Quantum 3.0: Quantum Learning, Quantum Heuristics and Beyond

Mrittunjoy Guha Majumdar1*

[1]Centre of Excellence for Quantum Technology UPES, Dehradun, India

Corresponding Author E-mail: mrittunjoy.majumdar@ddn.upes.ac.in

Abstract: Quantum learning paradigms address the question of how best to harness conceptual elements of quantum mechanics and information processing to improve operability and functionality of a computing system for specific tasks through experience. It is one of the fastest evolving framework, which lies at the intersection of physics, statistics and information processing, and is the next frontier for data sciences, machine learning and artificial intelligence. Progress in quantum learning paradigms is driven by multiple factors: need for more efficient data storage and computational speed, development of novel algorithms as well as structural resonances between specific physical systems and learning architectures. Given the demand for better computation methods for data-intensive processes in areas such as advanced scientific analysis and commerce as well as for facilitating more data-driven decision-making in education, energy, marketing, pharmaceuticals and healthcare, finance and industry.

Introduction: Quantum 1.0 was the revolutionary utilization of quantum resources for technology, primarily electrical and optical, such as transistors and optical masers, while Quantum 2.0 was about harnessing non-classical elements in the quantum formalism such as entanglement for information processing (1–15). Quantum mechanics have been used for undertaking information processing tasks such as teleportation, superdense coding, remote state preparation, quantum tomography, quantum cryptography and quantum key distribution, circuit-based and measurement-based quantum computing, quantum network coding and quantum internet, quantum random walks and quantum transduction (16–55). What can be called as Quantum 3.0 would be the harnessing of quantum resource and representation theory for a novel computational learning paradigm, along with quantum generalizations of existing classical computational learning models (56,57). This is the next frontier of exploration in the quantum realm and seeks to address the question: can we fundamentally reframe and re-envision learning models when the information-system and/or process is quantum mechanical in nature?

The term 'machine learning' was coined by Arthur Samuel in 1959, in the context of a 'looking- ahead' algorithm implemented on a classical computer for a game of checkers, realized using a neural net approach with randomly connected switching net as well as an approach involving a highly organized network designed to learn only specific things (58). In 1961, the punched-tape system known as Cybertron K-100 by Raytheon was developed as an early learning system that employed pattern recognition on sound samples, such as those from sonar signals, for learning (59). Basic pattern recognition, storage and referencing were the primary elements that supported the learning framework in the next few decades, such as with the work on feed forward networks with one layer of modifiable weights connecting input units to output units, in what could be called reflexive systems that could discover hidden relationships in data (60, 61). Machine learning addresses the twin questions of what fundamental information theoretic laws govern all learning systems and how could one construct computing systems that can improve their functioning through experience (62). For various applications, training based on a limited set of inputs and outputs obtained from

a system for specific empirical task seems to be more optimum than trying to predict the generalized response of the system in all circumstances.

Learning, in the context of information theory, describes the improvement of a yardstick or performance criterion for a system for a given task. For instance, we could seek to classify Iris flowers into three species: Iris adriatica, Iris taochia and Iris kemaonensis from a defining characteristic such as the cumulative length measurement of their petals and sepals. The yardstick index could be how accurately the classification takes place, and the training could be over a historical sample of irises from a phytologist's collection. We could also employ other criteria that have different penalties for incorrect classification. There are a number of algorithms for machine learning that are applicable to different kinds of datasets and problem types. Machine learning algorithms probe through various candidate models, based on training data and experience, to find the one that is optimum for the selected performance metric. The diversity of machine-learning algorithms can be segregated based on their manner of representation of candidate models (such as decision trees) and their manner of probing through the space of candidate models (such as evolutionary search methods).

The question is whether loading our phytological data into quantum states can make this classification better. We could create a parameterized quantum circuit or ansatz to do this, using a feature map. Any parameterized quantum circuit should be able to generate maximum number of quantum states in the Hilbert space and also have ability to entangle constituent qubits. It turns out that the variational quantum classifiers (VQCs) help classify the irises efficiently, using optimizers such as the constrained optimization by linear approximation optimizer (COBYLA). It is this interface of quantum mechanics and machine learning that provides various possibilities: machine learning with quantum computers, classical learning for quantum mechanical problems

and generalized quantum learning theory. Quantum machine learning brings together the learning capability and scalable nature of machine learning and the speed, efficiency and processing power of quantum computers. There have been several classical learning paradigms that have had quantum analogues proposed and realized: quantum neural networks, quantum reinforcement learning, quantum support vector machines, quantum linear regression, cluster detection and assignment algorithms, quantum principal component analysis and quantum decision tree classifier (63–70). In this comprehensive review of quantum learning paradigms, we will be looking at the spectrum of algorithms and protocols, before moving on to understanding the conceptual shift that this may provide when assessing learning theory from the perspective of quantum phenomenon.

Quantum Patterns: The underlying premise for the utility of the quantum realm for machine learning has to do with the emergence of classically anomalistic patterns in data. Among phenomena that are intrinsically quantum in nature, an interesting example is that of quantum revival patterns, where we find superposed and displaced replicas of the initial state of the system (71–74). Quantum Kerr systems have such a kaleidoscopic mode of evolving, which is found to be reproducible based on interference of trajectories in the classical phase space (75). In the quantum-classical limit, where Planck's constant $\bar{h} \rightarrow 0$, it is seen that observables behave differently, leading to the preservation or disappearance of significant variables and their stochastic tendencies, such as for oscillator coordinate and spin variable respectively (76). Even in the case of skyrmions, we cannot simulate quantum skyrmions on classical supercomputers due to fundamental limitations, particularly around quantum fluctuations (77). Quantum patterns are seen as periodic magnetic formations that are spontaneously formed in antiferromagnetic Bose-Einstein condensate (78). When ultracold atoms form quantum ferrofluids, we find a spontaneous development of coherent

quantum density patterns leading to the formation of a super-solid (79). When nanoscale lead films grown on silicon substrates are annealed to high temperatures, we observe intricate quantum patterns in the energy landscape (80). A beat pattern emerges in the quantum magnetoresistance of polar semiconductors without a centre of inversion symmetry, such as InAs/GaSb, when the carrier concentration is high (81). Patterns are ubiquitous in quantum chemistry, in properties such as chemical hardness and electronegativity (82).

Given the ability of classical computers to both produce and recognize classical statistical data, the extrapolation would be that systems that produce classically atypical data, like in quantum systems, would also be able to recognize such data. This gives us the possibility of harnessing the quantum analogue of complex classical pattern recognition. Pattern recognition compares data input with specific memorized patterns, as it processes this input. Quantum pattern recognition provides a number of advantages over its classical counterpart. For instance, if we assume a Hopfield network approach for pattern recognition, we do so by local optimization, while utilising a quantum approach such as adiabatic quantum computation does so using global optimization (83). A realization of this was using the liquid-state NMR technique, where some of the interesting insights included the possibility of representing a superposition of recognized patterns using a quantum neural register (84). The parallelism inherent in quantum phenomena such as entanglement facilitate the execution of subroutines, even with big data, and this is one of the primary reasons for quantum machine learning, in general, being regarded as being better than its classical counterpart. Among pattern recognition protocols, one-class classification is important due to its applications in areas where detection of abnormal data points vis-a-vis instances of a known class is needed, using techniques such as support vector machines and principal component analysis (85–90). This helps in addressing problems where we have imbalanced datasets, such as in metaheuristics, medical image datasets, manufacturing, high-energy physics and bioinformatics (91–98). Recently, a semi-supervised quantum one-class classification system known as Variational Quantum One-Class Classifier (VQOCC) was developed using (99). At the centre of this proposition is a quantum autoencoder, which is basically a circuit that undertakes compression of a quantum state onto a lesser number of qubits while retaining the information encapsulated in the initial state (100, 101). Quantum autoencoders have been applied to quantum data compression, quantum error correction, denoising of data and even for preserving entanglement (102–105).

An important question to be addressed here is what can be categorized as a quantum pattern. Is it just an arbitrary pattern of quantum randomness? Is it the byproduct of a quantum pro- cess, such as in cold exciton gases? Is it the quantum analogue of classical patterns in nature? Harney and Pirandola defined a quantum pattern as an m-mode coherent state undergoing local kary modulations (106). We can also define patterns in the information dynamics associated with quantum systems, such as in the case of quantum computing and communication systems (107). Examples would include basis encoding, quantum associative memory, amplitude encoding, angle encoding, QRAM encoding, quantum kernel estimator, variational quantum algorithm, variational quantum eigensolver, quantum associative memory, amplitude amplification, phase shift, oracle operations, quantum approximate optimization algorithm and quantum key exchange. Along with information theoretic arrangement and elements as well as quantum data types and quantum data structures, we can also specify entanglement structures to define the notion of quantum patterns (108). There are various kinds of entanglement patterns, from maximally entangled Dicke states to partially entangled cluster states. Even the outcome of measurements successively performed on an

open quantum system have a pattern because of the interaction between system and environment. The pattern encapsulates information on non- Markovian memory effect and the relaxation rates associated with the system (109). The primacy of quantum patterns can also be seen within phases of matter, which can be distinguished by distinct symmetry breaking instances - within the Landau theory, symmetry of ordering of constituents of a physical system differentiates one phase from others. In certain systems, such as chiral spin liquids, even without symmetry breaking, we have different characteristics due to what is known as topological order, which has been posited to describe entanglement in many-body systems (110).

Quantum Algorithmic Resource-Pool: The central question around assessing the possibilities in quantum machine learning is whether we have the algorithmic tools in the quantum domain for the same. A quantum algorithm is a succession of instructions for tackling a problem on a practical quantum computer (111). Cleve et al. highlighted that a common thread underlying all quantum algorithms can be as- certain when "quantum computation is viewed as multiparticle interference" (112). By and large, we have quantum search algorithms, quantum simulations of quantum systems on a quantum computer and algorithms premised on quantum implementations of the Fourier transform like Shor and Deutsch-Josza algorithm. Quantum supremacy or the enhancement of computational ability using quantum systems over that of classical counterparts has heralded the age of noisy intermediate-scale quantum (NISQ) technologies (36, 113). One of the earliest quan- tum algorithms came with by formulation of an algorithmic solution to a special case of the hidden subgroup problem by Peter Shor (114). Quantum Merlin-Arthur (QMA) completeness of a problem highlights that a supposed solution to it can be verified by a quantum computing system - a condition that has been extended to k unentangled provers in QMA(k) class problems (115, 116). When it comes to claims of quantum supremacy, we often speak of speedup, resource reduction, scaling and ability to address greater number as well as variety of problems. Quantum speedup can be contextually defined in terms of the asymptotic behavior of the ratio of the times taken by a specific classical and quantum algorithm for a particular problem when the size of the problem is made to be very large (117). Speedup can be provable, like in Grover's algorithm, while in some cases the enhancement due to the quantum effects is not obvious, like in quantum annealing (118, 119). Quantum resource optimization is seen in realization of randomness processing with quantum Bernoulli factories (120).

The reason for enhancement of computational power, when it comes to machine learning and data analysis, with quantum machine learning arises due to the inherent primacy, within quantum mechanics, of high-dimensional vector spaces and, more importantly, matrix transformations between vectors in such spaces, as is important for data analytics and machine learning methods. The speedup is especially pronounced when the matrices being dealt with have an element of sparseness or are low-rank matrices (121, 122). A classic example arises in a key tool used for machine learning - principal component analysis, which evaluates the eigensystem (and corresponding principal components) of the covariance matrix of data. Using a quantum random access memory, we initialize the vectors for a classical principal component analysis protocol onto quantum states and utilise the density matrix of the states (instead of the covariance matrix), which when exponentiated and operated on by a conditional SWAP operation for the quantum phase algorithm, yields the eigensystem that forms the premise of the principal component analysis (67). The algorithmic scaling is to the order of the square of the logarithm of the size of the system, in query as well as computational complexity. Essentially, the quantum principal component analysis scales exponentially faster than the classical counterpart.

While quantum principal component analysis is an unsupervised method that identifies patterns within higher-dimensional data for the reduction of data complexity with the retention of most of the information, there are various supervised learning algorithms that have had a quantum implementation (123). One such algorithm relates to quantum support vectors. Support vector machines undertake regression and classification using the delineation of feature vectors of the data into distinct classes around a hyperplane that is at maximum distance from the nearest data-points in either classes around it (124). If the data is not separable into distinct classes, we can use the kernel trick to project the data into a higher-dimensional space- a pursuit that by Cover's theorem can help in attaining separability in the data-set (125). Quantum algorithms such as Grover's algorithm are premised on binary classification. Even in cases of machine learning algorithms based on adiabatic quantum evolution, we define our approach in terms of a strong classifier that discerns whether a data-point is correct or erroneous based on a program specification (126). Like in the classical case, the quantum support vector machine relies on the definition of a quantum kernel that we can create by using a quantum feature map $\phi(\vec{x})$ between classical feature vectors $\vec{x}$ and a Hilbert space to help us obtain the kernel $K(\vec{x}_i, \vec{x}_j) = |\langle\phi(\vec{x}_i)|\phi(\vec{x}_j)\rangle|^2$. We can then expand the hyperplane in the general form: $f(\vec{x}) = \Sigma i\ \alpha_i y_i K(\vec{x}_i, \vec{x})$, where α_i denote bounded positive quantities and y_i the data labels. We can then use measures like the Gaussian or Rademacher Complexity to evaluate the classification error, based on this hyperplane definition and the data-set (127). While the quantum support vector machine paradigm defined using Grover's algorithm gives a quadratic speedup, the implementation using the least squares approach provides an exponential speedup over classical algorithms (64). In recent years, we have had quantum support vector machine implementations with the Newton method, amplitude estimation, gradient descent and using quantum annealers as well as variational quantum-circuitry (128–132).
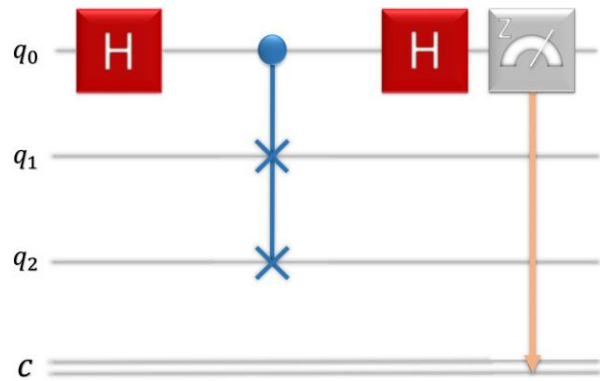

Figure 1: Swap Test

A natural extension of this comes with quantum k-nearest neighbour algorithms. Instead of a hyperplane as in the case of support vector machines, the underlying assumption of k-nearest neighbour methods is that the likelihood of two data-points that are proximal being of the same type is high (133). Labelled training vectors are used as reference for comparison of unlabelled testing vectors to determine k nearest train-state neighbours for the specific testing vector-state, whose label is ascertained using majority voting. In the quantum picture, we use the concept of fidelity of a specific testing state $|\psi\rangle$ with respect to multiple training states $|\phi_j\rangle$: $F_j = |\langle\psi|\phi_j\rangle|^2$. After taking an initial set of candidate neighbour states, we use quantum search algorithms to find other states and corresponding indices till the nearest neighbours are found. The functional way to obtain the fidelity is by using the swap test. In this test (Figure 1), we begin with two quantum states $|a\rangle$ and $|b\rangle$ along with an ancilla qubit initialized to $|0\rangle_{anc}$, on which we apply the Hadamard operation on the ancilla qubit. We thereafter apply a controlled swap operation on the ancilla qubit. A controlled swap operation is a gate where target qubits $|a\rangle$ and $|b\rangle$ are swapped if the control (in this case, the ancilla qubit) is in the state $|1\rangle$. Applying a second Hadamard gate on the ancilla gives us the state $|\chi\rangle = 1/2\ |0\rangle_{anc}(|a, b\rangle + |b, a\rangle) + 1/2\ |1\rangle_{anc}(|a, b\rangle - |b, a\rangle)$.

The measurement of the state $|0\rangle_{anc}$ has the associated probability $P(|0\rangle_{anc}) = 1/2 + 1/2$ $|\langle a|b\rangle|^2$. The inner product in this expression makes the overlap between the states primary: orthogonality gives us a probability of 0.5 while maximum overlaps gives a unity probability. This method can be used to find the distance between real-valued multi-dimensional vectors with the use of a quantum measurement (69). We can also undertake more rigorous pattern recognition between

binary strings $|a_1, a_2, ..., a_n\rangle$ and $|b_1, b_2, ..., b_n\rangle$, using an extension of the construction in the k-nearest cluster model by initializing a state $|\psi\rangle = |a_1, a_2, ..., a_n, b_1, b_2, ..., b_n\rangle \otimes \sqrt{1}\,(|0\rangle+|1\rangle)_{anc}$, undertaking a XOR operation between the respective $(a_k, b_k) \forall k$ before storing it in place of the $b_k \forall k$, and finally undertaking a Hadamard operation on the ancilla qubit (*134*). If we measure the ancilla qubit in the ground state $|0\rangle$anc, we obtain a resultant state whose amplitude has a scaling characteristic dependent on the Hamming distance between the binary strings.

Quantum Reinforcement Learning and Deep Learning: Going beyond supervised or unsupervised machine learning models, a major area of contemporary research has been in quantum-enhanced reinforcement learning, which is premised on the adaptive evolution of a quantum system based on reinforcement from a classical or quantum environment (135). The system receives percepts from the environment and undertakes actions. Unlike in conventional learning models, the learner in a reinforcement learning model has an influence on the state of the environment around it as much as it is influenced by it, thereby making it impossible to represent the environment in terms of a stationary memory. Both - the system and the environment, are stored as maps with memory, and the history of interactions between the two is the fundamental element in reinforcement learning. In the case of quantum reinforcement learning (QML), this history must be maintained in a quantum setting. In QML, we

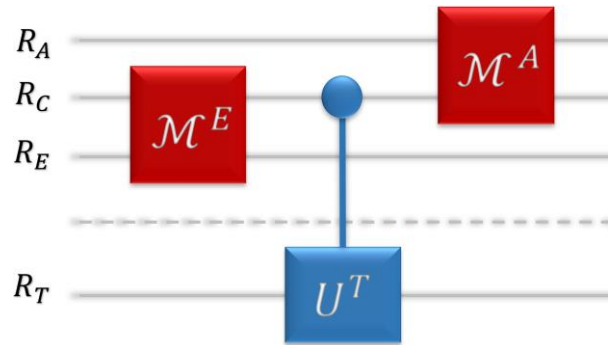define a Hilbert space for the percept and action states- HS and HA respectively.

Figure 2: Quantum Reinforcement Learning (QRL) is premised on testing agent-environment interaction where the percept and action states have distinct Hilbert spaces associated with them. The register RT is not controlled by the agent or by the environment.

The agent and environment both have memory registers (RA and RE respectively) to store the histories of the system-environment composite. We can model the interaction with a distinct Hilbert space HC and we can characterize the agent (environment) my a series of Completely Positive Trace Preserving (CPTP) maps {MiA}i ({MiE }i) that acts on a resultant register formed by concate- nation of the registers of system-interaction RARC (environment-interaction RCRE) systems. The performance of the system-environment is assessed against a figure of merit. Dunjko et al showed that a quantum agent will outperform a classical learning agent associated with a clas- sically delineated, controllable epochal environment against a particular figure of merit, if we were to consider a selected classical sporadic testing element (136). The quantum enhancement arises from being able to extract additional attributes from the environment for optimization of a classical agent. Our point of interest is in environments that are quantum in nature and that facilitate the preservation of superposition of percepts and actions.

When we are talking of quantum reinforcement learning, it is imperative to discuss what we mean by quantum accessibility. Quantum accessibility implies the utilization of access of the agent to the

environment for the simulation of an oracle Oq that behaves as follows:

$$|a_1, a_2, ..., a_k\rangle|x\rangle \xrightarrow{\mathcal{O}_q} |a_1, a_2, ..., a_k\rangle|x \oplus R(a_1, ..., a_k)\rangle$$

where R is the value of the 'reward' that the agent obtains once it successfully executes the actions a_1, a_2, ..., a_k, for some auxiliary state $|x\rangle$. In the cumulative process, we see the utilization of a quantum interaction using what is known as oracularization with the environment to undertake effective simulation of access to the oracle (which encapsulates important information about the environment), followed by the usage of this information for determining behavior that is optimum in the given environment. We can attain a greater insight into the oracularization process by understanding the import of an auxiliary state, which usually comprises of percept states $|s_1, s_2, ..., s_k\rangle$, the reward state $|R\rangle$ and $|e_1, e_2, ..., e_k\rangle$ that are auxiliary quantum subsystems kept within the environment. For the simulation of the oracle, these states need to be con- trolled and erased (137). To undertake this, the agent should be able to undertake modification of specific memory components of the environment with what are known as register scavenging and register hijacking operations, followed by a way to 'uncompute' with the implementation of the Hermitian adjoint of the net unitary (map) that the environment implements. There are two primary kinds of reinforcement learning paradigms with specific quantum realizations: value-based, like in the case of Q-learning, which is premised on the learning of a value function that guides how the system environment makes decisions at each time step, as well as policy- gradient methods that optimizes a policy (function) $\pi(a|s; \theta)$ using the parameter θ (138–141). An important concept in this regard is meta-learning, wherein the learning agent undertakes the identification of its meta-parameter configuration that is optimum so as to be able to facilitate the optimization of performance. This parameter meta-learning is done by having the learning agent monitor its performance, and utilising a metaparameter register as well as techniques like adaptive Bayesian quantum estimation (142). The

reduction of the agent-environment interaction to a unitary oracular query is not feasible when we consider memory effects. In general, the oracular element may vary temporally and in such scenarios amplifying amplitude using Grover-type methods may be a possible approach to undertake reinforced learning in reward spaces that have an increasing monotonicity in the success probability.

Deep learning has also seen a quantum expansion in recent years. Deep learning is premised on using artificial neural networks for discovering the representations required for the detection and classification of features from raw data (143, 144). The basic unit of a neural network is a perceptron or a single artificial neuron. Gallant described one of the first perceptron-based connectionist models where each cell i computes a single activation ui, which may be input to other cells or be an output of the network (145). Ricks et al. gave one of the earliest quantum neural network models based on implementation of quantum circuitry with gates whose weights are evolved through learning using quantum search as well as piecewise weight allocation (146). In each step, we have a density operator for the qubits representing the hidden states, which can can be extracted with any output ancilliaries using a partial trace operator and fed forward to the next layer of the neural network where the unitary transformations that encapsulate the action of perceptrons can be applied. There have been recent works on Bayesian approaches being implemented using quantum algorithms to learn Gaussian processes to train neural networks that are arbitrarily deep, with a quantum matrix inversion protocol being the core routine (147). In the quantum convolutional neural network presented by Cong et al (148), the modeling of the convolutional layer is in terms of a unitary transformation (on the state density of the input) that is quasi-local and is applied in error correction and phase recognition in quantum systems. The forward pass of the convolutional neural network can be computed as a convolutional product using quantum algorithmic tools even as gradient descent

methods can give us an idea about what are the relevant network parameters in the quantum context (149). There have also been hybrid models with quanvolutional layers in the network where we undertake local transformation of the data with several random quantum circuits in the set bounded-error quantum polynomial time (150). Recently, quanvolutional methods were used for physiological application such as for body part recognition using Hybrid Quantum Convolutional Neural Networks, although in this case the classical counterpart was found to have a greater validation accuracy by 0.5% (151).

Quantum Natural Language Processing, Quantum Advantage and the Path Forward Language is one of the earliest forms of representative evolution for humankind. Notwithstanding the tale of the Tower of Babel or the Gigantomachy myth, the diversity and complexity of language is of prime interest for computational representation and processing. Natural language processing is all about harnessing computational methods for the learning and production of content in human languages. We have come a long way from simply analysing the disparate linguistic forms and recognizing speech patterns to the creations of dialogues, translating speech from speech as well as the identification of emotional response of users towards services and products. The synchronic model of language aligns with the psycholinguistic understanding which highlights that language is dynamic (152). In any speech-based natural language pro- cessing (NLP) system, we have computation based on phonetic, phonemic and prosodic rules, besides the morphemic analysis which is as relevant for other kinds of NLPs as well (153–155). Higher levels of natural language processing include the lexical, the syntactic, the semantic, the discourse-oriented and the pragmatic (156). When it comes to natural language processing, we can have quantitative statistical approaches like the Hidden Markov models, connectionist approaches that are based on inter-related fundamental processing units as well as symbolic ap- proaches that lay emphasis on the

logic or rules that a linguistic framework encapsulates. Zeng and Coecke were among the first to give us a quantum version of natural language processing when they employed the closest vector problem for sentence similarity identification within the distributional compositional categorical (DisCoCat) model given by Coecke, Sadrzadeh and Clark (157–159). In the compositional distribution model, meanings are encapsulated in quantum states while quantum measurements give us the grammatical structure for a specific sentence and context (160). The embedding of the language in the vector space of quantum systems naturally leads to word-correlations represented by the vector geometry, even as we map diagrams formed from parsing of sentences to quantum variational circuits. Essential quantum properties such as superposition helps us in the modelling of uncertainties in the language while phenomena like entanglement help us in describing how semantics and syntax are com- posed and distributed. Recently, natural language processing experiments have been practically realized for over 100 sentences on the Noisy Intermediate-Scale Quantum (NISQ) computing system provided by IBM Quantum (161).

When we speak about quantum natural language processing, in particular, and machine learning, in general, the elephant in the room is: how much of an advantage do we obtain using quantum systems over classical ones? The 'quantum native' view of quantum natural language processing which posits that the vector-landscape provided by quantum systems is better suited for the linguistic-syntactic, semantic and pragmatic elements, is preliminary and unsubstantiated by practical realizations, with a critical shortcoming being the primary reliance on quantum RAMs, which are expensive and not yet empirically implemented. More broadly, there are many aspects of quantum computing that are often off-set by machine learning paradigms.

For instance, the size of inputs is usually small, such as in fault-tolerant algorithms, while there

could high-dimensional tensorial structures with numerous entries in machine learning. The problems in machine learning can be highly unstructured and complex while problems studied in quantum computing are structured, often with elements of regularity and periodicity. Theory is defined and delineated in quantum computing problems while empiricism is preeminent in machine learning, with modelling and interpolations being important in the latter. Quantum computing can have absolute benchmarks like the scaling of run-time while machine learning have more constructivist benchmarks between disparate models. Quantum advantage can be spoken of in terms of solvability, expressivity of the class of the model, size of training sample needed, generalizability and how the optimization landscape is structured (162–171). Due to the need to be able to capture the system dynamics in terms of quantum systems and circuitry, we can only gauge performance of machine learning algorithms in specific selected contexts and examples, with this selectivity preventing any generalizability of performance-based ad- vantage, if any. Such constraints also exist for readout and loading of data, when it comes to quantum examples. Performance parameters are also highly contextual and it is not straight-forward to assess if any 'quantum advantage' is due to the specific way in which we have selected benchmarking threshholds and hyper-parameters or whether they are actually observations made structurally. There may be a need to move towards studying Quantum 3.0, the revolution of quantum learning theory, as a paradigm in itself, without needing to resort to bylines towards 'quantum advantage' and 'quantum supremacy'. For example, classical intractability of kernels being bypassed using quantum tools can be an interesting pursuit but given that such kernels are not found to be utilizable in realizable machine learning, such an approach has no practical relevance. Today, we have moved beyond kernel methods with even more efficient ones like data re-uploading and explicit models (172, 173). We now have quantum machine learning models that have features that preserve privacy as well as those

that take into consideration the impact of decoherence (174, 175). The future is bright for quantum machine learning, even as the intersection of quantum mechanics, learning theory and information processing is explored to find novel methods of deploying artificial intelligence.

Competing Interests: The Author declares no Competing Financial or Non-Financial Interests.

Author Contributions: The author formulated the premise, undertook the analysis and wrote the paper.

References and Notes:

1. N. Bloembergen, Y.-R. Shen, Physical Review 133, A37 (1964).
2. T. L. Paoli, J. E. Ripper, Physical Review A 2, 2551 (1970).
3. K. Lehovec, Physical Review 75, 1100 (1949).
4. R. Hall, IRE Transactions on Electron Devices 7, 1 (1960).
5. D. Deutsch, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 400, 97 (1985).
6. R. P. Feynman, Optics news 11, 11 (1985).
7. A. Peres, Physical review A 32, 3266 (1985).
8. W. H. Zurek, Physical Review Letters 53, 391 (1984).
9. P. Benioff, Journal of statistical physics 22, 563 (1980).
10. D. P. DiVincenzo, Science 270, 255 (1995).
11. A. Barenco, et al., Physical review A 52, 3457 (1995).
12. P. W. Shor, Physical review A 52, R2493 (1995).
13. D. P. DiVincenzo, Physical Review A 51, 1015 (1995).
14. P. W. Shor, Proceedings 35th annual symposium on foundations of computer science (Ieee, 1994), pp. 124–134.
15. J. I. Cirac, P. Zoller, Physical review letters 74, 4091 (1995).
16. D. Bouwmeester, et al., Nature 390, 575 (1997).

17. A. Furusawa, et al., science 282, 706 (1998).

18. S. Pirandola, J. Eisert, C. Weedbrook, A. Furusawa, S. L. Braunstein, Nature photonics 9, 641 (2015).

19. M. Guha Majumdar, Pramana 95, 200 (2021).

20. C. Wang, F.-G. Deng, Y.-S. Li, X.-S. Liu, G. L. Long, Physical Review A 71, 044305 (2005).

21. X. S. Liu, G. L. Long, D. M. Tong, F. Li, Physical Review A 65, 022304 (2002).

22. A. Harrow, P. Hayden, D. Leung, Physical review letters 92, 187901 (2004).

23. C. H. Bennett, et al., Physical Review Letters 87, 077902 (2001).

24. B. Dakic´, et al., Nature Physics 8, 666 (2012).

25. C. H. Bennett, P. Hayden, D. W. Leung, P. W. Shor, A. Winter, IEEE Transactions on Information Theory 51, 56 (2005).

26. Y. Xia, J. Song, H.-S. Song, Journal of Physics B: Atomic, Molecular and Optical Physics 40, 3719 (2007).

27. T. Jullien, et al., Nature 514, 603 (2014).

28. S. T. Flammia, D. Gross, Y.-K. Liu, J. Eisert, New Journal of Physics 14, 095022 (2012).

29. N. Gisin, G. Ribordy, W. Tittel, H. Zbinden, Reviews of modern physics 74, 145 (2002).

30. G. Brassard, N. Lu¨tkenhaus, T. Mor, B. C. Sanders, Physical review letters 85, 1330 (2000).

31. V. Scarani, et al., Reviews of modern physics 81, 1301 (2009).

32. P. W. Shor, J. Preskill, Physical review letters 85, 441 (2000).

33. D. Gottesman, H.-K. Lo, N. Lutkenhaus, J. Preskill, International Symposium on Information Theory, 2004. ISIT 2004. Proceedings. (IEEE, 2004), p. 136.

34. A. Steane, Reports on Progress in Physics 61, 117 (1998).

35. J. L. O'brien, Science 318, 1567 (2007).

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

37. M. G. Majumdar, Journal of Quantum Information Science 8, 139 (2018).

38. B. Yurke, J. S. Denker, Physical Review A 29, 1419 (1984).

39. M. Hayashi, K. Iwama, H. Nishimura, R. Raymond, S. Yamashita, STACS 2007: 24th Annual Symposium on Theoretical Aspects of Computer Science, Aachen, Germany, February 22-24, 2007. Proceedings 24 (Springer, 2007), pp. 610–621.

40. C. Simon, Nature Photonics 11, 678 (2017).

41. H. J. Kimble, Nature 453, 1023 (2008).

42. S. Wehner, D. Elkouss, R. Hanson, Science 362, eaam9288 (2018).

43. M. G. Majumdar, S. S. Garani, Quantum Information Processing 20, 1 (2021).

44. Y. Aharonov, L. Davidovich, N. Zagury, Physical Review A 48, 1687 (1993).

45. J. Kempe, Contemporary Physics 44, 307 (2003).

46. B. C. Travaglione, G. J. Milburn, Physical Review A 65, 032310 (2002).

47. N. Shenvi, J. Kempe, K. B. Whaley, Physical Review A 67, 052307 (2003).

48. M. G. Majumdar, C. Chandrashekar, Journal of Physics B: Atomic, Molecular and Optical Physics 55, 045501 (2022).

49. N. Lauk, et al., Quantum Science and Technology 5, 020501 (2020).

50. M. T. Rakher, L. Ma, O. Slattery, X. Tang, K. Srinivasan, Nature Photonics 4, 786 (2010).

51. M. Mirhosseini, A. Sipahigil, M. Kalaee, O. Painter, Nature 588, 599 (2020).

52. M. Wu, E. Zeuthen, K. C. Balram, K. Srinivasan, Physical review applied 13, 014027 (2020).

53. R. W. Andrews, et al., Nature physics 10, 321 (2014).

54. T. Bagci, et al., Nature 507, 81 (2014).

55. M. G. Majumdar, C. Chandrashekar, Physics Letters A p. 128829 (2023).

56. J. Biamonte, et al., Nature 549, 195 (2017).

57. V. Dunjko, H. J. Briegel, Reports on Progress in Physics 81, 074001 (2018).

58. A. L. Samuel, IBM Journal of research and development 3, 210 (1959).

59. R. Scharff, The how and why Wonder Book of Robots and Electronic Brains (Wonder, 2017).

60. N. J. Nilsson, The mathematical foundations of learning machines (Morgan Kaufmann Publishers Inc., 1990).

61. R. O. Duda, P. E. Hart, et al., Pattern classification and scene analysis, vol. 3 (Wiley New York, 1973).

62. M. I. Jordan, T. M. Mitchell, Science 349, 255 (2015).

63. T. Menneer, A. Narayanan, Tech. Rep. R329 (1995).

64. P. Rebentrost, M. Mohseni, S. Lloyd, Physical review letters 113, 130503 (2014).

65. M. Schuld, I. Sinayskiy, F. Petruccione, Physical Review A 94, 022342 (2016).

66. S. Lu, S. L. Braunstein, Quantum information processing 13, 757 (2014).

67. S. Lloyd, M. Mohseni, P. Rebentrost, Nature Physics 10, 631 (2014).

68. S. Lloyd, IEEE transactions on information theory 28, 129 (1982).

69. S. Lloyd, M. Mohseni, P. Rebentrost, arXiv preprint arXiv:1307.0411 (2013).

70. K. A. McKiernan, E. Davis, M. S. Alam, C. Rigetti, arXiv preprint arXiv:1908.08054 (2019).

71. R. Vasseur, S. Parameswaran, J. Moore, Physical Review B 91, 140202 (2015).

72. R. W. Robinett, Physics reports 392, 1 (2004).

73. C. F. e Silva, A. E. Bernardini, Physical Review A 107, 042220 (2023).

74. M. Berry, I. Marzoli, W. Schleich, Physics World 14, 39 (2001).

75. G. M. Lando, R. O. Vallejos, G.-L. Ingold, A. M. O. de Almeida, Physical Review A 99, 042125 (2019).

76. S. Dumitru, E. I. Verriest, International Journal of Theoretical Physics 34, 1785 (1995).

77. V. V. Mazurenko, I. A. Iakovlev, O. M. Sotnikov, M. I. Katsnelson, Journal of the Physical Society of Japan 92, 081004 (2023).

78. J. Kronjäger, C. Becker, P. Soltan-Panahi, K. Bongs, K. Sengstock, Physical review letters 105, 090402 (2010).

79. J. Hertkorn, et al., Physical Review Research 3, 033125 (2021).

80. P. Czoschke, H. Hong, L. Basile, T.-C. Chiang, Physical review letters 93, 036103 (2004).

81. A. Rowe, J. Nehls, R. Stradling, R. Ferguson, Physical Review B 63, 201307 (2001).

82. M. V Putz, Current Physical Chemistry 1, 111 (2011).

83. T. Szandała, Procedia Computer Science 71, 68 (2015).

84. R. Neigovzen, J. L. Neves, R. Sollacher, S. J. Glaser, Physical Review A 79, 042321 (2009).

85. D. M. J. Tax (2002).

86. S. S. Khan, M. G. Madden, The Knowledge Engineering Review 29, 345 (2014).

87. S. S. Khan, M. G. Madden, Artificial Intelligence and Cognitive Science: 20th Irish Conference, AICS 2009, Dublin, Ireland, August 19-21, 2009, Revised Selected Papers 20 (Springer, 2010), pp. 188–197.

88. H. Hoffmann, Pattern recognition 40, 863 (2007).

89. B. Schölkopf, R. C. Williamson, A. Smola, J. Shawe-Taylor, J. Platt, Advances in neural information processing systems 12 (1999).

90. J. C. Platt, J. Shawe-Taylor, A. J. Smola, R. C. Williamson, et al., Technical Report MSR-T R-99–87, Microsoft Research (MSR) (1999).

91. F. S. Gharehchopogh, Artificial Intelligence Review 56, 5479 (2023).

92. L. Gao, L. Zhang, C. Liu, S. Wu, Artificial intelligence in medicine 108, 101935 (2020).

93. B. Krawczyk, M. Woźniak, F. Herrera, 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM) (IEEE, 2014), pp. 337–344.

94. X. Du, Journal of Process Control 83, 1 (2019).

95. J. Lee, Y. C. Lee, J. T. Kim, Journal of Manufacturing Systems 57, 357 (2020).

96. M. A. Md Ali, N. Badrud'din, H. Abdullah, F. Kemi, International Journal of Modern Physics A 35, 2050131 (2020).

97. E. J. Spinosa, A. C. de Carvalho, Advances in Bioinformatics and Computational Biology: Brazilian Symposium on Bioinformatics, BSB 2005, Sao Leopoldo, Brazil, July 27-29, 2005. Proceedings (Springer, 2005), pp. 54–64.

98. M. Yousef, M. D. Sac¸ar Demirci, W. Khalifa, J. Allmer, Advances in bioinformatics 2016 (2016).

99. G. Park, J. Huh, D. K. Park, Machine Learning: Science and Technology (2022).

100. J. Romero, J. P. Olson, A. Aspuru-Guzik, Quantum Science and Technology 2, 045001 (2017).

101. A. Pepper, N. Tischler, G. J. Pryde, Physical review letters 122, 060501 (2019).

102. C.-J. Huang, et al., Physical Review A 102, 032412 (2020).

103. D. F. Locher, L. Cardarelli, M. Mu¨ller, Quantum 7, 942 (2023).

104. D. Bondarenko, P. Feldmann, Physical review letters 124, 130502 (2020).

105. F. Zhou, et al., Applied Physics Letters 121, 134001 (2022).

106. C. Harney, S. Pirandola, PRX Quantum 3, 010311 (2022).

107. F. Leymann, Quantum Technology and Optimization Problems: First International Work- shop, QTOP 2019, Munich, Germany, March 18, 2019, Proceedings 1 (Springer, 2019), pp. 218–230.

108. S. Perdrix, Electronic Notes in Theoretical Computer Science 170, 125 (2007).

109. I. Luchnikov, S. Vintskevich, D. Grigoriev, S. Filippov, Physical review letters 124, 140502 (2020).

110. X.-G. Wen, Science 363, eaal3099 (2019).

111. A. Montanaro, npj Quantum Information 2, 1 (2016).

112. R. Cleve, A. Ekert, C. Macchiavello, M. Mosca, Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454, 339 (1998).

113. B. M. Terhal, Nature Physics 14, 530 (2018).

114. P. W. Shor, SIAM review 41, 303 (1999).

115. S. P. Jordan, D. Gosset, P. J. Love, Physical Review A 81, 032331 (2010).

116. Y.-K. Liu, M. Christandl, F. Verstraete, Physical review letters 98, 110503 (2007).

117. T. F. Rønnow, et al., science 345, 420 (2014).

118. V. Vedral, Foundations of Physics 40, 1141 (2010).

119. E. Crosson, D. Lidar, Nature Reviews Physics 3, 466 (2021).

120. R. B. Patel, T. Rudolph, G. J. Pryde, Science advances 5, eaau6668 (2019).

121. A. W. Harrow, A. Hassidim, S. Lloyd, Physical review letters 103, 150502 (2009).

122. B. D. Clader, B. C. Jacobs, C. R. Sprouse, Physical review letters 110, 250504 (2013).

123. M. Schuld, F. Petruccione, Supervised learning with quantum computers, vol. 17 (Springer, 2018).

124. M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, B. Scholkopf, IEEE Intelligent Systems and their applications 13, 18 (1998).

125. F. Samuelson, D. G. Brown, The 2011 International Joint Conference on Neural Networks (IEEE, 2011), pp. 1020–1026.

126. K. L. Pudenz, D. A. Lidar, Quantum information processing 12, 2027 (2013).

127. D. Anguita, S. Ridella, F. Rivieccio, R. Zunino, Neural Networks 16, 763 (2003).

128. R. Zhang, J. Wang, N. Jiang, H. Li, Z. Wang, Neural Networks 151, 376 (2022).

129. R. Zhang, J. Wang, N. Jiang, Z. Wang, Information Sciences 635, 25 (2023).

130. H. Li, N. Jiang, R. Zhang, Z. Wang, H. Wang, International Journal of Theoretical Physics 61, 92 (2022).

131. D. Willsch, M. Willsch, H. De Raedt, K. Michielsen, Computer physics communications 248, 107006 (2020).

132. M. Ezawa, Scientific Reports 12, 6758 (2022).

133. P. Cunningham, S. J. Delany, ACM computing surveys (CSUR) 54, 1 (2021).

134. M. Schuld, I. Sinayskiy, F. Petruccione, Contemporary Physics 56, 172 (2015).

135. D. Dong, C. Chen, H. Li, T.-J. Tarn, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 38, 1207 (2008).

136. V. Dunjko, J. M. Taylor, H. J. Briegel, 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (IEEE, 2017), pp. 282–287.

137. V. Dunjko, Y.-K. Liu, X. Wu, J. M. Taylor, arXiv preprint arXiv:1710.11160 (2017).

138. C. J. Watkins, P. Dayan, Machine learning 8, 279 (1992).

139. A. Skolik, S. Jerbi, V. Dunjko, Quantum 6, 720 (2022).

140. S. M. Kakade, Advances in neural information processing systems 14 (2001).

141. J. Yao, M. Bukov, L. Lin, Mathematical and Scientific Machine Learning (PMLR, 2020), pp. 605–634.

142. L. J. Fiderer, J. Schuff, D. Braun, Prx Quantum 2, 020303 (2021).

143. Y. LeCun, Y. Bengio, G. Hinton, nature 521, 436 (2015).

144. I. Goodfellow, Y. Bengio, A. Courville, Deep learning (MIT press, 2016)

145. S. I. Gallant, et al., IEEE Transactions on neural networks 1, 179 (1990).

146. B. Ricks, D. Ventura, Advances in neural information processing systems 16 (2003).

147. Z. Zhao, A. Pozas-Kerstjens, P. Rebentrost, P. Wittek, Quantum Machine Intelligence 1, 41 (2019).

148. I. Cong, S. Choi, M. D. Lukin, Nature Physics 15, 1273 (2019).

149. I. Kerenidis, J. Landman, A. Prakash, International Conference on Learning Representa- tions.

150. M. Henderson, S. Shakya, S. Pradhan, T. Cook, Quantum Machine Intelligence 2, 2 (2020).

151. P. Gohel, A. Chakraborty, K. R. JV, 2022 IEEE International Conference on Quantum Computing and Engineering (QCE) (IEEE, 2022), pp. 831–832.

152. J. Barr, Synchronic or Diachronic? (Brill, 1995), pp. 1–14.

153. T. Cho, Language and Linguistics Compass 10, 120 (2016).

154. S. Yolchuyeva, G. Ne´meth, B. Gyires-To´th, arXiv preprint arXiv:2004.06338 (2020).

155. S. C. Kak, International Journal of Approximate Reasoning 1, 117 (1987).

156. V. Raskin, Machine translation: Theoretical and methodological issues pp. 42–58 (1987).

157. W. Zeng, B. Coecke, arXiv preprint arXiv:1608.01406 (2016).

158. N. Wiebe, A. Kapoor, K. M. Svore, Quantum information and computation 15, 318 (2015).

159. S. Clark, B. Coecke, M. Sadrzadeh, Proceedings of the Second Quantum Interaction Sym- posium (QI-2008) (Citeseer, 2008), pp. 133–140.

160. R. Guarasci, G. De Pietro, M. Esposito, Applied Sciences 12, 5651 (2022).

161. R. Lorenz, A. Pearson, K. Meichanetzidis, D. Kartsaklis, B. Coecke, Journal of Artificial Intelligence Research 76, 1305 (2023).

162. Y. Liu, S. Arunachalam, K. Temme, Nature Physics 17, 1013 (2021).

163. J. Liu, et al., Machine Learning: Science and Technology 4, 025003 (2023).

164. J. J. Meyer, et al., PRX Quantum 4, 010328 (2023).

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

166. S. Arunachalam, R. De Wolf, The Journal of Machine Learning Research 19, 2879 (2018).

167. L. Banchi, J. Pereira, S. Pirandola, PRX Quantum 2, 040321 (2021).

168. Z. Krunic, F. F. Flo¨ther, G. Seegan, N. D. Earnest-Noble, O. Shehab, IEEE Transactions on Quantum Engineering 3, 1 (2022).

169. S. F. Ahmad, R. Rawat, M. Moharir, 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE) (IEEE, 2021), pp. 345–349.

170. J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, H. Neven, Nature communications 9, 4812 (2018).

171. X. Ge, R. Wu, H. Rabitz, Complex System Modeling and Simulation 1, 77 (2021).

172. A. Pe´rez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, J. I. Latorre, Quantum 4, 226 (2020).

173. V. Havl´ıcˇek, et al., Nature 567, 209 (2019).

174. W. M. Watkins, S. Y.-C. Chen, S. Yoo, Scientific Reports 13, 2453 (2023).

175. H. Liao, I. Convy, Z. Yang, K. B. Whaley, Quantum Machine Intelligence 5, 7 (2023).