matching (Zhang *et al.*, 2025a), and superposition-inspired spiking networks (Sun *et al.*, 2021).

Gate-inspired qubit neuron models update normalized amplitude-like hidden states via rotation-style transformations and produce probabilities through squared-amplitude readout (Ganjefar and Tofighi, 2018; Li *et al.*, 2013; Mori *et al.*, 2006). Subsequent variants extend the state space beyond qubits but remain best interpreted as constrained classical sequence models akin to modern structured parameterizations (Bakshi and Srinivasan, 2025). Quantum probability and density-matrix formalisms represent uncertainty via positive semidefinite operators and scoring hypotheses using quantum-inspired divergences (Leporini and Pastorello, 2022; Tiwari and Melucci, 2019; Yan *et al.*, 2021). These architectures sharpen QDL evaluation by enlarging the classical comparator set to include quantum-structured priors that require no quantum hardware.

F. Optimization as prerequisite in QDL

Training QDL models requires minimizing a loss function $\mathcal{L}(\boldsymbol{\theta})$ derived from quantum observables (Beer *et al.*, 2020; Liang *et al.*, 2021; McClean *et al.*, 2016; Peruzzo *et al.*, 2014; Romero and Aspuru-Guzik, 2021). Optimization proceeds via oracle access to shot-limited estimates of $\mathcal{L}$ and, when available, $\nabla\mathcal{L}$. On quantum hardware, device noise biases the objective, while finite-shot estimation induces stochasticity and drift induces nonstationarity, making optimizer choice a resource-allocation problem over evaluations, shots, classical compute, and wall-clock latency (Kungurtsev *et al.*, 2024; Scriva *et al.*, 2024). Optimizer comparisons require matched search and tuning budgets with reported evaluation and shot costs (Benítez-Buenache and Portell-Montserrat, 2025; Herbst *et al.*, 2024; Moussa *et al.*, 2024, 2022; Ostaszewski *et al.*, 2021).

1. Local optimizers

Local optimizers update parameters using information from a local neighborhood of the current iterate, e.g., gradients, local sampling, or a search distribution centered near current parameters, and thus can converge to nonglobal stationary points in high-dimensional, nonconvex landscapes. In QDL, the choice of local optimizer is governed less by mathematical differentiability than by the measurement overhead and stochastic variance inherent in the hybrid cost model. Stochastic loss and gradient estimates from the QPU feed a classical training stack for parameter updates and job submission (Benedetti *et al.*, 2019; Broughton *et al.*, 2021).

When gradient estimators are accessible, first-order methods are the natural choice. Gradient descent (Cauchy *et al.*, 1847; Curry, 1944) and stochastic variants such as Adam (Kingma and Ba, 2014) serve as standard baselines. However, their practical convergence behavior depends strongly on shot-noise variance and batching strategies (Scriva *et al.*, 2024). In low-noise regimes, quasi-Newton methods which approximate second-order curvature information, such as limited-memory BFGS (Liu and Nocedal, 1989), can provide strong local refinement but degrade rapidly under stochasticity. To stabilize gradient-based training, practitioners often employ symmetry constraints (Bian *et al.*, 2024), initialization heuristics (Mhiri *et al.*, 2025; Puig *et al.*, 2025), or operational strategies like classical surrogate pretraining followed by quantum fine-tuning (Dborin *et al.*, 2022). While these strategies successfully localize the search to trainable subspaces, they do not eliminate the underlying global nonconvexity.

Conversely, when gradients are prohibitively costly or unreliable, derivative-free search dominates. Lightweight local baselines, such as Nelder-Mead (Nelder and Mead, 1965) and pattern search (Hooke and Jeeves, 1961), remain viable despite their poor scalability with dimension. For broader exploration, population-based algorithms, such as particle swarm optimization (Bonyadi and Michalewicz, 2017; Kennedy and Eberhart, 1995), differential evolution (Storn and Price, 1997), and covariance matrix adaption evolution strategy (Hansen, 2006), propose candidates from population statistics and best-so-far solutions. Their global behavior re-

quires restarts or hybridization (Auger and Hansen, 2005; Muller *et al.*, 2009). In practice, derivative-free and surrogate-assisted routines are also used for outer-loop ansatz selection and tuning under shot-limited evaluations (Nguyen and Chen, 2022; Sagingalieva *et al.*, 2023b). Controlled benchmarks indicate that optimizer rankings are instance- and noise-regime dependent and sensitive to hyperparameter tuning (Bonet-Monroig *et al.*, 2023). Specialized applications further emphasize this nuance; for example, chemistry-motivated VQE variants benefit significantly from structure-aware routines that accelerate excitation-operator selection and parameter optimization (Jäger *et al.*, 2025).

2. Global optimizers

Global-exploration heuristics aim to reduce sensitivity to initialization by broader search, but for the nonconvex objectives they provide no general, finite-time global-optimality guarantees (Horst and Tuy, 2013; Stephens and Baritompa, 1998). Random search (Masri *et al.*, 1980; Munson and Rubin, 1959) provides a minimal-overhead baseline. A canonical model-based alternative is Bayesian optimization (Frazier, 2018; Garnett, 2023; Kushner, 1964; Shahriari *et al.*, 2016), which uses a surrogate, often a Gaussian process, and an acquisition rule such as expected improvement to balance exploration and exploitation (Anton *et al.*, 2024; Huang *et al.*, 2006; Jones *et al.*, 1998; Keane, 2006; Knowles, 2006; Močkus, 2005; Mockus and Mockus, 1991; Sóbester *et al.*, 2004). Bayesian optimization is attractive when evaluations are scarce and expensive, as in shot-limited circuit execution (Ji *et al.*, 2025b), but degrades in high dimensions unless strong structure is exploited. A Bayesian-related method is the data-based online nonlinear extremum-seeker (Verstraete *et al.*, 2015), which builds a surrogate model but exhibits more favorable scaling with dimensionality and iteration count than standard Bayesian optimizers. Finally, deterministic partitioning strategies such as multilevel coordinate search (Huyer and Neumaier, 1999) provide a complementary route to global exploration by iteratively subdividing and sampling promising regions.

3. Effects of noise and drift

Noise in QDL optimization arises from distinct mechanisms that should be separated operationally (Herbst, 2023). First, finite-shot measurement induces stochastic estimators of the implemented objective and can set the dominant precision cost (Scriva *et al.*, 2024). Second, physical device noise biases the implemented objective relative to the ideal target, so guarantees become oracle-assumption dependent (Kungurtsev *et al.*, 2024). Finally, hardware calibration drift introduces nonstationarity that complicates stopping and benchmarking under fixed budgets (Zhang *et al.*, 2023a). Accordingly, noise-aware optimization emphasizes variance control and explicit accounting of evaluation cost. Simple replication and batching with robust aggregation provides a simple baseline, while more advanced shot-adaptive dynamically couple the optimizer's step selection directly to the measurement precision (Kübler *et al.*, 2020; Sweke *et al.*, 2020).

A prominent stochastic approach is simultaneous perturbation stochastic approximation, which approximates gradients via two objective evaluations per iteration, independent of dimensionality in the number of objective calls per iteration, although estimator variance and required iteration counts can still grow with dimension (Spall, 1992, 1998; Wiedmann *et al.*, 2023). For surrogate-based globalization, noisy or constrained Bayesian optimization formulations can improve sample efficiency when evaluations are scarce (Letham *et al.*, 2017; Oliveira *et al.*, 2019). Alternatively, population-based heuristics can be adapted to navigate stochastic fitness landscapes, though their ultimate convergence performance remains highly sensitive to both the specific problem instance and the underlying parameter dimensionality (Arnold and Beyer, 2000, 2004; Beyer and Sendhoff, 2006, 2007; Levitan and Kauffman, 1995; Rakshit *et al.*, 2017; Richter *et al.*, 2020).

G. Theoretical advantages and fundamental limitations

Two fundamental questions determine the theoretical standing of any QDL proposal: what can be learned under a specified access-and-readout model, and what end-to-end resources are required to reach a target accuracy. Because both depend strictly on the framework $(\mathcal{D}, \mathcal{R}, \mathcal{A})$ established in Sec. II.C, an advantage claim is meaningful only when these parameters are fixed and the classical comparator class is explicitly stated. To operationalize this, we introduce a resource taxonomy separating consumable complexity axes from structural hypothesis-class constraints. Using this taxonomy, we analyze statistical learnability and complexity-theoretic separations, delimit provable quantum advantage against classical dequantization boundaries, and characterize trainability as a scaling constraint that directly couples expressivity to classical simulability.

1. Taxonomy of operational quantum advantage

Not all quantum advantages are alike: a reduction in sample complexity is a fundamentally different claim from a runtime speedup. Rigorous advantage statements operationalize Definition 3 by first fixing the physical access-and-readout model $\mathcal{A}$, which is the logical prior

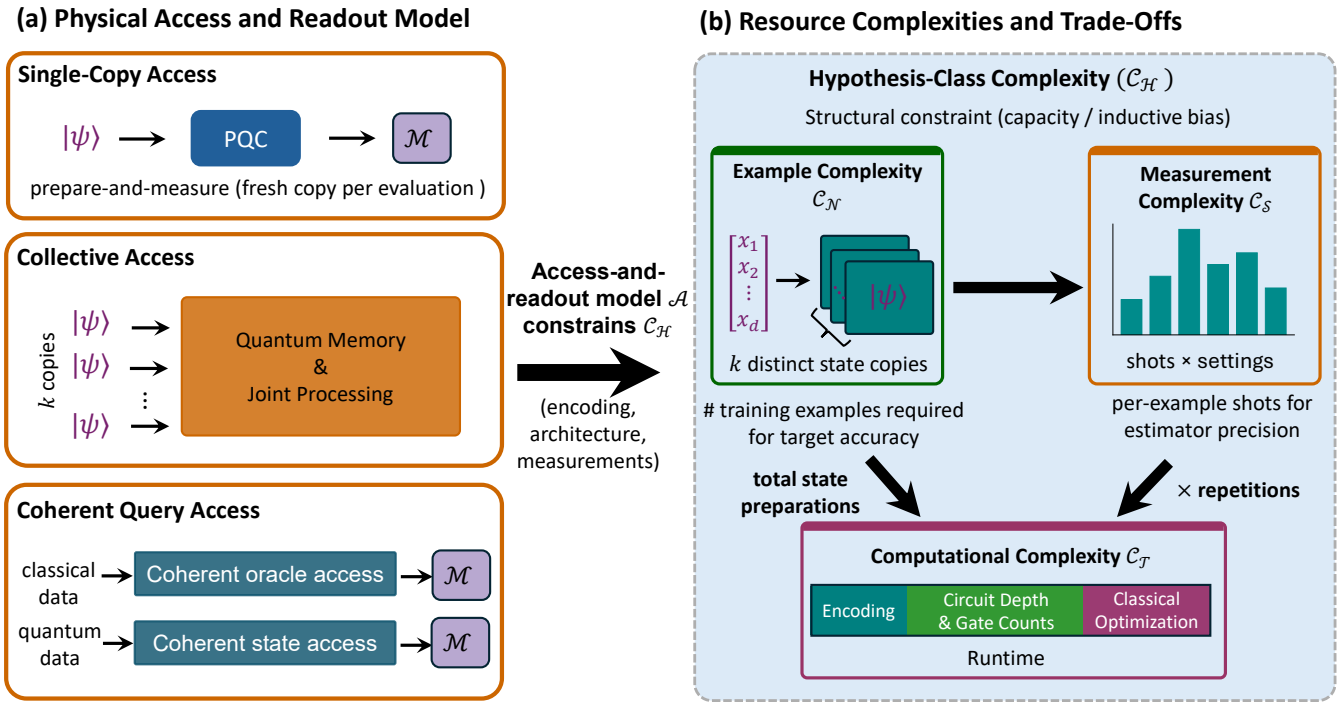

Figure 5 Resource complexities under an explicit access model $\mathcal{A}$. (a) three physical access regimes, including single-copy prepare-and-measure; collective access, where k copies of $|\psi\rangle$ per query are processed jointly via quantum memory (k denotes the number of copies jointly processed per example instance); and coherent oracle or QRAM-style state access. (b) $\mathcal{A}$ constrains hypothesis-class complexity $\mathcal{C}_\mathcal{H}$, which in turn mediates trade-offs among three consumable resource axes: example complexity $\mathcal{C}_\mathcal{N}$, measurement complexity $\mathcal{C}_\mathcal{S}$ (total circuit repetitions, i.e., shots × settings, typically incurred per training example for a fixed estimator precision), and computational complexity $\mathcal{C}_\mathcal{T}$. The number of total state preparations per objective evaluation is $\mathcal{C}_\mathcal{N} \times \mathcal{C}_\mathcal{S}$ and scales as $k\ \mathcal{C}_\mathcal{N} \times \mathcal{C}_\mathcal{S}$ under collective access.

of any resource comparison, and then separating consumable resource axes from structural constraints (Arunachalam and de Wolf, 2017; Chang and Cerezo, 2025; Huang *et al.*, 2021b). Figure 5 makes this explicit: the left panel catalogues three physically distinct access regimes, while the right panel maps the resulting consumable resources against the hypothesis-class constraint $\mathcal{C}_\mathcal{H}$ under $\mathcal{A}$. The access model must be stated before any resource is compared. We distinguish (i) single-copy access (prepare-and-measure), where each evaluation consumes a fresh state copy; (ii) collective access, which leverages quantum memory and joint processing to perform entangling measurements across k simultaneous copies; and (iii) coherent query access, which encompasses both coherent oracle access for classical data (e.g., via QRAM) and coherent state access for quantum data, methodologies that typically presuppose fault-tolerant capabilities. In addition, the readout model is part of the hypothesis class definition; changing what measurements are permitted can change the effective function class represented by the model.

Given a fixed $\mathcal{A}$, we distinguish three consumable resource complexities and one structural constraint:

(1) *Example complexity* $\mathcal{C}_\mathcal{N}$: the number of training examples or, for quantum data, distinct state copies, required to achieve prediction error at most ϵ with high probability under the stated access model.
(2) *Measurement complexity* $\mathcal{C}_\mathcal{S}$: the number of circuit repetitions required by the chosen estimator to evaluate losses, gradients, or observables to the precision needed for learning; operationally, it scales with "shots × settings" and is incurred per example.
(3) *Computational complexity* $\mathcal{C}_\mathcal{T}$: the end-to-end runtime resources for training and inference, including state preparation, circuit depth and gate counts, classical post-processing, and optimizer steps.
(4) *Hypothesis-class complexity* $\mathcal{C}_H$: a structural property of the induced hypothesis class determined by the encoding, architecture, and readout, which constrains the attainable trade-offs among $\mathcal{C}_N$, $\mathcal{C}_S$, and $\mathcal{C}_T$.

As indicated by the central arrow in Fig. 5, the access model $\mathcal{A}$ constrains $\mathcal{C}_H$ through what encodings, circuit families, and measurement procedures are physically permitted. Operationally, $\mathcal{C}_H$ captures capacity and inductive-bias constraints, and it can be instantiated via standard capacity proxies such as covering numbers

(Caro *et al.*, 2022), which bound the achievable trade-offs among $\mathcal{C}_N$, $\mathcal{C}_S$, and $\mathcal{C}_T$ under the stated access model (Abbas *et al.*, 2021; Vapnik, 1998). Throughout, $\mathcal{C}_H$ is not a consumable cost but a constraint on what error levels and trade-offs are attainable under the chosen model family; local geometric quantities used in trainability analyses are discussed below in Sec. III.G.5. These resource complexities realize the abstract resource contract $\mathcal{R}$ introduced in the computational symbiosis framework (Sec. II.C), with $\mathcal{C}_T$ and $\mathcal{C}_S$ as primary consumable resources and $\mathcal{C}_H$ constraining achievable trade-offs under the chosen access-and-readout model $\mathcal{A}$.

These quantities are not independent. Figure 5(b) makes explicit two common couplings: $\mathcal{C}_N$ contributes to the total number of state preparations required across training, while $\mathcal{C}_S$ determines the number of repetitions (shots, and possibly multiple settings) needed per evaluated quantity to control estimator precision. Consequently, in variational workflows, the end-to-end runtime $\mathcal{C}_T$ typically contains a term proportional to the number of state preparations multiplied by per-evaluation repetitions, plus classical overhead and the number of optimization steps. Conversely, reductions in $\mathcal{C}_S$ can require stronger measurements or additional quantum processing in the readout, thereby altering the admissible measurement map and the induced hypothesis class $\mathcal{C}_H$. Accordingly, improvements in $\mathcal{C}_T$ do not imply improvements in $\mathcal{C}_N$ or $\mathcal{C}_S$, and apparent gains along one axis can be offset by increases along another.

2. Statistical learnability

Statistical learnability formally quantifies the sample complexity $\mathcal{C}_N$ and measurement complexity $\mathcal{C}_S$ required to achieve generalization under $\mathcal{A}$, assuming the learning algorithm can successfully optimize the stated objective. Generalization is the foundational objective of any learning protocol: it provides the mathematical guarantee that a model achieving low empirical error on finite training data will also achieve a predictably small expected error on unseen data drawn from the same underlying distribution $\mathcal{D}$. This cleanly isolates a model's information-theoretic capacity from its *trainability*, which represents a fundamentally distinct bottleneck analyzed separately in Secs. III.F and III.G.5.

Generalization guarantees can be formalized in the probably approximately correct (PAC) framework through bounds on the sample complexity needed to achieve low error with high probability. Building on classical foundations (Kearns *et al.*, 1994), Arunachalam and De Wolf (2018) show that in broad distribution-independent settings the quantum and classical sample-complexity scalings match up to constant factors (see also (Arunachalam and de Wolf, 2017)), with details depending on the SQ-access model. Tight characterizations in PAC and agnostic models reinforce that coherent access to training examples in the stated learning model does not generically yield asymptotic improvements in the distribution-independent regime where the learner must succeed for all data distributions (Arunachalam and De Wolf, 2018). A related caution is that simulation hardness is not equivalent to learning hardness: depending on the data-generating and label model, a classical learner may achieve low prediction error via surrogate hypothesis classes or concentration structure even when exact simulation of the underlying quantum model is intractable (Huang *et al.*, 2021a; Schreiber *et al.*, 2023). Moreover, in the stated learning model and measurement interface, polynomial-time learning procedures given sample access to quantum outputs and an appropriate measurement interface exist for certain shallow circuit families (Huang *et al.*, 2024a). The central implication is that being quantum does not generically reduce data requirements without additional structure or stronger access assumptions. Nevertheless, separations can arise in distribution-specific regimes and under computational assumptions relative to restricted classical baselines. Sweke *et al.* (2021) construct distribution-learning tasks exhibiting a quantum-classical separation under cryptographic assumptions within the specified learning model, and Lewis *et al.* (2025) provide separations under explicit structural assumptions and restricted baselines. Beyond cryptographic constructions, Molteni *et al.* (2026) prove exponential quantum-classical separations for learning partially unknown observables from classical (measured-out) data in the PAC framework, and they delineate regimes in which the same observable-learning setting becomes efficiently classically learnable.

Generalization theory for QNNs remains incomplete in highly overparameterized regimes. Gil-Fuster *et al.* (2024a) show that QNNs can fit random quantum states and random labels, and they provide constructions in which QNNs can fit arbitrary labels to quantum states, behavior that is often poorly predicted under standard uniform convergence analyses used in current bounds. In parallel, recent theory has developed approximation and generalization-error bounds for PQCs and QNNs in structured settings (Manzano *et al.*, 2025; Ohno, 2025), bounds showing improved sample efficiency when the target lies in a low-complexity subspace of the induced feature space (Caro *et al.*, 2022), optimizer-dependent generalization bounds (Zhu *et al.*, 2025b), and tight analyses clarifying which parameter dependencies can and cannot appear in standard bounds (Wang and Wu, 2025). Moreover, geometric and information-theoretic perspectives relate generalization to the quantum Fisher information metric (Haug and Kim, 2024), margin-like notions (Hur and Park, 2024), and Rényi-information-based quantities (Banchi *et al.*, 2021).

For learning tasks targeting properties of quantum states or processes, the access model becomes decisive.

Huang *et al.* (2022, 2021b) show that exponential separations are possible between protocols that permit quantum memory and entangling joint measurements across multiple copies and protocols restricted to single-copy measurements for tasks such as predicting large families of observables and related property-learning problems. Oh *et al.* (2024) extend entanglement-enabled learning separations to bosonic continuous-variable channel-learning settings. These results establish an information-theoretic hierarchy within their stated models, but they rely on multi-copy coherent processing, which is substantially more demanding than single-copy prepare-and-measure access. Finally, for QDL workflows requiring classical outputs, measurement overhead is a central bottleneck, as detailed in Sec. II.A.4. Allowing additional quantum processing in the readout can enlarge the effective hypothesis class relative to strictly local readouts (Panadero *et al.*, 2024), while single-shot limitations can make measurement overhead information-theoretically unavoidable under the stated embedding and measurement model (Recio-Armengol *et al.*, 2025b). These considerations motivate the time-to-solution analysis for variational workflows under finite shots and noise.

3. Computational complexity

Computational complexity $\mathcal{C}_T$ concerns the runtime axis: the end-to-end resources required to train and deploy a specified model to target accuracy under an explicit access and readout model. Unlike statistical learnability, which is governed by information constraints on examples and measurements, computational separations ask whether quantum resources permit asymptotically faster learning or inference than any classical procedure restricted to the same access and readout interface. For generative workloads, this runtime axis naturally splits into learning, which estimates parameters under a specified loss function, and inference, which samples from the trained model. Recent work constructs families of generative quantum models that are argued to be efficiently trainable yet for which the inference task is classically hard to simulate under standard sampling-hardness assumptions, and reports beyond-classical demonstrations in this regime (Huang *et al.*, 2025a). Complementary limitations identify trade-offs between classical sampling hardness, often linked to anticoncentration, and average-case trainability for common estimator choices (Herbst *et al.*, 2025).

Unconditional separations are known in restricted circuit classes. Bravyi *et al.* (2018) establish separations between constant-depth quantum and constant-depth classical circuits for a relational task, demonstrating that shallow quantum computation can exceed classical shallow computation without unproven complexity assumptions. In a learning-theoretic direction, Pirnay *et al.* (2024) prove an unconditional separation in a PAC distribution-learning setting between constant-depth quantum and classical hypothesis classes, while Pirnay *et al.* (2023) obtain a superpolynomial quantum-classical separation for density modeling under standard cryptographic assumptions in a fault-tolerant regime.

For supervised learning, stronger separations typically rely on cryptographic or complexity assumptions. Liu *et al.* (2021b) construct classification problems tied to discrete-log-type hardness in which efficient quantum learning is possible while efficient classical learning is ruled out for specified baseline classes under the stated assumptions.

A complementary line isolates computational barriers and capabilities via oracle models that abstract broad families of learning algorithms. In the quantum statistical query (QSQ) model (Arunachalam *et al.*, 2020), the learner accesses expectation values of specified observables up to a tolerance, capturing "learning from statistics" in a quantum access setting. Du *et al.* (2021) relate noisy variational QNN training to the QSQ framework, motivating QSQ as a resource-aware lens for near-term learnability claims. Subsequent work sharpened the role of entanglement and statistics in learning hierarchies (Arunachalam *et al.*, 2023) and extended QSQ beyond predictors to learning quantum processes (Wadhwa and Doosti, 2025). However, SQ and QSQ lower bounds yield barriers that are robust to algorithmic details: even learning output distributions can be information-theoretically feasible yet computationally inaccessible to broad SQ-type learners. In particular, Hinsche *et al.* (2023) show hardness of distribution learning for circuit output distributions under minimal non-Clifford resources, and Nietner *et al.* (2025) prove average-case SQ hardness for learning output distributions of random local circuits across depth regimes.

A distinct route to runtime improvements targets the fully fault-tolerant regime, where QLA and differential-equation primitives can yield speedups for structured training dynamics under stringent input access, conditioning, and precision assumptions (Liu *et al.*, 2024). Closely related proposals for generative workloads potentially hinge on strong structure together with explicit assumptions about data loading and readout (Wang *et al.*, 2025d). Between NISQ heuristics and fully fault-tolerant algorithms lies an early fault-tolerant regime in which limited QEC extends feasible circuit depth by trading reduced effective noise against overhead (Dangwal *et al.*, 2025).

4. The classical boundary

Any claimed quantum advantage should be evaluated against the strongest classical algorithms available under the same access and readout interface. Tensor-network,

dequantization, and quantum-inspired baselines are reviewed in Sec. III.E; here we summarize their role as comparator families. For QDL on classical data, the boundary is often defined by classical surrogates that approximate the induced hypothesis class, classical simulation methods that approximate the quantum subroutine, and classical learning procedures that recover predictors without simulating the underlying quantum dynamics. Representative surrogate families include TN reductions, which can efficiently approximate variational circuits when the relevant entanglement complexity remains controlled (Shin *et al.*, 2024), and random-feature reductions, which can approximate families of quantum kernels or implicit feature maps when the induced feature structure admits efficient classical approximation (Sahebi *et al.*, 2025; Sweke *et al.*, 2025). The boundary is also set dynamically by advances in hardware; exascale-class state-vector simulations now reach 50 qubits (De Raedt *et al.*, 2025), reinforcing the need to evaluate the classical comparator against contemporary capabilities rather than historical baselines.

Entanglement alone is not a reliable proxy for computational hardness, and many state-of-the-art simulators are controlled by nonstabilizerness (the resource enabling universal quantum computation beyond Clifford gates), often called "magic". In this context, Leone *et al.* (2022) analyze stabilizer Rényi entropy (SRE) as a computable measure of distance from stabilizer structure to quantify non-Clifford resources, formalizing regimes in which near-stabilizer simulation can be effective even when entanglement is not small. Complementary approaches exploit operator structure, and in some settings noise structure, to accelerate simulation in restricted regimes. Pauli-propagation frameworks provide a representative example (González-García *et al.*, 2025; Rall *et al.*, 2019; Rudolph *et al.*, 2025).

Classical surrogates can also arise at the level of hypothesis classes. Sweke *et al.* (2025) characterize conditions under which random Fourier features yield an efficient classical surrogate for PQC-based regression models, in terms of the task, circuit-induced feature structure, and the approximation parameters. In such regimes, the PQC model admits an implicit feature-map description that can be approximated efficiently by classical random-feature methods. Accordingly, surpassing strong classical kernel baselines on classical data requires circuit-induced features that evade efficient random-feature approximations under the same access and readout interface. Quantum kernel advantages may persist in structured settings where the feature map encodes task symmetries that are difficult to capture for the chosen classical kernel class under the same interface (Wang *et al.*, 2021c), but recent work on neural quantum kernels (Rodriguez-Grasa *et al.*, 2025) and on the expressivity of embedding quantum kernels (Gil-Fuster *et al.*, 2024b) emphasizes that any such advantage requires balancing expressivity against inductive bias for generalization under finite data. Shadow-based protocols (Jerbi *et al.*, 2024) can further blur the boundary by enabling classical postprocessing and, in some settings, classical evaluation of learned predictors for many observables from a fixed set of quantum measurements after the quantum data-acquisition stage. As an illustrative boundary case study, topological data analysis admits explicit classical benchmarks and dequantizations that materially narrow the parameter window in which a superpolynomial advantage could plausibly persist under stated assumptions (Berry *et al.*, 2024; Gyurik *et al.*, 2022). Collectively, these results underscore that advantage claims depend on explicit access models and have to survive both simulation baselines and learning-based classical surrogates.

5. Trainability limits

This section focuses on whether a variational or hybrid workflow can achieve a target accuracy within a resource contract $\mathcal{R}$, accounting for finite-shot estimation, device noise, drift, compilation, and classical control overhead. Accordingly, advantage claims for trainable QDL modules should be stated in end-to-end terms that account explicitly for the shot budget $\mathcal{C}_S$, per-circuit latency, classical post-processing, and the number of optimization steps, rather than circuit depth or gate counts alone (Liu *et al.*, 2023a).

A central obstruction is unfavorable optimization geometry coupled with measurement-limited updates. Building upon the gradient variance suppression mechanisms defined in Sec. III.B.3, refined analyses underscore that blanket statements about barren plateaus are misleading without explicit qualifiers regarding cost locality, initialization, and architecture (Holmes *et al.*, 2022; Thanasilp *et al.*, 2023; Zhao and Gao, 2021). Vanishing gradients are not the sole failure mode: trap-dominated landscapes can obstruct optimization even when gradients do not concentrate in a barren-plateau sense (Anschuetz and Kiani, 2022; Nemkov *et al.*, 2025). Recent theory shows that wide-QNN losses and their derivatives admit Wishart-process limits (a random matrix theory framework for analyzing correlated Gaussian observables) under broad conditions, yielding architecture-dependent characterizations of gradient statistics and local minima beyond the original concentration picture (Anschuetz, 2025). Moreover, as formulated in Sec. III.B.3, gradient estimation is fundamentally measurement-limited, imposing an explicit statistical trade-off between ansatz expressibility and finite-shot measurement overhead.

Noise further limits variational performance. Under local noise models, limitations constrain the regimes in which variational algorithms can outperform efficient classical approximations as depth increases and noise ac-

cumulates (De Palma *et al.*, 2023). Any mitigation or robustness strategy is therefore assessed under $\mathcal{R}$, since reduced effective noise can shift costs to additional circuit repetitions and classical post-processing. A distinct and fundamental constraint is the inability to reuse unknown quantum information: limits on quantum backpropagation clarify why classical state-reuse assumptions do not carry over, reinforcing the need to include repeated state preparation and measurement in the end-to-end training cost (Abbas *et al.*, 2023).

Mitigating these limitations requires navigating the trainability-simulability tension discussed in Sec. II.C, aiming to reduce measurement demand and stabilize optimization while preserving a hypothesis class that is not trivially classically emulable under the stated $(\mathcal{R}, \mathcal{A})$. Structural routes achieve this by imposing inductive bias through symmetry or locality-respecting parameterizations (Schatzki *et al.*, 2024) and by using controllability- and overparameterization-based diagnostics to identify regimes with more favorable loss geometry (Larocca *et al.*, 2023). Algorithmic approaches instead redistribute computational load: train-classical-deploy-quantum paradigms shift optimization to classically tractable surrogates while reserving quantum hardware for inference under explicit hardness assumptions (Recio-Armengol *et al.*, 2025a). Alternatively, density-learning constructions (Coyle *et al.*, 2025) and complexity curricula (Recio-Armengol *et al.*, 2024) modify the model family itself to improve convergence.

IV. IMPLEMENTATIONS AND CHALLENGES

This section surveys enabling hardware platforms and software frameworks, reviews representative demonstrations, and analyzes dominant near-term constraints and failure modes.

A. The QDL implementation stack

The practical realization of QDL relies on continuous feedback between algorithm design and hardware constraints. Figure 6 organizes this implementation hierarchy into five distinct layers. A top-down perspective (L5 → L1) illustrates how high-level application demands drive lower-level specifications, while a bottom-up perspective (L1 → L5) traces the upward propagation of physical hardware constraints.

L1 (physical quantum resources) encompasses the qubit modality and the associated control electronics. Device properties, including qubit connectivity, coherence time, and native gate sets, bound the feasible circuit depth, while gate durations, measurements, active resets, and control latencies govern sampling rates. Calibration drift imposes temporal limits on iterative training workloads by restricting the window over which calibrations and noise models remain approximately stationary. L2 (quantum execution pipeline) translates abstract logical circuits into physical instruction schedules, enabling execution on hardware. This layer handles gate synthesis, routing, and qubit mapping (Ji *et al.*, 2025a). Compilation strategies that account for device calibration and noise properties can improve the fidelity of the realized circuit at a fixed logical depth. Some workflows bypass gate abstractions to optimize analog control pulses directly, treating waveform degrees of freedom as trainable parameters in a pulse-level control loop (Melo *et al.*, 2023; Tao *et al.*, 2025a,b). L3 (hybrid execution and training engine) orchestrates the iterative training loop, managing the latency-critical interplay between quantum circuit execution, gradient estimation, and classical parameter updates. These workloads benefit from low-latency integration of QPUs with classical high-performance computing for error management and statistical analysis, a workflow often termed quantum-centric supercomputing (Lanes *et al.*, 2025), in which QPUs act as tightly coupled coprocessors and performance can be dominated by orchestration overheads, such as latency, batching, and scheduling, in addition to circuit depth. Building upon the execution layers, L4 (QDL architecture design) specifies the concrete architecture, such as encoding schemes, ansätze, and measurement operators, to be trained. Finally, L5 (applications and problem formulation) defines the loss landscapes and evaluation protocols that drive requirements through the stack.

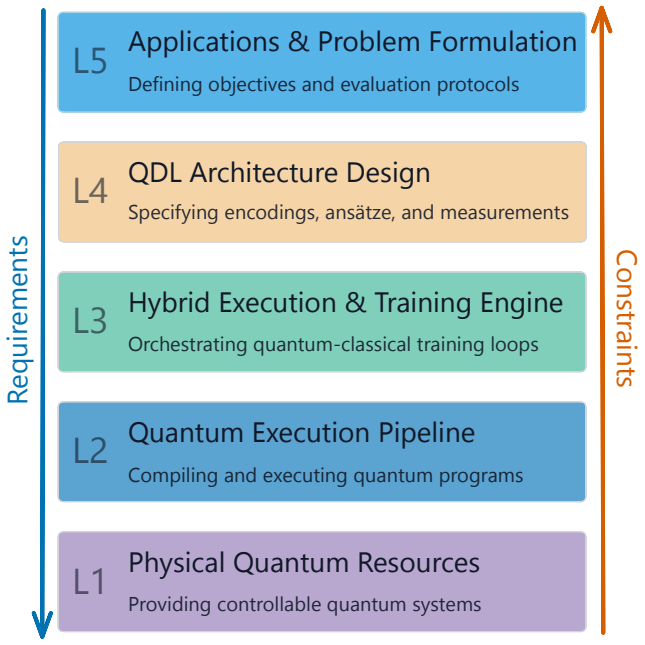

Figure 6 Five-layer implementation stack for QDL, illustrating top-down requirements flow and bottom-up constraint propagation.

Treating layers independently can obscure systems-level bottlenecks in QDL. The stack is characterized by complex interdependencies: top-down, application metrics (L5) dictate training budgets and compilation precision (L3/L2); bottom-up, the physical noise floor and calibration drift (L1) bound the feasible circuit depth and realizable model capacity at L4. For instance, an architecturally sound ansatz at L4 may become untrainable after mapping onto the restrictive topology or noise profile inherited from L1. A potential longer-term vision is a fully differentiable and data-driven stack where control, compilation, and architecture are optimized jointly. Recent work on pulse control (Li *et al.*, 2025a; Sarma and Hartmann, 2025), differentiable formulations of analog quantum computing and control problems (Clayton *et al.*, 2024; Leng *et al.*, 2022), and automated architecture search (Rapp *et al.*, 2025) indicates early progress toward such holistic optimization.

B. Quantum hardware platforms

The DiVincenzo criteria (DiVincenzo, 2000) enumerate the minimum physical prerequisites for scalable quantum computing: well-characterized qubits, state initialization, long coherence relative to control times, a universal gate set, and qubit-specific measurement. While several leading platforms satisfy these requirements at the component level (Acharya *et al.*, 2025; Wilhelm *et al.*, 2025), such as high-fidelity single- and two-qubit gates and readout, NISQ QDL implementations are limited by reliably executable depth due to system reliability and error rates insufficient for large-scale FTQC. Motivated by performance-centric roadmapping that ties hardware progress to end-to-end algorithmic success rates rather than qubit count alone (Barends and Wilhelm, 2025), we define four critical execution-centric metrics for QDL:
(1) Error-depth budget: gate and measurement infidelities, leakage, and coherence time (T_1 and T_2) relative to control times jointly bound the feasible effective depth of a QDL ansatz, while the width is strictly bounded by the physical qubit count and topology.
(2) Repetition rate and control latency: The repetition rate denotes the frequency at which quantum circuits are sampled (shots per second), while control latency represents the delay in the classical-quantum communication interface. In the context of QDL, these parameters dictate the total wall-clock training time, as the iterative nature of gradient estimation requires large numbers of circuit evaluations per epoch.
(3) Connectivity and routing overhead: Connectivity describes the topology of direct entangling interactions available between physical qubits. In QDL architecture design, sparse connectivity necessitates the insertion of auxiliary SWAP gates to realize non-local correlations (e.g., in strongly entangling layers); this routing overhead consumes the limited coherence budget and compounds the stochastic error per algorithmic-level operation (Ji *et al.*, 2025a; Murali *et al.*, 2019).
(4) Calibration stability: This represents the temporal invariance of hardware parameters, such as gate fidelities and resonant frequencies, over extended periods. For QDL workloads that span multiple hours or days, high stability is crucial to ensure a consistent loss landscape, preventing the optimization trajectory from being derailed by hardware drift.

Table II summarizes the characteristics most relevant to QDL feasibility and codesign opportunities. This diversity creates a rich algorithm-hardware codesign space in which QDL architectures, compilation strategies, and training protocols can be tailored to platform-specific execution constraints.

1. Superconducting circuits

Superconducting processors use lithographically fabricated superconducting circuits with Josephson junctions (Josephson, 1962, 1974) as nonlinear inductors, producing weakly anharmonic oscillators (e.g., transmons (Koch *et al.*, 2007)) addressable by microwave control (Clarke and Wilhelm, 2008; Gu *et al.*, 2017; Kjaergaard *et al.*, 2020; Krantz *et al.*, 2019). Circuit quantum electrodynamics (Blais *et al.*, 2021) enables microwave control and dispersive readout, with effective measurement fidelity depending on the chosen measurement basis (Ji *et al.*, 2026; Pommerening and DiVincenzo, 2020). The observation of macroscopic quantum tunneling (Devoret *et al.*, 1985) and energy quantization (Martinis *et al.*, 1985), recognized by the 2025 Nobel Prize in Physics (The Royal Swedish Academy of Sciences, 2025), established experimentally that engineered superconducting circuits can exhibit coherent quantum dynamics at macroscopic scales, an enabling prerequisite for modern qubit engineering and control (Nakamura *et al.*, 1999). Recent roadmaps emphasize performance-centric scaling, prioritizing end-to-end algorithmic success rates and operational stability over qubit count alone (Barends and Wilhelm, 2025).

This platform features fast gate operations (AbuGhanem, 2025) supporting high-throughput circuit evaluation, but sparse planar connectivity, which requires routing operations and increases circuit depth (Ji *et al.*, 2022). Hardware-supported gate sets combine single-qubit rotations with an entangling interaction such as CNOT or CZ (AbuGhanem, 2025). Pulse-level access supports hardware-tailored optimization and error mitigation (Ji *et al.*, 2023; Ji and Polian, 2024) and direct pulse-parameter optimization framed through quantum optimal control (Koch *et al.*, 2022; Magann *et al.*, 2021; Wittler *et al.*, 2021). Key limitations encompass two-qubit gate error, measurement error,

Table II Modality-level hardware characteristics most relevant to QDL feasibility and codesign. Connectivity refers to native two-qubit (or native entangling) interaction structure; QDL bottleneck denotes a typical dominant limitation for end-to-end hybrid training at current scales. CV: continuous-variable. For annealers, minor embedding refers to mapping a logical problem graph onto the fixed hardware graph.

Modality	**Primitive**	**Connectivity**	**QDL bottleneck**
Superconducting	digital gates	sparse 2D (routing)	routing overhead, drift/crosstalk
Trapped ions	digital gates	all-to-all (in-modular); modular links	wall-clock throughput
Neutral atoms	digital gates + analog	reconfigurable between shots	entangling error, atom loss
Photonics	optical sampling; CV	interferometric meshes; delay networks	optical loss, calibration
Annealers	analog sampling	fixed graph (minor embedding)	embedding, non-equilibrium bias
Semiconductor spin	digital gates	local coupling	variability, coupling range

crosstalk, and calibration drift, which can degrade long-running training stability. Data-efficient noise characterization can provide the context-specific error information targeting the interplay between algorithm and hardware to inform compiler decisions and routing heuristics (Ji *et al.*, 2025b).

The superconducting platform has been a primary experimental testbed for end-to-end variational QDL, including gradient-based training (Pan *et al.*, 2023a), quantum adversarial learning (Ren *et al.*, 2022), transfer learning (Mari *et al.*, 2020), and generative-model benchmarks (Hamilton *et al.*, 2019; Ristè *et al.*, 2017; Zoufal *et al.*, 2019). More recent advancements include the experimental demonstration of quantum continual learning (Zhang *et al.*, 2026a), addressing catastrophic forgetting in sequential tasks and parameter-efficient quantum anomaly detection for general image datasets (Wang *et al.*, 2025a).

2. Trapped ions

Trapped-ion processors confine ions in electromagnetic traps, using long-lived internal hyperfine or optical transitions as qubits (Bruzewicz *et al.*, 2019; Häffner *et al.*, 2008). Entangling operations are generated by laser-driven coupling to shared motional modes, so that collective phonons act as an interaction bus and Mølmer-Sørensen-type gates provide a standard native entangling primitive (Mølmer and Sørensen, 1999; Sørensen and Mølmer, 1999). Following the proposal by Cirac and Zoller (1995), early experiments demonstrated a fundamental two-qubit quantum logic gate in this platform (Monroe *et al.*, 1995). Demonstrations report single-ion coherence times exceeding one hour (Wang *et al.*, 2021a) and two-qubit gate fidelities at or above the 99.9% under optimized conditions (Strohm *et al.*, 2024). Current roadmaps emphasize fast mid-circuit measurement, measuring qubits during circuit execution rather than only at the end, and real-time feed-forward using measurement outcomes to control subsequent gates as central primitives for many QEC and lattice-surgery style protocols (Benito *et al.*, 2025; Paetznick *et al.*, 2024), while alternative measurement-free logical toolboxes are also being explored to mitigate slow measurement (Butt *et al.*, 2026).

All-to-all connectivity within a module eliminates SWAPs, enabling dense, highly expressive ansätze, but microsecond two-qubit gates yield lower throughput (Bruzewicz *et al.*, 2019; Häffner *et al.*, 2008), making the wall-clock time for large-batch training significantly higher. Dominant performance bottlenecks include motional heating, laser phase noise, and scaling-induced spectral crowding of the motional modes (Bruzewicz *et al.*, 2019; Strohm *et al.*, 2024). This platform has served as a critical testbed for high-fidelity, effectively all-to-all connectivity within modules and for QEC architecture studies. The platform's dense connectivity enables the implementation of complex entangling layers for generative modeling and classification tasks that exploit the native Mølmer-Sørensen structure (Chen *et al.*, 2024a). The ability to perform mid-circuit measurements and feed-forward can enable sample-efficient learning strategies and logical-layer training (Jones and Murali, 2025), potentially mitigating throughput constraints by enabling superior per-shot information density and error suppression. Benchmarks compare QNNs across trapped-ion and superconducting devices, showing strong noise sensitivity (Lakhdar-Hamina *et al.*, 2025).

3. Quantum annealers

Quantum annealers, introduced as an open-system computation model in Sec. II.A.2, implement programmable Ising-type Hamiltonians in superconducting flux-qubit hardware, enabling high-throughput sampling of near-ground-state configurations (Albash and Lidar, 2018; Hauke *et al.*, 2020; Kadowaki and Nishimori, 1998). Contemporary commercial implementations, such as D-Wave's Advantage2 featuring Zephyr topology (a specific qubit connectivity graph) with over 4,400 flux qubits, utilize thousands of superconducting flux qubits to physically realize these dynamics, providing a natural testbed for large-scale benchmarking of Ising optimization (Heidari *et al.*, 2024; Quinton *et al.*, 2025; Willsch *et al.*,

2022), simulation (King *et al.*, 2025; Vodeb *et al.*, 2025), and sampling tasks (Benedetti *et al.*, 2018; Cavallaro *et al.*, 2020; Perdomo-Ortiz *et al.*, 2018; Salloum *et al.*, 2024; Schulz *et al.*, 2025; Willsch *et al.*, 2020; Wilson *et al.*, 2021).

This platform offers high-throughput repeated sampling but operates on a fixed sparse hardware graph, so nonnative couplings require minor embedding (Pelofske, 2025; Willsch *et al.*, 2022; Yarkoni *et al.*, 2022). Open-system dynamics with thermal relaxation and schedule-dependent freeze-out (premature cessation of quantum dynamics before reaching ground state) yield an output distribution that can be modeled by an effective-temperature Gibbs form but exhibits systematic, problem- and schedule-dependent deviations (Amin, 2015; Marshall *et al.*, 2019; Nelson *et al.*, 2022). Utilization in learning workloads follows two distinct paradigms: treating annealers as approximate physical Gibbs samplers for training energy-based generative models (Adachi and Henderson, 2015; Benedetti *et al.*, 2017; Higham and Bedford, 2023) or as global optimizers for discriminative networks, where discrete network parameters are encoded into the Hamiltonian to minimize the training objective (Abel *et al.*, 2022; Zhang and Kamenev, 2025). Benchmarking studies emphasize sensitivity to embedding overhead and device-specific systematics (Pelofske, 2025). Credible performance claims therefore require controlling calibration drift and bias (Chancellor *et al.*, 2022), and validating that the realized sampling distribution is consistent with the intended target model for sampling-based uses, under the stated access model and matched classical baselines (Nelson *et al.*, 2022).

4. Photonics

Photonic quantum processors (Kok *et al.*, 2007; Lee *et al.*, 2025b; Romero and Milburn, 2025; Wang *et al.*, 2020) encode information in optical modes using discrete variables (DV) where information is encoded in countable photon numbers or continuous variables (CV) paradigms where information is encoded in field amplitude and phase quadratures. In DV approaches, qubits are realized via dual-rail (Alexander *et al.*, 2025) or time-bin encodings (Vagniluca *et al.*, 2020), while CV paradigms encode information in the canonical field quadratures such as amplitude and phase of optical bosonic modes. Linear-optical interferometers implement unitary mode transformations, but scalable universality requires an additional effective nonlinearity. DV schemes achieve entangling operations via measurement-induced nonlinearities with ancillas and feed-forward, whereas CV architectures require at least one non-Gaussian resource beyond Gaussian optics. Recent roadmaps emphasize integrated photonics and system-level codesign of sources, routing, and detection, with measurement-based and fusion-based architectures (Bartolucci *et al.*, 2023) providing prominent scaling roadmaps (Aghaee Rad *et al.*, 2025; Pérez-López and Torrijos-Morán, 2025).

The characteristics are high-rate sampling throughput and loss-limited effective depth. Time-multiplexed photonic samplers demonstrate fast end-to-end sampling in large interferometric networks, making photonics attractive for sampling-centric learning primitives (Madsen *et al.*, 2022). Connectivity is implemented through programmable interferometers. Spatial meshes provide flexible mode mixing over a fixed number of modes (Clements *et al.*, 2016), while time-multiplexed designs synthesize large structured interaction graphs via delay networks and component reuse (Madsen *et al.*, 2022); frequency-multiplexed architectures analogously realize large mode graphs through reconfigurable operations in the spectral domain (Lu *et al.*, 2023). The dominant constraint is photon loss, including coupling, propagation, and detection inefficiencies. In postselected DV settings, the probability that all carriers survive decays rapidly with circuit depth, directly limiting the feasible layer count for variational constructions (Arrazola *et al.*, 2021).

Photonics is particularly well-matched to generative (Salavrakos *et al.*, 2025; Sedrakyan *et al.*, 2024) and reservoir computing paradigms (Aadhi *et al.*, 2025; García-Beni *et al.*, 2023; Rambach *et al.*, 2025; Sakurai *et al.*, 2025; Zia *et al.*, 2025) that exploit fixed optical dynamics as high-dimensional feature generators. In these approaches, the optical quantum module is non-trainable or only weakly parameterized, and optimization is concentrated in a classical readout or lightweight outer loop, as exemplified by experimental photonic extreme learning machines (Pierangeli *et al.*, 2021; Suprano *et al.*, 2024) and boson sampling-enhanced predictors (Rambach *et al.*, 2025). Beyond sampling demonstrations, integrated photonic processors have enabled experimental quantum feature maps for kernel methods (Yin *et al.*, 2025). Recent progress also includes hybrid quantum-classical photonic neural networks with cavity-assisted interactions for universal logical processing (Basani *et al.*, 2025). Additional advancements encompass photonic QCNNs with adaptive state injection (Monbroussou *et al.*, 2025). Accordingly, credible performance claims should be formulated and evaluated end-to-end, accounting for loss, calibration drifts, and classical baselines under matched access models, an approach operationalized by recent community benchmarking efforts (Notton *et al.*, 2025).

5. Neutral atoms

Neutral atom processors (Bluvstein *et al.*, 2026; Saffman, 2025) use optical tweezers to trap hundreds to thousands of individual atoms in dynamically reconfig-

urable 2D or 3D geometries, encoding qubits in long-lived hyperfine ground states (Manetsch *et al.*, 2025; Pichard *et al.*, 2024; Saffman *et al.*, 2010; Scholl *et al.*, 2021). Entangling interactions are commonly mediated by Rydberg excitation, where strong van der Waals interactions yield a Rydberg blockade mechanism that enables fast multi-qubit correlations and two-qubit gates (Evered *et al.*, 2023; Jaksch *et al.*, 2000). The platform has demonstrated rapid scaling, including commercial implementations such as QuEra (Wurtz *et al.*, 2023), Pasqal (Scholl *et al.*, 2021), and Atom Computing (Wintersperger *et al.*, 2023). Notably, the continuous operation of coherent arrays at the $\sim$ 3000-qubit scale (Chiu *et al.*, 2025) and the trapping of over 6000 highly coherent tweezer-trapped qubits (Manetsch *et al.*, 2025) indicate rapid progress in system engineering alongside improvements in logical-level operation (Bluvstein *et al.*, 2024, 2026).

The characteristics are programmable geometry with distance-mediated interactions and cycle-time-limited end-to-end throughput. Rydberg-mediated entangling operations occur on microsecond timescales, but iteration rate is often dominated by state preparation, readout, and array assembly and rearrangement overheads (Pichard *et al.*, 2024; Wintersperger *et al.*, 2023). Reconfigurability is a key advantage. Atom positions can be adapted between circuit evaluations to better match a target interaction graph, reducing SWAP-driven routing overhead for structured architectures (Pichard *et al.*, 2024). Within a single execution, however, the effective interaction graph remains constrained by blockade radius, parallelization limits, and control crosstalk, so connectivity is highly programmable but not arbitrary at fixed time (Saffman, 2025). Uniquely, the platform supports both digital gate-model operation and analog Hamiltonian simulation, enabling hybrid digital-analog workflows that trade compiled depth for hardware-native correlation generation (Lu *et al.*, 2024). Dominant limitations for long training runs include finite Rydberg lifetime, laser amplitude and phase noise, atom loss (loss of trapped atoms from tweezers during execution) and recapture policies, and measurement-induced errors and drift across repeated cycles (Scholl *et al.*, 2021).

Neutral atoms are particularly well matched to QDL paradigms that exploit structured interaction graphs and analog dynamics as learning primitives. Graph-based learning via quantum evolution kernels uses programmable Rydberg-array dynamics to define kernels on graphs (Henry *et al.*, 2021), with recent extensions incorporating attributed graphs by encoding node and edge information through local detunings in the Rydberg Hamiltonian (Djellabi *et al.*, 2025). Reservoir computing finds a natural platform match here: the Rydberg array's complex analog dynamics act as a fixed, high-dimensional feature generator, with only the classical readout trained. Platform-specific realizations include foundational Rydberg-reservoir theory (Bravo *et al.*, 2022) and increasingly large-scale neutral-atom implementations (Kornjača *et al.*, 2024).

6. Semiconductor spin qubits

Semiconductor spin-qubit processors encode quantum information in spin degrees of freedom, spanning single-electron (or hole) spins in gate-defined quantum dots and dopant-based donors (Burkard *et al.*, 2023; Kane, 1998; Loss and DiVincenzo, 1998) as well as optically addressable defect-center spins such as nitrogen-vacancy (NV) centers in diamond (Atatüre *et al.*, 2018; Awschalom *et al.*, 2021; Fischer *et al.*, 2025). Semiconductor spin qubits leverage CMOS-aligned fabrication for high integration densities and cryogenic cointegration (Bartee *et al.*, 2025; Li *et al.*, 2018). Milestones include industrial multi-qubit modules (George *et al.*, 2024; Steinacker *et al.*, 2025), universal multi-qubit control (Edlbauer *et al.*, 2025; Philips *et al.*, 2022), and early QEC (Takeda *et al.*, 2022). Elevated-temperature operation has been demonstrated above 1K for electron spins (Huang *et al.*, 2024b; Petit *et al.*, 2022) and above 4K for hole spins (Camenzind *et al.*, 2022). In parallel, defect spins emphasize optical networking for modular scaling (Awschalom *et al.*, 2021).

For QDL, semiconductor spin qubits offer fast physical gate times with predominantly local connectivity and calibration-intensive control (Bartee *et al.*, 2025; Burkard *et al.*, 2023; Stano and Loss, 2022). Single-qubit gates use electron or electric-dipole spin resonance, while two-qubit gates employ nearest-neighbor exchange, achieving high fidelities (Burkard *et al.*, 2023; Noiri *et al.*, 2022; Stano and Loss, 2022). Nonlocal operations rely on shuttling (De Smet *et al.*, 2025; Nagai *et al.*, 2025) or resonator and photon links (Awschalom *et al.*, 2021; Burkard *et al.*, 2023). Hole spin qubits provide strong spin-orbit coupling, enabling fast all-electrical control. Recent work emphasizes operating points that mitigate charge-noise sensitivity while retaining strong driveability (Fanucchi *et al.*, 2025; Hendrickx *et al.*, 2024; John *et al.*, 2025; Secchi *et al.*, 2025). Defect spins offer remote entanglement and networking but face throughput limits from photon collection and rates (Atatüre *et al.*, 2018; Awschalom *et al.*, 2021). QDL demonstrations on spin-based platforms are presently concentrated in small-scale learning primitives and algorithmic subroutines (Pan *et al.*, 2014; Wang *et al.*, 2022; Wen *et al.*, 2019), rather than end-to-end trained deep variational models on scalable semiconductor arrays. Representative examples include quantum kernel estimation in nuclear magnetic resonance (Sabarad *et al.*, 2024) and subroutines such as quantum principal component analysis (Li *et al.*, 2021b; Xin *et al.*, 2021). Reservoir learning has also been explored, but much of the current evidence remains simulation- or theory-driven, including studies where dissipation shapes

performance (Mifune *et al.*, 2025).

C. Software tools and frameworks

This section reviews the software ecosystems supporting the construction, training, and deployment of QDL models, focusing on actively maintained open-source frameworks that provide practical interfaces for hybrid workflows. The landscape can be broadly categorized into general-purpose QML frameworks, hardware-optimized platforms, and photonic or continuous-variable specialized toolchains. Ten widely used QDL software frameworks are summarized in Table III.

General-purpose QML frameworks prioritize interoperability with classical ML libraries and support hybrid training loops. PennyLane (Bergholm *et al.*, 2022) offers a unified differentiable interface for quantum circuits, enabling seamless integration with PyTorch (Paszke *et al.*, 2019), TensorFlow/Keras (Abadi *et al.*, 2016), and JAX (Bradbury *et al.*, 2018). It combines classical autodifferentiation with quantum-specific gradient rules and provides analysis tools such as the quantum metric tensor for studying optimization landscapes. TensorFlow Quantum (Broughton *et al.*, 2021) embeds Google's Cirq (Developers, 2025) circuits as TensorFlow ops, allowing quantum subroutines to be trained within Keras pipelines using classical backpropagation and quantum gradient estimators. Qiskit Machine Learning (Sahin *et al.*, 2025) extends the Qiskit (Javadi-Abhari *et al.*, 2024) ecosystem with variational model builders and scikit-learn style interfaces (Buitinck *et al.*, 2013), while its TorchConnector enables gradient-based training via PyTorch. Similarly, sQUlearn (Kreplin *et al.*, 2025) provides scikit-learn compatible quantum classifiers and regressors, interfacing with Qiskit, PennyLane, and the Qulacs simulator (Suzuki *et al.*, 2021) for accelerated circuit evaluation.

Hardware-optimized and performance-oriented platforms emphasize efficient simulation and execution on heterogeneous hardware. NVIDIA's CUDA-Q (Slysz *et al.*, 2025) adopts a single-source programming model designed for quantum-classical co-execution in high-performance computing (HPC) environments, with native GPU acceleration. TensorCircuit-NG (Zhang *et al.*, 2023b) leverages just-in-time compilation and TN contraction, supporting automatic differentiation through JAX, TensorFlow, and PyTorch backends. Within the QPanda ecosystem (Dou *et al.*, 2022; Zou *et al.*, 2025), VQNet (Chen *et al.*, 2019) embeds trainable quantum operators into classical learning pipelines, and VQNet-2.0 (Bian *et al.*, 2023) extends this with hybrid automatic differentiation and cross-platform deployment capabilities.

Photonic and continuous-variable toolchains cater to non-qubit encodings. Strawberry Fields (Killoran *et al.*, 2019b) supports the design and simulation of Gaussian and non-Gaussian continuous-variable photonic circuits. Piquasso (Kolarovszki *et al.*, 2025) is a full-stack photonic programming and simulation platform supporting both discrete- and continuous-variable circuits, with optional high-performance backends and differentiable simulation, e.g., TensorFlow and JAX integration, to enable photonic QML prototyping and benchmarking. For discrete-variable photonics, Perceval (Heurtel *et al.*, 2023) provides a dedicated environment for linear-optical circuit construction and simulation, with hardware-oriented interfaces. More recently, DeepQuantum (He *et al.*, 2025) has emerged as a PyTorch-based platform that aims to unify hybrid training workflows across both qubit-based and photonic backends.

Across these frameworks, key practical considerations include simulation capabilities (GPU acceleration, HPC support, TN methods), hybrid workflow integration (support for Torch, JAX, or Keras), and quantum hardware support (circuit-based or photonic backends). Table III provides a detailed comparison along these dimensions, highlighting the trade-offs and specialization of each ecosystem.

There is also a supporting software layer for QDL, including compilation, optimization, and noise mitigation, which is increasingly central to hardware-facing QDL studies. Representative examples include compilers such as TKet (Sivarajah *et al.*, 2020), circuit optimization toolchains such as ZX-calculus (Kissinger and Van De Wetering, 2019), and backend-agnostic QEM libraries such as Mitiq (LaRose *et al.*, 2022). Benchmarking suites include Benchpress, which compares software stacks across large circuit families and device-motivated constraints (Nation *et al.*, 2025). The Munich Quantum Toolkit (Burgholzer *et al.*, 2025a,b) provides a modular suite of design-automation tools, spanning intermediate representations, compilation, verification, and benchmarking.

D. Experimental demonstrations

This section surveys QDL demonstrations on physical hardware, tracing a progression from early primitive validations to hierarchical NISQ training, scaling efforts, and the current frontier of practical advantage.

1. Early proofs-of-concept

Early experiments largely served to validate learning-relevant primitives and separations under controlled assumptions. Before variational hybrid training became standard, few-qubit photonic and nuclear magnetic resonance platforms demonstrated entanglement-assisted classification and support vector machine prim-

Table III Comparison of ten mainstream QDL software frameworks. The evaluation spans classical simulation capabilities, integration with classical libraries for hybrid workflows, and support for diverse quantum hardware paradigms. GPU-enabled: officially supported GPU-accelerated simulation backend; HPC-enabled: documented multi-node and distributed execution; TN: documented tensor network (TN) approximate simulators integrated in the stack.

Software	**Classical Simulation**			**Hybrid Workflow**			**QPU Integration**	
	GPU-enabled	HPC-enabled	TN	PyTorch	JAX	Keras	Circuit-based	Photonics
PennyLane	✓	✓	✓	✓	✓	✓	✓	✓
TensorFlow Quantum	✓	✓	×	×	×	✓	✓	×
Qiskit Machine Learning	✓	✓	✓	✓	×	×	✓	×
CUDA-Q	✓	✓	✓	✓	×	×	✓	×
sQULearn	✓	○	×	✓	×	×	✓	×
TensorCircuit-NG	✓	✓	✓	✓	✓	✓	✓	×
VQNet 2.0	✓	✓	×	○	×	×	✓	×
Strawberry Fields	✓	○	✓	✓	×	✓	×	✓
Perceval	×	×	×	×	×	×	×	✓
DeepQuantum	✓	○	✓	✓	×	×	✓	✓

✓ Supported ○ Limited × Unsupported

itives (Cai *et al.*, 2015; Li *et al.*, 2015). In parallel, early experiments began probing learning-relevant speedups (Lee *et al.*, 2019; Saggio *et al.*, 2021), typically as restricted-model query or communication improvements or constant-factor effects rather than end-to-end DL speedups.

A representative example is the oracle-based learning-parity-with-noise demonstration on a five-qubit superconducting processor by Ristè *et al.* (2017), which illustrates a query-complexity separation on hardware within the corresponding oracle model. Although far from data-driven DL, such results helped clarify that the *meaning* of advantage depends on the access model and the precisely fixed classical baseline required for falsifiability. In parallel, small-scale experiments explored perceptron-like and classifier primitives on gate-based devices. For example, Tacchino *et al.* (2019) experimentally tested a quantum-perceptron concept on a small processor; these demonstrations primarily test end-to-end training interfaces and readout-induced nonlinearities, rather than providing performance evidence on realistic datasets. Concurrently, early hardware demonstrations of variational generative modeling (Hamilton *et al.*, 2019; Hu *et al.*, 2019) and kernel-style classification (Havlíček *et al.*, 2019; Peters *et al.*, 2021) established the practical hybrid training stack and highlighted sensitivity to shot noise and device imperfections.

2. Training hierarchical models on NISQ processors

Once hybrid training became the default interface, the central experimental challenge shifted to building *multi-layer* models without losing trainability under finite-shot gradient estimation and device noise. Skolik *et al.* (2021) proposed and tested in simulation a layerwise training strategy, in which only subsets of parameters are updated while earlier layers are held fixed; such schemes can improve optimization behavior in practice.

A notable experimental milestone is the demonstration that layered QNNs can be trained with gradient-based updates on superconducting hardware using hardware-in-the-loop differentiation, with the forward process executed on hardware and the backward process implemented on a classical computer (Pan *et al.*, 2023b). However, this approach does not constitute a scalable, fully quantum backpropagation primitive. Extending it to wider or deeper models would require either large-scale classical simulation of the backward pass or measurement-intensive reconstruction of intermediate quantum information, both of which can scale exponentially with subsystem size in the worst case (Abbas *et al.*, 2023). These works sharpen the practical questions for experimental QDL: how depth, shot budgets, noise, and drift trade off against achievable performance and reproducibility.

3. Pushing the frontiers: scale, depth, and robustness

Recent experiments have pushed QDL along three axes: scale (qubit count or modes), executable depth, and robustness. On the scale axis, reservoir-style approaches are particularly attractive because the quantum substrate is not trained by circuit-parameter gradients; instead, fixed quantum dynamics provide a high-dimensional feature map, and only a classical readout is optimized. This avoids quantum-gradient evaluation characteristic of variational models under finite-shot noise, though repeated evaluations are still required to

estimate reservoir features and fit the readout. A prominent example is the large-scale experimental demonstration of quantum reservoir computing on a neutral-atom analog quantum computer, scaling to systems of up to 108 qubits (Kornjača *et al.*, 2024). Note that this is an analog-reservoir regime rather than deep variational circuit training. While such reservoir pipelines are not end-to-end trained variational deep networks, they provide a concrete and practically relevant route by which large qubit counts can contribute functionally to temporal learning tasks.

Executable depth records are best interpreted primarily as hardware capability rather than as QDL performance evidence; for example, deep coherent circuit demonstrations such as quantum signal processing sequences inform depth headroom and control limits (Bu *et al.*, 2025). Within learning-centric experiments, depth remains constrained by a combined budget of decoherence, coherent control errors, and sampling noise, which motivates architectures and training protocols that are explicitly depth-aware.

Robustness has also emerged as a critical figure of merit. Empirical studies have reported training instability and sensitivity to hardware noise and finite-shot sampling overhead in variational deep RL workflows on superconducting platforms (Franz *et al.*, 2023). Separately, early experimental work has demonstrated that quantum models can exhibit adversarial vulnerability in close analogy to classical networks (Ren *et al.*, 2022), motivating systematic robustness benchmarking (Zhang *et al.*, 2026b). Mitigation strategies include depth reduction via approximate data encoding (West *et al.*, 2024) and finite-shot noise reduction via variance-regularized training objectives (Kreplin and Roth, 2024).

4. Benchmarking and reproducibility

As experimental demonstrations mature, benchmarking has become a dominant methodological bottleneck. The central difficulty is that nascent quantum models are often compared against classical baselines that are not equivalently tuned, or under mismatched computational budgets. Without a disclosed hyperparameter-search protocol and training-time budget, performance comparisons are rarely informative (Bowles *et al.*, 2024). Benchmark-quality studies should therefore report the resource contract alongside the full hyperparameter and wall-clock budget used to tune both quantum and classical models. At a minimum, a credible pipeline-level advantage claim requires: (i) a functional-contribution test, e.g., ablation showing the quantum subroutine changes the outcome beyond what is explained by classical preprocessing and noise; (ii) budget-matched comparison to tuned classical baselines, including transparent hyperparameter and training-time budgets; (iii) repeated runs across calibrations to quantify variability; and (iv) explicit costing of verification, especially as experiments approach beyond-simulability regimes where full classical replication is unavailable and only restricted, property-based certificates may be feasible under stated assumptions.

While most current QDL experiments remain in or near classically reproducible regimes, hidden code sampling provides a concrete example of such property-based verification logic by checking statistical properties of output distributions without full simulation (Deshpande *et al.*, 2025). Recent benchmarking efforts further bridge this gap, including time-series prediction studies that compare against strong classical sequence baselines (Fellner *et al.*, 2025), reinforcement-learning comparisons conducted in controlled simulation settings (Zare and Boroushaki, 2025), and large-scale hardware image-classification experiments that aim to provide more systematic, reproducible benchmarks (Gharibyan *et al.*, 2025). However, reproducibility remains hindered by device drift and calibration variability, underscoring the need for explicit characterization of time-dependent error modes and drift-aware protocols to substantiate (Buonaiuto *et al.*, 2024; Proctor *et al.*, 2020). Platform diversity expands accessible paradigms and benchmarking targets, yet it also amplifies noncomparable claims absent standardized resource models and reporting conventions (Lall *et al.*, 2025).

5. Toward practical quantum advantage

The experimental arc reveals a consistent pattern: successive demonstrations have widened the operational envelope of QDL, yet to our knowledge, no published experiment simultaneously satisfies all six conditions of Definition 3 under an explicit, budget-matched classical comparator at a scale beyond the best available classical simulation. The limiting factor is not a lack of isolated positive results, but the scarcity of claims that are simultaneously falsifiable, budget-matched, and fully disclosed, and reproducible across calibrations within a single, explicitly stated resource contract. Closing this gap requires treating verification as a first-class experimental deliverable: as experiments move beyond regimes where full classical replication is feasible, the evidentiary standard must shift from end-to-end simulation to protocol-level certification under clearly specified access-and-readout model assumptions.

Within this constrained landscape, near-term pathways to practical advantage diverge strictly based on the data modality. For *quantum-native data* and *structured generative* tasks, state-preparation overheads are either absent or offset by sampling hardness, allowing theoretical separations to be grounded directly in information-theoretic constraints. Conversely, for *classical-data QDL*,

severe encoding bottlenecks dictate that the most plausible routes to advantage are systems-level and targeted rather than monolithic. Representative loci include: (i) amortized inference-time benefits, where a fixed quantum subroutine is executed at high repetition rate and the relevant metric is throughput at fixed accuracy; (ii) weakly trained or gradient-free quantum feature dynamics, such as fixed-circuit or reservoir-style primitives, that trade expressivity for reduced gradient-estimation overhead, while still requiring explicit sampling budgets; and (iii) hardware codesign, where device-native primitives, including midcircuit measurement and reset, analog blocks, and parallelism, are integrated with compilation, mitigation, and classical postprocessing to reduce constant-factor overheads in an end-to-end contract. These divergent application domains are surveyed in Secs. V.A and V.B, while the corresponding trajectory toward scalable, fault-tolerant implementations is developed in Sec. VI.B.

E. Challenges and opportunities

Scalable advantage claims for deep quantum learners are constrained by both hardware resource budgets and the scaling of optimization and statistical estimation costs. Hardware constraints bound compiled depth and executable two-qubit layers, per-iteration measurement budgets, and throughput within calibration-drift windows. Moreover, many QDL architectures face scaling barriers, most prominently barren plateaus and input-output bottlenecks, that can render training or inference resource-prohibitive.

1. Hardware constraints

While general-purpose quantum metrics are well-documented (Lall *et al.*, 2025), QDL workloads are constrained by four coupled budgets: effective compiled depth D_{eff}, shot time t_{shot}, iteration latency T_{iter}, and stability window τ_{stable}. The depth budget D_{eff} is defined by the compiled circuit on the target topology, comprising both the trainable ansatz and the data-encoding subcircuit. For common feature-map and data re-uploading strategies, encoding depth often scales with input dimension and can consume a substantial fraction of the coherence budget before processing begins. Consequently, the maximum depth is bounded by accumulated gate errors and decoherence (Sec. IV.B).

Wall-clock feasibility is governed by finite throughput and system latencies. The iteration time T_{iter} decomposes as:

$$T_{\text{iter}} \approx S t_{\text{shot}} + t_{\text{control}} + t_{\text{compile}} + t_{\text{load}} + t_{\text{classical}}, \quad (17)$$

where S is the total number of circuit executions (shots) per optimization iteration. t_{control} and t_{compile} denote overhead from the control stack and circuit compilation, respectively. t_{load} represents the bandwidth-bound latency required to transfer classical features to quantum control parameters, and $t_{\text{classical}}$ encapsulates the compute-bound classical workload, including bitstring post-processing, gradient estimation, and optimizer updates.

Calibration drift renders the implemented channel time-dependent over a stability window (τ_{stable}). Rigorous QDL protocols therefore require time-resolved diagnostics, such as interleaved reference circuits, to detect and quantify drift during training (Proctor *et al.*, 2020, 2025). Fault tolerance mitigates depth constraints via logical encoding but shifts bottlenecks to syndrome-extraction bandwidth and decoding latency without guaranteeing stable learning-loop performance.

These budgets imply that robust QDL progress is often constrained by end-to-end systems feasibility: the full training loop should remain executable within D_{eff}, t_{shot}, T_{iter}, and τ_{stable}. At a minimum, studies are recommended to report: (i) compiled circuit depth, including two-qubit layer count and compilation overhead; (ii) the number of trainable parameters, objective terms, and the estimator used per term; (iii) total QPU calls and shots per optimization step with wall-clock cost, including any gradient-estimation overhead; (iv) drift and calibration management during training; and (v) the classical resources and baselines used for training and evaluation.

2. Scalability challenges

Scalability means increasing qubit number, circuit depth, and model capacity while keeping training and inference resources competitive with tuned classical baselines under comparable access assumptions. Three coupled barriers dominate feasibility in practice: (i) optimization signal degradation, including shot-noise-limited gradients and barren-plateau regimes, (ii) limited scaling of error-mitigation strategies along iterative learning trajectories, and (iii) input-output overheads from data loading and finite-shot readout.

As established in Sec. III.B.3, highly expressive ansätze can exhibit barren plateaus. Operationally, this can imply rapidly increasing shot requirements to resolve gradients at fixed precision. Crucially for hardware implementations, this scaling barrier is severely compounded by realistic local noise models (Singkanipa and Lidar, 2025; Wang *et al.*, 2021b) and necessitates refined Lie-algebraic diagnostics (Larocca *et al.*, 2025; Ragone *et al.*, 2024) to identify scalable regimes. Moreover, improved trainability guarantees need not imply computational advantage, and can coincide with regimes admitting efficient classical descriptions (Cerezo *et al.*, 2025). Mitigations based on circuit and objective structure (Pesah *et al.*, 2021), initialization (Grant *et al.*, 2019), and non-

local parametrization (Broers and Mathey, 2024) can improve trainability, and can also be aided by measurement design (Sack *et al.*, 2022), but they need to be evaluated jointly with the induced resource budgets and with matched baselines (see Secs. III.B.3 and III.G.4). Landscape statistics can also change qualitatively with parameter count: overparameterization phase transitions have been analyzed and can improve trainability in specific model families, but they come with increased optimization and measurement burdens (Larocca *et al.*, 2023), and noise can materially modify these regimes (García-Martín *et al.*, 2024; Mele *et al.*, 2024).

As introduced in Sec. II.A.5, QEM can incur sampling overheads that grow rapidly with circuit depth for broad classes of noise models (Cai *et al.*, 2023; Takagi *et al.*, 2022). In QDL, these overheads severely compound with the repeated measurement demands of iterative training (Schumann *et al.*, 2024; Wang *et al.*, 2024a, 2021b). Recent lower bounds sharpen this limitation by showing that, for broad classes of noise models and mitigation protocols, the required sampling overhead can scale very unfavorably, even when mitigation is otherwise well-defined (Quek *et al.*, 2024; Takagi *et al.*, 2023)

Input-output constraints can dominate scaling. The encoding and access-model overheads are developed in Sec. III.A. Data-encoding choices can also materially affect trainability by controlling induced ansatz expressibility and entanglement, tightening the coupling between input design and optimization scaling (Kashif and Al-Kuwari, 2023). For many classical tasks, data-loading costs can dominate total runtime, reducing or eliminating any net advantage from quantum processing (Aaronson, 2015). On the output side, classical information extraction is limited by shot noise. Estimating expectation values and gradients to precision ϵ requires $\Omega(\epsilon^{-2})$ samples for bounded-variance estimators, with additional factors depending on the number of measured terms and the chosen gradient estimation strategy (Schuld and Killoran, 2019). These constraints delimit regimes where scalable advantage is most plausible, including quantum-native data and objectives that avoid classical-to-quantum loading (Huang *et al.*, 2022), and structured generative settings in which learning can be efficient while sampling and inference are argued to remain classically hard, as reported experimentally on a 68-qubit superconducting processor (Huang *et al.*, 2025a).

V. APPLICATIONS AND COMPARATIVE ANALYSIS

This section surveys QDL application domains and then synthesizes a cross-domain protocol for quantum-classical comparison using those domains as empirical evidence.

A. Application domains

QDL applications are categorized by data regime, distinguishing classical inputs processed by hybrid subroutines from inherently quantum data accessed coherently, as shown in Fig. 7. In classical regimes, evaluation emphasizes empirical feasibility and parameter efficiency constrained by encoding overheads and shot noise. Conversely, in quantum-data settings, provable separations emerge for specific task families under explicit coherent or collective access models, with classical comparators restricted to information obtainable from measurement outcomes. Scientific discovery, typically involving both regimes, bridges these extremes by imposing rigorous resource accounting on expensive forward models, motivating the four-pillar comparison protocol in Sec. V.B.

1. Image classification

Image classification is widely used as a diagnostic and comparative benchmark for QDL (Kharsa *et al.*, 2023; Senokosov *et al.*, 2024). Classical inputs and mature classical vision architectures provide well-understood baselines. Quantum pipeline costs are often dominated by data encoding and sampling, making this domain a controlled testbed for quantum feature maps, inductive biases, and finite-shot trainability under device noise. Rigorous evaluations require carefully tuned classical baselines and, where relevant, quantum-inspired classical surrogates (e.g., TN or kernel methods) under matched resource contracts. Reported empirical improvements are, at present, more often explained by baseline choice, inductive-bias effects, and hybridization rather than by a demonstrated end-to-end quantum advantage under budget-matched tuning.

Early pipelines downsample images into patches and encode features into PQC parameters (Farhi and Neven, 2018; Grant *et al.*, 2018; Li *et al.*, 2020a), which weaken or only partially preserve spatial inductive biases and can make performance strongly dependent on the chosen feature map (Havlíček *et al.*, 2019; Pérez-Salinas *et al.*, 2020; Schuld *et al.*, 2021). Many QDL models draw architectural inspiration from classical CNNs that are well established for vision tasks. QCNNs formally defined in Sec. III.B.1 impose multiscale convolution-pooling structure (Cong *et al.*, 2019; Herrmann *et al.*, 2022; Hur *et al.*, 2022) and can avoid barren-plateau behavior in certain locality-structured settings (Pesah *et al.*, 2021), but, as discussed in Sec. III.E.1, under additional locality or limited-entanglement assumptions, they may also admit efficient classical simulation in broad regimes (Bermejo *et al.*, 2024).

Hybrid CNN pipelines insert small PQCs into classical vision stacks (Fan *et al.*, 2024a; Liu *et al.*, 2021a; Long *et al.*, 2025). Closely related are quanvolutional pipelines

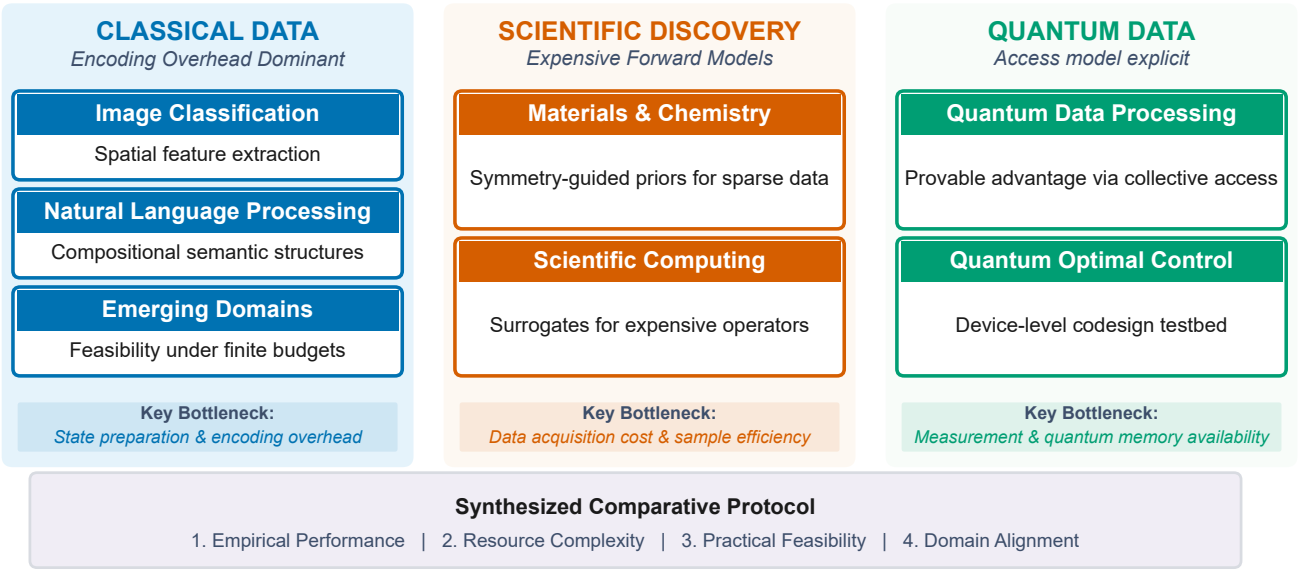

Figure 7 Taxonomy of QDL application regimes by data interface. Classical data pipelines embed classical features into quantum circuits and are typically bottlenecked by state-preparation overheads. Scientific discovery integrates quantum modules into label-expensive workflows with costly forward models for both classical and quantum data, where sample efficiency is paramount. Quantum data pipelines operate directly on quantum-native states or processes under an explicit access model, with bottlenecks set by the readout interface and the availability of coherent quantum memory. The lower panel summarizes the four-pillar evaluation protocol used in Sec. V.B.

(Henderson *et al.*, 2020a) defined in Sec. III.B.1 in which shallow patchwise circuits produce spatial feature maps consumed by classical backbones, including recent extensions with trainable kernels (Bhatia *et al.*, 2023; Kashif and Shafique, 2025), and transfer-learning schemes that compress images via classical embeddings before quantum refinement (Azevedo *et al.*, 2022; Kim *et al.*, 2023a; Mari *et al.*, 2020; Sugunapriya and Markkandan, 2025). However, for classical-image workloads at currently accessible sizes, these architectures typically employ small numbers of qubits and shallow circuits, and are therefore often compatible with efficient classical simulation in the studied parameter regimes. Alternative models for image classification aiming to explore less conventional quantum learning paradigms have been proposed, such as quantum extreme learning machines (De Lorenzis *et al.*, 2025) or transformer-like proposals (Cherrat *et al.*, 2024; Guo *et al.*, 2024b; Xue *et al.*, 2025).

Most studies use traditional vision datasets such as MNIST (LeCun *et al.*, 1998), Fashion-MNIST (Xiao *et al.*, 2017), and CIFAR (Krizhevsky *et al.*, 2009). A smaller study considers domain-specific scientific images such as high-energy physics experiments (Chen *et al.*, 2022c) and medical imaging (Chow, 2025; Landman *et al.*, 2022; Mathur *et al.*, 2021). While select studies report parameter efficiency compared to small classical networks (Liu *et al.*, 2025a), the stability, reproducibility, and magnitude of reported gains are often sensitive to baseline choice and hyperparameter tuning (Basilewitsch *et al.*, 2025; Tasnim *et al.*, 2025). To date, no study has established a scalable and reproducible quantum advantage over state-of-the-art classical vision baselines under an explicit, budget-matched resource contract for practically relevant image-classification workloads, and performance is frequently limited by data-loading costs, sampling noise, and device errors.

2. Natural language processing

NLP (Chowdhary, 2020) studies discriminative and generative modeling of classical text data. Quantum NLP refers to NLP pipelines where the quantum component changes the learning or inference interface in a way that is explicit in the resource contract (Widdows *et al.*, 2024a). A technically specific motivation is provided by compositional semantic formalisms in which grammatical structure induces multilinear maps. For such models, sentence meaning reduces to a TN contraction compilable into circuits. This transparency enables falsifiable questions about inductive bias and parameter efficiency (Guarasci *et al.*, 2022; Nausheen *et al.*, 2025; Pallavi and Prasanna Kumar, 2025; Varmantchaonala *et al.*, 2024; Widdows *et al.*, 2024b). Given Transformer dominance (Devlin *et al.*, 2019; Vaswani *et al.*, 2017) and classical NLP's strong scaling laws, near-term quantum NLP is best viewed as a platform for methodology development and controlled benchmarking rather than as a competitor

to large-scale classical models.

In compositional distributional semantics, the DisCoCat framework (Coecke *et al.*, 2010; Grefenstette *et al.*, 2011) maps grammatical derivations to tensor-diagram contractions that define sentence meanings. In quantum implementations, these diagrams can be compiled into parameterized circuits (e.g., via ZX-calculus-based rewrites) (Coecke *et al.*, 2020; Meichanetzidis *et al.*, 2020) and trained end-to-end using dedicated toolchains such as lambeq (Kartsaklis *et al.*, 2021). Demonstrations include small-scale sentence classification experiments (Lorenz *et al.*, 2023; Meichanetzidis *et al.*, 2023; Silver *et al.*, 2024) and natural language generation (Karamlou *et al.*, 2022) on superconducting quantum processors, and scalable proposals on trapped ions (Duneau *et al.*, 2024). Related graph-based variants incorporate PQC-based attention within message-passing NLP architectures (Aktar *et al.*, 2025). An alternative route is composition-agnostic, mapping text to quantum feature states for kernel estimation (Alexander and Widdows, 2022). Related proposals study attention-inspired PQCs primitives (Chen and Kuo, 2025b; Li *et al.*, 2024; Tomal *et al.*, 2025) and hybrid modules for parameter-efficient large language model (LLM) adaptation (Kong *et al.*, 2025). Circuit-level analyses of transformer subroutines and attention computation (Gao *et al.*, 2023; Liao and Ferrie, 2024) are conditional on explicit access models.

Quantum-inspired NLP uses quantum probability formalisms as classical modeling primitives, representing queries or documents as density matrices and fitting matching models via information-theoretic criteria such as entropy-based objectives (Sordoni *et al.*, 2014; Sordoni and Nie, 2014; Sordoni *et al.*, 2013). Representative modern variants develop density-matrix-based sentence representations for classification and matching (Shi *et al.*, 2024, 2023; Zhang *et al.*, 2018, 2019), introduce explicit state-evolution updates for language modeling and sentiment dynamics (Fan *et al.*, 2024b; Yan *et al.*, 2024b), and explore entanglement-inspired embeddings that encode nonclassical correlations across tokens (Chen *et al.*, 2023b).

3. Materials design and quantum chemistry

Materials design and quantum chemistry (McArdle *et al.*, 2020; Sajjan *et al.*, 2022) are natural target domains for QDL and are often highlighted as potentially favorable because high-fidelity classical data acquisition is costly and frequently yields small or sparse labeled datasets, making inductive bias and parameter efficiency especially important. Although inputs are encoded in classical representations such as atomic structures, graphs, or descriptors, the supervised targets are inherently quantum-mechanical observables, including energies, forces, spectra, and response properties governed by an electronic Hamiltonian. The dominant bottleneck is the data-generation budget, so methodological claims are most meaningful under matched supervision budgets and accuracy targets.

The quantum contribution is typically introduced through a quantum-parameterized hypothesis class such as a PQC layer (Benedetti *et al.*, 2019) or a quantum kernel estimator (Havlíček *et al.*, 2019; Schuld, 2021; Schuld and Killoran, 2019), used as a module within a hybrid model and evaluated against compute-matched baselines including molecular GNNs (Gilmer *et al.*, 2017), SchNet-class models (Schütt *et al.*, 2017), and equivariant interatomic potentials (Batzner *et al.*, 2022). Key confounders are data-loading assumptions, the strength of symmetry-aware baselines, and the gap between chemically relevant accuracy thresholds and finite-shot noise.

A primary scientific difficulty is maintaining chemical accuracy across high-dimensional configurational and chemical spaces, particularly in strongly correlated regimes where bias and generalization are fragile (McArdle *et al.*, 2020). Classical neural wavefunction ansätze such as FermiNet (Pfau *et al.*, 2020) and PauliNet (Hermann *et al.*, 2020) provide strong references, while on quantum hardware wavefunction-centric hybrids target electronic-structure observables via variational objectives, most prominently VQE (McClean *et al.*, 2016; Peruzzo *et al.*, 2014). Beyond VQE, Xia and Kais (2018) explores an RBM wavefunction trained with quantum-assisted optimization in a small-scale setting. Inverse-design workflows optimize a target functional over many-body states prepared on a quantum simulator and then use Hamiltonian learning to extract a geometrically local parent Hamiltonian (Kokail *et al.*, 2026).

In property-prediction pipelines, data loading remains a recurring obstacle: mapping high-dimensional descriptors to few-qubit devices can dominate cost unless access assumptions are explicit (Giovannetti *et al.*, 2008; Tang, 2021). Reddy and Bhattacherjee (2021) combine latent-space compression with a quantum regression module, which functions as a specialized inductive-bias component and should be compared under matched budgets. Graph-based inductive biases remain central in molecular ML, and quantum-embedded GNN pipelines have been explored in resource-limited regimes (Chen *et al.*, 2021; Lu *et al.*, 2025; Verdon *et al.*, 2019b). Reported gains often emphasize parameter efficiency or stability and remain baseline-dependent. Recent work explores data-efficient models (Hagelueken *et al.*, 2025) and early drug-discovery prototypes (Ghazi Vakili *et al.*, 2025), but large-scale chemically accurate quantum advantage remains unestablished.

4. Scientific computing

Scientific computing focuses on data-driven surrogates for accurate but expensive first-principles simulations. High-fidelity supervision is the limiting resource, favoring hypothesis classes that exploit structural priors such as locality, conservation laws, and symmetries. QDL approaches include surrogate modeling for quantum mechanical problems, quantum model discovery from sparse data, and operator learning with expensive forward models (Chen *et al.*, 2024d; Xiao *et al.*, 2025; Ye *et al.*, 2024). Quantum contributions typically take the form of quantum-parametrized hypothesis classes or subroutines within classical training loops. Quantum surrogates are most informatively assessed when benchmarked against structure-aware classical baselines, particularly physics-informed neural networks (Raissi *et al.*, 2019) and neural operators (Kovachki *et al.*, 2023; Li *et al.*, 2020b; Lu *et al.*, 2021).

Scientific workloads impose stringent constraints because high-fidelity supervision often requires expensive forward solves and calibration, which limits data-generation budgets even when synthetic sampling is possible. Thus, data efficiency and robustness often matter as much as predictive accuracy. Existing hybrid demonstrations emphasize feasibility studies for surrogates of partial differential equation dynamics (Chen *et al.*, 2024d; Ye *et al.*, 2024). More structured proposals target operator learning (Xiao *et al.*, 2025), where the objective is to approximate maps between function spaces. Related exploratory directions include hybrid quantum physics-informed solvers and quantum neural operators (Leong *et al.*, 2025; Wang *et al.*, 2025b). Additionally, many-body Hamiltonians serve as natural testbeds for evaluating these quantum surrogates, given their intrinsic alignment with quantum-mechanical state spaces. Classical neural network quantum states provide strong baselines (Carleo and Troyer, 2017; Gao and Duan, 2017). QDL proposals introduce PQC-based parametrizations and/or quantum-assisted sampling strategies within this variational-learning setting (Gardas *et al.*, 2018; Zhang *et al.*, 2025b).

High-energy physics offers realistic benchmarking with large structured classical datasets and stringent calibration requirements (Di Meglio *et al.*, 2024), making it a demanding benchmarking environment for end-to-end evaluation. Quantum pipelines apply collider-style analyses with circuit feature maps and kernel estimators (Wu *et al.*, 2021a), while quantum-inspired classical methods, including TN classifiers, are highly competitive (Felser *et al.*, 2021). These comparisons highlight the importance of domain-informed inductive bias and careful benchmarking. This domain is methodologically demanding for QDL. Defensible conclusions require rigorous baseline tuning, uncertainty-aware evaluation, and explicit end-to-end resource accounting.

5. Quantum data processing

As formalized in the access-model taxonomy of Sec. III.G.1, quantum-data workflows are distinguished most sharply by the access-and-readout interface: (i) *measure-first* protocols that map each copy to a classical record and learn by classical postprocessing, and (ii) *coherent or collective* protocols that retain quantum information (e.g., via quantum memory) to enable joint measurements or coherent processing across copies. General learnability bounds and access-model separations are treated in Sec. III.G.2; here we focus on application-level instantiations of these interfaces.

Operationally, the quantum contribution lies in exploiting quantum resources at the interface, including measurement-efficient acquisition protocols with explicit sample-complexity guarantees (Chen *et al.*, 2024c; Grier *et al.*, 2024; Huang *et al.*, 2020), quantum-native compression (Ma *et al.*, 2024; Pepper *et al.*, 2019; Romero *et al.*, 2017), or fixed quantum dynamical reservoirs (Palacios *et al.*, 2024; Senanian *et al.*, 2024; Zhu *et al.*, 2025a). Kernel methods belong here only when the inputs are quantum-native and the kernel is estimated from quantum access; kernels for classical inputs are treated under classical-data applications.

Full reconstruction of generic n-qubit states requires resources exponential in n in the worst case, so efficient protocols require either additional structure promises (e.g., low rank) or weaker learning goals (Gross *et al.*, 2010). PAC-style learning can instead target prediction of measurement outcomes without full reconstruction (Aaronson, 2007). Representative task families that exploit coherent or collective interfaces include collective-measurement protocols for learning tasks (Chen *et al.*, 2022a; Huang *et al.*, 2022), unsupervised classification of quantum states (Sentís *et al.*, 2019), and separations highlighting limitations of fixed-measurement measure-first pipelines in supervised learning (Gyurik *et al.*, 2025). A scalable photonic implementation has experimentally reported a task-specific quantum learning advantage in sample complexity under an explicit photonic access model, illustrating this hierarchy in practice (Liu *et al.*, 2025e).

The classical shadows framework (Aaronson, 2018; Grier *et al.*, 2024; Huang *et al.*, 2020; Jerbi *et al.*, 2024) introduced in Sec. II.A.4 is a canonical *measure-first* strategy for predicting many observables from randomized single-copy measurements via classical postprocessing. By constructing a compact classical record from randomized measurements, it enables amortized evaluation across downstream queries with explicit sample-complexity control (Grier *et al.*, 2024; Huang *et al.*, 2020; Jerbi *et al.*, 2024). In the worst case, shadow tomography without quantum memory provably requires $\Omega(\min(M, 2^n))$ copies to estimate M target observables (Chen *et al.*, 2022a) to a fixed accuracy, where n is the

number of qubits. This lower bound clarifies when a fixed classical record cannot simultaneously support all downstream queries. Related constructions extend these ideas to resource-constrained characterization and learning of quantum processes, with costs set by both measurement budgets and circuit resources (Levy *et al.*, 2024). Complementary analyses clarify when classical postprocessing alone suffices to recover task-relevant structure and when additional quantum access is information-theoretically necessary (De Palma *et al.*, 2024).

When data admit a low-complexity model class, generative learning replaces worst-case reconstruction with trainable representations from incomplete measurements (Carrasquilla *et al.*, 2019; Torlai *et al.*, 2018). Practical performance depends on optimization stability and evaluation under realistic shot budgets, as examined in recent trainability and benchmarking studies (Hibat-Allah *et al.*, 2024; Rudolph *et al.*, 2024). Related directions include learning dynamics and system identification (Gentile *et al.*, 2021; Hangleiter *et al.*, 2024), quantum-dynamics modeling (Liu *et al.*, 2022b), representation learning (Jayakumar *et al.*, 2024), compression (Ma *et al.*, 2024; Pepper *et al.*, 2019; Romero *et al.*, 2017), and measurement-efficient online learning with fixed quantum dynamical reservoirs (Palacios *et al.*, 2024; Senanian *et al.*, 2024; Zhu *et al.*, 2025a). Quantum data processing is the application regime where the access-and-readout model can be stated most explicitly. Practical impact therefore hinges on realistic interface assumptions, including whether quantum memory or only measured outcomes are available, measurement budgets, and rigorous resource-matched benchmarking.

6. Quantum optimal control

Quantum optimal control designs time-dependent controls that steer quantum dynamics toward a target objective, and is traditionally addressed with classical gradient-based methods and, more recently, RL approaches (Khaneja *et al.*, 2005; Koch *et al.*, 2022). As noted in the superconducting platform discussion of Sec. IV.B.1, pulse-level access enables systematic translations between control parameterizations and circuit ansätze (Magann *et al.*, 2021), a correspondence that connects hardware-level control with variational QDL training. Accordingly, classical optimal-control and RL pipelines define the comparator stack: empirical claims are most interpretable under matched experimental-call budgets, hardware constraints, and evaluation metrics, such as time-to-target fidelity or robustness. Standard classical benchmarks include GRAPE (Khaneja *et al.*, 2005), GOAT (Machnes *et al.*, 2018), PEPR (Heimann *et al.*, 2025), and Krotov-type updates (Konnov and Krotov, 1999; Reich *et al.*, 2012; Tannor *et al.*, 1992), while RL has been used for robust pulse design in hardware-facing settings (Baum *et al.*, 2021; Bukov *et al.*, 2018; Li *et al.*, 2025a; Niu *et al.*, 2019; Sarma and Hartmann, 2025; Sivak *et al.*, 2022). Pulse-level control can also act as an alternative to gate-level compilation in variational workflows, trading discrete gate synthesis for directly optimized control waveforms and therefore requiring pulse-level resource accounting (de Keijzer *et al.*, 2023; Leng *et al.*, 2022; Meirom and Frankel, 2023; Melo *et al.*, 2023).

In this review, we focus on QOC as a QDL application when the policy or control ansatz is quantum-parametrized and trained in a measurement-limited closed loop where updates are driven by measured cost signals. Variational-circuit policies have been proposed for continuous-action control via quantum deterministic policy-gradient variants (Wu *et al.*, 2025b), and related work explores quantum enhancements to deep RL in large action spaces (Jerbi *et al.*, 2021b). At the hardware level, pulse-parametrized learning frameworks enable direct optimization of control waveforms within variational learning stacks (Liang *et al.*, 2022). Global pulse parametrizations, such as Fourier-basis schedules, can reduce gradient-variance concentration relative to local piecewise-constant controls (Broers and Mathey, 2024), resulting in a mitigation of barren plateaus. This is related to the control-theoretic diagnostics that link trainability to controllability through the dynamical Lie algebra generated by the ansatz (Larocca *et al.*, 2022a). Hybrid VQAs have also been proposed explicitly for quantum control objectives, providing a direct bridge between VQA methodology and QOC benchmarks (Huang *et al.*, 2025c). Under this scope criterion, QOC provides a device-level testbed for quantum-classical codesign under explicit end-to-end resource accounting. Outside of it, QOC primarily supplies comparator families and evaluation protocols for learning-based claims.

7. Emerging domains

QDL has been explored in a growing set of emerging sociotechnical and engineering domains, including finance (Cherrat *et al.*, 2023; Herman *et al.*, 2023), cyberphysical systems (Ajagekar and You, 2021; Ansere *et al.*, 2024; Kaseb *et al.*, 2024; Schetakis *et al.*, 2025; Yi, 2025), healthcare (Gupta *et al.*, 2025; Hossain *et al.*, 2024; Sagingalieva *et al.*, 2023a), security- and privacy-constrained inference (Yu *et al.*, 2023), logistics (Correll *et al.*, 2023), robotics (Hohenfeld *et al.*, 2022; Yan *et al.*, 2024a), and distributed learning (Kwak *et al.*, 2023; Nguyen *et al.*, 2025). Representative industrial vision tasks provide an additional testbed beyond toy datasets (Yang and Sun, 2022). Across these areas, the dominant role of such studies is best understood as using the application as a testbed, employing realistic domain problems to stress-test QDL methodology rather than claiming domain-specific advantage. In line with the pat-

terns highlighted for image classification and natural language processing, they primarily probe how hybrid QDL pipelines behave under realistic deployment constraints, rather than establishing domain-specific, scalable quantum advantage.

Despite diverse problem settings, a common architectural pattern recurs. Most works embed a shallow PQC as a learnable nonlinear module, often a feature map or compact classifier head, inside an otherwise classical workflow that provides the main representational capacity. This convergence largely reflects near-term feasibility, including depth and shot budgets, as well as the practicality of integrating PQCs into existing DL toolchains. Consequently, reported benefits are typically finite budget effects, such as parameter reductions, shifts in constant factors, or comparable performance under tight resource ceilings, rather than asymptotic improvements (Bowles *et al.*, 2024).

A recurring limitation in QDL studies is the difficulty of generalizing empirical conclusions across works, as evaluation practices and baselines are not yet standardized and are domain-specific. Many studies use domain-aligned objectives, for example, robustness, latency, privacy, or verification cost, but the resulting comparisons can be sensitive to comparator choice, tuning budgets, and resource accounting. This underscores the need for stronger benchmark discipline in emerging domains: clearly specified evaluation contracts and matched baselines are essential to distinguish genuine quantum utility from artifacts of experimental design. These issues motivate the comparative perspective of Sec. V.B, where empirical QDL results are examined under explicit benchmarking and resource matching protocols.

B. Comparative analysis with classical DL

Cross-domain evidence in Sec. V.A suggests four recurrent evaluation bottlenecks: (i) comparator sensitivity under matched tuning, dominant in vision and language; (ii) end-to-end resource accounting including encoding and sampling, prominent in chemistry and scientific pipelines; (iii) hardware-facing feasibility and verification cost, prominent in control and security; and (iv) structural alignment between the quantum component and the data, often decisive for quantum-native settings.

Motivated by Schuld and Killoran (2022)'s critique of single-goal quantum advantage framings, we translate the explicit-contract specification of Definition 3 (Sec. II.C; see also Sec. III.G.1) into four practical reporting pillars (Table IV). While theoretical separations and fundamental limitations are treated in Sec. III.G, here we focus on experimental controls that most strongly influence empirical conclusions:
(1) *Empirical performance* operationalizes the task and data model $\mathcal{D}$ by specifying how quality is measured under the stated distribution and protocol.
(2) *Resource complexity* operationalizes the end-to-end resource contract $\mathcal{R}$ by specifying cost accounting and budget ceilings.
(3) *Practical feasibility* operationalizes the implementability and validation clauses of Definition 3 by documenting hardware- and verification-facing constraints at the target scale.
(4) *Domain alignment* operationalizes the access-and-readout model $\mathcal{A}$ by checking that the quantum component's structural and interface assumptions match the task's information pathway and the stated comparator class.

We treat these four pillars as jointly necessary conditions for interpreting finite-size quantum-classical comparisons. Passing Pillars 1–3 without Pillar 4 is not, by itself, evidence that an observed gain is attributable to quantum-mechanical resources under the declared $(\mathcal{D}, \mathcal{R}, \mathcal{A})$. Absent controlled scaling under matched budgets, empirical results primarily serve as proof-of-principle demonstrations informing feasibility, robustness, and inductive-bias hypotheses. To maximize interpretability, empirical studies are most interpretable when they report: baseline families and inclusion rationale; preprocessing and feature pipeline on both sides; hyperparameter search space and allocated tuning budgets; a common stopping rule (e.g., time-to-target under $\mathcal{R}$) with learning curves; and negative controls that isolate the quantum block (e.g., parameter-matched classical modules, randomized features, or classical surrogates.) Claims of improvement are conditional on the declared comparator stack and contract $(\mathcal{D}, \mathcal{R}, \mathcal{A})$, interpreted alongside domain-specific bottlenecks identified in Sec. V.A.

1. Empirical performance

Building on the reporting criteria of Sec. IV.D.4, empirical performance asks whether a QDL model improves task-relevant quality under matched evaluation and comparable tuning effort relative to strong classical baselines. On current benchmark suites, outcomes can be highly sensitive to preprocessing choices and hyperparameter optimization, and reported performance often reflects these design decisions as much as architectural differences (Bowles *et al.*, 2024; Moussa *et al.*, 2024). Moreover, the relevant classical comparator set is not static as the DL software-hardware stack evolves (Jouppi *et al.*, 2017). In addition to the elements required by Definition 3, a reproducible comparison benefits from disclosing the full feature pipeline, reporting the hyperparameter-optimization budget used for each model class, and summarizing variability across random seeds and data splits with uncertainty intervals. When efficiency is part of the claim, results are most interpretable when summarized under

Table IV A four-pillar protocol for rigorous quantum-classical comparison in finite-size empirical studies. Claims are evaluated under an explicit access and readout interface and an end-to-end resource contract $\mathcal{R}$ (Definition 3). Mechanisms underlying limitations are treated in Sec. III.G.

Pillar	Reporting axes	Primary failure modes
1. Empirical performance *Does it improve quality under matched evaluation?*	– Task-specific metric and calibration – Learning curves versus data and compute – Robustness across seeds and splits – Sample-efficiency diagnostics	– Under-tuned or outdated baselines – Unmatched hyperparameter optimization – Hidden pre or postprocessing – Selective reporting of best runs
2. Resource complexity *Is it more efficient under end-to-end accounting?*	– Total shots and QPU calls – Wall-clock breakdown and time-to-target – Energy and monetary cost boundary – Compiled depth and two-qubit counts – Classical coprocessing and communication	– Reporting circuit metrics without wall-clock cost – Incommensurate access or readout interfaces or training interfaces (e.g., classical backpropagation) – State-preparation and encoding omitted – Missing classical surrogate checks
3. Practical feasibility *Can it be trained and executed reliably on quantum hardware?*	– Training stability and failure rate – Verification method and its cost – Inference throughput and latency – Compilation overheads	– Optimizer stagnation and seed sensitivity – Shot-noise dominated updates – Unsupported intractability
4. Domain alignment *Is the quantum component structurally matched to the task?*	– Architecture class and symmetry constraints – Modality and interface realism – Structure-sensitive ablations and controls – Inductive-bias diagnostics	– Access-model mismatch to the deployment setting – Baselines that ignore the same structure – Claims of inductive bias without controls

a fixed stopping rule, for example time-to-target quality under the stated contract $\mathcal{R}$. Beyond a single operating point, learning curves versus dataset size and training budget, together with ablations that replace the quantum block by parameter-matched classical components or randomized feature maps, help isolate which aspects of the pipeline contribute materially to observed gains. Recent large-scale benchmarking indicates that, on standard classical datasets, reported advantages are often regime dependent and can diminish under broader tuning, preprocessing control, or dataset coverage (Bowles *et al.*, 2024); see also time-series case studies in (Fellner *et al.*, 2025). Quantum-kernel benchmarking similarly shows sensitivity to classical optimization choices and dataset selection (Alvarez-Estevez, 2025; Schnabel and Roth, 2025). In this light, finite-size improvements on classical benchmarks are often most informatively interpreted as probes of inductive bias, feasibility, and robustness, rather than as evidence for sustained performance separation.

2. Resource complexity

Resource complexity asks whether an observed finite-size improvement persists once end-to-end costs are accounted for under the same access and readout interface. For classical baselines, relevant costs include wall-clock time, energy consumption, monetary cost, memory footprint, and accelerator utilization, all of which can shift rapidly as hardware and software coevolve (Henderson *et al.*, 2020b; Jouppi *et al.*, 2017; Patterson *et al.*, 2021; Schwartz *et al.*, 2020; Strubell *et al.*, 2020). For quantum pipelines, the operational object of comparison is the end-to-end cost to reach a target quality under a fixed contract $\mathcal{R}$, including state preparation and encoding, circuit execution and compilation overheads, measurement, and outer-loop orchestration. For energy accounting, $\mathcal{R}$ typically specifies the system boundary, for instance at device or facility level, and whether cooling and control are included, and the measurement or estimation protocol (Cordier *et al.*, 2025; Lorenz *et al.*, 2025). Laid out at the theory level, energy can itself be treated as a complexity resource; in a query-complexity framework, Meier and Yamasaki (2025) prove an exponential energy-consumption advantage for Simon's problem and propose criteria for experimental demonstration. Empirical resource claims are most naturally interpreted against contemporary classical capabilities rather than static historical baselines, with explicit documentation of the chosen comparator stack. Small-scale empirical comparisons further illustrate the sensitivity of energy-to-solution to the chosen boundary and workload (Desdentado *et al.*, 2024).

A practical accounting decomposes the pipeline into four coupled contributions:
(i) *Input cost* comprises classical preprocessing and the end-to-end latency of embedding data into the quantum model; this is particularly relevant whenever classical feature compression is used to make encoding feasible.
(ii) *Compute cost* captures the per-circuit footprint after compilation, e.g., depth, two-qubit gate counts, and compilation time, ideally anchored to measured (or explicitly justified) execution latency on the target backend.
(iii) *Output cost* accounts for the total number of circuit repetitions used in training and inference, together with estimator details and the achieved statistical uncertainty at the reported operating point; this is often a dominant term in kernel-style estimators and expectation-value training loops (Bowles *et al.*, 2024; Schnabel and Roth, 2025).
(iv) *Loop overhead* aggregates the outer optimization costs, including total QPU calls, optimizer steps, and classical postprocessing and communication, which can dominate wall-clock time in hybrid workflows even when per-circuit execution is fast (Guerreschi and Smelyanskiy, 2017).

Claims of resource advantage are most informative when supported by an end-to-end breakdown of time and energy, and by time-to-target and energy-to-target curves under a fixed stopping rule and declared measurement protocol. Any surrogate baseline used is ideally stated explicitly, and access-model assumptions are typically declared as part of $\mathcal{A}$ and reported alongside the contract $\mathcal{R}$.

3. Practical feasibility

Practical feasibility, applying the verification and hardware constraints of Secs. IV.D.4 and IV.E.1 to the evaluation context, asks whether a QDL architecture can be trained and executed reliably on available hardware, and whether inference can be performed at acceptable cost under the intended deployment contract. This pillar is empirical by design, focusing on what is commonly measured and reported for hardware-facing claims to be interpretable. First, studies that invoke limited classical verifiability or conditional intractability generally benefit from specifying the verification strategy and its cost. This includes the simulator or emulator class used for validation, the observable set being checked, the statistical confidence reported, and the verification budget relative to the claimed advantage. Without such specification, statements about hardness or unverifiability can be difficult to contextualize methodologically. Second, training feasibility is often characterized by stability rather than only best-achieved performance. Empirical papers often report failure rates across seeds, sensitivity to initialization and stopping rules, and a gradient signal proxy at the stated shot budgets, together with time-to-target distributions rather than single-run outcomes. In addition, the training procedure itself is shaped by hardware realism: computational tools commonly used in simulated training of QDL models, most notably backpropagation and related gradient evaluation techniques, are fundamentally incompatible with quantum hardware due to the no-cloning theorem and measurement collapse, meaning gradients are instead estimated through resource-intensive repeated circuit executions or hybrid classical-quantum loops. When training is performed on hardware, drift sensitivity and device-side throughput constraints are often reported in comparable terms, since these effects directly enter the time-to-solution contract. Third, deployment-facing claims generally include inference-time costs. Quantum models typically produce stochastic measurement outcomes; inference therefore incurs a shots-per-prediction cost that is reported alongside throughput and latency. Single-shot or few-shot prediction regimes, when claimed, are most interpretable when accompanied by explicit accuracy–confidence tradeoffs under the stated measurement model (Recio-Armengol *et al.*, 2025b). Feasibility claims that omit inference overhead are correspondingly harder to compare to deterministic classical inference baselines.

4. Domain alignment

Domain alignment asks whether the proposed quantum component is structurally matched to the task and to the data interface, and therefore whether the benchmark is capable of testing a plausible quantum contribution. For classical-data domains, the dominant barriers are often encoding overhead and the strength of pretrained classical baselines; for quantum-native domains, the dominant barriers are access-model realism and measurement budgets. Accordingly, inductive-bias arguments can be viewed as design hypotheses that are validated by structure-sensitive controls and baselines.

Data modality and interface realism form a primary consideration. For high-dimensional classical inputs, hybrid designs and classical feature compression are often required to make encoding feasible, and the cost of this compression enters the end-to-end comparison (Mari *et al.*, 2020). For quantum-native data, coherent interfaces can be decisive; empirical claims in this regime typically state the access and readout model and benchmark against classical comparators restricted to the same measurement interface (Cho and Kim, 2024; Huang *et al.*, 2022; Liu *et al.*, 2025e).

Structure and symmetry provide another lens for alignment. QDL is particularly well motivated when architectures enforce known structure such as locality or group symmetries, restricting the effective hypothesis class to task-relevant functions and often improving stability and

data efficiency in structured regimes (Larocca *et al.*, 2022b; Meyer *et al.*, 2023; Nguyen *et al.*, 2024). In such settings, alignment is commonly established empirically by baselines that incorporate the same structure and by causal ablations that break the purported prior, rather than by comparing to generic unstructured baselines.

Label scarcity and acquisition cost offer a further dimension. In data-rich regimes such as large-scale vision and language with billions of training examples, large pretrained classical models impose a high bar; conversely, in label-expensive scientific regimes such as materials and chemistry, the premium often shifts toward sample efficiency, uncertainty quantification, and robust generalization, making targeted structure and access-model choices more consequential (McArdle *et al.*, 2020).

VI. CONCLUSIONS AND FUTURE DIRECTIONS

A. Summary of key findings

The evolution of QDL is shaped by three coupled tensions that separate finite-size, proof-of-principle improvements from scalable advantage: (i) expressivity-trainability trade-off, (ii) trainability-simulability constraint, and (iii) quantum-classical interface and certification bottleneck. The expressivity-trainability trade-off (Holmes *et al.*, 2022) becomes more diagnostic when recast in physically grounded controllability language: the dynamical Lie algebra generated by an ansatz's implementable generators provides a principled handle on which directions in Hilbert space are actually reachable and on how "effective expressivity" correlates with loss concentration and training costs (Kazi *et al.*, 2025; Ragone *et al.*, 2024). This refines the barren-plateau reasoning: exponentially suppressed gradients are well established for broad classes of global objectives and highly expressive circuit ensembles, yet the onset and severity of gradient suppression depend sensitively on symmetry structure, cost locality, noise, and the measurement procedure used to obtain gradients (Larocca *et al.*, 2025; McClean *et al.*, 2018). Accordingly, trainability functions as a scaling constraint rather than a purely algorithmic inconvenience, and it naturally motivates structured inductive biases, such as symmetry-preserving parameterizations, locality-aware architectures, and localized cost constructions, that reduce concentration without silently changing the learning task (Cerezo *et al.*, 2022).

Mitigating gradient pathologies, however, tightens the second tension: the same structural restrictions that improve trainability can also move a quantum model toward regimes that admit efficient classical simulation. Locality, shallow depth, restricted connectivity and constrained generator sets can suppress concentration effects (Cerezo *et al.*, 2021b; Pesah *et al.*, 2021), but these constraints often align with known simulation or dequantization mechanisms, shrinking the space of candidate quantum advantages (Bermejo *et al.*, 2024; Shin *et al.*, 2024). In this sense, quantum-inspired classical algorithms, including tensor-network models and dequantized linear-algebraic algorithms (Tang, 2019, 2021), are not merely "baselines" but essential foils: they delimit which gains plausibly require device-native quantum resources rather than classical structure exploitation. Credible empirical comparisons therefore hinge on contracts that fix the task specification, access assumptions, and comparator families tightly enough to prevent implicit asymmetries in preprocessing, tuning budgets, or oracle access (Schuld and Killoran, 2022).

The third tension unifies what is often treated as a grab bag of practical issues: data access, readout, and verification are coupled by the quantum-classical interface. For classical data, coherent state preparation and task-aligned preprocessing can dominate end-to-end cost and can change which resource axis is actually limiting (Bowles *et al.*, 2024; Zhang and Yuan, 2024). This observation helps explain why near-term claims are typically most defensible in quantum-native data regimes, where the dominant overheads shift from loading classical information to measurement, control, and stability constraints (Huang *et al.*, 2022). It also clarifies why "verification" cannot be treated as an afterthought: as experiments move beyond classically reproducible regimes, the strongest classical cross-checks degrade, and evaluation increasingly relies on protocol-level validation and statistically meaningful tests under the stated access model, rather than full classical reproduction of outputs. A complementary strategy is to pivot from deterministic data-access pipelines to sampling-centric objectives, treating quantum devices as native samplers for families of distributions that are conjectured to be classically hard to sample under matched access assumptions (Huang *et al.*, 2025a).

The empirical record emphasizes that methodology and systems integration are not peripheral details but part of the scientific content of any advantage claim. Reported performance becomes interpretable only when shot budgets, mitigation choices, optimizer settings, hyperparameter search spaces, reproducibility under drift, and tuned baselines are specified together ,so that conclusions are attributable to the declared contract rather than to hidden degrees of freedom. Error mitigation can enable meaningful finite-size demonstrations, but its bias-variance and sampling-cost tradeoffs typically steepen with circuit depth and target accuracy, suggesting that sustained accuracy and reliability for large-scale, deeply iterative learning loops will increasingly align with fault-tolerant error correction rather than mitigation alone. Consequently, the most robust pathway to practical advantage is to navigate these coupled tensions through explicit, verifiable contracts; to target problem classes in which data interface and readout demands

match realistic access models; and to articulate fault-tolerant roadmaps in which scaling claims are tied to auditable cost accounting and validation procedures.

B. Future directions and roadmap

Progress in QDL requires simultaneous, rather than isolated, advances across four dimensions: (i) hardware scale, encompassing the number, quality, and connectivity of qubits; (ii) error management, spanning strategies from mitigation to full correction; (iii) algorithmic sophistication and trainability, reflecting the depth and inductive bias of deployable QDL architectures; and (iv) verification rigor, addressing the certification of results in regimes beyond classical simulation. Figure 8 maps these dimensions across three operational regimes, explicitly tracking the dominant bottlenecks that constrain end-to-end utility. The core thesis is that expanding qubit counts without commensurate advances in error management and verification limits effective circuit depth to classically reproducible or too shallow regimes. Verification needs to be elevated to a first-class design constraint, parallel to hardware and algorithm design.

Near-term: NISQ hardware and verification-aware hybrid workflows.—The near-term regime is defined by NISQ hardware (Preskill, 2018) with limited depth and throughput, where hybrid quantum-classical models compete with strong quantum-inspired classical baselines. Error management relies on QEM (Cai *et al.*, 2023); however, because QEM techniques can incur large, and in some settings exponential, sampling overheads as a function of circuit depth and/or target precision, their viability is strictly bounded by the total shot budget and device drift. A central methodological challenge is the verification-simulation gap: as experiments enter regimes that are difficult to classically simulate, verification need to be treated as part of the resource contract rather than as an afterthought. Accordingly, near-term priorities include: (i) designing trainability-aware architectures with explicit inductive bias (e.g., locality or symmetry) (Meyer *et al.*, 2023; Ragone *et al.*, 2024); (ii) tightening algorithm-hardware codesign to match ansätze, measurement strategies, and gradient estimators to device topology and calibrated noise (Ji *et al.*, 2025a); and (iii) institutionalizing rigorous, multi-platform benchmarking against tuned classical baselines with disclosed budgets (Bowles *et al.*, 2024). Applications prioritize quantum-native data, such as learning from physical experiments or quantum sensors, bypassing the prohibitive cost of state preparation (Gentile *et al.*, 2021; Guo *et al.*, 2024c; de Leon *et al.*, 2021; McArdle *et al.*, 2020; Merchant *et al.*, 2023; Sajjan *et al.*, 2022). The open question in QDL is whether a practical quantum advantage can be demonstrated and certified under realistic end-to-end constraints, without collapsing into either unverified beyond-simulability behavior or a classically reproducible surrogate.

Mid-term: early fault-tolerant systems and QEC-aware training.—The mid-term horizon targets early fault-tolerant quantum computers, with operational budgets expanding to approximately 10^6 to 10^9 elementary operations. This regime enables deeper circuits and longer training horizons but introduces the QEC overhead and decoding latencies as the primary bottlenecks, including the logical cost ratio (physical-to-logical qubits) and the real-time syndrome-decoding latency (Eisert and Preskill, 2025; Ryan-Anderson *et al.*, 2021). Research imperatives therefore center on (i) realizing low-overhead codes, such as quantum low-density parity-check codes (Breuckmann and Eberhardt, 2021), on hardware with sufficient connectivity (Gottesman, 2013; Komoto and Kasai, 2025; Tillich and Zémor, 2014); (ii) integrating decoding and control into high-throughput hybrid training loops; and (iii) developing QEC-aware training protocols that maintain trainability under residual logical noise by balancing gradient variance against syndrome costs. Theoretically, because worst-case complexity guarantees are often inaccessible, certification may increasingly rely on conditional average-case hardness evidence under carefully stated computational assumptions (Huang *et al.*, 2025b). The central open question is whether a net QDL advantage survives the full fault-tolerant resource contract, including overheads and latency, while remaining verifiable beyond classical reach.

Long-term: FASQ and the frontier of quantum intelligence.—The long-term horizon encompasses fault-tolerant application-scale quantum (FASQ) systems (Preskill, 2025), with operation budgets exceeding 10^9–10^{12} operations. In this regime, error management aims for scalable fault tolerance, suppressing logical error rates below task-dependent tolerances at the cost of substantial overhead rather than eliminating them. This capability could enable the realization of autonomous quantum agents that integrate quantum sensors, quantum memory, and deep networks with minimal classical intermediation. With gate error rates no longer the primary constraint, the bottleneck transitions to data access and state preparation, where QRAM or encoding costs may dominate for classical data. Consequently, long-horizon priorities accordingly emphasize (i) data-frugal QDL protocols operating directly on quantum data or compressed classical encodings; (ii) fault-tolerant trainability primitives, such as parameter-shift-compatible logical gate decompositions and gradient estimators that account for QEC overhead, to enable end-to-end optimization at application scale; (iii) quantitative frameworks for rigorously bounding the complexity of quantum advantage and quantum generalization in hybrid systems, clarifying how quantum models interact with next-generation classical systems. A key open question is whether certain QDL advantages exist that defy classical certification, so

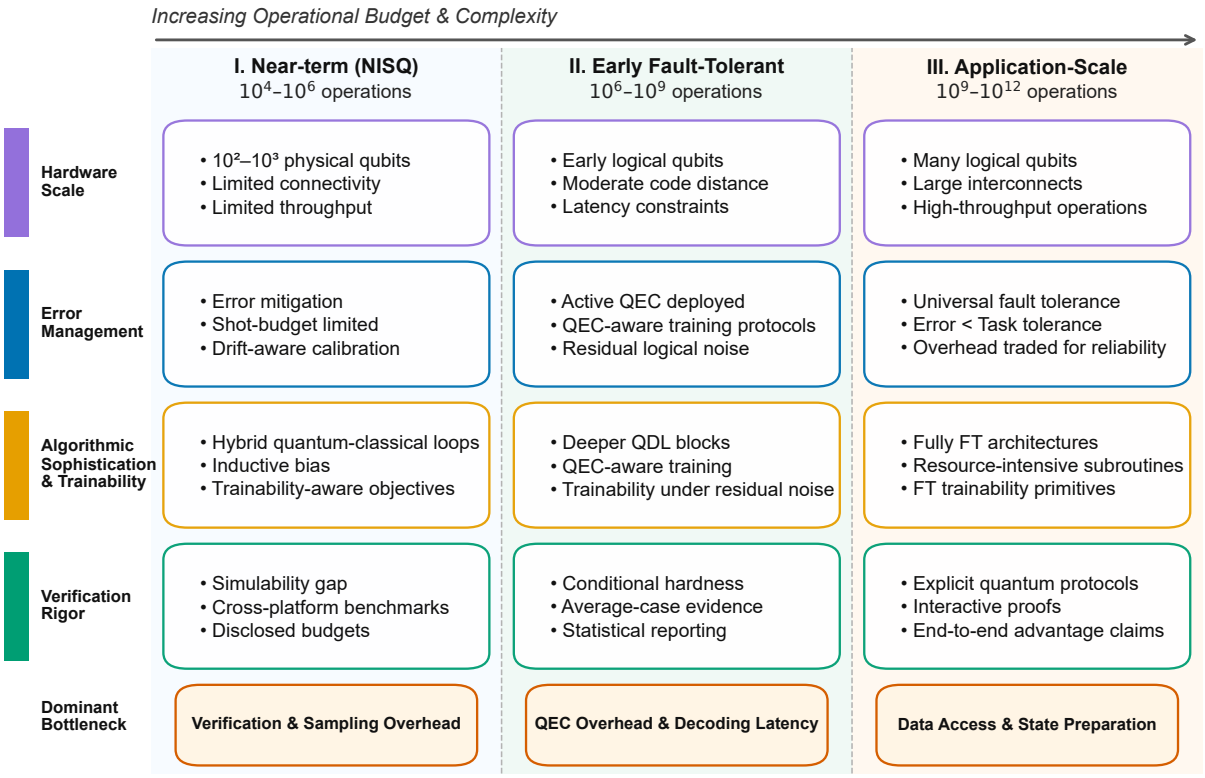

Figure 8 Strategic roadmap for QDL, illustrating the coupled evolution of four dimensions: hardware scale, error management, algorithmic sophistication, and verification rigor. We distinguish three capability regimes by heuristic operation budgets and error-handling assumptions (Eisert and Preskill, 2025; Preskill, 2025). Here, an "elementary operation" means one primitive physical action in a single compiled circuit execution (one QPU call), counting each native single-qubit gate, two-qubit gate, and single-qubit measurement once (summed over all qubits). Under this convention: (i) the near-term NISQ regime spans $\sim 10^4$–10^6 elementary operations per circuit execution, dominated by hybrid training and error mitigation; (ii) an early fault-tolerant regime spans $\sim 10^6$–10^9, enabling deeper circuits and explicitly QEC-aware training protocols; and (iii) an application-scale fault-tolerant regime spans $\sim 10^9$–10^{12}, where fully fault-tolerant quantum deep architectures and resource-intensive algorithmic subroutines become feasible.

that verification would require explicitly quantum protocols under an explicit contract rather than purely classical postprocessing.

Across all three regimes, this roadmap is intended as a constraint-driven guide rather than a prescriptive promise. The defining scientific question is no longer merely whether larger devices will enable more complex models; rather, it is whether we can establish advantage claims that are end-to-end, resource-accounted, and strictly verifiable under realistic access, noise, and measurement assumptions. Achieving this demands coordinated progress across quantum information, learning theory, systems engineering, and those specific application domains where quantum-native data and objectives arise naturally. Over the past decade, the field has transitioned from theoretical speculation into a rigorous discipline characterized by reproducible experiments, sharpened theory, and emerging design principles. The coming era will determine whether these principles cohere into a reliable, operational advantage. This endeavor, at the confluence of quantum computing and artificial intelligence, constitutes a highly ambitious scientific pursuit, one that fundamentally probes the physical limits of computation and the mechanics of intelligence.

ACKNOWLEDGMENTS

The authors thank Lourens van Niekerk and Peter Röseler for insightful suggestions. Y.J. acknowledges support from the programme "Profilbildung 2022", an initiative of the Ministry of Culture and Science of the State of North Rhine-Westphalia, within the framework of the project "Quantum-based Energy Grids (QuGrids)". Z.-Y. Chen acknowledges support from the National Key Research and Development Program of China (Grant No. 2023YFB4502500). M.R. was supported by the Dieter Schwarz foundation through the

Fraunhofer Heilbronn Research and Innovation Centers HNFIZ. O.A. and M.K. acknowledge the support from DLR through funds provided by BMFTR (50WM2347, 50MW2247) and by the State of Berlin within the Pro FIT program under Grant No. 10206829. D.W. acknowledges support from the project Jülich UNified Infrastructure for Quantum computing (JUNIQ) that has received funding from the Federal Ministry of Research, Technology and Space (BMFTR) and the Ministry of Culture and Science of the State of North Rhine-Westphalia. C.S. was supported by the European Union's Horizon 2020 Research and Innovation Programme, grant agreement 101147319 (EBRAINS 2.0 Project), the Helmholtz Association port-folio theme "Supercomputing and Modeling for the Human Brain", and the Helmholtz Association's Initiative and Networking Fund through the Helmholtz International BigBrain Analytics and Learning Laboratory (HIBALL) under the Helmholtz International Lab grant agreement InterLabs-0015, and by HELMHOLTZ IMAGING, a platform of the Helmholtz Information & Data Science Incubator [X-BRAIN, grant number: ZT-I-PF-4-061]. L.M. acknowledges support by the ERDF of the European Union and by "Fonds of the Hamburg Ministry of Science, Research, Equalities and Districts (BWFGB)", as well as funding by the Cluster of Excellence "Advanced Imaging of Matter" (EXC 2056) Project No.390715994.

REFERENCES

Aadhi, A., L. Di Lauro, B. Fischer, P. Dmitriev, I. Alamgir, C. Mazoukh, N. Perron, E. A. Viktorov, A. V. Kovalev, A. Eshaghi, *et al.* (2025), Nat. Commun. **17** (1225), 1225.

Aaronson, S. (2007), Proc. A **463** (2088), 3089.

Aaronson, S. (2015), Nat. Phys. **11** (4), 291.

Aaronson, S. (2018), in *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018 (Association for Computing Machinery, New York, NY, USA) p. 325–338.

Aaronson, S., and D. Gottesman (2004), Phys. Rev. A **70**, 052328.

Abadi, M., P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.* (2016), in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pp. 265–283.

Abbas, A., R. King, H.-Y. Huang, W. J. Huggins, R. Movassagh, D. Gilboa, and J. R. McClean (2023), in *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23 (Curran Associates Inc., Red Hook, NY, USA) pp. 44792–44819.

Abbas, A., D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner (2021), Nat. Comput. Sci. **1** (6), 403.

Abedi, E., S. Beigi, and L. Taghavi (2023), Quantum **7**, 989.

Abel, S., J. C. Criado, and M. Spannowsky (2022), Phys. Rev. A **106**, 022601.

AbuGhanem, M. (2025), The Journal of Supercomputing **81** (5), 687.

Acampora, G., A. Ambainis, N. Ares, L. Banchi, P. Bhardwaj, D. Binosi, G. A. D. Briggs, T. Calarco, V. Dunjko, J. Eisert, *et al.* (2025), "Quantum computing and artificial intelligence: status and perspectives," arXiv:2505.23860.

Acharya, R., D. A. Abanin, L. Aghababaie-Beni, I. Aleiner, T. I. Andersen, M. Ansmann, *et al.* (2025), Nature **638** (8052), 920.

Ackley, D. H., G. E. Hinton, and T. J. Sejnowski (1985), Cognitive Science **9** (1), 147.

Adachi, S. H., and M. P. Henderson (2015), "Application of quantum annealing to training of deep neural networks," arXiv:1510.06356.

Advani, M. S., A. M. Saxe, and H. Sompolinsky (2020), Neural Networks **132**, 428.

Aghaee Rad, H., T. Ainsworth, R. N. Alexander, B. Altieri, M. F. Askarani, R. Baby, L. Banchi, B. Q. Baragiola, J. E. Bourassa, R. S. Chadwick, *et al.* (2025), Nature **638** (8052), 912.

Aharonov, D., and M. Ben-Or (1997), in *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '97 (Association for Computing Machinery, New York, NY, USA) p. 176–188.

Aharonov, D., W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev (2004), in *45th Annual IEEE Symposium on Foundations of Computer Science*, pp. 42–51.

Ajagekar, A., and F. You (2021), Appl. Energy **303**, 117628.

Aktar, S., A. Bärtschi, A.-H. A. Badawy, and S. Eidenbenz (2025), "Quantum graph transformer for nlp sentiment classification," arXiv:2506.07937.

Albash, T., and D. A. Lidar (2018), Rev. Mod. Phys. **90**, 015002.

Alchieri, L., D. Badalotti, P. Bonardi, and S. Bianco (2021), Quantum Mach. Intell. **3** (2), 28.

Alexander, A., and D. Widdows (2022), in *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*, pp. 355–361.

Alexander, K., A. Benyamini, D. Black, D. Bonneau, S. Burgos, B. Burridge, H. Cable, G. Campbell, G. Catalano, A. Ceballos, *et al.* (2025), Nature **641** (8064), 876.

Alexeev, Y., M. H. Farag, T. L. Patti, M. E. Wolf, N. Ares, A. Aspuru-Guzik, S. C. Benjamin, Z. Cai, S. Cao, C. Chamberland, *et al.* (2025), Nat. Commun. **16** (10829), 10829.

Alvarez-Estevez, D. (2025), IEEE Transactions on Quantum Engineering **6**, 1.

Amico, L., R. Fazio, A. Osterloh, and V. Vedral (2008), Rev. Mod. Phys. **80**, 517.

Amin, M. H. (2015), Phys. Rev. A **92**, 052323.

Amin, M. H., E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko (2018), Phys. Rev. X **8**, 021050.

Anai, K., S. Ikehara, Y. Yano, D. Okuno, and S. Takeda (2024), Phys. Rev. A **110**, 022404.

Anschuetz, E. R. (2025), "A unified theory of quantum neural network loss landscapes," arXiv:2408.11901.

Anschuetz, E. R., A. Bauer, B. T. Kiani, and S. Lloyd (2023), Quantum **7**, 1189.

Anschuetz, E. R., and B. T. Kiani (2022), Nat. Commun. **13** (7760), 7760.

Ansere, J. A., E. Gyamfi, V. Sharma, H. Shin, O. A. Dobre, and T. Q. Duong (2024), IEEE Transactions on Wireless Communications **23** (6), 6221.

Anton, O., V. A. Henderson, E. Da Ros, I. Sekulic, S. Burger, P.-I. Schneider, and M. Krutzik (2024), Mach. Learn.: Sci. Technol. **5** (2), 025022.

Arnold, D. V., and H.-G. Beyer (2000), in *Parallel Problem Solving from Nature PPSN VI*, edited by M. Schoenauer, K. Deb, G. Rudolph, X. Yao, E. Lutton, J. J. Merelo, and H.-P. Schwefel (Springer Berlin Heidelberg, Berlin, Heidel-

berg) pp. 39–48.
Arnold, D. V., and H.-G. Beyer (2004), IEEE Transactions on Automatic Control **49** (4), 617.
Arrasmith, A., M. Cerezo, P. Czarnik, L. Cincio, and P. J. Coles (2021), Quantum **5**, 558.
Arrazola, J. M., V. Bergholm, K. Brádler, T. R. Bromley, M. J. Collins, I. Dhand, A. Fumagalli, T. Gerrits, A. Goussev, L. G. Helt, *et al.* (2021), Nature **591** (7848), 54.
Arrazola, J. M., A. Delgado, B. R. Bardhan, and S. Lloyd (2020), Quantum **4**, 307.
Arunachalam, S., and R. De Wolf (2018), J. Mach. Learn. Res. **19** (1), 2879–2878.
Arunachalam, S., A. B. Grilo, and H. Yuen (2020), "Quantum statistical query learning," arXiv:2002.08240.
Arunachalam, S., V. Havlíček, and L. Schatzki (2023), in *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23 (Curran Associates Inc., Red Hook, NY, USA).
Arunachalam, S., and R. de Wolf (2017), SIGACT News **48** (2), 41–67.
Arute, F., K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, *et al.* (2019), Nature **574** (7779), 505.
Atatüre, M., D. Englund, N. Vamivakas, S.-Y. Lee, and J. Wrachtrup (2018), Nat. Rev. Mater. **3** (5), 38.
Auger, A., and N. Hansen (2005), in *2005 IEEE congress on evolutionary computation*, Vol. 2 (IEEE) pp. 1769–1776.
Awschalom, D., K. K. Berggren, H. Bernien, S. Bhave, L. D. Carr, P. Davids, S. E. Economou, D. Englund, A. Faraon, M. Fejer, *et al.* (2021), PRX Quantum **2**, 017002.
Azevedo, V., C. Silva, and I. Dutra (2022), Quantum Mach. Intell. **4** (1), 5.
Babbush, R., R. King, S. Boixo, W. Huggins, T. Khattar, G. H. Low, J. R. McClean, T. O'Brien, and N. C. Rubin (2025), "The grand challenge of quantum applications," arXiv:2511.09124.
Bahri, Y., E. Dyer, J. Kaplan, J. Lee, and U. Sharma (2024), Proc. Natl. Acad. Sci. U.S.A. **121** (27), e2311878121.
Bahri, Y., J. Kadmon, J. Pennington, S. S. Schoenholz, J. Sohl-Dickstein, and S. Ganguli (2020), Annual review of condensed matter physics **11** (1), 501.
Bakshi, K., and K. Srinivasan (2025), Scientific Reports **15** (1), 28711.
Balewski, J., M. G. Amankwah, R. Van Beeumen, E. W. Bethel, T. Perciano, and D. Camps (2024), Scientific Reports **14** (1), 3435.
Banchi, L., and G. E. Crooks (2021), Quantum **5**, 386.
Banchi, L., J. Pereira, and S. Pirandola (2021), PRX Quantum **2**, 040321.
Bangar, S., L. Sunny, K. Yeter-Aydeniz, and G. Siopsis (2025), Quantum Mach. Intell. **7** (1), 33.
Barenco, A., C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter (1995), Phys. Rev. A **52**, 3457.
Barends, R., and F. K. Wilhelm (2025), "Performance-centric roadmap for building a superconducting quantum computer," arXiv:2506.23178.
Bartee, S. K., W. Gilbert, K. Zuo, K. Das, T. Tanttu, C. H. Yang, N. Dumoulin Stuyck, S. J. Pauka, R. Y. Su, W. H. Lim, *et al.* (2025), Nature **643** (8071), 382.
Barthe, A., and A. Pérez-Salinas (2024), Quantum **8**, 1523.
Bartolucci, S., P. Birchall, H. Bombín, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, *et al.* (2023), Nat. Commun. **14** (912), 912.
Basani, J. R., M. Y. Niu, and E. Waks (2025), npj Quantum Inf. **11** (142), 142.
Basheer, A., Y. Feng, C. Ferrie, S. Li, and H. Pashayan (2025), AAAI **39** (15), 15498.
Basilewitsch, D., J. F. Bravo, C. Tutschku, and F. Struckmeier (2025), Quantum Mach. Intell. **7** (2), 110.
Bassey, J., L. Qian, and X. Li (2021), "A survey of complex-valued neural networks," arXiv:2101.12249.
Battaglia, P. W., J. B. Hamrick, V. Bapst, A. Sanchez-Gonzalez, V. Zambaldi, M. Malinowski, A. Tacchetti, D. Raposo, A. Santoro, R. Faulkner, *et al.* (2018), "Relational inductive biases, deep learning, and graph networks," arXiv:1806.01261.
Batzner, S., A. Musaelian, L. Sun, M. Geiger, J. P. Mailoa, M. Kornbluth, N. Molinari, T. E. Smidt, and B. Kozinsky (2022), Nat. Commun. **13** (2453), 2453.
Baum, Y., M. Amico, S. Howell, M. Hush, M. Liuzzi, P. Mundada, T. Merkh, A. R. Carvalho, and M. J. Biercuk (2021), PRX Quantum **2**, 040324.
Bausch, J. (2020), in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20 (Curran Associates Inc., Red Hook, NY, USA).
Bausch, J., A. W. Senior, F. J. H. Heras, T. Edlich, A. Davies, M. Newman, C. Jones, K. Satzinger, M. Y. Niu, S. Blackwell, *et al.* (2024), Nature **635** (8040), 834.
Baydin, A. G., B. A. Pearlmutter, A. A. Radul, and J. M. Siskind (2018), Journal of machine learning research **18** (153), 1.
Beer, K., D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann, D. Scheiermann, and R. Wolf (2020), Nat. Commun. **11** (808), 808.
Belkin, M., D. Hsu, S. Ma, and S. Mandal (2019), Proc. Natl. Acad. Sci. U.S.A. **116** (32), 15849.
Benedetti, M., E. Lloyd, S. Sack, and M. Fiorentini (2019), Quantum Sci. Technol. **4** (4), 043001.
Benedetti, M., J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz (2017), Phys. Rev. X **7**, 041052.
Benedetti, M., J. Realpe-Gómez, and A. Perdomo-Ortiz (2018), Quantum Sci. Technol. **3** (3), 034007.
Bengua, J. A., H. N. Phien, and H. D. Tuan (2015), in *2015 IEEE International Congress on Big Data* (IEEE) pp. 669–672.
Benioff, P. (1980), Journal of statistical physics **22** (5), 563.
Benítez-Buenache, A., and Q. Portell-Montserrat (2025), Quantum Mach. Intell. **7** (2), 85.
Benito, C., A. R. Vasquez, J. Home, K. K. Mehta, T. Monz, M. Müller, and A. Bermudez (2025), "Scaling roadmap for modular trapped-ion qec and lattice-surgery teleportation," arXiv:2512.20435.
Berezutskii, A., M. Liu, A. Acharya, R. Ellerbrock, J. Gray, R. Haghshenas, Z. He, A. Khan, V. Kuzmin, D. Lyakh, *et al.* (2025), Nat. Rev. Phys. **7** (10), 581.
Bergholm, V., J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. Akash-Narayanan, A. Asadi, *et al.* (2022), "Pennylane: Automatic differentiation of hybrid quantum-classical computations," arXiv:1811.04968.
Bermejo, P., P. Braccia, M. S. Rudolph, Z. Holmes, L. Cincio, and M. Cerezo (2024), "Quantum convolutional neural networks are (effectively) classically simulable," arXiv:2408.12739.
Berry, D. W., Y. Su, C. Gyurik, R. King, J. Basso, A. D. T. Barba, A. Rajput, N. Wiebe, V. Dunjko, and R. Babbush (2024), PRX Quantum **5**, 010319.

Beyer, H.-G., and B. Sendhoff (2006), in *2006 IEEE International Conference on Evolutionary Computation*, pp. 1346–1353.

Beyer, H.-G., and B. Sendhoff (2007), Computer Methods in Applied Mechanics and Engineering **196** (33), 3190.

Bhatia, A. S., S. Kais, and M. A. Alam (2023), Quantum Sci. Technol. **8** (4), 045032.

Biamonte, J., P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd (2017), Nature **549** (7671), 195.

Bian, H., Z. Jia, M. Dou, Y. Fang, L. Li, Y. Zhao, H. Wang, Z. Zhou, W. Wang, W. Zhu, *et al.* (2023), "Vqnet 2.0: A new generation machine learning framework that unifies classical and quantum," arXiv:2301.03251.

Bian, K., S. Zhang, F. Meng, W. Zhang, and O. Dahlsten (2024), Phys. Rev. A **110**, 022406.

Blais, A., A. L. Grimsmo, S. M. Girvin, and A. Wallraff (2021), Rev. Mod. Phys. **93**, 025005.

Bloch, F. (1946), Phys. Rev. **70**, 460.

Bluvstein, D., S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, *et al.* (2024), Nature **626** (7997), 58.

Bluvstein, D., A. A. Geim, S. H. Li, S. J. Evered, J. P. Bonilla Ataides, G. Baranes, A. Gu, T. Manovitz, M. Xu, M. Kalinowski, *et al.* (2026), Nature **649** (8095), 39.

Bommasani, R., D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill, *et al.* (2022), "On the opportunities and risks of foundation models," arXiv:2108.07258.

Bonet-Monroig, X., H. Wang, D. Vermetten, B. Senjean, C. Moussa, T. Bäck, V. Dunjko, and T. E. O'Brien (2023), Phys. Rev. A **107**, 032407.

Bonyadi, M. R., and Z. Michalewicz (2017), Evolutionary Computation **25** (1), 1.

Bordelon, B., A. Atanasov, and C. Pehlevan (2025), J. Stat. Mech.: Theory Exp. **2025** (8), 084002.

Bowles, J., S. Ahmed, and M. Schuld (2024), "Better than classical? the subtle art of benchmarking quantum machine learning models," 2403.07059.

Bowles, J., D. Wierichs, and C.-Y. Park (2025), Quantum **9**, 1873.

Bradbury, J., R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, *et al.* (2018), .

Bradshaw, Z. P., E. N. Evans, M. Cook, and M. L. LaBorde (2025), Phys. Rev. Appl. **23**, 044007.

Brandão, F. G. S. L., A. W. Harrow, and M. Horodecki (2016), Commun. Math. Phys. **346** (2), 397.

Bravo, R. A., K. Najafi, X. Gao, and S. F. Yelin (2022), PRX Quantum **3**, 030325.

Bravyi, S., D. Gosset, and R. König (2018), Science **362** (6412), 308.

Bravyi, S., and A. Kitaev (2005), Phys. Rev. A **71**, 022316.

Breuckmann, N. P., and J. N. Eberhardt (2021), PRX Quantum **2**, 040101.

Broers, L., and L. Mathey (2024), Phys. Rev. Res. **6**, 013076.

Bronstein, M. M., J. Bruna, T. Cohen, and P. Veličković (2021), "Geometric deep learning: Grids, groups, graphs, geodesics, and gauges," arXiv:2104.13478.

Broughton, M., G. Verdon, T. McCourt, A. J. Martinez, J. H. Yoo, S. V. Isakov, P. Massey, R. Halavati, M. Y. Niu, A. Zlokapa, *et al.* (2021), "Tensorflow quantum: A software framework for quantum machine learning," arXiv:2003.02989.

Bruzewicz, C. D., J. Chiaverini, R. McConnell, and J. M. Sage (2019), Applied Physics Reviews **6** (2), 021314.

Bu, J.-T., L. Zhang, Z. Yu, J.-B. Wang, W.-Q. Ding, W.-F. Yuan, B. Wang, H.-J. Du, W.-J. Chen, L. Chen, *et al.* (2025), Phys. Rev. Appl. **23**, 034073.

Buitinck, L., G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, *et al.* (2013), "Api design for machine learning software: experiences from the scikit-learn project," arXiv:1309.0238.

Bukov, M., A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta (2018), Phys. Rev. X **8**, 031086.

Buonaiuto, G., F. Gargiulo, G. De Pietro, M. Esposito, and M. Pota (2024), Quantum Mach. Intell. **6** (1), 9.

Burgholzer, L., J. Echavarria, P. Hopf, Y. Stade, D. Rovara, L. Schmid, E. Kaya, B. Mete, M. N. Farooqi, M. Chung, *et al.* (2025a), "The munich quantum software stack: Connecting end users, integrating diverse quantum technologies, accelerating hpc," arXiv:2509.02674.

Burgholzer, L., Y. Stade, T. Peham, and R. Wille (2025b), Journal of Open Source Software **10** (108), 7478.

Burkard, G., T. D. Ladd, A. Pan, J. M. Nichol, and J. R. Petta (2023), Rev. Mod. Phys. **95**, 025003.

Butt, F., I. Pogorelov, R. Freund, A. Steiner, M. Meyer, T. Monz, and M. Müller (2026), Nat. Commun. **17** (995), 995.

Cai, X.-D., D. Wu, Z.-E. Su, M.-C. Chen, X.-L. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan (2015), Phys. Rev. Lett. **114**, 110504.

Cai, Z., R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien (2023), Rev. Mod. Phys. **95**, 045005.

Camenzind, L. C., S. Geyer, A. Fuhrer, R. J. Warburton, D. M. Zumbühl, and A. V. Kuhlmann (2022), Nat. Electron. **5** (3), 178.

Carleo, G., and M. Troyer (2017), Science **355** (6325), 602.

Caro, M. C., E. Gil-Fuster, J. J. Meyer, J. Eisert, and R. Sweke (2021), Quantum **5**, 582.

Caro, M. C., H.-Y. Huang, M. Cerezo, K. Sharma, A. Sornborger, L. Cincio, and P. J. Coles (2022), Nat. Commun. **13** (4919), 4919.

Carrasquilla, J., G. Torlai, R. G. Melko, and L. Aolita (2019), Nature Machine Intelligence **1** (3), 155–161.

Cauchy, A., *et al.* (1847), Comp. Rend. Sci. Paris **25** (1847), 536.

Cavallaro, G., D. Willsch, M. Willsch, K. Michielsen, and M. Riedel (2020), in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1973–1976.

Cerezo, M., A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, *et al.* (2021a), Nat. Rev. Phys. **3** (9), 625.

Cerezo, M., M. Larocca, D. García-Martín, N. L. Diaz, P. Braccia, E. Fontana, M. S. Rudolph, P. Bermejo, A. Ijaz, S. Thanasilp, *et al.* (2025), Nat. Commun. **16** (7907), 7907.

Cerezo, M., A. Sone, T. Volkoff, L. Cincio, and P. J. Coles (2021b), Nat. Commun. **12** (1791), 1791.

Cerezo, M., G. Verdon, H.-Y. Huang, L. Cincio, and P. J. Coles (2022), Nat. Comput. Sci. **2** (9), 567.

Chancellor, N., P. J. D. Crowley, T. Durić, W. Vinci, M. H. Amin, A. G. Green, P. A. Warburton, and G. Aeppli (2022), npj Quantum Inf. **8** (73), 73.

Chang, S. Y., and M. Cerezo (2025), "A primer on quantum machine learning," arXiv:2511.15969.

Chen, C.-S., and E.-J. Kuo (2025a), "Quantum adaptive self-attention for quantum transformer models," arXiv:2504.05336.

Chen, C.-S., and E.-J. Kuo (2025b), "Quantum-enhanced natural language generation: A multi-model framework with hybrid quantum-classical architectures," arXiv:2508.21332.

Chen, F., Q. Zhao, L. Feng, C. Chen, Y. Lin, and J. Lin (2025), Neural Networks **185**, 107123.

Chen, J., S. Cheng, H. Xie, L. Wang, and T. Xiang (2018), Phys. Rev. B **97**, 085104.

Chen, J.-S., E. Nielsen, M. Ebert, V. Inlek, K. Wright, V. Chaplin, A. Maksymov, E. Páez, A. Poudel, P. Maunz, and J. Gamble (2024a), Quantum **8**, 1516.

Chen, L., T. Li, Y. Chen, X. Chen, M. Wozniak, N. Xiong, and W. Liang (2024b), Connection Science .

Chen, S., J. Cotler, H.-Y. Huang, and J. Li (2022a), in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 574–585.

Chen, S., J. Cotler, H.-Y. Huang, and J. Li (2023a), Nat. Commun. **14** (6001), 6001.

Chen, S., W. Gong, and Q. Ye (2024c), in *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 1086–1105.

Chen, S. Y.-C., C.-M. Huang, C.-W. Hsing, H.-S. Goan, and Y.-J. Kao (2022b), Mach. Learn.: Sci. Technol. **3** (1), 015025.

Chen, S. Y.-C., T.-C. Wei, C. Zhang, H. Yu, and S. Yoo (2021), "Hybrid quantum-classical graph convolutional network," arXiv:2101.06189.

Chen, S. Y.-C., T.-C. Wei, C. Zhang, H. Yu, and S. Yoo (2022c), Phys. Rev. Res. **4**, 013231.

Chen, S. Y.-C., C.-H. H. Yang, J. Qi, P.-Y. Chen, X. Ma, and H.-S. Goan (2020a), IEEE Access **8**, 141007.

Chen, T., S. Kornblith, M. Norouzi, and G. Hinton (2020b), in *Proceedings of the 37th International Conference on Machine Learning (ICML)*, pp. 1597–1607.

Chen, Y. (2022), IEEE Access **10**, 20240.

Chen, Y., Y. Pan, and D. Dong (2023b), IEEE Transactions on Cybernetics **53** (6), 3467–3478.

Chen, Z.-Y., T.-Y. Ma, C.-C. Ye, L. Xu, W. Bai, L. Zhou, M.-Y. Tan, X.-N. Zhuang, X.-F. Xu, Y.-J. Wang, *et al.* (2024d), Comput. Methods Appl. Mech. Eng. **432**, 117428.

Chen, Z.-Y., C. Xue, S.-M. Chen, and G.-P. Guo (2019), "Vqnet: Library for a quantum-classical hybrid neural network," arXiv:1901.09133.

Cheng, S., L. Wang, and P. Zhang (2021), Phys. Rev. B **103**, 125117.

Cherrat, E. A., I. Kerenidis, N. Mathur, J. Landman, M. Strahm, and Y. Y. Li (2024), Quantum **8**, 1265.

Cherrat, E. A., S. Raj, I. Kerenidis, A. Shekhar, B. Wood, J. Dee, S. Chakrabarti, R. Chen, D. Herman, S. Hu, P. Minssen, R. Shaydulin, Y. Sun, R. Yalovetzky, and M. Pistoia (2023), Quantum **7**, 1191.

Chia, N.-H., A. P. Gilyén, T. Li, H.-H. Lin, E. Tang, and C. Wang (2022), Journal of the ACM **69** (5), 1.

Childs, A. M., Y. Su, M. C. Tran, N. Wiebe, and S. Zhu (2021), Phys. Rev. X **11**, 011020.

Chinzei, K., Q. H. Tran, Y. Endo, and H. Oshima (2024), "Resource-efficient equivariant quantum convolutional neural networks," arXiv:2410.01252.

Chinzei, K., S. Yamano, Q. H. Tran, Y. Endo, and H. Oshima (2025), npj Quantum Inf. **11** (79), 79.

Chittem, M. B., and A. Kumar Pradhan (2025), in *2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS)*, pp. 84–89.

Chiu, N.-C., E. C. Trapp, J. Guo, M. H. Abobeih, L. M. Stewart, S. Hollerith, P. L. Stroganov, M. Kalinowski, A. A. Geim, S. J. Evered, *et al.* (2025), Nature **646** (8087), 1075.

Chizat, L., E. Oyallon, and F. Bach (2019), *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 10.5555/3454287.3454551.

Cho, G., and D. Kim (2024), Nat. Commun. **15** (7552), 7552.

Chow, J. C. L. (2025), Algorithms **18** (3), 156.

Chowdhary, K. R. (2020), *Fundamentals of Artificial Intelligence*, SpringerLink , 603.

Chrisley, R. (1995), in *New directions in cognitive science: Proceedings of the international symposium, Saariselka*, Vol. 4.

Cichocki, A., N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic (2016), Foundations and Trends in Machine Learning **9** (4-5), 249.

Cirac, J. I., D. Pérez-García, N. Schuch, and F. Verstraete (2021), Rev. Mod. Phys. **93**, 045003.

Cirac, J. I., and P. Zoller (1995), Phys. Rev. Lett. **74**, 4091.

Clarke, J., and F. K. Wilhelm (2008), Nature **453** (7198), 1031.

Clayton, C., J. Leng, G. Yang, Y.-L. Qiao, M. C. Lin, and X. Wu (2024), in *Proceedings of the 38th International Conference on Neural Information Processing Systems*, NIPS '24 (Curran Associates Inc., Red Hook, NY, USA).

Clements, W. R., P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley (2016), Optica **3** (12), 1460.

Coecke, B., G. de Felice, K. Meichanetzidis, and A. Toumi (2020), "Foundations for near-term quantum natural language processing," arXiv:2012.03755.

Coecke, B., M. Sadrzadeh, and S. Clark (2010), "Mathematical foundations for a compositional distributional model of meaning," arXiv:1003.4394.

Cohen, N., O. Sharir, and A. Shashua (2016), "On the expressive power of deep learning: A tensor analysis," arXiv:1509.05009.

Cohen, T., and M. Welling (2016), in *International Conference on Machine Learning* (PMLR) pp. 2990–2999.

Cong, I., S. Choi, and M. D. Lukin (2019), Nat. Phys. **15** (12), 1273.

Coopmans, L., and M. Benedetti (2024), Communications Physics **7** (1), 274.

Cordier, S., K. Thibault, M.-L. Arpin, and B. Amor (2025), Quantum Sci. Technol. **10** (2), 025058.

Correll, R., S. J. Weinberg, F. Sanches, T. Ide, and T. Suzuki (2023), Adv. Quantum Technol. **6** (7), 2200183.

Cotler, J., H.-Y. Huang, and J. R. McClean (2021), "Revisiting dequantization and quantum advantage in learning tasks," arXiv:2112.00811.

Coyle, B., S. Raj, N. Mathur, E. A. Cherrat, N. Jain, S. Kazdaghli, and I. Kerenidis (2025), npj Quantum Inf. **11** (172), 172.

Curry, H. B. (1944), Quarterly of Applied Mathematics **2** (3), 258.

Cybenko, G. (1989), Math. Control Signals Systems **2** (4), 303.

Dallaire-Demers, P.-L., and N. Killoran (2018), Phys. Rev. A **98**, 012324.

Dangwal, S., S. Vittal, L. M. Seifert, F. T. Chong, and G. S. Ravi (2025), in *Proceedings of the 52nd Annual International Symposium on Computer Architecture*, ISCA '25 (Association for Computing Machinery, New York, NY,

USA) p. 1417–1431.
Dborin, J., F. Barratt, V. Wimalaweera, L. Wright, and A. G. Green (2022), Quantum Sci. Technol. **7** (3), 035014.
De Lorenzis, A., M. Casado, M. Estarellas, N. Lo Gullo, T. Lux, F. Plastina, A. Riera, and J. Settino (2025), Phys. Rev. Appl. **23**, 044024.
De Palma, G., T. Klein, and D. Pastorello (2024), J. Math. Phys. **65** (9), 092201.
De Palma, G., M. Marvian, C. Rouzé, and D. S. França (2023), PRX Quantum **4**, 010309.
De Raedt, H., F. Jin, D. Willsch, M. Willsch, N. Yoshioka, N. Ito, S. Yuan, and K. Michielsen (2019), Comput. Phys. Commun. **237**, 47 .
De Raedt, H., J. Kraus, A. Herten, V. Mehta, M. Bode, M. Hrywniak, K. Michielsen, and T. Lippert (2025), "Universal quantum simulation of 50 qubits on europes first exascale supercomputer harnessing its heterogeneous cpu-gpu architecture," arXiv:2511.03359.
De Smet, M., Y. Matsumoto, A.-M. J. Zwerver, L. Tryputen, S. L. de Snoo, S. V. Amitonov, S. R. Katiraee-Far, A. Sammak, N. Samkharadze, Ö. Gül, *et al.* (2025), Nat. Nanotechnol. **20** (7), 866.
Dell'Anna, F., R. Gómez-Lurbe, A. Pérez, and E. Ercolessi (2025), Phys. Rev. A **112**, 022612.
Deng, D.-L., X. Li, and S. Das Sarma (2017), Phys. Rev. X **7**, 021021.
Deng, J., W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei (2009), in *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248–255.
Desdentado, E., C. Calero, M. Á. Moraga, M. Serrano, and F. García (2024), Journal of Systems and Software **217**, 112165.
Deshpande, A., B. Fefferman, S. Ghosh, M. Gullans, and D. Hangleiter (2025), "Peaked quantum advantage using error correction," arXiv:2510.05262.
Deshpande, A., M. Hinsche, K. Najafi, K. Sharma, R. Sweke, and C. Zoufal (2024), "Dynamic parameterized quantum circuits: expressive and barren-plateau free," arXiv:2411.05760.
Deutsch, D. (1985), Proc. A. **400** (1818), 97.
Developers, C. (2025), "Cirq," .
Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2019), "Bert: Pre-training of deep bidirectional transformers for language understanding," arXiv:1810.04805.
Devoret, M. H., J. M. Martinis, and J. Clarke (1985), Phys. Rev. Lett. **55**, 1908.
Di Meglio, A., K. Jansen, I. Tavernelli, C. Alexandrou, S. Arunachalam, C. W. Bauer, K. Borras, S. Carrazza, A. Crippa, V. Croft, *et al.* (2024), PRX Quantum **5**, 037001.
Dieks, D. (1982), Phys. Lett. A **92** (6), 271.
Dirac, P. A. M. (1930), *The principles of quantum mechanics*, 27 (Oxford university press).
DiVincenzo, D. P. (1995), Phys. Rev. A **51**, 1015.
DiVincenzo, D. P. (2000), Fortschr. Phys. **48** (9-11), 771.
Djellabi, M., M. Hecker, and S. Acheche (2025), "Attributed-graphs kernel implementation using local detuning of neutral-atoms rydberg hamiltonian," arXiv:2509.09421.
Dong, D., C. Chen, H. Li, and T.-J. Tarn (2008), IEEE Trans. Syst. Man Cybern. Part B Cybern. **38** (5), 1207.
Dongfen, L., Y. Yuan, Z. Hu, Y. Sun, and Q. Xiang (2025), Phys. Scr. **100** (7), 075117.
Dou, M., T. Zou, Y. Fang, J. Wang, D. Zhao, L. Yu, B. Chen, W. Guo, Y. Li, Z. Chen, *et al.* (2022), "Qpanda: high-performance quantum computing framework for multiple application scenarios," arXiv:2212.14201.
Dou, T., G. Zhang, and W. Cui (2023), J. Franklin Inst. **360** (11), 7438.
Du, Y., M.-H. Hsieh, T. Liu, and D. Tao (2020), Phys. Rev. Res. **2**, 033125.
Du, Y., M.-H. Hsieh, T. Liu, S. You, and D. Tao (2021), PRX Quantum **2**, 040337.
Du, Y., T. Huang, S. You, M.-H. Hsieh, and D. Tao (2022), npj Quantum Inf. **8** (62), 62.
Du, Y., X. Wang, N. Guo, Z. Yu, Y. Qian, K. Zhang, M.-H. Hsieh, P. Rebentrost, and D. Tao (2025), "Quantum machine learning: A hands-on tutorial for machine learning practitioners and researchers," arXiv:2502.01146.
Duneau, T., S. Bruhn, G. Matos, T. Laakkonen, K. Saiti, A. Pearson, K. Meichanetzidis, and B. Coecke (2024), "Scalable and interpretable quantum natural language processing: an implementation on trapped ions," arXiv:2409.08777.
Dunjko, V., and H. J. Briegel (2018), Rep. Prog. Phys. **81** (7), 074001.
Dunjko, V., J. M. Taylor, and H. J. Briegel (2016), Phys. Rev. Lett. **117**, 130501.
Edlbauer, H., J. Wang, A. M. S.-E. Huq, I. Thorvaldson, M. T. Jones, S. H. Misha, W. J. Pappas, C. M. Moehle, Y.-L. Hsueh, H. Bornemann, *et al.* (2025), Nature **648** (8094), 569.
Eisert, J. (2021), Phys. Rev. Lett. **127**, 020501.
Eisert, J., M. Cramer, and M. B. Plenio (2010), Rev. Mod. Phys. **82**, 277.
Eisert, J., and J. Preskill (2025), "Mind the gaps: The fraught road to quantum advantage," arXiv:2510.19928.
Endo, S., Z. Cai, S. C. Benjamin, and X. Yuan (2021), J. Phys. Soc. Jpn. **90** (3), 032001.
Evenbly, G. (2018), Phys. Rev. B **98**, 085155.
Evenbly, G., and G. Vidal (2009), Phys. Rev. B **79**, 144108.
Evered, S. J., D. Bluvstein, M. Kalinowski, S. Ebadi, T. Manovitz, H. Zhou, S. H. Li, A. A. Geim, T. T. Wang, N. Maskara, *et al.* (2023), Nature **622** (7982), 268.
Faehrmann, P. K., M. Steudtner, R. Kueng, M. Kieferova, and J. Eisert (2022), Quantum **6**, 806.
Fan, F., Y. Shi, T. Guggemos, and X. X. Zhu (2024a), IEEE Transactions on Neural Networks and Learning Systems **35** (12), 18145.
Fan, Z., J. Zhang, P. Zhang, Q. Lin, Y. Li, and Y. Qian (2024b), Information Processing & Management **61** (4), 103741.
Fanucchi, E., G. Forghieri, A. Secchi, P. Bordone, and F. Troiani (2025), Phys. Rev. B **111**, 205409.
Farhi, E., and H. Neven (2018), "Classification with quantum neural networks on near term processors," arXiv:1802.06002.
Fellner, T., D. Kreplin, S. Tovey, and C. Holm (2025), "Quantum vs. classical: A comprehensive benchmark study for predicting time series with variational quantum machine learning," arXiv:2504.12416.
Felser, T., M. Trenti, L. Sestini, A. Gianelle, D. Zuliani, D. Lucchesi, and S. Montangero (2021), npj Quantum Inf. **7** (111), 111.
Feynman, R. P. (1982), International Journal of Theoretical Physics **21** (6/7), 467.
Feynman, R. P. (1985), Optics News **11** (2), 11.
Fischer, J., Y. Herrmann, C. F. J. Wolfs, S. Scheijen, M. Ruf,

and R. Hanson (2025), Nat. Commun. **16** (11680), 11680.
Fitzek, D., R. S. Jonsson, W. Dobrautz, and C. Schäfer (2024), Quantum **8**, 1313.
Fowler, A. G., M. Mariantoni, J. M. Martinis, and A. N. Cleland (2012), Phys. Rev. A **86**, 032324.
Franz, M., L. Wolf, M. Periyasamy, C. Ufrecht, D. D. Scherer, A. Plinge, C. Mutschler, and W. Mauerer (2023), J. Franklin Inst. **360** (17), 13822.
Frazier, P. I. (2018), in *Recent advances in optimization and modeling of contemporary problems* (Informs) pp. 255–278.
Friedrich, L., and J. Maziero (2023), Scientific reports **13** (1), 9978.
Fujii, K., and K. Nakajima (2017), Phys. Rev. Appl. **8**, 024030.
Fukushima, K. (1980), Biological cybernetics **36** (4), 193.
Ganjefar, S., and M. Tofighi (2018), Neurocomputing **291**, 175.
Gao, X., and L.-M. Duan (2017), Nat. Commun. **8** (662), 662.
Gao, Y., Z. Song, X. Yang, and R. Zhang (2023), "Fast quantum algorithm for attention computation," arXiv:2307.08045.
García-Beni, J., G. L. Giorgi, M. C. Soriano, and R. Zambrini (2023), Phys. Rev. Appl. **20**, 014051.
García-Martín, D., M. Larocca, and M. Cerezo (2024), Phys. Rev. Res. **6**, 013295.
García-Martín, D., M. Larocca, and M. Cerezo (2025), Nat. Phys. **21** (7), 1153.
Gardas, B., M. M. Rams, and J. Dziarmaga (2018), Phys. Rev. B **98**, 184304.
Garnett, R. (2023), *Bayesian Optimization* (Cambridge University Press).
Gaudet, C. J., and A. S. Maida (2018), in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8.
Gebhart, V., R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Banchi, F. Marquardt, L. Pezzè, and C. Bonato (2023), Nat. Rev. Phys. **5** (3), 141.
Gentile, A. A., B. Flynn, S. Knauer, N. Wiebe, S. Paesani, C. E. Granade, J. G. Rarity, R. Santagati, and A. Laing (2021), Nat. Phys. **17** (7), 837.
Gentinetta, G., D. Sutter, C. Zoufal, B. Fuller, and S. Woerner (2023), in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01, pp. 256–262.
George, H. C., M. T. Madzik, E. M. Henry, A. J. Wagner, M. M. Islam, F. Borjans, E. J. Connors, J. Corrigan, M. Curry, M. K. Harper, *et al.* (2024), Nano Lett. 10.1021/acs.nanolett.4c05205.
Gharibian, S., and F. Le Gall (2022), in *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2022 (Association for Computing Machinery, New York, NY, USA) p. 19–32.
Gharibyan, H., H. Karapetyan, T. Sedrakyan, P. Subasic, V. P. Su, R. H. Tanin, and H. Tepanyan (2025), "Quantum image classification: Experiments on utility-scale quantum computers," arXiv:2504.10595.
Ghazi Vakili, M., C. Gorgulla, J. Snider, A. Nigam, D. Bezrukov, D. Varoli, A. Aliper, D. Polykovsky, K. M. Padmanabha Das, H. Cox Iii, *et al.* (2025), Nat. Biotechnol. **43** (12), 1954.
Gidney, C., M. Newman, P. Brooks, and C. Jones (2025), Nat. Commun. **16** (4498), 4498.
Gil-Fuster, E., J. Eisert, and C. Bravo-Prieto (2024a), Nat. Commun. **15** (2277), 2277.
Gil-Fuster, E., J. Eisert, and V. Dunjko (2024b), Mach. Learn.: Sci. Technol. **5** (2), 025003.
Gil Vidal, F. J., and D. O. Theis (2020), Front. Phys. **8**, 473896.
Gilmer, J., S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl (2017), in *Proceedings of the 34th International Conference on Machine Learning - Volume 70*, ICML'17 (JMLR.org) p. 1263–1272.
Gilyén, A., Y. Su, G. H. Low, and N. Wiebe (2019) (Association for Computing Machinery, New York, NY, USA) p. 193–204.
Giovannetti, V., S. Lloyd, and L. Maccone (2008), Phys. Rev. Lett. **100**, 160501.
Girardi, F., and G. De Palma (2025), Commun. Math. Phys. **406** (4), 92.
Glasser, I., N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac (2018), Phys. Rev. X **8**, 011006.
Goh, M. L., M. Larocca, L. Cincio, M. Cerezo, and F. Sauvage (2025), Phys. Rev. Res. **7**, 033266.
Gonthier, J. F., M. D. Radin, C. Buda, E. J. Doskocil, C. M. Abuan, and J. Romero (2022), Phys. Rev. Res. **4**, 033154.
Gonzalez-Conde, J., T. W. Watts, P. Rodriguez-Grasa, and M. Sanz (2024), Quantum **8**, 1297.
González-García, G., J. I. Cirac, and R. Trivedi (2025), Quantum **9**, 1730.
Goodfellow, I., Y. Bengio, and A. Courville (2016), *Deep Learning* (MIT Press) `http://www.deeplearningbook.org`.
Gottesman, D. (1998), arXiv:quant-ph/9807006.
Gottesman, D. (2009), "An introduction to quantum error correction and fault-tolerant quantum computation," arXiv:0904.2557.
Gottesman, D. (2013), Quantum Inf. Comput. **14**, 1338.
Grant, E., M. Benedetti, S. Cao, A. Hallam, J. Lockhart, V. Stojevic, A. G. Green, and S. Severini (2018), npj Quantum Inf. **4** (65), 65.
Grant, E., L. Wossnig, M. Ostaszewski, and M. Benedetti (2019), Quantum **3**, 214.
Grasedyck, L., D. Kressner, and C. Tobler (2013), GAMM-Mitteilungen. **36** (1), 53.
Grefenstette, E., M. Sadrzadeh, S. Clark, B. Coecke, and S. Pulman (2011), *Proceedings of the Ninth International Conference on Computational Semantics*, IWCS '11, 125–134.
Gresch, A., and M. Kliesch (2025), Nat. Commun. **16** (689), 689.
Grier, D., H. Pashayan, and L. Schaeffer (2024), Quantum **8**, 1373.
Gross, D., Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert (2010), Phys. Rev. Lett. **105**, 150401.
Grover, L. K. (1996), in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '96 (Association for Computing Machinery, New York, NY, USA) p. 212–219.
Grover, L. K. (1997), Phys. Rev. Lett. **79**, 325.
Gu, X., A. F. Kockum, A. Miranowicz, Y.-x. Liu, and F. Nori (2017), Phys. Rep. **718-719**, 1.
Guarasci, R., G. De Pietro, and M. Esposito (2022), Applied Sciences **12** (11), 10.3390/app12115651.
Guerreschi, G. G., and M. Smelyanskiy (2017), "Practical optimization for hybrid quantum-classical algorithms," arXiv:1701.01450.
Gundlach, H., H. Kukina, J. Lynch, and N. Thompson (2025), "Quantum deep learning still needs a quantum leap," arXiv:2511.01253.

Guo, N., K. Mitarai, and K. Fujii (2024a), Phys. Rev. Res. **6**, 043227.

Guo, N., Z. Yu, M. Choi, A. Agrawal, K. Nakaji, A. Aspuru-Guzik, and P. Rebentrost (2024b), "Quantum linear algebra is all you need for transformer architectures," arXiv:2402.16714.

Guo, Z., A. Khan, V. S. Sheng, S. Jabeen, and Z. Pan (2025), Quantum Sci. Technol. **10** (3), 035054.

Guo, Z., R. Li, X. He, J. Guo, and S. Ju (2024c), MGE Advances **2** (4), e73.

Gupta, R. S., C. E. Wood, T. Engstrom, J. D. Pole, and S. Shrapnel (2025), npj Digital Med. **8** (237), 237.

Gyurik, C., C. Cade, and V. Dunjko (2022), Quantum **6**, 855.

Gyurik, C., R. Molteni, and V. Dunjko (2025), in *Proceedings of the 42nd International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 267, edited by A. Singh, M. Fazel, D. Hsu, S. Lacoste-Julien, F. Berkenkamp, T. Maharaj, K. Wagstaff, and J. Zhu (PMLR) pp. 21530–21545.

Häffner, H., C. F. Roos, and R. Blatt (2008), Phys. Rep. **469** (4), 155.

Hadfield, C., S. Bravyi, R. Raymond, and A. Mezzacapo (2022), Commun. Math. Phys. **391** (3), 951.

Hagelueken, M., M. F. Huber, and M. Roth (2025), New J. Phys. **27** (5), 054508.

Hamilton, K. E., E. F. Dumitrescu, and R. C. Pooser (2019), Phys. Rev. A **99**, 062323.

Han, Z.-Y., J. Wang, H. Fan, L. Wang, and P. Zhang (2018), Phys. Rev. X **8**, 031012.

Hangleiter, D., I. Roth, J. Fuksa, J. Eisert, and P. Roushan (2024), Nat. Commun. **15** (9595), 9595.

Hann, C. T., G. Lee, S. Girvin, and L. Jiang (2021), PRX Quantum **2**, 020311.

Hansen, N. (2006), "The cma evolution strategy: A comparing review," in *Towards a New Evolutionary Computation: Advances in the Estimation of Distribution Algorithms*, edited by J. A. Lozano, P. Larrañaga, I. Inza, and E. Bengoetxea (Springer Berlin Heidelberg, Berlin, Heidelberg) pp. 75–102.

Harrow, A. W., A. Hassidim, and S. Lloyd (2009), Phys. Rev. Lett. **103**, 150502.

Harrow, A. W., and R. A. Low (2009), Commun. Math. Phys. **291** (1), 257.

Harvey, C., R. Yeung, and K. Meichanetzidis (2025), Scientific Reports **15** (1), 7155.

Haug, T., and M. S. Kim (2024), Phys. Rev. Lett. **133**, 050603.

Hauke, P., H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver (2020), Rep. Prog. Phys. **83** (5), 054401.

Havlíček, V., A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta (2019), Nature **567** (7747), 209.

Hayashi, A., A. Sakurai, W. J. Munro, and K. Nemoto (2025), Phys. Rev. A **111**, 022431.

He, J.-J., K.-M. Hu, Y.-Z. Zhu, G.-J. Yan, S.-Y. Liang, X. Zhao, D. Wang, F.-X. Guo, Z.-F. Lan, X.-W. Shang, *et al.* (2025), "Deepquantum: A pytorch-based software platform for quantum machine learning and photonic quantum computing," arXiv:2512.18995.

He, K., X. Zhang, S. Ren, and J. Sun (2016), in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.

Heidari, S., M. J. Dinneen, and P. Delmas (2024), Future Gener. Comput. Syst. **160**, 54.

Heimann, N., L. Broers, and L. Mathey (2025), Phys. Rev. Res. **7**, 013101.

Henderson, M., S. Shakya, S. Pradhan, and T. Cook (2020a), Quantum Mach. Intell. **2** (1), 2.

Henderson, P., J. Hu, J. Romoff, E. Brunskill, D. Jurafsky, and J. Pineau (2020b), J. Mach. Learn. Res. **21** (1), 10.5555/3455716.3455964.

Hendrickx, N. W., L. Massai, M. Mergenthaler, F. J. Schupp, S. Paredes, S. W. Bedell, G. Salis, and A. Fuhrer (2024), Nat. Mater. **23** (7), 920.

Henry, L.-P., S. Thabet, C. Dalyac, and L. Henriet (2021), Phys. Rev. A **104**, 032416.

Herbst, S. (2023), *Beyond 0's and 1's: Exploring the impact of noise, data encoding, and hyperparameter optimization in quantum machine learning*, Ph.D. thesis (Technische Universität Wien).

Herbst, S., I. Brandić, and A. Pérez-Salinas (2025), arXiv:2512.24801.

Herbst, S., V. De Maio, and I. Brandic (2024), in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01, pp. 1478–1489.

Herman, D., C. Googin, X. Liu, Y. Sun, A. Galda, I. Safro, M. Pistoia, and Y. Alexeev (2023), Nat. Rev. Phys. **5** (8), 450.

Hermann, J., Z. Schätzle, and F. Noé (2020), Nat. Chem. **12** (10), 891.

Herrmann, J., S. M. Llima, A. Remm, P. Zapletal, N. A. McMahon, C. Scarato, F. Swiadek, C. K. Andersen, C. Hellings, S. Krinner, *et al.* (2022), Nat. Commun. **13** (4144), 4144.

Heurtel, N., A. Fyrillas, G. d. Gliniasty, R. Le Bihan, S. Malherbe, M. Pailhas, E. Bertasi, B. Bourdoncle, P.-E. Emeriau, R. Mezher, L. Music, N. Belabas, B. Valiron, P. Senellart, S. Mansfield, and J. Senellart (2023), Quantum **7**, 931.

Hibat-Allah, M., M. Mauri, J. Carrasquilla, and A. Perdomo-Ortiz (2024), Communications Physics **7** (1), 68.

Higham, C. F., and A. Bedford (2023), Scientific reports **13** (1), 3939.

Hinsche, M., M. Ioannou, A. Nietner, J. Haferkamp, Y. Quek, D. Hangleiter, J.-P. Seifert, J. Eisert, and R. Sweke (2023), Phys. Rev. Lett. **130**, 240602.

Hinton, G. E. (2002), Neural Comput. **14** (8), 1771.

Hinton, G. E., S. Osindero, and Y.-W. Teh (2006), Neural computation **18** (7), 1527.

Ho, J., A. Jain, and P. Abbeel (2020), *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, 10.5555/3495724.3496298.

Hoffmann, J., S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford, D. d. L. Casas, L. A. Hendricks, J. Welbl, A. Clark, *et al.* (2022), "Training compute-optimal large language models," arXiv:2203.15556.

Hohenfeld, H., D. Heimann, F. Wiebe, and F. Kirchner (2022), "Quantum deep reinforcement learning for robot navigation tasks," arXiv:2202.12180.

Holevo, A. S. (1973), J. Multivariate Anal. **3** (4), 337.

Holmes, Z., K. Sharma, M. Cerezo, and P. J. Coles (2022), PRX Quantum **3**, 010313.

Hooke, R., and T. A. Jeeves (1961), J. ACM **8** (2), 212–229.

Hopfield, J. J. (1982), Proc. Natl. Acad. Sci. U.S.A. **79** (8), 2554.

Hornik, K., M. Stinchcombe, and H. White (1989), Neural Networks **2** (5), 359.

Horodecki, R., P. Horodecki, M. Horodecki, and K. Horodecki (2009), Rev. Mod. Phys. **81**, 865.

Horst, R., and H. Tuy (2013), *Global optimization: Deterministic approaches* (Springer Science & Business Media).

Hossain, S., S. Umer, R. K. Rout, and H. A. Marzouqi (2024), IEEE Transactions on Neural Systems and Rehabilitation Engineering **32**, 1556.

Hu, L., S.-H. Wu, W. Cai, Y. Ma, X. Mu, Y. Xu, H. Wang, Y. Song, D.-L. Deng, C.-L. Zou, *et al.* (2019), Sci. Adv. **5** (1), 10.1126/sciadv.aav2761.

Huang, D., T. T. Allen, W. I. Notz, and R. A. Miller (2006), Structural and Multidisciplinary Optimization **32** (5), 369.

Huang, H.-Y., M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean (2022), Science **376** (6598), 1182.

Huang, H.-Y., M. Broughton, N. Eassa, H. Neven, R. Babbush, and J. R. McClean (2025a), "Generative quantum advantage for classical and quantum problems," arXiv:2509.09033.

Huang, H.-Y., M. Broughton, M. Mohseni, R. Babbush, S. Boixo, H. Neven, and J. R. McClean (2021a), Nat. Commun. **12** (2631), 2631.

Huang, H.-Y., S. Choi, J. R. McClean, and J. Preskill (2025b), "The vast world of quantum advantage," arXiv:2508.05720.

Huang, H.-Y., R. Kueng, and J. Preskill (2020), Nat. Phys. **16** (10), 1050.

Huang, H.-Y., R. Kueng, and J. Preskill (2021b), Phys. Rev. Lett. **126**, 190505.

Huang, H.-Y., Y. Liu, M. Broughton, I. Kim, A. Anshu, Z. Landau, and J. R. McClean (2024a), in *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, STOC 2024 (Association for Computing Machinery, New York, NY, USA) p. 1343–1351.

Huang, J. Y., R. Y. Su, W. H. Lim, M. Feng, B. van Straaten, B. Severin, W. Gilbert, N. Dumoulin Stuyck, T. Tanttu, S. Serrano, *et al.* (2024b), Nature **627** (8005), 772.

Huang, T., J.-J. Zhu, and Z.-Y. Ni (2025c), "Optimal control by variational quantum algorithms," arXiv:2505.23373.

Huggins, W., P. Patil, B. Mitchell, K. B. Whaley, and E. M. Stoudenmire (2019), Quantum Sci. Technol. **4** (2), 024001.

Huggins, W. J., J. R. McClean, N. C. Rubin, Z. Jiang, N. Wiebe, K. B. Whaley, and R. Babbush (2021), npj Quantum Inf. **7** (23), 23.

Hur, T., I. F. Araujo, and D. K. Park (2024), Phys. Rev. A **110**, 022411.

Hur, T., L. Kim, and D. K. Park (2022), Quantum Mach. Intell. **4** (1), 3.

Hur, T., and D. K. Park (2024), "Understanding generalization in quantum machine learning with margins," arXiv:2411.06919.

Huyer, W., and A. Neumaier (1999), J. Global Optim. **14** (4), 331.

Incudini, M., M. Grossi, A. Mandarino, S. Vallecorsa, A. Di Pierro, and D. Windridge (2023), IEEE Transactions on Quantum Engineering **4** (01), 1.

Ivakhnenko, A. G., and V. G. Lapa (1965), *Cybernetic Predicting Devices* (CCM Information Corporation).

Jacot, A., F. Gabriel, and C. Hongler (2018), in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18 (Curran Associates Inc., Red Hook, NY, USA) p. 8580–8589.

Jaderberg, B., L. W. Anderson, W. Xie, S. Albanie, M. Kiffner, and D. Jaksch (2022), Quantum Sci. Technol. **7** (3), 035005.

Jaderberg, B., A. A. Gentile, Y. A. Berrada, E. Shishenina, and V. E. Elfving (2024), Phys. Rev. A **109**, 042421.

Jaderberg, B., G. Pennington, K. V. Marshall, L. W. Anderson, A. Agarwal, L. P. Lindoy, I. Rungger, S. Mensa, and J. Crain (2026), Phys. Rev. Res. **8**, 013081.

Jäger, J., T. N. Kaldenbach, M. Haas, and E. Schultheis (2025), Communications Physics **8** (1), 418.

Jahromi, S. S., and R. Orús (2024), Scientific Reports **14** (1), 19017.

Jaksch, D., J. I. Cirac, P. Zoller, S. L. Rolston, R. Côté, and M. D. Lukin (2000), Phys. Rev. Lett. **85**, 2208.

Jansen, S., M.-B. Ruskai, and R. Seiler (2007), J. Math. Phys. **48** (10), 102111.

Jaques, S., and A. G. Rattew (2025), Quantum **9**, 1922.

Javadi-Abhari, A., M. Treinish, K. Krsulich, C. J. Wood, J. Lishman, J. Gacon, S. Martiel, P. D. Nation, L. S. Bishop, A. W. Cross, B. R. Johnson, and J. M. Gambetta (2024), "Quantum computing with qiskit," arXiv:2405.08810.

Jayakumar, A., M. Vuffray, and A. Y. Lokhov (2024), Phys. Rev. Res. **6**, 033201.

Jerbi, S., C. Gyurik, S. Marshall, H. Briegel, and V. Dunjko (2021a), *Advances in Neural Information Processing Systems*, **34**, 28362.

Jerbi, S., C. Gyurik, S. C. Marshall, R. Molteni, and V. Dunjko (2024), Nat. Commun. **15** (5676), 5676.

Jerbi, S., L. M. Trenkwalder, H. Poulsen Nautrup, H. J. Briegel, and V. Dunjko (2021b), PRX Quantum **2**, 010328.

Ji, Y., S. Brandhofer, and I. Polian (2022), in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE Computer Society, Los Alamitos, CA, USA) pp. 204–214.

Ji, Y., X. Chen, I. Polian, and Y. Ban (2025a), Phys. Rev. Appl. **23**, 034022.

Ji, Y., S. Kirchhoff, and F. K. Wilhelm (2026), "Exploring the fidelity of flux qubit measurement in different bases via quantum flux parametron," arXiv:2510.24082.

Ji, Y., K. F. Koenig, and I. Polian (2023), Phys. Rev. A **108**, 022610.

Ji, Y., K. F. Koenig, and I. Polian (2024), Phys. Rev. A **110**, 032421.

Ji, Y., and I. Polian (2024), Entropy **26** (7), 586.

Ji, Y., M. Roth, D. A. Kreplin, I. Polian, and F. K. Wilhelm (2025b), "Data-efficient quantum noise modeling via machine learning," arXiv:2509.12933.

Jia, Z.-A., B. Yi, R. Zhai, Y.-C. Wu, G.-C. Guo, and G.-P. Guo (2019), Adv. Quantum Technol. **2** (7-8), 1800077.

John, V., C. X. Yu, B. van Straaten, E. A. Rodríguez-Mena, M. Rodríguez, S. D. Oosterhout, L. E. A. Stehouwer, G. Scappucci, M. Rimbach-Russ, S. Bosco, *et al.* (2025), Nat. Commun. **16** (10560), 10560.

Jones, D. R., M. Schonlau, and W. J. Welch (1998), Journal of Global optimization **13** (4), 455.

Jones, S., and P. Murali (2025), "Architecting scalable trapped ion quantum computers using surface codes," arXiv:2510.23519.

Josephson, B. D. (1962), Phys. Lett. **1** (7), 251.

Josephson, B. D. (1974), Rev. Mod. Phys. **46**, 251.

Jouppi, N. P., C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, *et al.* (2017), in *Proceedings of the 44th Annual International Symposium on Computer Architecture*, ISCA '17 (Association for Computing Machinery, New York, NY, USA) p. 1–12.

Jozsa, R., and N. Linden (2003), Proceedings: Mathematical, Physical and Engineering Sciences **459** (2036), 2011.

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *et al.* (2021), Nature **596** (7873), 583.

Kabir, M., M. Kaosar, H. Laga, and F. Sohel (2025), Quantum Mach. Intell. **7** (1), 51.

Kadowaki, T., and H. Nishimori (1998), Phys. Rev. E **58**, 5355.

Kak, S. C. (1995), Advances in imaging and electron physics **94**, 259.

Kandala, A., A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta (2017), Nature **549** (7671), 242.

Kane, B. E. (1998), Nature **393** (6681), 133.

Kang, H., Y. Kim, E. Adermann, M. Sevior, and M. Usman (2025), Quantum Sci. Technol. **11** (1), 015021.

Kaplan, J., S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, S. Gray, A. Radford, J. Wu, and D. Amodei (2020), "Scaling Laws for Neural Language Models," arXiv:2001.08361.

Karamlou, A., M. Pfaffhauser, and J. Wootton (2022), in *Proceedings of the 15th International Conference on Natural Language Generation*, edited by S. Shaikh, T. Ferreira, and A. Stent (Association for Computational Linguistics, Waterville, Maine, USA and virtual meeting) pp. 267–277.

Kartsaklis, D., I. Fan, R. Yeung, A. Pearson, R. Lorenz, A. Toumi, G. de Felice, K. Meichanetzidis, S. Clark, and B. Coecke (2021), "lambeq: An efficient high-level python library for quantum nlp," arXiv:2110.04236.

Kaseb, Z., M. Möller, G. T. Balducci, P. Palensky, and P. P. Vergara (2024), Electr. Power Syst. Res. **235**, 110677.

Kashif, M., and S. Al-Kuwari (2023), International Journal of Parallel, Emergent and Distributed Systems **38** (5), 362.

Kashif, M., and S. Al-Kuwari (2024), EPJ Quantum Technol. **11** (1), 4.

Kashif, M., and M. Shafique (2025), Scientific Reports **15** (1), 21764.

Kazi, S., M. Larocca, M. Farinati, P. J. Coles, M. Cerezo, and R. Zeier (2025), PRX Quantum **6**, 040345.

Keane, A. J. (2006), AIAA journal **44** (4), 879.

Kearns, M., Y. Mansour, D. Ron, R. Rubinfeld, R. E. Schapire, and L. Sellie (1994), in *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing*, STOC '94 (Association for Computing Machinery, New York, NY, USA) p. 273–282.

de Keijzer, R., O. Tse, and S. Kokkelmans (2023), Quantum **7**, 908.

Kempkes, M., A. Ijaz, E. Gil-Fuster, C. Bravo-Prieto, J. Spiegelberg, E. van Nieuwenburg, and V. Dunjko (2025), "Double descent in quantum kernel methods," arXiv:2501.10077.

Kennedy, J., and R. Eberhart (1995), in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Vol. 4, pp. 1942–1948 vol.4.

Kerenidis, I., J. Landman, and A. Prakash (2019), "Quantum algorithms for deep convolutional neural networks," arXiv:1911.01117.

Kerenidis, I., and A. Prakash (2016), "Quantum recommendation systems," arXiv:1603.08675.

Khan, T. M., and A. Robles-Kelly (2020), IEEE Access **8**, 219275.

Khaneja, N., T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser (2005), J. Magn. Reson. **172** (2), 296, 15649756.

Kharsa, R., A. Bouridane, and A. Amira (2023), Neurocomputing **560**, 126843.

Killoran, N., T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd (2019a), Phys. Rev. Res. **1**, 033063.

Killoran, N., J. Izaac, N. Quesada, V. Bergholm, M. Amy, and C. Weedbrook (2019b), Quantum **3**, 129.

Kim, J., J. Huh, and D. K. Park (2023a), Neurocomputing **555**, 126643.

Kim, Y., A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, *et al.* (2023b), Nature **618** (7965), 500.

King, A. D., A. Nocera, M. M. Rams, J. Dziarmaga, R. Wiersema, W. Bernoudy, J. Raymond, N. Kaushal, N. Heinsdorf, R. Harris, *et al.* (2025), Science **388** (6743), 199.

Kingma, D. P., and J. Ba (2014), "Adam: A method for stochastic optimization," arXiv:1412.6980.

Kipf, T. N., and M. Welling (2017), "Semi-supervised classification with graph convolutional networks," arXiv:1609.02907.

Kissinger, A., and J. Van De Wetering (2019), "Pyzx: Large scale automated diagrammatic reasoning," arXiv:1904.04735.

Kjaergaard, M., M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver (2020), Annual Review of Condensed Matter Physics **11** (1), 369.

Klusch, M., J. Lässig, D. Müssig, A. Macaluso, and F. K. Wilhelm (2024), Künstl. Intell. **38** (4), 257.

Knill, E., R. Laflamme, and W. Zurek (1996), "Threshold accuracy for quantum computation," arXiv:quant-ph/9610011.

Knill, E., R. Laflamme, and W. H. Zurek (1998), Science **279** (5349), 342.

Knowles, J. (2006), IEEE Transactions on Evolutionary Computation **10** (1), 50.

Koch, C. P., U. Boscain, T. Calarco, G. Dirr, S. Filipp, S. J. Glaser, R. Kosloff, S. Montangero, T. Schulte-Herbrüggen, D. Sugny, *et al.* (2022), EPJ Quantum Technol. **9** (1), 19.

Koch, J., T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf (2007), Phys. Rev. A **76**, 042319.

Koczor, B., and S. C. Benjamin (2022), Phys. Rev. A **106**, 062416.

Kok, P., W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn (2007), Rev. Mod. Phys. **79**, 135.

Kokail, C., P. E. Dolgirev, R. van Bijnen, D. Gonzalez-Cuadra, M. D. Lukin, and P. Zoller (2026), "Inverse quantum simulation for quantum material design," arXiv:2601.12239.

Kolarovszki, Z., T. Rybotycki, P. Rakyta, Á. Kaposi, B. Poór, S. Jóczik, D. T. R. Nagy, H. Varga, K. H. El-Safty, G. Morse, M. Oszmaniec, T. Kozsik, and Z. Zimborás (2025), Quantum **9**, 1708.

Kolotouros, I., and P. Wallden (2024), Quantum **8**, 1503.

Komoto, D., and K. Kasai (2025), npj Quantum Inf. **11** (154), 154.

Kong, X., L. Li, Z. Chen, C. Xue, X. Xu, H. Liu, Y. Wu, Y. Fang, H. Fang, K. Chen, Y. Yang, M. Dou, and G. Guo (2025), "Quantum-enhanced llm efficient fine tuning," arXiv:2503.12790.

Konnov, A. I., and V. F. Krotov (1999), Avtomatika i Telemekhanika (10), 77.

Kornjača, M., H.-Y. Hu, C. Zhao, J. Wurtz, P. Weinberg, M. Hamdan, A. Zhdanov, S. H. Cantu, H. Zhou, R. A. Bravo, *et al.* (2024), "Large-scale quantum reservoir learn-

ing with an analog quantum computer," arXiv:2407.02553.
Kovachki, N., Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar (2023), Journal of Machine Learning Research **24** (89), 1.
Krantz, P., M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver (2019), Applied Physics Reviews **6** (2), 021318.
Kreplin, D. A., and M. Roth (2024), Quantum **8**, 1385.
Kreplin, D. A., M. Willmann, J. Schnabel, F. Rapp, M. Hagelüken, and M. Roth (2025), IEEE Software **42** (5), 65.
Krizhevsky, A., G. Hinton, *et al.* (2009), .
Krizhevsky, A., I. Sutskever, and G. E. Hinton (2012), in *Advances in Neural Information Processing Systems*, Vol. 25, edited by F. Pereira, C. Burges, L. Bottou, and K. Weinberger (Curran Associates, Inc.).
Kübler, J. M., A. Arrasmith, L. Cincio, and P. J. Coles (2020), Quantum **4**, 263.
Kübler, J. M., S. Buchholz, and B. Schölkopf (2021), *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, 10.5555/3540261.3541230.
Kungurtsev, V., G. Korpas, J. Marecek, and E. Y. Zhu (2024), Quantum **8**, 1495.
Kurkin, A., K. Shen, S. Pielawa, H. Wang, and V. Dunjko (2025), "Universality and kernel-adaptive training for classically trained, quantum-deployed generative models," arXiv:2510.08476.
Kushner, H. J. (1964), Journal of Basic Engineering **86** (1), 97.
Kwak, Y., W. J. Yun, J. P. Kim, H. Cho, J. Park, M. Choi, S. Jung, and J. Kim (2023), ICT Express **9** (3), 486.
Kwun, G., B. Zhang, and Q. Zhuang (2025), Phys. Rev. A **111**, 032610.
Kyriienko, O., and V. E. Elfving (2021), Phys. Rev. A **104**, 052417.
Lakhdar-Hamina, D., X. Liu, R. Barney, S. H. Miller, A. M. Green, N. M. Linke, and V. Galitski (2025), "Benchmarking a tunable quantum neural network on trapped-ion and superconducting hardware," arXiv:2507.21222.
Lall, D., A. Agarwal, W. Zhang, L. Lindoy, T. Lindström, S. Webster, S. Hall, N. Chancellor, P. Wallden, R. Garcia-Patron, *et al.* (2025), "A review and collection of metrics and benchmarks for quantum computers: definitions, methodologies and software," arXiv:2502.06717.
Lamichhane, P., and D. B. Rawat (2025), IEEE Access.
Landman, J., N. Mathur, Y. Y. Li, M. Strahm, S. Kazdaghli, A. Prakash, and I. Kerenidis (2022), Quantum **6**, 881.
Lanes, O., M. Beji, A. D. Corcoles, C. Dalyac, J. M. Gambetta, L. Henriet, A. Javadi-Abhari, A. Kandala, A. Mezzacapo, C. Porter, *et al.* (2025), "A framework for quantum advantage," arXiv:2506.20658.
Larocca, M., P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo (2022a), Quantum **6**, 824.
Larocca, M., N. Ju, D. García-Martín, P. J. Coles, and M. Cerezo (2023), "Theory of overparametrization in quantum neural networks," .
Larocca, M., F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo (2022b), PRX Quantum **3**, 030341.
Larocca, M., S. Thanasilp, S. Wang, K. Sharma, J. Biamonte, P. J. Coles, L. Cincio, J. R. McClean, Z. Holmes, and M. Cerezo (2025), Nat. Rev. Phys. **7** (4), 174.
LaRose, R., A. Mari, S. Kaiser, P. J. Karalekas, A. A. Alves, P. Czarnik, M. El Mandouh, M. H. Gordon, Y. Hindy, A. Robertson, P. Thakre, M. Wahl, D. Samuel, R. Mistri, M. Tremblay, N. Gardner, N. T. Stemen, N. Shammah, and W. J. Zeng (2022), Quantum **6**, 774.
Le Gall, F. (2025), Comput. Complexity **34** (1), 2.
LeCun, Y., Y. Bengio, and G. Hinton (2015), Nature **521** (7553), 436.
LeCun, Y., B. Boser, J. S. Denker, D. Henderson, R. E. Howard, and W. Hubbard (1989), Neural Comput. **1** (4), 541.
LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner (1998), Proceedings of the IEEE **86** (11), 2278.
LeCun, Y., S. Chopra, R. Hadsell, M. Ranzato, F. Huang, *et al.* (2006), Predicting structured data **1** (0).
Lee, C., I. F. Araujo, D. Kim, J. Lee, S. Park, J.-Y. Ryu, and D. K. Park (2025a), Front. Phys. **13**, 1529188.
Lee, J., S. Omkar, Y. S. Teo, S.-H. Lee, H. Kwon, M. S. Kim, and H. Jeong (2025b), Newton **0** (0), 10.1016/j.newton.2025.100359.
Lee, J.-S., J. Bang, S. Hong, C. Lee, K. H. Seol, J. Lee, and K.-G. Lee (2019), Phys. Rev. A **99**, 012313.
Leng, J., Y. Peng, Y.-L. Qiao, M. C. Lin, and X. Wu (2022), *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, 10.5555/3600270.3600610.
de Leon, N. P., K. M. Itoh, D. Kim, K. K. Mehta, T. E. Northup, H. Paik, B. S. Palmer, N. Samarth, S. Sangtawesin, and D. W. Steuerman (2021), Science **372** (6539), 10.1126/science.abb2823.
Leone, L., S. F. E. Oliviero, and A. Hamma (2022), Phys. Rev. Lett. **128**, 050402.
Leong, F. Y., W.-B. Ewe, S. B. Q. Tran, Z. Zhang, and J. Y. Khoo (2025), Comput. Fluids **301**, 106782.
Leporini, R., and D. Pastorello (2022), Scientific Reports **12** (1), 8781.
Letcher, A., S. Woerner, and C. Zoufal (2024), Quantum **8**, 1484.
Letham, B., B. Karrer, G. Ottoni, and E. Bakshy (2017), Bayesian Analysis **14**, 10.1214/18-BA1110.
Levine, Y., O. Sharir, N. Cohen, and A. Shashua (2019), Phys. Rev. Lett. **122**, 065301.
Levitan, B., and S. Kauffman (1995), Mol. Diversity **1** (1), 53.
Levy, R., D. Luo, and B. K. Clark (2024), Phys. Rev. Res. **6**, 013029.
Lewis, L., D. Gilboa, and J. R. McClean (2025), Nat. Commun. **17** (1341), 1341.
Leymann, F., and J. Barzen (2020), Quantum Sci. Technol. **5** (4), 044007.
Leyton-Ortega, V., A. Perdomo-Ortiz, and O. Perdomo (2021), Quantum Mach. Intell. **3** (1), 17.
Li, G., X. Zhao, and X. Wang (2024), Sci. China Inf. Sci. **67** (4), 142501.
Li, J., X. Yang, X. Peng, and C.-P. Sun (2017), Phys. Rev. Lett. **118**, 150503.
Li, P., H. Xiao, F. Shang, X. Tong, X. Li, and M. Cao (2013), Neurocomputing **117**, 81.
Li, Q., D. Gkoumas, A. Sordoni, J.-Y. Nie, and M. Melucci (2021a), AAAI **35** (15), 13270.
Li, R., L. Petit, D. P. Franke, J. P. Dehollain, J. Helsen, M. Steudtner, N. K. Thomas, Z. R. Yoscovits, K. J. Singh, S. Wehner, *et al.* (2018), Sci. Adv. **4** (7), 10.1126/sciadv.aar3960.
Li, S., Y. Fan, X. Li, X. Ruan, Q. Zhao, Z. Peng, R.-B. Wu, J. Zhang, and P. Song (2025a), npj Quantum Inf. **11** (124),

124.
Li, Y., R.-G. Zhou, R. Xu, J. Luo, and W. Hu (2020a), Quantum Sci. Technol. **5** (4), 044003.
Li, Z., Z. Chai, Y. Guo, W. Ji, M. Wang, F. Shi, Y. Wang, S. Lloyd, and J. Du (2021b), Sci. Adv. **7** (34), 10.1126/sciadv.abg2589.
Li, Z., X. Fu, L. Meng, and R. Du (2025b), Scientific Reports **15** (1), 32111.
Li, Z., N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar (2020b), "Fourier neural operator for parametric partial differential equations," arXiv:2010.08895.
Li, Z., X. Liu, N. Xu, and J. Du (2015), Phys. Rev. Lett. **114**, 140504.
Liang, Y., W. Peng, Z.-J. Zheng, O. Silvén, and G. Zhao (2021), Neural Networks **143**, 133.
Liang, Z., H. Wang, J. Cheng, Y. Ding, H. Ren, Z. Gao, Z. Hu, D. S. Boning, X. Qian, S. Han, W. Jiang, and Y. Shi (2022), in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, pp. 556–565.
Liao, H.-J., J.-G. Liu, L. Wang, and T. Xiang (2019), Phys. Rev. X **9**, 031041.
Liao, Y., and C. Ferrie (2024), "Gpt on a quantum computer," arXiv:2403.09418.
Linnainmaa, S. (1970), *The representation of the cumulative rounding error of an algorithm as a Taylor expansion of the local rounding errors*, Ph.D. thesis (Master's Thesis, Univ. Helsinki).
Liu, A., C. Wen, and J. Wang (2025a), Adv. Quantum Technol. **8** (10), 2400703.
Liu, D. C., and J. Nocedal (1989), Math. Program. **45** (1), 503.
Liu, H., T. Hur, S. Zhang, L. Che, X. Long, X. Wang, K. Huang, Y.-a. Fan, Y. Zheng, Y. Feng, Y. Zhou, J. Ng, X. Nie, D. K. Park, and D. Lu (2025b), Phys. Rev. Lett. **135**, 080603.
Liu, H.-Y., Z.-Y. Chen, T.-P. Sun, C. Xue, Y.-C. Wu, and G.-P. Guo (2023a), "Can variational quantum algorithms demonstrate quantum advantages? time really matters," arXiv:2307.04089.
Liu, H.-Y., T.-P. Sun, Y.-C. Wu, Y.-J. Han, and G.-P. Guo (2023b), New J. Phys. **25** (1), 013039.
Liu, J., K. H. Lim, K. L. Wood, W. Huang, C. Guo, and H.-L. Huang (2021a), Sci. China Phys. Mech. Astron. **64** (9), 290311.
Liu, J., M. Liu, J.-P. Liu, Z. Ye, Y. Wang, Y. Alexeev, J. Eisert, and L. Jiang (2024), Nat. Commun. **15** (434), 434.
Liu, J., F. Tacchino, J. R. Glick, L. Jiang, and A. Mezzacapo (2022a), PRX Quantum **3**, 030323.
Liu, J., F. Wilde, A. A. Mele, X. Jin, L. Jiang, and J. Eisert (2025c), Phys. Rev. A **111**, 052441.
Liu, J.-G., and L. Wang (2018), Phys. Rev. A **98**, 062324.
Liu, W., Y. Zhu, Y. Zha, Q. Wu, L. Jian, and Z. Liu (2025d), IEEE Transactions on Pattern Analysis and Machine Intelligence **47** (11), 10329.
Liu, Y., S. Arunachalam, and K. Temme (2021b), Nat. Phys. **17** (9), 1013.
Liu, Z., L.-M. Duan, and D.-L. Deng (2022b), Phys. Rev. Res. **4**, 013097.
Liu, Z.-H., R. Brunel, E. E. B. Østergaard, O. Cordero, S. Chen, Y. Wong, J. A. H. Nielsen, A. B. Bregnsbo, S. Zhou, H.-Y. Huang, *et al.* (2025e), Science **389** (6767), 1332.
Lloyd, S. (1996), Science **273** (5278), 1073.
Lloyd, S., M. Mohseni, and P. Rebentrost (2013), "Quantum algorithms for supervised and unsupervised machine learning," arXiv:1307.0411.
Lloyd, S., M. Mohseni, and P. Rebentrost (2014), Nat. Phys. **10** (9), 631.
Lloyd, S., M. Schuld, A. Ijaz, J. Izaac, and N. Killoran (2020), "Quantum embeddings for machine learning," arXiv:2001.03622.
Lloyd, S., and C. Weedbrook (2018), Phys. Rev. Lett. **121**, 040502.
Lockwood, O., and M. Si (2020), in *Proceedings of the Sixteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, AIIDE'20 (AAAI Press).
Long, C., M. Huang, X. Ye, Y. Futamura, and T. Sakurai (2025), Scientific Reports **15** (1), 31780.
Lorenz, J. M., T. Monz, J. Eisert, D. Reitzner, F. Schopfer, F. Barbaresco, K. Kurowski, W. van der Schoot, T. Strohm, J. Senellart, *et al.* (2025), "Systematic benchmarking of quantum computers: status and recommendations," arXiv:2503.04905.
Lorenz, R., A. Pearson, K. Meichanetzidis, D. Kartsaklis, and B. Coecke (2023), Journal of Artificial Intelligence Research **76**, 1305–1342.
Loss, D., and D. P. DiVincenzo (1998), Phys. Rev. A **57**, 120.
Low, G. H., and I. L. Chuang (2019), Quantum **3**, 163.
Lu, H.-H., H.-H. Lu, M. Liscidini, A. L. Gaeta, A. M. Weiner, J. M. Lukens, and J. M. Lukens (2023), Optica **10** (12), 1655.
Lu, J. Z., L. Jiao, K. Wolinski, M. Kornjača, H.-Y. Hu, S. Cantu, F. Liu, S. F. Yelin, and S.-T. Wang (2024), Quantum Sci. Technol. **10** (1), 015038.
Lu, L., P. Jin, G. Pang, Z. Zhang, and G. E. Karniadakis (2021), Nature machine intelligence **3** (3), 218.
Lu, M., L. Du, Z. Cui, Y. Zhao, Q. Yan, J. Zhao, Y. Li, M. Dou, Q. Wang, Y.-C. Wu, *et al.* (2025), J. Chem. Inf. Model. **65** (15), 8057.
Lu, Z., H. Pu, F. Wang, Z. Hu, and L. Wang (2017), in *Advances in Neural Information Processing Systems*, Vol. 30, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.).
Lubinski, T., C. Granade, A. Anderson, A. Geller, M. Roetteler, A. Petrenko, and B. Heim (2022), Front. Phys. **10**, 940293.
Ma, H., G. J. Mooney, I. R. Petersen, L. C. L. Hollenberg, and D. Dong (2024), npj Quantum Inf. **10** (86), 86.
MacCormack, I., C. Delaney, A. Galda, N. Aggarwal, and P. Narang (2022), Phys. Rev. Res. **4**, 013117.
Machnes, S., E. Assémat, D. Tannor, and F. K. Wilhelm (2018), Phys. Rev. Lett. **120**, 150401.
Madsen, L. S., F. Laudenbach, M. Falamarzi. Askarani, F. Rortais, T. Vincent, J. F. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, M. J. Collins, *et al.* (2022), Nature **606** (7912), 75.
Magann, A. B., C. Arenz, M. D. Grace, T.-S. Ho, R. L. Kosut, J. R. McClean, H. A. Rabitz, and M. Sarovar (2021), PRX Quantum **2** (1), 010101.
Maier, B., M. Spannowsky, and S. Williams (2025), "Continuous-variable photonic quantum extreme learning machines for fast collider-data selection," arXiv:2510.13994.
Mande, N. S., and C. Shao (2025), Quantum **9**, 1593.
Manetsch, H. J., G. Nomura, E. Bataille, X. Lv, K. H. Leung, and M. Endres (2025), Nature **647** (8088), 60.

Manin, Y. I. (1980), "Vychislimoe i nevychislimoe (computable and noncomputable),".

Manzano, A., D. Dechant, J. Tura, and V. Dunjko (2025), Quantum **9**, 1658.

Mari, A., T. R. Bromley, J. Izaac, M. Schuld, and N. Killoran (2020), Quantum **4**, 340.

Marshall, J., D. Venturelli, I. Hen, and E. G. Rieffel (2019), Phys. Rev. Appl. **11**, 044083.

Martinis, J. M., M. H. Devoret, and J. Clarke (1985), Phys. Rev. Lett. **55**, 1543.

Martyn, J. M., Z. M. Rossi, A. K. Tan, and I. L. Chuang (2021), PRX Quantum **2**, 040203.

Masri, S., G. Bekey, and F. Safford (1980), Applied Mathematics and Computation **7** (4), 353.

Mathur, N., J. Landman, Y. Y. Li, M. Strahm, S. Kazdaghli, A. Prakash, and I. Kerenidis (2021), "Medical image classification via quantum neural networks," arXiv:2109.01831.

McArdle, S., S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan (2020), Rev. Mod. Phys. **92**, 015003.

McClean, J. R., S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven (2018), Nat. Commun. **9** (4812), 4812.

McClean, J. R., J. Romero, R. Babbush, and A. Aspuru-Guzik (2016), New Journal of Physics **18** (2), 023023.

Mei, S., A. Montanari, and P.-M. Nguyen (2018), Proc. Natl. Acad. Sci. U.S.A. **115** (33), E7665.

Meichanetzidis, K., S. Gogioso, G. De Felice, N. Chiappori, A. Toumi, and B. Coecke (2020), "Quantum natural language processing on near-term quantum computers," arXiv:2005.04147.

Meichanetzidis, K., A. Toumi, G. de Felice, and B. Coecke (2023), Quantum Mach. Intell. **5** (1), 10.

Meier, F., and H. Yamasaki (2025), PRX Energy **4**, 023008.

Meirom, D., and S. H. Frankel (2023), Front. Quantum Sci. Technol. **2**, 1273581.

Melchor Hernandez, A., F. Girardi, D. Pastorello, and G. De Palma (2025), Ann. Henri Poincaré , 1.

Mele, A. A., A. Angrisani, S. Ghosh, S. Khatri, J. Eisert, D. S. França, and Y. Quek (2024), "Noise-induced shallow circuits and absence of barren plateaus," arXiv:2403.13927.

Melko, R. G., G. Carleo, J. Carrasquilla, and J. I. Cirac (2019), Nat. Phys. **15** (9), 887.

Melo, A., N. Earnest-Noble, and F. Tacchino (2023), Quantum **7**, 1130.

Menickelly, M., Y. Ha, and M. Otten (2023), Quantum **7**, 949.

Menneer, T., and A. Narayanan (1995), in *Proceedings of the Neural Information Processing Systems*, Vol. 95, pp. 27–30.

Merchant, A., S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon, and E. D. Cubuk (2023), Nature **624** (7990), 80.

Meyer, J. J., M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert (2023), PRX Quantum **4**, 010328.

Meyer, S., F. Scala, F. Tacchino, and A. Lucchi (2025), "Trainability of quantum models beyond known classical simulability," arXiv:2507.06344.

Mhiri, H., R. Puig, S. Lerch, M. S. Rudolph, T. Chotibut, S. Thanasilp, and Z. Holmes (2025), "A unifying account of warm start guarantees for patches of quantum landscapes," arXiv:2502.07889.

Mifune, S., T. Kanao, and T. Tanamoto (2025), Japanese Journal of Applied Physics **64** (4), 04SP08.

Minervini, M., D. Patel, and M. M. Wilde (2025), Phys. Rev. A **112**, 022424.

Mishra, N., M. Kapil, H. Rakesh, A. Anand, N. Mishra, A. Warke, S. Sarkar, S. Dutta, S. Gupta, A. Prasad Dash, *et al.* (2020), *Data Management, Analytics and Innovation*, SpringerLink , 101.

Mitarai, K., M. Negoro, M. Kitagawa, and K. Fujii (2018), Phys. Rev. A **98**, 032309.

Mnih, V., K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, *et al.* (2015), Nature **518** (7540), 529.

Močkus, J. (2005), in *Optimization Techniques IFIP Technical Conference Novosibirsk, July 1–7, 1974* (Springer, Berlin, Germany) pp. 400–404.

Mockus, J. B., and L. J. Mockus (1991), J. Optim. Theory Appl. **70** (1), 157.

Mølmer, K., and A. Sørensen (1999), Phys. Rev. Lett. **82**, 1835.

Molteni, R., C. Gyurik, and V. Dunjko (2026), npj Quantum Inf. **12** (19), 19.

Monbroussou, L., B. Polacchi, V. Yacoub, E. Caruccio, G. Rodari, F. Hoch, G. Carvacho, N. Spagnolo, T. Giordani, M. Bossi, *et al.* (2025), *Advanced Photonics, Vol. 7, Issue 6*, Adv. Photonics **7** (6), 066012.

Monroe, C., D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland (1995), Phys. Rev. Lett. **75**, 4714.

Montanez-Barrera, J., Y. Ji, M. R. von Spakovsky, D. E. B. Neira, and K. Michielsen (2025), "Optimizing qaoa circuit transpilation with parity twine and swap network encodings," arXiv:2505.17944.

Moreau, G. S., L. Pisani, M. Profir, C. Podda, L. Leoni, and G. Cao (2025), Adv. Quantum Technol. **8** (10), 2400716.

Mori, K., T. Isokawa, N. Kouda, N. Matsui, and H. Nishimura (2006), in *The 2006 IEEE International Joint Conference on Neural Network Proceedings* (IEEE) pp. 16–21.

Moussa, C., Y. J. Patel, V. Dunjko, T. Bäck, and J. N. van Rijn (2024), Mach. Learn. **113** (4), 1941.

Moussa, C., J. N. van Rijn, T. Bäck, and V. Dunjko (2022), in *Discovery Science: 25th International Conference, DS 2022, Montpellier, France, October 10–12, 2022, Proceedings* (Springer-Verlag, Berlin, Heidelberg) p. 32–46.

Mujal, P., R. Martínez-Peña, J. Nokkala, J. García-Beni, G. L. Giorgi, M. C. Soriano, and R. Zambrini (2021), Adv. Quantum Technol. **4** (8), 2100027.

Mukhopadhyay, P. (2025), Scientific Reports **15** (1), 11002.

Muller, C. L., B. Baumgartner, and I. F. Sbalzarini (2009), in *2009 IEEE Congress on Evolutionary Computation*, pp. 2685–2692.

Munson, J. K., and A. I. Rubin (1959), IRE Transactions on Electronic Computers **EC-8** (2), 200.

Murali, P., J. M. Baker, A. Javadi-Abhari, F. T. Chong, and M. Martonosi (2019) (Association for Computing Machinery, New York, NY, USA) p. 1015–1029.

Nagai, R., T. Takemoto, Y. Wachi, and H. Mizuno (2025), "Digital-controlled method of conveyor-belt spin shuttling in silicon for large-scale quantum computation," arXiv:2502.20955.

Nagy, D. T. R., C. Czabán, B. Bakó, P. Hága, Z. Kallus, and Z. Zimborás (2025), Quantum Mach. Intell. **7** (2), 88.

Nakaji, K., H. Tezuka, and N. Yamamoto (2023), Quantum Sci. Technol. **9** (1), 015022.

Nakaji, K., S. Uno, Y. Suzuki, R. Raymond, T. Onodera, T. Tanaka, H. Tezuka, N. Mitsuda, and N. Yamamoto (2022), Phys. Rev. Res. **4**, 023136.

Nakamura, Y., Yu. A. Pashkin, and J. S. Tsai (1999), Nature **398** (6730), 786.

Nakkiran, P., G. Kaplun, Y. Bansal, T. Yang, B. Barak, and

I. Sutskever (2021), J. Stat. Mech.: Theory Exp. **2021** (12), 124003.
Nation, P. D., A. A. Saki, S. Brandhofer, L. Bello, S. Garion, M. Treinish, and A. Javadi-Abhari (2025), Nat. Comput. Sci. **5** (5), 427.
Nausheen, F., K. Ahmed, and M. I. Khan (2025), "Quantum natural language processing: A comprehensive review of models, methods, and applications," arXiv:2504.09909.
Nelder, J. A., and R. Mead (1965), The Computer Journal **7** (4), 308.
Nelson, J., M. Vuffray, A. Y. Lokhov, T. Albash, and C. Coffrin (2022), Phys. Rev. Appl. **17**, 044046.
Nemkov, N. A., E. O. Kiktenko, and A. K. Fedorov (2025), Phys. Rev. A **111**, 012441.
Nguyen, D. C., M. R. Uddin, S. Shaon, R. Rahman, O. Dobre, and D. Niyato (2025), "Quantum federated learning: A comprehensive survey," arXiv:2508.15998.
Nguyen, N., and K.-C. Chen (2022), IEEE Access **10**, 41444.
Nguyen, Q. T., L. Schatzki, P. Braccia, M. Ragone, P. J. Coles, F. Sauvage, M. Larocca, and M. Cerezo (2024), PRX Quantum **5**, 020328.
Nielsen, M. A., and I. L. Chuang (2010), *Quantum computation and quantum information* (Cambridge university press).
Nietner, A., M. Ioannou, R. Sweke, R. Kueng, J. Eisert, M. Hinsche, and J. Haferkamp (2025), Quantum **9**, 1883.
Niu, M. Y., S. Boixo, V. N. Smelyanskiy, and H. Neven (2019), npj Quantum Inf. **5** (33), 33.
Noiri, A., K. Takeda, T. Nakajima, T. Kobayashi, A. Sammak, G. Scappucci, and S. Tarucha (2022), Nature **601** (7893), 338.
Notton, C., V. Apostolou, A. Senellart, A. Walsh, D. Wang, Y. Xie, S. Yang, I. Mejdoub, O. Zouhry, K.-C. Chen, *et al.* (2025), "Establishing baselines for photonic quantum machine learning: Insights from an open, collaborative initiative," arXiv:2510.25839.
Novikov, A., D. Podoprikhin, A. Osokin, and D. Vetrov (2015), *Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'15, 442–450.
Oh, C., S. Chen, Y. Wong, S. Zhou, H.-Y. Huang, J. A. H. Nielsen, Z.-H. Liu, J. S. Neergaard-Nielsen, U. L. Andersen, L. Jiang, and J. Preskill (2024), Phys. Rev. Lett. **133**, 230604.
Ohno, H. (2025), Quantum Inf. Process. **24** (11), 365.
Oliveira, R., L. Ott, and F. Ramos (2019), in *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, Vol. 89, edited by K. Chaudhuri and M. Sugiyama (PMLR) pp. 1177–1184.
Orka, N. A., M. A. Awal, P. Liò, G. Pogrebna, A. G. Ross, and M. A. Moni (2025), Artificial Intelligence Review **58** (5), 134.
Ortiz Marrero, C., M. Kieferová, and N. Wiebe (2021), PRX Quantum **2**, 040316.
Orús, R. (2014), Annals of physics **349**, 117.
Oseledets, I. V. (2011), SIAM Journal on Scientific Computing **33** (5), 2295.
Ostaszewski, M., E. Grant, and M. Benedetti (2021), Quantum **5**, 391.
Paetznick, A., M. da Silva, C. Ryan-Anderson, J. Bello-Rivas, J. Campora III, A. Chernoguzov, J. Dreiling, C. Foltz, F. Frachon, J. Gaebler, *et al.* (2024), "Demonstration of logical qubits and repeated error correction with better-than-physical error rates," arXiv:2404.02280.
Palacios, A., R. Martínez-Peña, M. C. Soriano, G. L. Giorgi, and R. Zambrini (2024), Communications Physics **7** (1), 369.
Pallavi, G., and R. Prasanna Kumar (2025), Front. Comput. Sci. **7**, 1464122.
Pan, J., Y. Cao, X. Yao, Z. Li, C. Ju, H. Chen, X. Peng, S. Kais, and J. Du (2014), Phys. Rev. A **89**, 022313.
Pan, X., X. Cao, W. Wang, Z. Hua, W. Cai, X. Li, H. Wang, J. Hu, Y. Song, D.-L. Deng, *et al.* (2023a), npj Quantum Inf. **9** (18), 18.
Pan, X., Z. Lu, W. Wang, Z. Hua, Y. Xu, W. Li, W. Cai, X. Li, H. Wang, Y.-P. Song, *et al.* (2023b), Nat. Commun. **14** (4006), 4006.
Panadero, I., Y. Ban, H. Espinós, R. Puebla, J. Casanova, and E. Torrontegui (2024), Scientific Reports **14** (1), 31669.
Pappalardo, A., P.-E. Emeriau, G. de Felice, B. Ventura, H. Jaunin, R. Yeung, B. Coecke, and S. Mansfield (2025), Phys. Rev. A **111**, 032429.
Parcollet, T., M. Morchid, and G. Linarès (2020), Artif. Intell. Rev. **53** (4), 2957.
Parcollet, T., M. Ravanelli, M. Morchid, G. Linarès, C. Trabelsi, R. De Mori, and Y. Bengio (2018), "Quaternion recurrent neural networks," arXiv:1806.04418.
Paszke, A., S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, *et al.* (2019), "Pytorch: an imperative style, high-performance deep learning library," in *Proceedings of the 33rd International Conference on Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, USA).
Patterson, D., J. Gonzalez, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. So, M. Texier, and J. Dean (2021), "Carbon Emissions and Large Neural Network Training," arXiv:2104.10350.
Patti, T. L., K. Najafi, X. Gao, and S. F. Yelin (2021), Phys. Rev. Res. **3**, 033090.
Pérez-López, D., and L. Torrijos-Morán (2025), npj Nanophoton. **2** (32), 32.
Pelofske, E. (2025), Quantum Sci. Technol. **10** (2), 025025.
Pepper, A., N. Tischler, and G. J. Pryde (2019), Phys. Rev. Lett. **122**, 060501.
Peral-García, D., J. Cruz-Benito, and F. J. García-Peñalvo (2024), Computer Science Review **51**, 100619.
Perdomo-Ortiz, A., M. Benedetti, J. Realpe-Gómez, and R. Biswas (2018), Quantum Sci. Technol. **3** (3), 030502.
Peres, A. (2002), *Quantum theory: concepts and methods* (Springer).
Perez-Garcia, D., F. Verstraete, M. M. Wolf, and J. I. Cirac (2007), Quantum Info. Comput. **7** (5), 401–430.
Pérez-Salinas, A., A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre (2020), Quantum **4**, 226.
Peruzzo, A., J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien (2014), Nat. Commun. **5** (4213), 4213.
Pesah, A., M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles (2021), Phys. Rev. X **11**, 041011.
Peters, E., J. Caldeira, A. Ho, S. Leichenauer, M. Mohseni, H. Neven, P. Spentzouris, D. Strain, and G. N. Perdue (2021), npj Quantum Inf. **7** (161), 161.
Peters, E., and M. Schuld (2023), Quantum **7**, 1210.
Petit, L., M. Russ, G. H. Eenink, W. I. Lawrie, J. S. Clarke, L. M. Vandersypen, and M. Veldhorst (2022), Communications Materials **3** (1), 82.

Pfau, D., J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes (2020), Phys. Rev. Res. **2**, 033429.

Philips, S. G. J., M. T. Madzik, S. V. Amitonov, S. L. de Snoo, M. Russ, N. Kalhor, C. Volk, W. I. L. Lawrie, D. Brousse, L. Tryputen, *et al.* (2022), Nature **609** (7929), 919.

Pichard, G., D. Lim, E. Bloch, J. Vaneecloo, L. Bourachot, G.-J. Both, G. Mériaux, S. Dutartre, R. Hostein, J. Paris, B. Ximenez, A. Signoles, A. Browaeys, T. Lahaye, and D. Dreon (2024), Phys. Rev. Appl. **22**, 024073.

Pierangeli, D., D. Pierangeli, D. Pierangeli, G. Marcucci, C. Conti, C. Conti, and C. Conti (2021), Photonics Res. **9** (8), 1446.

Pirnay, N., S. Jerbi, J.-P. Seifert, and J. Eisert (2024), "An unconditional distribution learning advantage with shallow quantum circuits," arXiv:2411.15548.

Pirnay, N., R. Sweke, J. Eisert, and J.-P. Seifert (2023), Phys. Rev. A **107**, 042416.

Pommerening, J. C., and D. P. DiVincenzo (2020), Phys. Rev. A **102**, 032623.

Power, A., Y. Burda, H. Edwards, I. Babuschkin, and V. Misra (2022), "Grokking: Generalization beyond overfitting on small algorithmic datasets," arXiv:2201.02177.

Preskill, J. (1997), "Fault-tolerant quantum computation," arXiv:quant-ph/9712048.

Preskill, J. (2018), Quantum **2**, 79.

Preskill, J. (2025), ACM Transactions on Quantum Computing **6** (3), 10.1145/3723153.

Proctor, T., M. Revelle, E. Nielsen, K. Rudinger, D. Lobser, P. Maunz, R. Blume-Kohout, and K. Young (2020), Nat. Commun. **11** (5396), 5396.

Proctor, T., K. Young, A. D. Baczewski, and R. Blume-Kohout (2025), Nat. Rev. Phys. **7** (2), 105.

Puig, R., M. Drudis, S. Thanasilp, and Z. Holmes (2025), PRX Quantum **6**, 010317.

Purushothaman, G., and N. Karayiannis (1997), IEEE Transactions on Neural Networks **8** (3), 679.

Puterman, M. L. (1994), *Markov Decision Processes* (John Wiley & Sons, Ltd., Chichester, England, UK).

Quek, Y., D. Stilck França, S. Khatri, J. J. Meyer, and J. Eisert (2024), Nat. Phys. **20** (10), 1648.

Quinton, F. A., P. A. S. Myhr, M. Barani, P. Crespo del Granado, and H. Zhang (2025), Scientific Reports **15** (1), 12733.

Ragone, M., B. N. Bakalov, F. Sauvage, A. F. Kemper, C. Ortiz Marrero, M. Larocca, and M. Cerezo (2024), Nat. Commun. **15** (7172), 7172.

Raissi, M., P. Perdikaris, and G. E. Karniadakis (2019), J. Comput. Phys. **378**, 686.

Rajak, A., S. Suzuki, A. Dutta, and B. K. Chakrabarti (2022), Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **381** (2241), 20210417.

Rakshit, P., A. Konar, and S. Das (2017), Swarm and Evolutionary Computation **33**, 18.

Rall, P., D. Liang, J. Cook, and W. Kretschmer (2019), Phys. Rev. A **99**, 062337.

Rambach, M., A. Roy, A. Gilchrist, A. Sakurai, W. J. Munro, K. Nemoto, and A. G. White (2025), "Photonic quantum-accelerated machine learning," arXiv:2512.08318.

Ramsauer, H., B. Schäfl, J. Lehner, P. Seidl, M. Widrich, T. Adler, L. Gruber, M. Holzleitner, M. Pavlović, G. K. Sandve, *et al.* (2020), "Hopfield networks is all you need," arXiv:2008.02217.

Rapp, F., D. A. Kreplin, M. F. Huber, and M. Roth (2025), Mach. Learn.: Sci. Technol. **6** (1), 015041.

Rath, M., and H. Date (2024), EPJ Quantum Technol. **11** (1), 72.

Rattew, A. G., P.-W. Huang, N. Guo, L. Pira, and P. Rebentrost (2025), "Accelerating inference for multilayer neural networks with quantum computers," arXiv:2510.07195.

Rebentrost, P., M. Mohseni, and S. Lloyd (2014), Phys. Rev. Lett. **113**, 130503.

Recio-Armengol, E., S. Ahmed, and J. Bowles (2025a), "Train on classical, deploy on quantum: scaling generative quantum machine learning to a thousand qubits," arXiv:2503.02934.

Recio-Armengol, E., J. Eisert, and J. J. Meyer (2025b), Phys. Rev. A **111**, 042420.

Recio-Armengol, E., F. J. Schreiber, J. Eisert, and C. Bravo-Prieto (2024), arXiv:2411.11954.

Reddy, P., and A. B. Bhattacherjee (2021), Mach. Learn.: Sci. Technol. **2** (2), 025019.

Reich, D. M., M. Ndong, and C. P. Koch (2012), The Journal of Chemical Physics **136** (10), 104103.

Ren, W., W. Li, S. Xu, K. Wang, W. Jiang, F. Jin, X. Zhu, J. Chen, Z. Song, P. Zhang, *et al.* (2022), Nat. Comput. Sci. **2** (11), 711.

Richter, J., J. Shi, J.-J. Chen, J. Rahnenführer, and M. Lang (2020), in *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, GECCO '20 (Association for Computing Machinery, New York, NY, USA) p. 877–885.

Rieser, H.-M., F. Köster, and A. P. Raulf (2023), Proc. A **479** (2275), 10.1098/rspa.2023.0218.

Ristè, D., M. P. da Silva, C. A. Ryan, A. W. Cross, A. D. Córcoles, J. A. Smolin, J. M. Gambetta, J. M. Chow, and B. R. Johnson (2017), npj Quantum Inf. **3** (16), 16.

Rodríguez-Díaz, F., D. Gutiérrez-Avilés, A. Troncoso, and F. Martínez-Álvarez (2025), ACM Comput. Surv. **58** (4), 10.1145/3764582.

Rodriguez-Grasa, P., Y. Ban, and M. Sanz (2025), Phys. Rev. Res. **7**, 023269.

Romero, J., and A. Aspuru-Guzik (2021), Adv. Quantum Technol. **4** (1), 2000003.

Romero, J., and G. Milburn (2025), "Photonic quantum computing," .

Romero, J., J. P. Olson, and A. Aspuru-Guzik (2017), Quantum Sci. Technol. **2** (4), 045001.

Rosenblatt, F. (1958), Psychol. Rev. **65** (6), 386, 13602029.

Rosenkranz, M., E. Brunner, G. Marin-Sanchez, N. Fitzpatrick, S. Dilkes, Y. Tang, Y. Kikuchi, and M. Benedetti (2025), Quantum **9**, 1703.

Ruderman, D. L., and W. Bialek (1994), Phys. Rev. Lett. **73**, 814.

Rudolph, M. S., T. Jones, Y. Teng, A. Angrisani, and Z. Holmes (2025), "Pauli propagation: A computational framework for simulating quantum systems," arXiv:2505.21606.

Rudolph, M. S., S. Lerch, S. Thanasilp, O. Kiss, O. Shaya, S. Vallecorsa, M. Grossi, and Z. Holmes (2024), npj Quantum Inf. **10** (116), 116.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Nature **323** (6088), 533.

Ryan-Anderson, C., J. G. Bohnet, K. Lee, D. Gresh, A. Hankin, J. P. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N. C. Brown, T. M. Gatterman, S. K. Halit, K. Gilmore, J. A. Gerber, B. Neyenhuis, D. Hayes, and R. P. Stutz (2021), Phys. Rev. X **11**, 041058.

Sabarad, V., V. Varma, and T. Mahesh (2024), "Experimental

machine learning with classical and quantum data via nmr quantum kernels," arXiv:2412.09557.
Sack, S. H., R. A. Medina, A. A. Michailidis, R. Kueng, and M. Serbyn (2022), PRX Quantum **3**, 020365.
Saffman, M. (2025), "Quantum computing with atomic qubit arrays: confronting the cost of connectivity," arXiv:2505.11218.
Saffman, M., T. G. Walker, and K. Mølmer (2010), Rev. Mod. Phys. **82**, 2313.
Saggio, V., B. E. Asenbeck, A. Hamann, T. Strömberg, P. Schiansky, V. Dunjko, N. Friis, N. C. Harris, M. Hochberg, D. Englund, *et al.* (2021), Nature **591** (7849), 229.
Sagingalieva, A., M. Kordzanganeh, N. Kenbayev, D. Kosichkina, T. Tomashuk, and A. Melnikov (2023a), Cancers **15** (10), 10.3390/cancers15102705.
Sagingalieva, A., M. Kordzanganeh, A. Kurkin, A. Melnikov, D. Kuhmistrov, M. Perelshtein, A. Melnikov, A. Skolik, and D. V. Dollen (2023b), Quantum Mach. Intell. **5** (2), 38.
Sahebi, M., A. Barthe, Y. Suzuki, Z. Holmes, and M. Grossi (2025), "On dequantization of supervised quantum machine learning via random fourier features," arXiv:2505.15902.
Sahin, M. E., E. Altamura, O. Wallis, S. P. Wood, A. Dekusar, D. A. Millar, T. Imamichi, A. Matsuo, and S. Mensa (2025), "Qiskit machine learning: an open-source library for quantum machine learning tasks at scale on quantum hardware and classical simulators," arXiv:2505.17756.
Sajjan, M., J. Li, R. Selvarajan, S. H. Sureshbabu, S. S. Kale, R. Gupta, V. Singh, and S. Kais (2022), Chem. Soc. Rev. **51** (15), 6475.
Sakurai, A., A. Hayashi, W. J. Munro, and K. Nemoto (2025), Optica Quantum **3** (3), 238.
Salavrakos, A., T. Sedrakyan, J. Mills, S. Mansfield, and R. Mezher (2025), Phys. Rev. A **111**, L030401.
Salloum, H., A. Salloum, M. Mazzara, and S. Zykov (2024), Procedia Computer Science **246**, 3285.
Sander, L. C., N. A. McMahon, P. Zapletal, and M. J. Hartmann (2025), Phys. Rev. Res. **7**, L042032.
Sarma, B., and M. J. Hartmann (2025), Phys. Rev. Appl. **23**, 014015.
Scardapane, S., S. Van Vaerenbergh, A. Hussain, and A. Uncini (2020), IEEE Transactions on Emerging Topics in Computational Intelligence **4** (2), 140.
Schatzki, L., M. Larocca, Q. T. Nguyen, F. Sauvage, and M. Cerezo (2024), npj Quantum Inf. **10** (12), 12.
Schetakis, N., P. Bonfini, N. Alisoltani, K. Blazakis, S. I. Tsintzos, A. Askitopoulos, D. Aghamalyan, P. Fafoutellis, and E. I. Vlahogianni (2025), Scientific Reports **15** (1), 19400.
Schnabel, J., and M. Roth (2025), Quantum Mach. Intell. **7** (1), 58.
Scholl, P., M. Schuler, H. J. Williams, A. A. Eberharter, D. Barredo, K.-N. Schymik, V. Lienhard, L.-P. Henry, T. C. Lang, T. Lahaye, *et al.* (2021), Nature **595** (7866), 233.
Schollwöck, U. (2011), Annals of physics **326** (1), 96.
Schreiber, F. J., J. Eisert, and J. J. Meyer (2023), Phys. Rev. Lett. **131**, 100803.
Schuch, N., M. M. Wolf, F. Verstraete, and J. I. Cirac (2007), Phys. Rev. Lett. **98**, 140506.
Schuld, M. (2021), "Supervised quantum machine learning models are kernel methods," arXiv:2101.11020.
Schuld, M., V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran (2019), Phys. Rev. A **99**, 032331.
Schuld, M., A. Bocharov, K. M. Svore, and N. Wiebe (2020a), Phys. Rev. A **101**, 032308.
Schuld, M., K. Brádler, R. Israel, D. Su, and B. Gupt (2020b), Phys. Rev. A **101**, 032314.
Schuld, M., and N. Killoran (2019), Phys. Rev. Lett. **122**, 040504.
Schuld, M., and N. Killoran (2022), PRX Quantum **3**, 030101.
Schuld, M., and F. Petruccione (2018), SpringerLink .
Schuld, M., and F. Petruccione (2021), *Machine Learning with Quantum Computers* (Springer International Publishing, Cham, Switzerland).
Schuld, M., I. Sinayskiy, and F. Petruccione (2014), Quantum Inf. Process. **13** (11), 2567.
Schuld, M., I. Sinayskiy, and F. Petruccione (2015), Contemporary Physics **56** (2), 172.
Schuld, M., R. Sweke, and J. J. Meyer (2021), Phys. Rev. A **103**, 032430.
Schulz, S., D. Willsch, and K. Michielsen (2025), Quantum **9**, 1898.
Schumann, M., F. K. Wilhelm, and A. Ciani (2024), Quantum Sci. Technol. **9** (4), 045019.
Schütt, K. T., P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller (2017), *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, 992–1002.
Schwartz, R., J. Dodge, N. A. Smith, and O. Etzioni (2020), Commun. ACM **63** (12), 54–63.
Scriva, G., N. Astrakhantsev, S. Pilati, and G. Mazzola (2024), Phys. Rev. A **109**, 032408.
Secchi, A., G. Forghieri, P. Bordone, D. Loss, S. Bosco, and F. Troiani (2025), "Hole-spin qubits in germanium beyond the single-particle regime," arXiv:2505.02449.
Sedrakyan, T., T. Sedrakyan, and A. Salavrakos (2024), Optica Quantum **2** (6), 458.
Senanian, A., S. Prabhu, V. Kremenetski, S. Roy, Y. Cao, J. Kline, T. Onodera, L. G. Wright, X. Wu, V. Fatemi, *et al.* (2024), Nat. Commun. **15** (7490), 7490.
Senokosov, A., A. Sedykh, A. Sagingalieva, B. Kyriacou, and A. Melnikov (2024), Mach. Learn.: Sci. Technol. **5** (1), 015040.
Sentís, G., A. Monràs, R. Muñoz Tapia, J. Calsamiglia, and E. Bagan (2019), Phys. Rev. X **9**, 041029.
Shahriari, B., K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas (2016), Proceedings of the IEEE **104** (1), 148.
Shen, K., A. Kurkin, A. Pérez-Salinas, E. Shishenina, V. Dunjko, and H. Wang (2025), npj Quantum Inf. **11** (178), 178.
Shi, J., T. Chen, W. Lai, S. Zhang, and X. Li (2024), IEEE Transactions on Cybernetics **54** (10), 5973.
Shi, J., Z. Li, W. Lai, F. Li, R. Shi, Y. Feng, and S. Zhang (2023), IEEE Transactions on Knowledge and Data Engineering **35** (4), 4335.
Shi, Y.-Y., L.-M. Duan, and G. Vidal (2006), Phys. Rev. A **74**, 022320.
Shin, S., Y. S. Teo, and H. Jeong (2023), Phys. Rev. A **107**, 012422.
Shin, S., Y. S. Teo, and H. Jeong (2024), Phys. Rev. Res. **6**, 023218.
Shingu, Y., Y. Seki, S. Watabe, S. Endo, Y. Matsuzaki, S. Kawabata, T. Nikuni, and H. Hakoshima (2021), Phys. Rev. A **104**, 032413.
Shor, P. (1996), in *Proceedings of 37th Conference on Foundations of Computer Science*, pp. 56–65.
Shor, P. W. (1994), in *Proceedings 35th annual symposium on*

foundations of computer science (Ieee) pp. 124–134.
Shor, P. W. (1995), Phys. Rev. A **52**, R2493.
Shor, P. W. (1997), SIAM Journal on Computing **26** (5), 1484.
Silver, D., A. Ranjan, R. Achutha, T. Patel, and D. Tiwari (2024), in *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis*, SC '24 (IEEE Press).
Sim, S., P. D. Johnson, and A. Aspuru-Guzik (2019), Adv. Quantum Technol. **2** (12), 1900070.
Simoncelli, E. P., and B. A. Olshausen (2001), Annu. Rev. Neurosci. **24:1193-216.**, 10.1146/annurev.neuro.24.1.1193.
Simonyan, K., and A. Zisserman (2014), "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556.
Singkanipa, P., and D. A. Lidar (2025), Quantum **9**, 1617.
Sivak, V. V., A. Eickbusch, H. Liu, B. Royer, I. Tsioutsios, and M. H. Devoret (2022), Phys. Rev. X **12**, 011059.
Sivarajah, S., S. Dilkes, A. Cowtan, W. Simmons, A. Edgington, and R. Duncan (2020), Quantum Sci. Technol. **6** (1), 014003.
Skolik, A., M. Cattelan, S. Yarkoni, T. Bäck, and V. Dunjko (2023a), npj Quantum Inf. **9** (47), 47.
Skolik, A., S. Jerbi, and V. Dunjko (2022), Quantum **6**, 720.
Skolik, A., S. Mangini, T. Bäck, C. Macchiavello, and V. Dunjko (2023b), EPJ Quantum Technol. **10** (1), 8.
Skolik, A., J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib (2021), Quantum Mach. Intell. **3** (1), 5.
Slysz, M., P. Rydlichowski, K. Kurowski, O. Bacarreza, E. C. Gomez, Z. Chandani, B. Heim, P. Khalate, W. R. Clements, and J. Fletcher (2025), "Hybrid classical-quantum supercomputing: A demonstration of a multi-user, multi-qpu and multi-gpu environment," arXiv:2508.16297.
Smaldone, A. M., G. W. Kyro, and V. S. Batista (2023), Quantum Mach. Intell. **5** (2), 41.
Sóbester, A., S. J. Leary, and A. J. Keane (2004), Structural and multidisciplinary optimization **27** (5), 371.
Sohail, M. A., M. Heidari, and S. S. Pradhan (2025), npj Quantum Inf. **11** (109), 109.
Sohl-Dickstein, J., E. A. Weiss, N. Maheswaranathan, and S. Ganguli (2015), in *Proceedings of the 32nd International Conference on Machine Learning - Volume 37*, ICML'15 (JMLR.org) p. 2256–2265.
Song, Y., J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole (2021), "Score-based generative modeling through stochastic differential equations," arXiv:2011.13456.
Sordoni, A., Y. Bengio, and J.-Y. Nie (2014), in *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, AAAI'14 (AAAI Press) p. 1586–1592.
Sordoni, A., and J.-Y. Nie (2014), "Looking at vector space and language models for ir using density matrices," arXiv:1401.1732.
Sordoni, A., J.-Y. Nie, and Y. Bengio (2013), in *Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '13 (Association for Computing Machinery, New York, NY, USA) p. 653–662.
Sørensen, A., and K. Mølmer (1999), Phys. Rev. Lett. **82**, 1971.
Spall, J. (1992), IEEE Transactions on Automatic Control **37** (3), 332.
Spall, J. (1998), IEEE Transactions on Aerospace and Electronic Systems **34** (3), 817.
Stano, P., and D. Loss (2022), Nat. Rev. Phys. **4** (10), 672.
Steane, A. M. (1996), Phys. Rev. Lett. **77**, 793.
Steinacker, P., N. Dumoulin Stuyck, W. H. Lim, T. Tanttu, M. Feng, S. Serrano, A. Nickl, M. Candido, J. D. Cifuentes, E. Vahapoglu, *et al.* (2025), Nature **646** (8083), 81.
Stephens, C. P., and W. Baritompa (1998), J. Optim. Theory Appl. **96** (3), 575.
Stokes, J., J. Izaac, N. Killoran, and G. Carleo (2020), Quantum **4**, 269.
Storn, R., and K. Price (1997), J. Global Optim. **11** (4), 341.
Stoudenmire, E. M. (2018), Quantum Sci. Technol. **3** (3), 034003.
Stoudenmire, E. M., and D. J. Schwab (2016), *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, 4806–4814.
Strohm, T., K. Wintersperger, F. Dommert, D. Basilewitsch, G. Reuber, A. Hoursanov, T. Ehmer, D. Vodola, and S. Luber (2024), "Ion-based quantum computing hardware: Performance and end-user perspective," arXiv:2405.11450.
Strubell, E., A. Ganesh, and A. McCallum (2020), AAAI **34** (09), 13693.
Sugunapriya, A., and S. Markkandan (2025), Scientific Reports **15**, 38551.
Sun, Y., Y. Zeng, and T. Zhang (2021), iScience **24** (8), 102880.
Suprano, A., D. Zia, L. Innocenti, S. Lorenzo, V. Cimini, T. Giordani, I. Palmisano, E. Polino, N. Spagnolo, F. Sciarrino, G. M. Palma, A. Ferraro, and M. Paternostro (2024), Phys. Rev. Lett. **132**, 160802.
Sutton, R. S., and A. G. Barto (2018), *Reinforcement Learning: An Introduction*, 2nd ed. (MIT Press, Cambridge, MA).
Sutton, R. S., D. McAllester, S. Singh, and Y. Mansour (1999), *Proceedings of the 13th International Conference on Neural Information Processing Systems*, NIPS'99, 1057–1063.
Suzuki, M. (1976a), Commun. Math. Phys. **51** (2), 183.
Suzuki, M. (1976b), Progress of theoretical physics **56** (5), 1454.
Suzuki, M. (1985), J. Math. Phys. **26** (4), 601.
Suzuki, Y., Y. Kawase, Y. Masumura, Y. Hiraga, M. Nakadai, J. Chen, K. M. Nakanishi, K. Mitarai, R. Imai, S. Tamiya, *et al.* (2021), Quantum **5**, 559.
Sweke, R., E. Recio-Armengol, S. Jerbi, E. Gil-Fuster, B. Fuller, J. Eisert, and J. J. Meyer (2025), Quantum **9**, 1640.
Sweke, R., J.-P. Seifert, D. Hangleiter, and J. Eisert (2021), Quantum **5**, 417.
Sweke, R., F. Wilde, J. Meyer, M. Schuld, P. K. Faehrmann, B. Meynard-Piganeau, and J. Eisert (2020), Quantum **4**, 314.
Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich (2015), in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9.
Tacchino, F., C. Macchiavello, D. Gerace, and D. Bajoni (2019), npj Quantum Inf. **5** (26), 26.
Takagi, R., S. Endo, S. Minagawa, and M. Gu (2022), npj Quantum Inf. **8** (114), 114.
Takagi, R., H. Tajima, and M. Gu (2023), Phys. Rev. Lett. **131**, 210602.
Takeda, K., A. Noiri, T. Nakajima, T. Kobayashi, and S. Tarucha (2022), Nature **608** (7924), 682.

Tang, E. (2019), in *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019 (Association for Computing Machinery, New York, NY, USA) p. 217–228.

Tang, E. (2021), Phys. Rev. Lett. **127**, 060503.

Tang, E. (2022), Nat. Rev. Phys. **4** (11), 692.

Tang, Y., and J. Yan (2022), in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22 (Curran Associates Inc., Red Hook, NY, USA).

Tannor, D. J., V. Kazakov, and V. Orlov (1992), in *Time-Dependent Quantum Molecular Dynamics* (Springer, Boston, MA, Boston, MA, USA) pp. 347–360.

Tao, H.-X., J. Hu, and R.-B. Wu (2025a), Quantum Mach. Intell. **7** (2), 109.

Tao, H.-X., X. Wang, and R.-B. Wu (2025b), "On the design of expressive and trainable pulse-based quantum machine learning models," arXiv:2508.05559.

Tasnim, T., A. Saha, M. Rahman, and F. Wu (2025), in *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)* (IEEE) pp. 00203–00208.

Terhal, B. M. (2015), Rev. Mod. Phys. **87**, 307.

Thanasilp, S., S. Wang, M. Cerezo, and Z. Holmes (2024), Nat. Commun. **15** (5200), 5200.

Thanasilp, S., S. Wang, N. A. Nghiem, P. Coles, and M. Cerezo (2023), Quantum Mach. Intell. **5** (1), 21.

The Royal Swedish Academy of Sciences, (2024), "Scientific background on the nobel prize in physics 2024: Foundational discoveries and inventions that enable machine learning with artificial neural networks," .

The Royal Swedish Academy of Sciences, (2025), "Scientific background on the Nobel Prize in Physics 2025: Macroscopic quantum mechanical tunnelling and energy quantisation in an electric circuit," .

Theis, T. N., and H.-S. P. Wong (2017), Computing in science & engineering **19** (2), 41.

Tian, J., X. Sun, Y. Du, S. Zhao, Q. Liu, K. Zhang, W. Yi, W. Huang, C. Wang, X. Wu, *et al.* (2023), IEEE Transactions on Pattern Analysis and Machine Intelligence **45** (10), 12321.

Tillich, J.-P., and G. Zémor (2014), IEEE Transactions on Information Theory **60** (2), 1193.

Tilly, J., H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, *et al.* (2022), Phys. Rep. **986**, 1.

Tiwari, P., and M. Melucci (2019), IEEE Access **7**, 42354.

Tomal, S., A. A. Shafin, D. Bhattacharjee, M. Amin, and R. S. Shahir (2025), "Quantum-enhanced attention mechanism in nlp: A hybrid classical-quantum approach," arXiv:2501.15630.

Torlai, G., G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo (2018), Nat. Phys. **14** (5), 447.

Trabelsi, C., O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal (2017), "Deep complex networks," arXiv:1705.09792.

Trugenberger, C. A. (2001), Phys. Rev. Lett. **87**, 067901.

Tunyasuvunakool, K., J. Adler, Z. Wu, T. Green, M. Zielinski, A. Žídek, A. Bridgland, A. Cowie, C. Meyer, A. Laydon, *et al.* (2021), Nature **596** (7873), 590.

Ullah, U., and B. Garcia-Zapirain (2024), IEEE Access **12**, 11423.

Vagniluca, I., B. Da Lio, D. Rusca, D. Cozzolino, Y. Ding, H. Zbinden, A. Zavatta, L. K. Oxenløwe, and D. Bacco (2020), Phys. Rev. Appl. **14**, 014051.

Vapnik, V. N. (1998), *Statistical Learning Theory* (Wiley).

Vargas-Calderón, V. (2025), Quantum Mach. Intell. **7** (2), 65.

Varmantchaonala, C. M., J. L. K. E. Fendji, J. Schöning, and M. Atemkeng (2024), IEEE Access **12**, 99578.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin (2017), in *Advances in Neural Information Processing Systems*, Vol. 30, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Curran Associates, Inc.).

Ventura, D., and T. Martinez (2000), Inform. Sci. **124** (1), 273.

Verdon, G., J. Marks, S. Nanda, S. Leichenauer, and J. Hidary (2019a), "Quantum hamiltonian-based models and the variational quantum thermalizer algorithm," arXiv:1910.02071.

Verdon, G., T. McCourt, E. Luzhnica, V. Singh, S. Leichenauer, and J. Hidary (2019b), "Quantum graph neural networks," arXiv:1909.12264.

Verstraete, F., and J. I. Cirac (2004), arXiv:cond-mat/0407066.

Verstraete, F., V. Murg, and J. I. Cirac (2008), Advances in physics **57** (2), 143.

Verstraete, F., D. Porras, and J. I. Cirac (2004), Phys. Rev. Lett. **93**, 227205.

Verstraete, H. R. G. W., S. Wahls, J. Kalkman, and M. Verhaegen (2015), Opt. Lett. **40** (24), 5722.

Verteletskyi, V., T.-C. Yen, and A. F. Izmaylov (2020), The Journal of chemical physics **152** (12).

Vidal, G. (2007), Phys. Rev. Lett. **99**, 220405.

Vodeb, J., J.-Y. Desaules, A. Hallam, A. Rava, G. Humar, D. Willsch, F. Jin, M. Willsch, K. Michielsen, and Z. Papić (2025), Nat. Phys. **21**, 386.

Von Neumann, J. (1955), *Mathematical Foundations of Quantum Mechanics* (Princeton university press).

Wadhwa, C., and M. Doosti (2025), Quantum **9**, 1739.

Wall, M. L., M. R. Abernathy, and G. Quiroz (2021), Phys. Rev. Res. **3**, 023010.

Wang, J., F. Sciarrino, A. Laing, and M. G. Thompson (2020), Nat. Photonics **14** (5), 273.

Wang, M., J. Jiang, and P. V. Coveney (2025a), Phys. Rev. Res. **7**, 043094.

Wang, P., C.-Y. Luan, M. Qiao, M. Um, J. Zhang, Y. Wang, X. Yuan, M. Gu, J. Zhang, and K. Kim (2021a), Nat. Commun. **12** (233), 233.

Wang, R., Z. Xia, G. Yan, and J. Yan (2025b), in *Forty-second International Conference on Machine Learning*.

Wang, S., P. Czarnik, A. Arrasmith, M. Cerezo, L. Cincio, and P. J. Coles (2024a), Quantum **8**, 1287.

Wang, S., E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles (2021b), Nat. Commun. **12** (6961), 6961.

Wang, X., Y. Du, Y. Luo, and D. Tao (2021c), Quantum **5**, 531.

Wang, X., Z. Lin, L. Che, H. Chen, and D. Lu (2022), Adv. Quantum Technol. **5** (8), 2200005.

Wang, X., and R. Wu (2025), "Tight generalization bound for supervised quantum machine learning," arXiv:2510.24348.

Wang, X., Y. Zhang, S. Hazra, T. Li, C. Shao, and S. Chakraborty (2025c), "Randomized quantum singular value transformation," arXiv:2510.06851.

Wang, Y., Y. Alexeev, L. Jiang, F. T. Chong, and J. Liu (2024b), npj Quantum Inf. **10** (71), 71.

Wang, Y., R. Jiang, Y. Fan, X. Jia, J. Eisert, J. Liu, and J.-P. Liu (2025d), "Towards efficient quantum algorithms for diffusion probability models," arXiv:2502.14252.
Wang, Y., and J. Liu (2024), Rep. Prog. Phys. **87** (11), 116402.
Wang, Y., B. Qi, C. Ferrie, and D. Dong (2024c), Phys. Rev. Appl. **22**, 054005.
Wen, J., Z. Huang, D. Cai, and L. Qian (2024), Communications Physics **7** (1), 220.
Wen, J., X. Kong, S. Wei, B. Wang, T. Xin, and G. Long (2019), Phys. Rev. A **99**, 012320.
West, M. T., A. C. Nakhl, J. Heredge, F. M. Creevey, L. C. L. Hollenberg, M. Sevior, and M. Usman (2024), Intelligent Computing **3**, 10.34133/icomputing.0100.
Widdows, D., W. Aboumrad, D. Kim, S. Ray, and J. Mei (2024a), KI-Künstliche Intelligenz **38** (4), 293.
Widdows, D., A. Alexander, D. Zhu, C. Zimmerman, and A. Majumder (2024b), Annals of Mathematics and Artificial Intelligence **92** (5), 1249–1272.
Wiebe, N., D. Braun, and S. Lloyd (2012), Phys. Rev. Lett. **109**, 050505.
Wiebe, N., A. Kapoor, and K. M. Svore (2015), "Quantum deep learning," arXiv:1412.3489.
Wiedmann, M., M. Hölle, M. Periyasamy, N. Meyer, C. Ufrecht, D. D. Scherer, A. Plinge, and C. Mutschler (2023), in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01, pp. 450–456.
Wierichs, D., C. Gogolin, and M. Kastoryano (2020), Phys. Rev. Res. **2**, 043246.
Wierichs, D., J. Izaac, C. Wang, and C. Y.-Y. Lin (2022), Quantum **6**, 677.
Wiersema, R., A. F. Kemper, B. N. Bakalov, and N. Killoran (2025), Phys. Rev. Res. **7**, 013148.
Wilhelm, F. K., R. Steinwandt, P. Lageyre, and S. Kirchhoff (2025), *The status of quantum computer development*, Technical Report Project 283 (Federal Office for Information Security (BSI), Bonn, Germany) version 2.2.
Williams, R. J. (1992), Mach. Learn. **8** (3), 229.
Willsch, D., M. Willsch, H. De Raedt, and K. Michielsen (2020), Comput. Phys. Commun. **248**, 107006.
Willsch, D., M. Willsch, C. D. Gonzalez Calaza, F. Jin, H. De Raedt, M. Svensson, and K. Michielsen (2022), Quantum Inf. Process. **21**, 141.
Wilson, M., T. Vandal, T. Hogg, and E. G. Rieffel (2021), Quantum Mach. Intell. **3** (1), 19.
Wintersperger, K., F. Dommert, T. Ehmer, A. Hoursanov, J. Klepsch, W. Mauerer, G. Reuber, T. Strohm, M. Yin, and S. Luber (2023), EPJ Quantum Technol. **10** (1), 32.
Wittler, N., F. Roy, K. Pack, M. Werninghaus, A. S. Roy, D. J. Egger, S. Filipp, F. K. Wilhelm, and S. Machnes (2021), Phys. Rev. Appl. **15**, 034080.
Wootters, W. K., and W. H. Zurek (1982), Nature **299** (5886), 802.
Wu, D., R. Rossi, F. Vicentini, and G. Carleo (2023), Phys. Rev. Res. **5**, L032001.
Wu, J.-C., Q. Ye, D.-L. Deng, and L.-W. Yu (2025a), Phys. Rev. Lett. **135**, 227301.
Wu, S., S. Jin, D. Wen, D. Han, and X. Wang (2025b), Quantum **9**, 1660.
Wu, S. L., S. Sun, W. Guan, C. Zhou, J. Chan, C. L. Cheng, T. Pham, Y. Qian, A. Z. Wang, R. Zhang, *et al.* (2021a), Phys. Rev. Res. **3**, 033221.
Wu, Y., W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, *et al.* (2021b), Phys. Rev. Lett. **127**, 180501.
Wu, Y., J. Yao, P. Zhang, and H. Zhai (2021c), Phys. Rev. Res. **3**, L032049.
Wurtz, J., A. Bylinskii, B. Braverman, J. Amato-Grill, S. H. Cantu, F. Huber, A. Lukin, F. Liu, P. Weinberg, J. Long, *et al.* (2023), "Aquila: Quera's 256-qubit neutral-atom quantum computer," arXiv:2306.11727.
Xia, R., and S. Kais (2018), Nat. Commun. **9** (4195), 4195.
Xiao, H., K. Rasul, and R. Vollgraf (2017), "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms," arXiv:1708.07747.
Xiao, P., M. Zheng, A. Jiao, X. Yang, and L. Lu (2025), Quantum **9**, 1761.
Xiao, T., X. Zhai, J. Huang, J. Fan, and G. Zeng (2024), Communications Physics **7** (1), 276.
Xin, T., L. Che, C. Xi, A. Singh, X. Nie, J. Li, Y. Dong, and D. Lu (2021), Phys. Rev. Lett. **126**, 110502.
Xue, C., Z.-Y. Chen, X.-N. Zhuang, Y.-J. Wang, T.-P. Sun, J.-C. Wang, H.-Y. Liu, Y.-C. Wu, Z.-L. Wang, and G.-P. Guo (2025), "End-to-end quantum vision transformer: Towards practical quantum speedup in large-scale models," arXiv:2402.18940.
Yan, F., A. M. Iliyasu, N. Li, A. S. Salama, and K. Hirota (2024a), Quantum Mach. Intell. **6** (2), 86.
Yan, K., P. Lai, and Y. Wang (2024b), in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, edited by K. Duh, H. Gomez, and S. Bethard (Association for Computational Linguistics, Mexico City, Mexico) pp. 2112–2121.
Yan, P., L. Li, M. Jin, and D. Zeng (2021), Neurocomputing **445**, 276.
Yang, Y.-F., and M. Sun (2022), in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2313–2322.
Yarkoni, S., E. Raponi, T. Bäck, and S. Schmitt (2022), Rep. Prog. Phys. **85** (10), 104001.
Ye, C.-C., N.-B. An, T.-Y. Ma, M.-H. Dou, W. Bai, D.-J. Sun, Z.-Y. Chen, and G.-P. Guo (2024), Phys. Fluids **36** (12), 127111.
Yi, H. (2025), IEEE Access **13**, 156628.
Yin, Z., I. Agresti, G. de Felice, D. Brown, A. Toumi, C. Pentangelo, S. Piacentini, A. Crespi, F. Ceccarelli, R. Osellame, *et al.* (2025), Nat. Photonics **19** (9), 1020.
Yu, H., X. Ren, C. Zhao, S. Yang, and J. McCann (2023), Scientific Reports **13** (1), 19130.
Zaman, K., T. Ahmed, M. A. Hanif, A. Marchisio, and M. Shafique (2025), in *Grid, Cloud, and Cluster Computing; Quantum Technologies; and Modeling, Simulation and Visualization Methods* (Springer, Cham, Switzerland) pp. 102–115.
Zare, A., and M. Boroushaki (2025), EPJ Quantum Technol. **12** (1), 74.
Zeguendry, A., Z. Jarir, and M. Quafafou (2023), Entropy **25** (2), 287.
Zhang, C., Z. Lu, L. Zhao, S. Xu, W. Li, K. Wang, J. Chen, Y. Wu, F. Jin, X. Zhu, *et al.* (2026a), npj Quantum Inf. **12** (28), 28.
Zhang, C., Z. Su, Q. Li, D. Song, and P. Tiwari (2025a), Neural Networks **182**, 106915.
Zhang, H., and A. Kamenev (2025), Communications Physics **8** (1), 478.
Zhang, H., and Q. Zhao (2025), "A survey of quantum trans-

formers: Technical approaches, challenges and outlooks," arXiv:2504.03192.
Zhang, H.-F., Z.-Y. Chen, P. Wang, L.-L. Guo, T.-L. Wang, X.-Y. Yang, R.-Z. Zhao, Z.-A. Zhao, S. Zhang, L. Du, *et al.* (2026b), "Experimental robustness benchmark of quantum neural network on a superconducting quantum processor," arXiv:2505.16714.
Zhang, J., H. Wang, G. S. Ravi, F. T. Chong, S. Han, F. Mueller, and Y. Chen (2023a), in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 01, pp. 1062–1073.
Zhang, P., J. Niu, Z. Su, B. Wang, L. Ma, and D. Song (2018) (AAAI Press).
Zhang, S.-X., J. Allcock, Z.-Q. Wan, S. Liu, J. Sun, H. Yu, X.-H. Yang, J. Qiu, Z. Ye, Y.-Q. Chen, *et al.* (2023b), Quantum **7**, 912.
Zhang, X.-M., T. Li, and X. Yuan (2022), Phys. Rev. Lett. **129**, 230504.
Zhang, X.-M., and X. Yuan (2024), npj Quantum Inf. **10** (42), 42.
Zhang, Y., D. Song, P. Zhang, X. Li, and P. Wang (2019), Applied Intelligence **49** (8), 3093–3108.
Zhang, Z., Y. Li, and X. Xu (2025b), "Quantum-enhanced neural networks for quantum many-body simulations," arXiv:2501.12130.
Zhao, C., and X.-S. Gao (2021), Quantum **5**, 466.
Zhao, Z., A. Pozas-Kerstjens, P. Rebentrost, and P. Wittek (2019), Quantum Mach. Intell. **1** (1), 41.
Zhu, C., P. J. Ehlers, H. I. Nurdin, and D. Soh (2025a), Phys. Rev. Res. **7**, 023290.
Zhu, C., H. Yao, Y. Liu, and X. Wang (2025b), Quantum Mach. Intell. **7** (2), 92.
Zia, D., L. Innocenti, G. Minati, S. Lorenzo, A. Suprano, R. Di Bartolo, N. Spagnolo, T. Giordani, V. Cimini, G. M. Palma, *et al.* (2025), Sci. Adv. **11** (50), 10.1126/sciadv.ady7987.
Ziyin, L., and M. Ueda (2023), Phys. Rev. Res. **5**, 043243.
Zlokapa, A., H. Neven, and S. Lloyd (2021), "A quantum algorithm for training wide and deep classical neural networks," arXiv:2107.09200.
Zou, T., Y. Fang, J. Wang, M. Dou, J. Fu, Z. Zhao, S. Zhao, L. Yu, D. Zhao, Z. Chen, *et al.* (2025), "Qpanda3: A high-performance software-hardware collaborative framework for large-scale quantum-classical computing integration," arXiv:2504.02455.
Zoufal, C., A. Lucchi, and S. Woerner (2019), npj Quantum Inf. **5** (103), 103.
Zoufal, C., A. Lucchi, and S. Woerner (2021), Quantum Mach. Intell. **3** (1), 7.