Quantum Machine Learning

Jacob Biamonte,[1,2,*] Peter Wittek,[3,4,†] Nicola Pancotti,[5,‡] Patrick Rebentrost,[6,§] Nathan Wiebe,[7,¶] and Seth Lloyd[8,**]

[1] *Quantum Complexity Science Initiative Department of Physics, University of Malta, MSD 2080 Malta*
[2] *Institute for Quantum Computing University of Waterloo, Waterloo, N2L 3G1 Ontario, Canada*
[3] *ICFO-The Institute of Photonic Sciences Castelldefels (Barcelona), 08860 Spain*
[4] *University of Borås Borås, 501 90 Sweden*
[5] *Max Planck Insitute of Quantum Optics Hans-Kopfermannstr. 1, D-85748 Garching, Germany*
[6] *Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, MA 02139*
[7] *Station Q Quantum Architectures and Computation Group, Microsoft Research, Redmond WA 98052*
[8] *Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge MA 02139 USA*

Recent progress implies that a crossover between machine learning and quantum information processing benefits both fields. Traditional machine learning has dramatically improved the benchmarking and control of experimental quantum computing systems, including adaptive quantum phase estimation and designing quantum computing gates. On the other hand, quantum mechanics offers tantalizing prospects to enhance machine learning, ranging from reduced computational complexity to improved generalization performance. The most notable examples include quantum enhanced algorithms for principal component analysis, quantum support vector machines, and quantum Boltzmann machines. Progress has been rapid, fostered by demonstrations of midsized quantum optimizers which are predicted to soon outperform their classical counterparts. Further, we are witnessing the emergence of a physical theory pinpointing the fundamental and natural limitations of learning. Here we survey the cutting edge of this merger and list several open problems.

Machine learning has fundamentally changed the way humans interact with and relate to data. Applications range from self-driving cars to intelligent agents capable of exceeding the best humans at *Jeopardy* and *Go*. These applications exhibit large data sets and push current algorithms and computational resources to their limit. Information is fundamentally governed by the laws of physics. The laws are quantum mechanical at the scales of present day information processing technology, in contrast to the more familiar 'classical' physics at the human scale. The interface of quantum physics and machine learning naturally goes both ways: machine learning algorithms find application in understanding and controlling quantum systems and, on the other hand, quantum computational devices promise enhancement of the performance of machine learning algorithms for problems beyond the reach of classical computing.

Machine learning is rapidly being employed for the benchmarking, control, and harnessing of quantum effects [1–9]. State-of-the art quantum experiments in optical or solid state systems have recently reached sizes where optimization methods face unprecedented data-intensive landscapes. In addition, machine learning was employed in a variety of related fields, e.g. the discovery of the Higgs Boson [10], molecular energy prediction trained using databases of known energy spectra [11] and gravitational wave detection [12]. In the computing realm, this progress allows experimental breakthroughs which probe the threshold of producing the first practical quantum computer [13], which in turn enables quantum enhanced versions of these very same learning algorithms.

Quantum information has shown promising algorithmic developments leading to quantum speedups of computational problems such as prime number factoring and searching an unstructured database. The underlying algorithmic toolbox allows extensions to problems relevant for machine learning and artificial intelligence. Recently, it was shown that quantum mechanics offers physical resources to enhance machine learning with quantum algorithms [14–21]. Quantum-enhanced versions of classical machine learning algorithms include least-squares fitting, support vector machines, principal component analysis, and deep learning. Challenges that have to be addressed in this emerging field is the input of classical data into the quantum device, the efficient processing of the data, and subsequent readout of classically relevant information.

Beyond quantum algorithms for machine learning, there has been progress in developing a physics based the-

* jacob.biamonte@qubit.org; www.QuamPlexity.org
† peterwittek.com
‡ nicola.pancotti@mpq.mpg.de
§ pr@patrickre.com
¶ nawiebe@microsoft.com
** slloyd@mit.edu

ory pinpointing the fundamental and natural limitations of learning, quantum enhanced learning algorithms and the employment of learning algorithms to better control and harness these same quantum effects [14–21]. Quantum information theory sets the stage to understand how fundamental laws of nature impact the ability of physical agents to learn. The cutting edge of the intersection of machine learning and quantum physics is reviewed here. We explain how the above areas interact and we list several open problems that are of contemporary research interest.

CONTENTS

I. CLASSICAL LEARNING IN QUANTUM SYSTEMS

Recent decades have seen a concerted effort to design, develop, benchmark, and control systems operating in a quantum regime. Such systems range from condensed phase systems such as Bose-Einstein condensates, quantum clocks, and quantum computers in optical, solid-state, and other environments. For quantum computers, the goal is to achieve 'quantum supremacy' when a quantum computer outperforms a conventional computer for a particular problem. Classical learning algorithms were recently employed for several building blocks needed in such a quantum computational device. This is particularly timely as the data size of these problems now makes exhaustive and greedy approaches either impossible or at best, highly non-optimal. Quantum computing gates can be optimized using machine learning and evolutionary algorithms. In addition, analyzing the data output from measurement of even small quantum devices benefits from modern data-processing algorithms.

a. **Learning about quantum systems.** Experimental quantum systems must be characterized and benchmarked under laboratory conditions in order for them to be controlled. A tantamount task is then to find a model (a.k.a. effective) Hamiltonian of the system and to determine properties of the present noise sources. By computing likelihood functions in an adaptation of Bayesian inference, Wiebe et al. [22–25] found that quantum Hamiltonian learning can be performed using realistic resources such as depolarizing noise. Wiebe et al. [24] further provides empirical evidence that their learning algorithm will find an approximation that is maximally close to the true model when facing cases where the hypothetical model lacks terms present in the actual one. This suggests that even imperfect quantum resources may be valuable when applying learning methods to characterize quantum systems.

Sasaki et al. [1, 26] pioneered the approach framing the classification of unknown quantum states as a form of supervised learning. The authors considered semiclassical and fully coherent quantum strategies, proving that the latter is optimal [1, 26]. Bisio et al. [2] considered learning a unitary transformation from a finite number of examples. The best strategy for learning a unitary involves a double optimization that requires both an optimal input state—akin to active learning in the classical theory of statistical learning—and an optimal measurement, thus this protocol is incoherent and enables induction in the classical sense [2]. In a separate study, Bisio et al. [3] derived a learning algorithm for arbitrary von Neumann measurements such that, differently from the learning of unitary gates, the optimal algorithm for learning of quantum measurements was not able to be parallelized, and required quantum memory for the storage of quantum information [3].

The authors in [4] also devised a quantum learning machine for binary classification of qubit states that does not require/in no need of a quantum memory. The required classical memory was found to grow only logarithmically with the number of training qubits [4]. The binary discrimination problem was considered in [5] specifically for the case of coherent states of light. They found that a global measurement, performed jointly over the signal and the training set, enhances identification rates compared to learning strategies based on first estimating the unknown amplitude by means of Gaussian measurements on the training set, followed by an adaptive discrimination procedure on the signal [5].

Concept drift is an essential problem in machine learning: it refers to shifts in the distribution that is being sampled and learned [27]. A similar problem in quantum mechanics is detecting the change point, that is, identifying when a source changes its output quantum state(s). The work of [9] constructs strategies for measuring the particles individually and provides an answer as soon as a new particle is emitted by the source, replicating the overall scheme of online learning. The authors also show that these strategies underperform the optimal strategy, which is a global measurement. Sasaki et al. [1, 26] pioneered this approach by framing the classification of unknown quantum states as a form of supervised learning. The authors considered semiclassical and fully coherent quantum strategies, proving that the latter is optimal [1, 26]. Learning the 'community structure' of quantum states and walks was considered in [28] by means of maximizing modularity with hierarchical clustering.

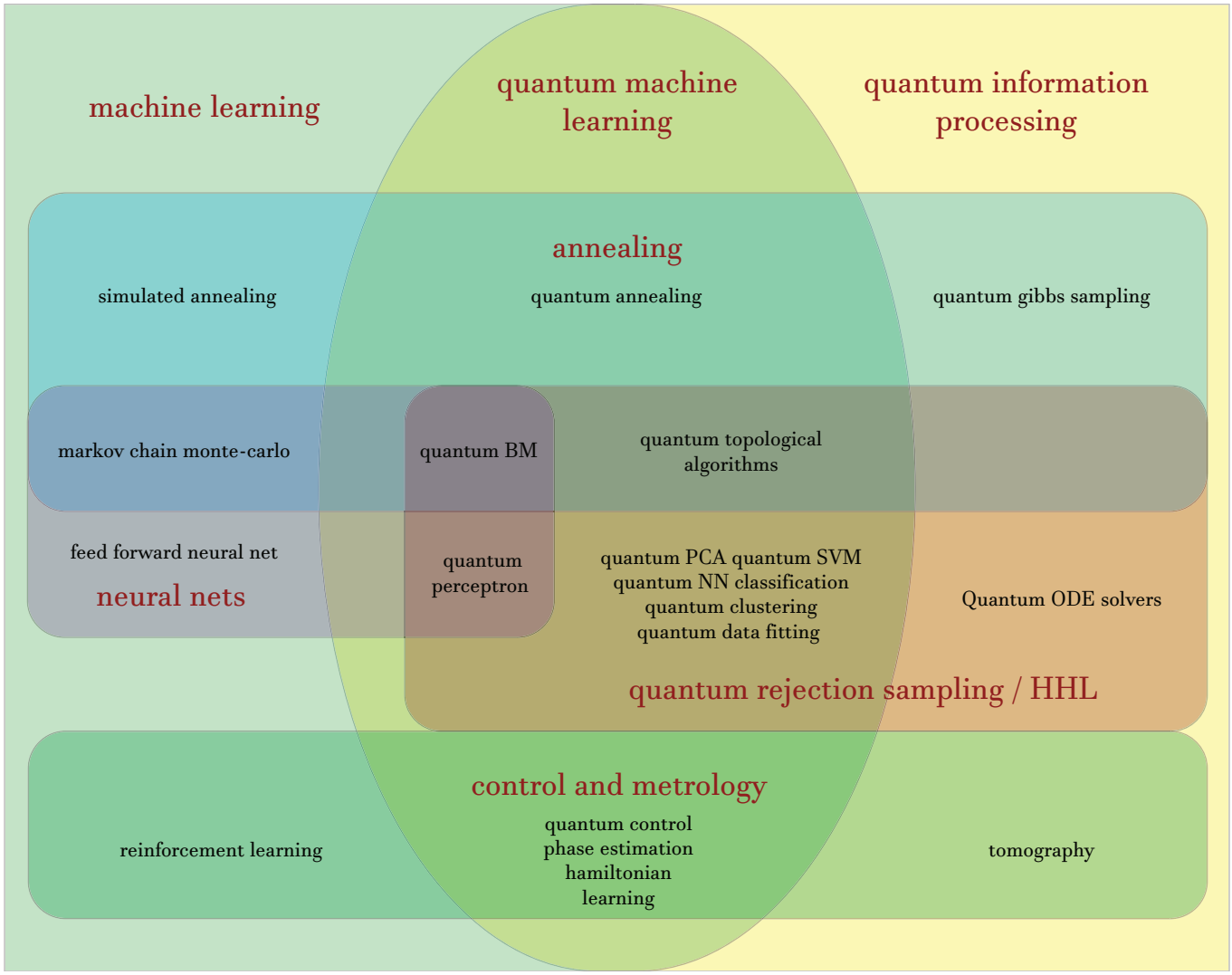

FIG. 1. **Conceptual depiction of mutual crossovers between quantum and traditional machine learning.**

b. **Controlling quantum systems** Learning methods have also seen ample success in developing control sequences to optimize interferometric 'quantum phase estimation' which is a key quantum algorithmic building block [29, 30] that appears in quantum simulation algorithms and elsewhere [31], used as a key component in [32] in a proposal for a quantum perceptron. Having employed heuristic global optimization algorithms, Hentschel and Sanders [29] optimized many-particle adaptive quantum metrology in a reinforcement learning scenario. Later Lovett et al. [30] extended their procedure to several challenges including phase estimation and coined quantum walks. Palittapongarnpim et al. [33] optimized this latter approach by orders of magnitude while also improving on noise tolerance and robustness.

A similar heuristic methodology has been developed to create quantum gates (a challenge for several decades in the development of quantum computation and information science) [34–37]. In the presence of noise and by adapting a differential evolution scheme, Zahedinejad, Ghosh and Sanders [34] considered nearest-neighbor-coupled superconducting artificial atoms and employed supervised learning, resulting in gate fidelity above 99.9% and hence reaching an accepted threshold for fault-tolerant quantum computing. In a separate study [35], Zahedinejad, Ghosh and Sanders developed a quantum-control procedure to construct a single-shot Toffoli gate (a crucial building block of a universal quantum computer), again reaching gate fidelity above 99.9%. Using an alternative approach, Banchi, Pancotti and Bose [36] also realized a Toffoli gate without time-dependent control using the natural dynamics of a quantum network. Las et al. [38] used genetic algorithms to reduce digital and experimental errors in quantum gates. The authors [38] added ancillary qubits to design a modular gate

made out of imperfect gates, so that their fidelity is interestingly greater than the fidelity of any of the constituent gates. To realize quantum gates, memories and protocols, contemporary methods to develop dynamical decoupling sequences (a leading method to protect quantum states from decoherence) can also be surpassed using recurrent neural networks—see for instance August and Ni [39].

Common to these approaches in quantum gate design is that they work in a supervised learning setting, in contrast to the quantum adaptive phase estimation which is closer to control theory and uses reinforcement learning. One can also exploit reinforcement learning in gate-based quantum systems. For instance, Tiersch, Ganahl and Briegel [40] laid out a path for adaptive controllers based on intelligent agents for quantum information tasks, illustrating how to adapt to measurement directions while corresponding to an external stray field of unknown magnitude in a fixed direction can be overcome—which they then applied to a measurement-based algorithm for Grover's search [40]. Mavadia et al. also used a reinforcement learning scheme to predict and compensate for qubit decoherence [41].

Other quantum algorithms directly involve ideas from machine learning in their basic operation. Most notably, the iterative phase estimation algorithm uses concepts from machine learning to infer eigenvalues of a given unitary operator. These techniques allow the algorithm to be run using fewer qubits and also using far less experimental time than previous methods. This approach, originally proposed by Kitaev, was further refined by Higgins, Berry et al [42, 43] who explored the use of adaptive methods to optimally learn the unknown eigenphase. Such use of adaptive policies to learn and infer eigenphases was pioneered by Hentschell and Sanders [29]. Wiebe and Granade provided efficient alternative methods to policy based phase estimation methods by using a form of adaptive Bayesian inference, itself based on assumed density filtering [44]. These works illustrate that the process of data extraction from quantum algorithms can be meaningfully influenced by ideas from machine learning.

Future applications of supervised machine learning to tackle noise, tailor gates and develop core quantum information processing building blocks is a direction of tantamount importance. Reinforcement learning in quantum control should also be further explored—see Rosi et al. [45] for a prime example. Furthermore, quantum walks— representing an established model that captures essential physics behind many natural and synthetic phenomena, and proven to provide a universal model of quantum computation—were briefly touched upon here [28, 30]. To date however, comparatively little work [28, 30, 46] has been done towards a merger with machine learning, providing an interesting avenue of open problems for future research.

c. **Learning properties of quantum and statistical physics.** Classical machine learning has recently unveiled properties of quantum and related statistical systems, such as critical points of phase transitions [47] or expectation values of observables [48], and can be employed in other related simulation tasks [38, 49] leading to applications in several fields facing many-body problems.

Making use of Google's deep-learning 'TensorFlow' library [50], Carrasquilla and Melko [47] developed a learning procedure capable of determining the current phase of matter of a quantum system. The work is based on a standard feed-forward neural network (for proposals that realize neural networks in quantum dots, see [51, 52]), and showed that it can be trained to detect multiple types of order parameters directly from raw state configurations sampled with Monte Carlo methods. Interestingly this particular network in the work [47] is not aware of the model Hamiltonian which generated the data, or the length of the interactions. This analysis outputs nontrivial results for a large variety of models, ranging from the classical Ising model to Coulomb phases and topological phases [47].

A simple recurrent neural network, a so-called Boltzmann machine, is able to faithfully reproduce expectation values by creating a large set of configurations via Monte Carlo sampling from the partition function of an Ising Hamiltonian at different temperatures [48]. Those configurations are then used to train, test and validate the Boltzmann machine. Once the learning has converged, characteristic physical properties—such as energy, magnetization and specific heat—are computed. Near the transition point, one appears to experience more difficult learning when the associated number of neurons in the network are required to achieve the same level of precision [48].

Choosing a Boltzmann machine with hidden variables as an ansatz for the wave function, Carleo and Troyer [49] address the many-body problem—central to physics, materials science and chemistry—through a search method for a lowest-energy state. Such a function is then trained via a pseudo-gradient descent algorithm originally designed for Monte Carlo simulations in chemistry. Furthermore Carleo and Troyer [49] challenged their result, comparing it against tensor network algorithms (see Section II 0 g) in both one and two dimensions, and concluding that their own method systematically improves the best known variational states for 2D finite lattice systems. Deng et al. [53] extend this idea to topological states with long-range entanglement, showing analytically that a topological ground state can be represented *exactly* by a *short-range* Restricted Boltzmann Machine.

II. QUANTUM ENHANCED LEARNING

Quantum mechanics can enhance machine learning in two different ways. First, a quantum computational device could perform machine learning algorithms for problems beyond the reach of classical computers. We discuss recent developments in quantum techniques for big data, adiabatic optimization, and Gibbs sampling. Sec-

ond, techniques developed in quantum theory can improve machine learning algorithms. In this context, we discuss tensor networks, renormalization, and Bayesian networks.

d. **Quantum techniques for big data.** Extremely large data sets have become widespread and regularly analysed to reveal patterns, trends, and associations, ranging from many areas of physical sciences to human behavior and economics. As quantum physics offers certain enhancements in the storage and processing of information, a clear research track is to develop and tailor these quantum methods to apply to problems when facing 'big data' sets [14, 15, 18–20, 48, 49, 54–59].

A quantum speedup is characterized in several different ways. One characterization is by the query complexity, that is the number of queries to the information storage medium for the classical or quantum algorithm, respectively. The storage medium can be more abstractly considered to be an oracle and the algorithmic speedup is relative to that oracle [60]. Another way of characterizing performance is the gate complexity, counting the number of elementary gates, say single and two qubit gates, required to obtain the desired results. Many recent quantum algorithms for machine learning rely on two main types of speedups. First, amplitude amplification is commonly used to quadratically reduce the number of samples needed in sampling algorithms. Specifically, if N samples would be required on average in a sampling algorithm then amplitude amplification can be used to reduce this to $O(\sqrt{N})$ samples on average. Grover search problem is a well known example of amplitude amplification, and so such quadratic speedups are often called "Grover-like". Second, other types of are speedups are related to prime number factoring and finding eigenvalues and eigenvectors of large matrices. This speedup is enabled by quantum phase estimation, quantum Fourier transform, and quantum simulation methods. In many cases, the number of quantum gates is proportional to $O(\log N)$ for preparing a quantum state encoding eigenvalues of an $N \times N$ matrix and the associated eigenstates, while classically $O(N)$ operations are required to find eigenvalues and eigenvectors.

Early work by Ventura and Martinez [61] applied quantum computing to training associative memories that built on discrete Grover's search. Their modification allows storing only a few patterns in a superposition, and the retrieval protocol receives the most similar ones to a given new instance. Grover's search can be used for discrete optimization, and Anguita et al. [62] applied this variant to train support vector machines. Their idea was later generalized to create building blocks of learning algorithms using Grover's search [54, 63]. Common to these approaches is discretization of the search space to achieve a quadratic speedup over classical counterparts. By a similar technique, [64] proves rigorous bounds on the learning capacity of a quantum perceptron.

Harrow, Hassidim and Lloyd [14] provided a quantum algorithm to solve linear systems (in which given a matrix A and a vector b, one is faced with finding a vector x such that $Ax = b$). Matrix inversion represents a commonly employed subroutine in data science and learning algorithms. In their variant of the problem [14], one does not need to know the solution x itself, but rather an approximation of the expectation value of some operator associated with x. They recovered an exponential improvement over the best known classical algorithms when A is sparse and 'well conditioned' [14]. By developing a state preparation routine that can initialize generic states, Clader, Jacobs and Sprouse [15] show how elementary ancilla measurements can be used to calculate quantities of interest, and hence integrate a quantum-compatible preconditioner which expands the number of problems that can achieve exponential speedup over classical linear system solvers for constant precision solutions. They further demonstrated that their algorithm can be used to compute the electromagnetic scattering cross section of an arbitrary target exponentially faster than the best known classical algorithm [15]. Building on these linear systems results, a quantum algorithm discovered by Wiebe, Braun and Lloyd efficiently determines the quality of a least-squares fit over an exponentially large data set [16]. They further suggest that in many cases their algorithms can also efficiently find a concise function that approximates the data to be fitted and bound the approximation error [16], particularly when the data is sparse. Wang [65] uses singular value decomposition for the same purpose, replacing sparsity by a low-rank condition employing the quantum principal component analysis of Lloyd et al. [20]. Keeping the same assumption, Schuld et al. [66] developed a protocol for predicting labels for new points in regression.

A quantum algorithm for the support vector machine based on matrix inversion was provided by Rebentrost, Mohseni and Lloyd [18]. Relying on a least-squares formulation of the support vector machine, this algorithm was shown to have run time logarithmic in the number of features and training examples for both training of the classifier, and the classification of new data. In cases when classical sampling algorithms terminate in polynomial time, an exponential quantum speed-up in queries to the training data can be achieved. Central to their quantum algorithm [18] is a non-sparse matrix exponentiation technique for efficient matrix inversion of the training data inner-product (kernel) matrix.

Returning to the problem of supervised vs. unsupervised learning, Lloyd, Mohseni and Rebentrost [17] discovered quantum machine learning algorithms for cluster assignment and cluster finding— providing a polynomial speedup over sampling based classical methods for k–means clustering [17, 19].

Finding nearest-neighbors is an association problem faced in data-analysis—some of these classical methods have been applied to determine the so called community structure of quantum transport problems [28]. Finding nearest-neighbors on a quantum computer was addressed with a quantum algorithm discovered by Wiebe, Kapoor

and Svore in [19]. Central to the algorithm are several subroutines for computing distance metrics such as the inner product and Euclidean distance. Careful analysis revealed that even in the worst case, the quantum algorithms offer polynomial reductions in query complexity over classical sampling based methods.

In [20], Lloyd, Mohseni and Rebentrost devised a quantum algorithm for principal component analysis of an unknown low-rank density matrix. The main idea is to take multiple copies of a possibly unknown density matrix and apply it as a Hamiltonian to another quantum state. As in quantum tomography, such a density matrix can be prepared from an arbitrary quantum process not necessarily involving QRAM. This allows large eigenvalues and corresponding eigenvectors of the density matrix to be computed. If constant precision is required, this method can accomplish the task by using exponentially fewer accesses to the training data than any existing classical algorithm. In an oracular (or QRAM) setting, this effort was later extended to the singular value decomposition of non-sparse low-rank and non-positive matrices, and applied to the Procrustes problem of finding the best orthogonal matrix mapping one matrix into another [67]. Moreover, if class labels are also available, linear discriminant analysis is more advantageous than principal component analysis. Cong and Duan [68] adapted the quantum algorithm for solving linear equations [14] to achieve an exponential reduction in the number of queries made to the data for this task as well. These scenarios are special cases of manifold learning algorithms, where it is assumed that the data points lie on some high-dimensional manifold. Principal component analysis and singular value decomposition ensure a global optimum, but often one is more interested in the topology of the data instances, such as connected components and voids. Lloyd, Garnerone and Zanardi [55] designed quantum algorithms for the approximation of Betti numbers from the combinatorial Laplacian for a type of topological manifold learning known as persistent homology. Their algorithm provided an exponential speedup for computing constant precision approximations to Betti numbers relative to the current best known classical algorithms. Dridi and Alghassi [69] also use quantum annealing for homology computation. While the empirical results of their algorithm look encouraging, more work is needed to assess whether their approach truly can give an exponential speedup.

Quantum mechanics was also shown in [6] to provide a speed-up for reinforcement learning. A large class of learning agents was introduced, for which a quadratic boost in learning efficiency over their classical analogues was recovered [6]. Development of learning agents in quantum environments was further considered in [7, 8]. In [7] classical agents were 'upgraded' to their quantum counterparts by a nested process of adding coherent control, where the focus was on implementation in ion traps. Further, in [8] the authors analyze the types of classically specified environments which allow for quantum enhancements in learning. They conclude that if the agent has quantum resources while the environment is classical, the only improvements can be in terms of computational complexity, and they show scenarios for a quadratic speedup by Grover-like protocols [8].

A challenge facing the application of many of these methods for big data is the fact that the training set of classical data must be loaded into the quantum computer, a step that can dominate the cost of the algorithm in some cases [70]. This issue does not, however, occur if the data are provided via an efficient quantum subroutine or a pre–trained generative model. The alternative solution is to load the data into a QRAM, which is a low depth circuit for accessing data in quantum superposition. Work is ongoing to engineer inexpensive QRAMs in both existing [71] and fault-tolerant hardware [72], as well as benchmarking the performance of QRAM enabled algorithms against massively parallel classical machine learning algorithms.

e. **Adiabatic quantum optimization.** Adiabatic quantum computing relies on the idea of embedding a problem instance into a physical system, such that the system's lowest energy configuration stores the problem instance solution [73]. Recent experimental progress has resulted in annealers with hundreds of spins [74]—detailed further in Section III.

These annealers make use of a logical Ising model, providing an immediate connection to Hopfield neural networks [75], as well as many other models phrased in terms of energy minimization of the Ising model. Indeed, at the heart of many learning algorithms is a constrained optimization problem, which can be restated as an energy minimization problem of an Ising model.

Adiabatic quantum optimization relies on a physical process to estimate the ground state energy of the Ising model—resembling the widely used global optimization heuristic that exploits both thermal fluctuations and quantum tunneling to find the global energy minimum of a system—see figure 2 **A**. In other words, given a discrete nonconvex optimization problem, we are able to find the global optimum as long as we meet the criteria of the adiabatic theorem that drives the physical process [76]. Adding non-commuting (so called, xx) interactions to the Ising model is known to render it universal [77] for adiabatic quantum computation—yet programming this universal model and understanding its connection to machine learning is an open problem.

Denchev et al. developed robust, regularized boosting algorithms using quantum annealing [78, 79]. Dulny III and Kim [80] used a similar methodology in a range of tasks, including natural language processing and testing for linear separability, whereas Pudenz and Lidar [81] applied it to anomaly detection. Learning the structure of a probabilistic graphical model, for instance that of a Bayesian network, is a notoriously hard task: O'Gorman et al. [82] address this difficulty by quantum annealing. They map the posterior-probability scoring function on graphs to the Ising model: n random variables map to

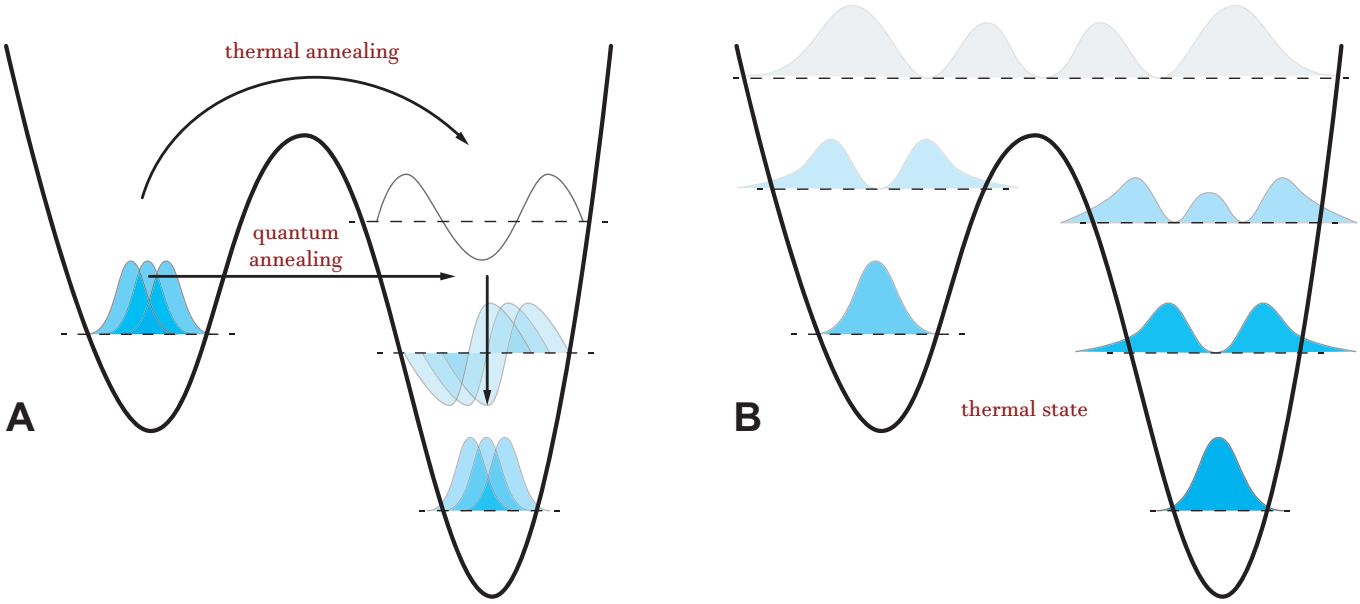

FIG. 2. **A quantum state tunnels when approaching a resonance point before decoherence induces thermalization**. **A**. A quantum state must traverse a local minimum in thermal annealing whereas a coherent quantum state can tunnel when brought close to resonance. **B**. Coherent effects decay through interaction with an environment, resulting in a probability distribution in occupancy of a systems energy levels following a Gibbs distribution.

$\mathcal{O}(n^2)$ qubits.

Limited connectivity in current quantum annealers is a recurrent problem in developing quantum optimization algorithms. Zaribafiyan et. al [83] devised a generic, efficient graph-minor embedding methods to address this issue. Following a similar line of thought, Diridi and Alghassi [69] designed a quantum annealing algorithm for manifold learning, more specifically, for the computation of homology of a large set of data points (see also Section II 0 d). Benedetti et al. [84] in turn developed an embedding of arbitrary pairwise connectivity to train a maximum entropy model with at most quadratic number of qubits required to represent the nodes of the original graph.

f. **Gibbs sampling.** Current quantum annealing technology is seldom guaranteed to provide the global optimum. Instead, the energy levels after repeated steps of annealing approximately follow a Gibbs distribution—see figure 2 **B**. Addressing the correct embedding on the connectivity graph and estimating the temperature can be used for training Boltzmann machines [85, 86]. Such machines appear in several variants in this review [48, 49, 58, 85–89], and are simple generative neural networks consisting of hidden and visible nodes. Typically, classical methods focus on training restricted Boltzmann machines that only have connectivity between the adjacent layers of hidden and visible units but not within a layer. Deep belief networks, extensively used in speech and image recognition, can be formed by stacking many restricted Boltzmann machines. Exact training of Boltzmann machines requires Gibbs sampling, but given the computational complexity thereof, a heuristic algorithm called contrastive divergence can be employed. While such contrastive divergence often suffices for machine learning, it can fail to converge to the solution provided by exact training and also it cannot be used directly to efficiently train non–restricted Boltzmann machines. Quantum Gibbs sampling replaces this heuristic.

Wiebe et al. [88] developed a Gibbs state preparation and sampling protocol, also with the objective of training deep belief networks. They achieve polynomial improvements in computational complexity relative to its classical analogue, and in some cases superpolynomial speedups relative to contrastive divergence training. Furthermore, their state preparation procedure is not specific to any topology for Boltzmann machines, which allows deep networks that are not stacked restricted Boltzmann machines to be accurately trained. Further advances in Gibbs state preparation methods [89] open the door to improved methods for training other graph topologies, such as Markov logic networks [90].

Taking these ideas further, Amin et al. [58] suggest an approach based on quantum Boltzmann distribution of a transverse-field Ising Hamiltonian as the fundamental model for the data. While modest advantages are seen for small networks, more work is needed to understand the power that such quantum models possess.

g. **Learning tensor networks and renormalization.** Tensor networks have taken a central role in modern quantum physics due to their ease of use as a graphical language to describe and pictorially reason about quantum circuits and protocols, renormalization, and numerical tensor contraction and simulation algorithms. These tensor network algorithms are known to efficiently

represent the low-energy wave function for a vast family of Hamiltonians and they offer a means to efficiently simulate classes of quantum circuits. At the core of the algorithms we find something roughly similar to principle component analysis, yet in the case of quantum systems singular value decomposition is applied recursively to factor stationary states, and often sequentially repeated when modeling time-dependence. Accordingly, various geometrical constructions have been shown to offer advantages when compressing the data required to represent different classes of quantum states—see figure 1. These methods apply well to certain classical systems and problems—for instance, to counting—and can readily be merged with machine learning techniques to generally enhance their applicability [91–94].

Anandkumar et al. [92] considers parameter estimation for a wide class of latent variable models—including Gaussian mixture models, hidden Markov models, and Latent Dirichlet Allocation—which exploit a certain tensor network structure found in their low-order observable moments.

The aforementioned latent variable models are shallow in the sense that there are not many layers of parameters to be identified—deep learning architectures are the direct opposite of this way of thinking. Mehta and Schwab [93] merges ideas of mapping from the variational renormalization group, first introduced by Kadanoff [95], and deep learning models based on restricted Boltzmann machines so that the results indicate that deep learning algorithms might be viewed as employing a generalized renormalization group-like scheme to learn relevant features from data. Also considering data analysis, [94] relies on the matrix product state (MPS) decomposition as a computational tool to likewise extract features of multidimensional data.

Probabilistic graphical models (see the next section) can also form a hierarchical model akin to a deep learning architecture. [91] compared ideas behind the renormalization group—of which certain tensor network methods represent the modern incarnation—and such models. The multiscale entanglement renormalization ansatz (MERA network) was converted into a learning algorithm based on a hierarchical Bayesian model. Under the assumption that the distribution to be learned is fully characterized by local correlations, this algorithm [91] does not require sampling.

h. **Causality and Bayesian Networks.** Probabilistic graphical models—including Bayesian networks and Markov networks, and their special topological variants such as hidden Markov models, Ising models, and Kalman filters—offer compactness of representation, capturing the sparsity structure of the model and independence conditions among correlations. The graph structure of these encompasses the qualitative properties of the distribution. While causation does not directly appear in the representation, the edge structure of the graph indicates influence, opening the door to applications which rely on determining causation of a given effect.

In the absence of specific graph topologies, two challenges are addressed: (i) structure and weight learning from correlations, and (ii) inference based on the structure and partial observation (a.k.a. clamping) of nodes. The first problem is akin to the training phase of other machine learning approaches. The second phase is about applying the learned model. The computational complexity is typically negligible of this second phase in other forms of learning, but for Bayesian and Markov networks, this probabilistic inference is #P-hard.

The starting point is correlations—this is already problematic if we study quantum correlations which may not have a definite causal order [96], as it has been experimentally probed [97]. Furthermore, intuition of cause and effect fails with quantum correlations, making causal discovery a challenge [98]. In classical Bayesian networks, the d-separation theorem clearly validates whether a given set of correlations can match a given directed graph structure. Progress has been made on the quantum generalization of the d-deparation theorem [56, 99], but the fully general case is still an open issue.

To solve the first phase of learning such a model, an adiabatic quantum optimization method was introduced to learn the structure of a classical Bayesian network [82]. In this case, all correlations are classical.

For probabilistic inference, an early effort used complex-valued probability amplitudes [100]. The problems above were bypassed by requiring that the amplitudes factorize according to classical conditions, restricting the case to pure states. Leifer and Poulin developed an inexact belief propagation protocol for mixed states [101].

If we restrict the topology, learning weights and probabilistic inference can be polynomial or even linear in the classical sense/case. This is the case with hidden Markov models. Monràs and Winter [57] consider the situation where the hidden variables are quantum, and ask whether there is a quantum instrument that could reproduce the observable dynamics. Cholewa et al. [102] assume a series of classical observations with underlying quantum dynamics, and generalize known classical algorithms for training hidden Markov networks.

III. QUANTUM LEARNING EXPERIMENTS

Quantum learning algorithms have been realized in a host of experimental systems and cover a range of applications. Brunner et al. used photonics to demonstrate simultaneous spoken digit and speaker recognition and chaotic time-series prediction at data rates beyond a gigabyte per second [103]. They were able to identify all digits with very low classification errors and perform chaotic time-series prediction with 10% error. Using photonics, a classifier was realized in [104] which worked on up to eight-dimensional vectors. Facing a nonlinear photonic delay system [105], classically employed methods of gra-

dient descent with backpropagation through time were employed on a model of the system to optimize input encoding. Physical experiments obtained show that the input encodings result in significantly better performance than the common reservoir computing approach [105]. Also in nonlinear photonics [106], demonstrated an all-optical linear classifier capable of learning the classification boundary iteratively from training data through a feedback rule. Lau et al. [21] proposed building blocks for learning algorithms in continuous-variable quantum systems with a matching photonic implementation, but experiments to test this are lying still ahead.

Neural networks have been realized using liquid state nuclear magnetic resonance (NMR), by Hopfield neural networks with simulated adiabatic quantum computation to recover basic pattern recognition tasks [107]. Handwriting recognition was also explored by an NMR test bench in [108], enabling the recognition of standard character fonts from a set with two candidates and hence realizing a 'quantum support vector machine'.

Defaulting on a chain of trapped ions, [109] simulated a neural network with induced long range interactions. The storage capacity of such a network was possible to control by changing the external trapping potential.

Quantum annealing for machine learning by superconducting systems is lead by the technology developed by D-Wave Systems. An early demonstration focused on regularized boosting with a nonconvex objective function in a classification task [110]. The optimization problem was discretized and mapped to a quadratic unconstrained binary problem. Subsequent work developed this idea of discretization and mapping [78, 79, 82, 111]. Since the quantum annealer suffers from a number of implementation issues that deviate from the underlying theory, in general it cannot be guaranteed that by the end of the annealing process one will be able to read out the ground state. The distribution of readouts of the final state approximates a Gibbs distribution. This opened the way to training Boltzmann machines [58, 85–87].

IV. FRONTIERS IN QUANTUM MACHINE LEARNING

Quantum computers can outperform classical ones in some machine learning tasks, but the full scope of their power is unknown. Aïmeur et al. [112] asked what we could expect by various combinations of quantum and classical data and objectives, and there are a few results in terms of complexity bounds for quantum machine learning [113–115]. Servedio and Gortler [113] establish a polynomial relationship between the number of quantum versus classical queries required for certain classes of learning problems [113]. Their work implies that, while the sampling complexity of broad classes of quantum and classical machine learning algorithms are polynomially equivalent, their computational complexities need not be. This idea was also confirmed by Arunachalam and de Wolf [115], where it was found that quantum and classical sample complexity are equal up to constant factors in both the probably approximately correct vs. agnostic models. Arunachalam and de Wolfs' results imply that, when restricted to sample complexity, the classical and quantum cases are equal up to constant factors. Despite these differences, the ability of quantum computers to expediently search for the most informative samples can (in some cases) polynomially improve the expected number of samples necessary to learn a concept [64]. Thus characterizing the statistical efficiency of quantum machine learning algorithms remains an open problem.

Another aspect of learning theory, model complexity is central to establishing bounds on generalization performance. Monràs et al. [114] prove that a supervised inductive learning protocol always splits into a training and testing phase, even with quantum resources, and thus establishes the theoretical foundations for defining model complexity. In a similar vein, Wiebe and Granade [116] ask whether a logarithmically small quantum system can learn at all in the sense of Bayesian updates, and their answer is affirmative if the system has access to classical memory but can be negative otherwise. This is because information stored in a quantum posterior distribution cannot be extracted without destroying the information that it spent so long learning. This indicates that properties of quantum mechanics, such as the no-cloning theorem, challenge us to think more broadly what it means for a quantum system to learn about its surroundings.

A clear current use of machine learning in quantum physics builds on the dramatic success these techniques have had in learning to control experimental quantum systems [24, 29, 30, 34–36, 39–41, 117, 118]. On the other hand, a viable contemporary approach to quantum enhanced machine learning is running a quantum annealer, built for example from artificial flux-based qubits [74] as a subroutine to solve optimization problems. This offers an advantage inasmuch as problems can be stated in terms of objective functions which can then be minimized remotely, with the further advantage that mid-size versions of the technology are already available [74]. Although quantum optimizers have seen dramatic increases in scalability (at the time of this writing $\sim$ 2000 manufactured spins are available on a single chip), quantum supremacy—in which a quantum device outperforms any existing classical device for a comparable algorithmic task—has not yet been achieved.

Quantum computers, however, are thought to cross the quantum supremacy threshold and offer dramatic runtime reductions for several classes of problems that can be used as subroutines in machine learning methods [14, 15, 18–20, 48, 49, 54–58]. Yet at the same time, certain powerful and central ideas in the theory of machine learning—such as neural computing [119], quantum generalizations of Bayesian networks [56] and quantum causal models [99]—seem to necessitate an increased effort to understand several fundamental questions attached to their quantum mechanical generaliza-

tions [56, 57, 96, 99].

From the papers and research directions reviewed here, we see that machine learning and quantum computing became intertwined on many different levels: quantum control using classical reinforcement learning, learning unitaries with optimal strategies, and speedup in various learning algorithms are notable examples. This leads to a frontier where both quantum computing and machine intelligence will co-evolve, and they will become enabling technologies for each other.

ACKNOWLEDGMENTS

JDB acknowledges IARPA and the Foundational Questions Institute (FQXi) for financial support. PW acknowledges financial support from the ERC (Consolidator Grant QITBOX), MINECO (Severo Ochoa grant SEV-2015-0522 and FOQUS), Generalitat de Catalunya (SGR 875), and Fundació Privada Cellex. NP acknowledges ExQM (Exploring Quantum Matter) for financial support. SL and PR were supported by ARO and AFOSR.

[1] Masahide Sasaki, Alberto Carlini, and Richard Jozsa. Quantum template matching. *Phys. Rev. A*, 64:022317, July 2001.

[2] Alessandro Bisio, Giulio Chiribella, Giacomo Mauro D'Ariano, Stefano Facchini, and Paolo Perinotti. Optimal quantum learning of a unitary transformation. *Phys. Rev. A*, 81:032324, March 2010.

[3] Alessandro Bisio, Giacomo Mauro D'Ariano, Paolo Perinotti, and Michal Sedlák. Quantum learning algorithms for quantum measurements. *Phys. Lett. A*, 375(39):3425–3434, 2011.

[4] G. Sentís, J. Calsamiglia, R. Muñoz-Tapia, and E. Bagan. Quantum learning without quantum memory. *Sci. Rep.*, 2, October 2012. Article.

[5] Gael Sentís, Mădălin Guţă, and Gerardo Adesso. Quantum learning of coherent states. *EPJ Quantum Technology*, 2(1):17, 2014.

[6] Giuseppe Davide Paparo, Vedran Dunjko, Adi Makmal, Miguel Angel Martin-Delgado, and Hans J. Briegel. Quantum speedup for active learning agents. *Phys. Rev. X*, 4:031002, July 2014.

[7] Vedran Dunjko, Nicolai Friis, and Hans J. Briegel. Quantum-enhanced deliberation of learning agents using trapped ions. *New J. Phys.*, 17(2):023006, 2015.

[8] Vedran Dunjko, Jacob M. Taylor, and Hans J. Briegel. Quantum-enhanced machine learning. *Physical Review Letters*, 117(13), September 2016.

[9] Gael Sentís, Emilio Bagan, John Calsamiglia, Giulio Chiribella, and Ramon Muñoz Tapia. Quantum change point. *Phys. Rev. Lett.*, 117(15), October 2016.

[10] Pierre Chiappetta, Pietro Colangelo, P. De Felice, Giuseppe Nardulli, and Guido Pasquariello. Higgs search by neural networks at LHC. *Phys. Lett. B*, 322(3):219–223, 1994.

[11] Matthias Rupp. Machine learning for quantum mechanics in a nutshell. *Int. J. Quantum Chem.*, 115(16):1058–1073, 2015.

[12] Rahul Biswas, Lindy Blackburn, Junwei Cao, Reed Essick, Kari Alison Hodge, Erotokritos Katsavounidis, Kyungmin Kim, Young-Min Kim, Eric-Olivier Le Bigot, Chang-Hwan Lee, John J. Oh, Sang Hoon Oh, Edwin J. Son, Ye Tao, Ruslan Vaulin, and Xiaoge Wang. Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data. *Phys. Rev. D*, 88:062003, September 2013.

[13] Rami Barends, Julian Kelly, Anthony Megrant, Andrzej Veitia, Daniel Sank, Evan Jeffrey, Ted C. White, Josh Mutus, Austin G. Fowler, Brooks Campbell, Yu Chen, Zijun Chen, Ben Chiaro, Andrew Dunsworth, Charles Neill, Peter O'Malley, Pedram Roushan, Amit Vainsencher, Jim Wenner, Alexander N. Korotkov, Andrew N. Cleland, and John M. Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, 508(7497):500–503, April 2014. Letter.

[14] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103:150502, October 2009.

[15] B. David Clader, Bart C. Jacobs, and Chad R. Sprouse. Preconditioned quantum linear system algorithm. *Phys. Rev. Lett.*, 110:250504, June 2013.

[16] Nathan Wiebe, Daniel Braun, and Seth Lloyd. Quantum algorithm for data fitting. *Phys. Rev. Lett.*, 109:050505, August 2012.

[17] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. *arXiv:1307.0411*, July 2013.

[18] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Phys. Rev. Lett.*, 113:130503, September 2014.

[19] Nathan Wiebe, Ashish Kapoor, and Krysta M. Svore. Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning. *Quantum Info. Comput.*, 15(3-4):316–356, March 2015.

[20] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nat. Phys.*, 10(9):631–633, September 2014. Letter.

[21] Hoi-Kwan Lau, Raphael Pooser, George Siopsis, and Christian Weedbrook. Quantum machine learning over infinite dimensions. *arXiv:1603.06222*, March 2016.

[22] Christopher E Granade, Christopher Ferrie, Nathan Wiebe, and David G. Cory. Robust online Hamiltonian learning. *New J. Phys.*, 14(10):103013, 2012.

[23] Nathan Wiebe, Christopher Granade, Christopher Fer-

rie, and David G. Cory. Hamiltonian learning and certification using quantum resources. *Phys. Rev. Lett.*, 112:190501, May 2014.

[24] Nathan Wiebe, Christopher Granade, Christopher Ferrie, and David Cory. Quantum Hamiltonian learning using imperfect quantum resources. *Phys. Rev. A*, 89:042314, April 2014.

[25] Nathan Wiebe, Christopher Granade, and D G Cory. Quantum bootstrapping via compressed quantum Hamiltonian learning. *New J. Phys.*, 17(2):022005, 2015.

[26] Masahide Sasaki and Alberto Carlini. Quantum learning and universal quantum matching machine. *Phys. Rev. A*, 66:022303, August 2002.

[27] Geoffrey I. Webb, Roy Hyde, Hong Cao, Hai Long Nguyen, and Francois Petitjean. Characterizing concept drift. *arXiv:1511.03816*, December 2015.

[28] Mauro Faccin, Piotr Migdał, Tomi H. Johnson, Ville Bergholm, and Jacob D. Biamonte. Community detection in quantum complex networks. *Phys. Rev. X*, 4:041012, October 2014.

[29] Alexander Hentschel and Barry C. Sanders. Machine learning for precise quantum measurement. *Phys. Rev. Lett.*, 104:063603, February 2010.

[30] Neil B. Lovett, Cécile Crosnier, Martí Perarnau-Llobet, and Barry C. Sanders. Differential evolution for many-particle adaptive quantum metrology. *Phys. Rev. Lett.*, 110:220501, May 2013.

[31] Benjamin P. Lanyon, James D. Whitfield, Geoff G. Gillet, Michael E. Goggin, Marcelo P. Almeida, Ivan Kassal, Jacob D. Biamonte, Masoud Mohseni, Ben J. Powell, Marco Barbieri, Alán Aspuru-Guzik, and Andrew G. White. Towards quantum chemistry on a quantum computer. *Nat. Chem.*, 2(2):106–111, February 2010.

[32] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. Simulating a perceptron on a quantum computer. *Phys. Lett. A*, 379(7):660–663, March 2015.

[33] Pantita Palittapongarnpim, Peter Wittek, and Barry C. Sanders. Controlling adaptive quantum phase estimation with scalable reinforcement learning. In *Proceedings of ESANN-16, 24th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, pages 327–332, April 2016.

[34] Ehsan Zahedinejad, Joydip Ghosh, and Barry C. Sanders. Designing high-fidelity single-shot three-qubit gates: A machine learning approach. *arXiv:1511.08862*, December 2015.

[35] Ehsan Zahedinejad, Joydip Ghosh, and Barry C. Sanders. High-fidelity single-shot Toffoli gate via quantum control. *Phys. Rev. Lett.*, 114:200502, 2015.

[36] Leonardo Banchi, Nicola Pancotti, and Sougato Bose. Quantum gate learning in qubit networks: Toffoli gate without time-dependent control. *npj Quantum Inf.*, 2:16019, July 2016.

[37] Pantita Palittapongarnpim, Peter Wittek, Ehsan Zahedinejad, and Barry C. Sanders. Learning in quantum control: High-dimensional global optimization for noisy quantum dynamics. *arXiv:1607.03428*, July 2016.

[38] Urtzi Las Heras, Unai Alvarez-Rodriguez, Enrique Solano, and Mikel Sanz. Genetic algorithms for digital quantum simulations. *Phys. Rev. Lett.*, 116:230504, June 2016.

[39] Moritz August and Xiaotong Ni. Using recurrent neural networks to optimize dynamical decoupling for quantum memory. *arXiv:1604.00279*, April 2016.

[40] Markus Tiersch, E. J. Ganahl, and Hans J. Briegel. Adaptive quantum computation in changing environments using projective simulation. *Sci. Rep.*, 5:12874, August 2015. Article.

[41] Sandeep Mavadia, Virginia Frey, Jarrah Sastrawan, Stephen Dona, and Michael J. Biercuk. Prediction and real-time compensation of qubit decoherence via machine-learning. *arXiv:1604.03991*, April 2016.

[42] Brendon L. Higgins, Dominic W. Berry, Stephen D. Bartlett, Howard M. Wiseman, and Geoff J. Pryde. Entanglement-free Heisenberg-limited phase estimation. *Nature*, 450(7168):393–396, November 2007.

[43] Dominic W. Berry, Brendon L. Higgins, Stephen D. Bartlett, Morgan W. Mitchell, Geoff J. Pryde, and Howard M. Wiseman. How to perform the most accurate possible phase measurements. *Phys. Rev. A*, 80:052114, November 2009.

[44] Nathan Wiebe and Chris Granade. Efficient Bayesian phase estimation. *Phys. Rev. Lett.*, 117:010503, June 2016.

[45] S. Rosi, A. Bernard, Nicole Fabbri, Leonardo Fallani, Chiara Fort, Massimo Inguscio, Tommasco Calarco, and Simone Montangero. Fast closed-loop optimal control of ultracold atoms in an optical lattice. *Phys. Rev. A*, 88:021601, August 2013.

[46] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. Quantum walks on graphs representing the firing patterns of a quantum neural network. *Phys. Rev. A*, 89:032333, March 2014.

[47] Juan Carrasquilla and Roger G. Melko. Machine learning phases of matter. *arXiv:1605.01735*, May 2016.

[48] Giacomo Torlai and Roger G. Melko. Learning thermodynamics with Boltzmann machines. *arXiv:1606.02718*, June 2016.

[49] Giuseppe Carleo and Matthias Troyer. Solving the quantum many-body problem with artificial neural networks. *arXiv:1606.02318*, June 2016.

[50] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zhang. TensorFlow: A system for large-scale machine learning. *arXiv:1605.08695*, May 2016.

[51] Elizabeth C. Behrman, John Niemel, James E. Steck, and Steve R. Skinner. A quantum dot neural network. In *Proceedings of PhysComp-96, 4th Workshop on Physics of Computation*, pages 22–28, 1996.

[52] Mikhail V. Altaisky, Nadezhda N. Zolnikova, Natalia E. Kaputkina, Victor A. Krylov, Yurii E. Lozovik, and Nikesh S. Dattani. Towards a feasible implementation of quantum neural networks using quantum dots. *Appl. Phys. Lett.*, 108(10), 2016.

[53] Dong-Ling Deng, Xiaopeng Li, and S. Das Sarma. Exact machine learning topological states. *arXiv:1609.09060*, September 2016.

[54] Esma Aïmeur, Gilles Brassard, and Sébastien Gambs. Quantum speed-up for unsupervised learning. *Machine Learning*, 90(2):261–287, 2013.

[55] Seth Lloyd, Silvano Garnerone, and Paolo Zanardi.

Quantum algorithms for topological and geometric analysis of data. *Nat. Commun.*, 7, January 2016. Article.

[56] Joe Henson, Raymond Lal, and Matthew F Pusey. Theory-independent limits on correlations from generalized Bayesian networks. *New J. Phys.*, 16(11):113043, 2014.

[57] Alex Monràs and Andreas Winter. Quantum learning of classical stochastic processes: The completely positive realization problem. *J. Math. Phys.*, 57(1):015219, 2016.

[58] Mohammad H. Amin, Evgeny Andriyash, Jason Rolfe, Bohdan Kulchytskyy, and Roger Melko. Quantum Boltzmann machine. *arXiv:1601.02036*, January 2016.

[59] Keisuke Fujii and Kohei Nakajima. Harnessing disordered quantum dynamics for machine learning. *arXiv:1602.08159*, February 2016.

[60] Scott Aaronson. Bqp and the polynomial hierarchy. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 141–150. ACM, 2010.

[61] Dan Ventura and Tony Martinez. Quantum associative memory. *Inform. Sciences*, 124(1):273–296, 2000.

[62] Davide Anguita, Sandro Ridella, Fabio Rivieccio, and Rodolfo Zunino. Quantum optimization for training support vector machines. *Neural Netw.*, 16(5–6):763–770, 2003. Advances in Neural Networks Research: {IJCNN} '03.

[63] Esma Aïmeur, Gilles Brassard, and Sébastien Gambs. Quantum clustering algorithms. In *Proceedings of ICML-07, 24th International Conference on Machine Learning*, pages 1–8, Corvallis, OR, USA, June 2007.

[64] Nathan Wiebe, Ashish Kapoor, and Krysta M Svore. Quantum perceptron models. *arXiv:1602.04799*, February 2016.

[65] Guoming Wang. Quantum algorithms for curve fitting. *arXiv:1402.0660*, 2014.

[66] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. Prediction by linear regression on a quantum computer. *Phys. Rev. A*, 94(2), August 2016.

[67] Patrick Rebentrost, Adrian Steffens, and Seth Lloyd. Quantum singular value decomposition of non-sparse low-rank matrices. *arXiv:1607.05404*, July 2016.

[68] Iris Cong and Luming Duan. Quantum discriminant analysis for dimensionality reduction and classification. *New J. Phys.*, 18(7):073011, July 2016.

[69] Raouf Dridi and Hedayat Alghassi. Homology computation of large point clouds using quantum annealing. *arXiv:1512.09328*, December 2015.

[70] Scott Aaronson. Read the fine print. *Nat. Phys.*, 11(4):291–293, April 2015. Commentary.

[71] Francesco De Martini, Vittorio Giovannetti, Seth Lloyd, Lorenzo Maccone, Eleonora Nagali, Linda Sansoni, and Fabio Sciarrino. Experimental quantum private queries with linear optics. *Phys. Rev. A*, 80:010302, July 2009.

[72] Srinivasan Arunachalam, Vlad Gheorghiu, Tomas Jochym-O'Connor, Michele Mosca, and Priyaa Varshinee Srinivasan. On the robustness of bucket brigade quantum RAM. *New J. Phys.*, 17(12):123010, 2015.

[73] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda. A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science*, 292(5516):472–475, 2001.

[74] Mark W. Johnson, Mohammad H. S. Amin, Suzanne Gildert, Trevor Lanting, Firas Hamze, Neil Dickson, R. Harris, Andrew J. Berkley, Jan Johansson, Paul Bunyk, Erin M. Chapple, C. Enderud, Jeremy P. Hilton, Kamran Karimi, Eric Ladizinsky, N. Ladizinsky, T. Oh, Ilya Perminov, C. Rich, M. C. Thom, Elena Tolkacheva, Colin J. S. Truncik, Sergey Uchaikin, J. Wang, B. Wilson, and Geordie Rose. Quantum annealing with manufactured spins. *Nature*, 473(7346):194–198, May 2011.

[75] John J Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.*, 79(8):2554–2558, 1982.

[76] Donny Cheung, Peter Høyer, and Nathan Wiebe. Improved error bounds for the adiabatic approximation. *J. Phys. A: Math. Theor.*, 44(41):415302, 2011.

[77] Jacob D. Biamonte and Peter J. Love. Realizable Hamiltonians for universal adiabatic quantum computers. *Phys. Rev. A*, 78:012352, July 2008.

[78] Vasil S. Denchev, Nan Ding, S.V.N. Vishwanathan, and Hartmut Neven. Robust classification with adiabatic quantum optimization. In *Proceedings of ICML-2012, 29th International Conference on Machine Learning*, June 2012.

[79] Vasil S. Denchev, Nan Ding, Shin Matsushima, S. V. N. Vishwanathan, and Hartmut Neven. Totally corrective boosting with cardinality penalization. *arXiv:1504.01446*, April 2015.

[80] Joseph Dulny III and Michael Kim. Developing quantum annealer driven data discovery. *arXiv:1603.07980*, March 2016.

[81] Kristen L. Pudenz and Daniel A. Lidar. Quantum adiabatic machine learning. *Quantum Inf. Process.*, 12:2027–2070, May 2013.

[82] Bryan A. O'Gorman, Alejandro Perdomo-Ortiz, Ryan Babbush, Alán Aspuru-Guzik, and Vadim Smelyanskiy. Bayesian network structure learning using quantum annealing. *The European Physical Journal Special Topics*, 224(1):163–188, 2015.

[83] Arman Zaribafiyan, Dominic J. J. Marchand, and Seyed Saeed Changiz Rezaei. Systematic and deterministic graph-minor embedding for cartesian products of graphs. *arXiv:1602.04274*, February 2016.

[84] Marcello Benedetti, John Realpe-Gómez, Rupak Biswas, and Alejandro Perdomo-Ortiz. Quantum-assisted learning of graphical models with arbitrary pairwise connectivity. *arXiv:1609.02542*, September 2016.

[85] Steven H. Adachi and Maxwell P. Henderson. Application of quantum annealing to training of deep neural networks. *arXiv:1510.06356*, November 2015.

[86] Marcello Benedetti, John Realpe-Gómez, Rupak Biswas, and Alejandro Perdomo-Ortiz. Estimation of effective temperatures in quantum annealers for sampling applications: A case study with possible applications in deep learning. *Phys. Rev. A*, 94:022308, August 2016.

[87] Alejandro Perdomo-Ortiz, Bryan O'Gorman, Joseph Fluegemann, Rupak Biswas, and Vadim N. Smelyanskiy. Determination and correction of persistent biases in quantum annealers. *Sci. Rep.*, 6:18628, January 2016. Article.

[88] Nathan Wiebe, Ashish Kapoor, and Krysta M. Svore. Quantum deep learning. *arXiv:1412.3489*, 2014.

[89] Anirban Narayan Chowdhury and Rolando D. Somma. Quantum algorithms for Gibbs sampling and hitting-time estimation. *arXiv:1603.02940*, March 2016.

[90] Peter Wittek and Christian Gogolin. Quantum enhanced inference in markov logic networks.

arXiv:1611.08104, November 2016.

[91] Cédric Bény. Deep learning and the renormalization group. *arXiv:1301.3124*, January 2013.

[92] Anima Anandkumar, Rong Ge, Daniel Hsu, Sham M. Kakade, and Matus Telgarsky. *Tensor Decompositions for Learning Latent Variable Models*, pages 19–38. Springer International Publishing, Cham, October 2015.

[93] Pankaj Mehta and David J. Schwab. An exact mapping between the variational renormalization group and deep learning. *arXiv:1410.3831*, November 2014.

[94] Johann A. Bengua, Ho N. Phien, Hoang D. Tuan, and Minh N. Do. Matrix product state for feature extraction of higher-order tensors. *arXiv:1503.00516*, March 2015.

[95] Leo P. Kadanoff, Anthony Houghton, and Mehmet C. Yalabik. Variational approximations for renormalization group transformations. *J. Stat. Phys.*, 14(2):171–203, 1976.

[96] Ognyan Oreshkov, Fabio Costa, and Caslav Brukner. Quantum correlations with no causal order. *Nat. Commun.*, 3:1092, October 2012.

[97] Giulia Rubino, Lee A. Rozema, Adrien Feix, Mateus Araújo, Jonas M. Zeuner, Lorenzo M. Procopio, Časlav Brukner, and Philip Walther. Experimental verification of an indefinite causal order. *arXiv:1608.01683*, August 2016.

[98] Christopher J. Wood and Robert S. Spekkens. The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning. *New J. Phys.*, page 033002, 2015.

[99] Jacques Pienaar and Časlav Brukner. A graph-separation theorem for quantum causal models. *New J. Phys.*, 17(7):073020, 2015.

[100] Robert R. Tucci. Quantum Bayesian Nets. *International Journal of Modern Physics B*, 09(03):295–337, 1995.

[101] Matt. S Leifer and David Poulin. Quantum graphical models and belief propagation. *Ann. Phys.*, 323(8):1899–1946, 2008.

[102] MichałCholewa, Piotr Gawron, Przemysław Głomb, and Dariusz Kurzyk. Quantum hidden Markov models based on transition operation matrices. *arXiv:1503.08760*, March 2015.

[103] Daniel Brunner, Miguel C. Soriano, Claudio R. Mirasso, and Ingo Fischer. Parallel photonic information processing at gigabyte per second data rates using transient states. *Nat. Commun.*, 4:1364, January 2013.

[104] Xin-Dong Cai, Dian Wu, Zu-En Su, Ming-Cheng Chen, Xi-Ling Wang, Li Li, Nai-Le Liu, Chao-Yang Lu, and Jian-Wei Pan. Entanglement-based machine learning on a quantum computer. *Phys. Rev. Lett.*, 114:110504, March 2015.

[105] Michiel Hermans, Miguel C. Soriano, Joni Dambre, Peter Bienstman, and Ingo Fischer. Photonic delay systems as machine learning implementations. *J. Mach. Learn. Res.*, 16(1):2081–2097, January 2015.

[106] Nikolas Tezak and Hideo Mabuchi. A coherent perceptron for all-optical learning. *EPJ Quantum Technol.*, 2(1), April 2015.

[107] Rodion Neigovzen, Jorge L. Neves, Rudolf Sollacher, and Steffen J. Glaser. Quantum pattern recognition with liquid-state nuclear magnetic resonance. *Phys. Rev. A*, 79:042321, April 2009.

[108] Zhaokai Li, Xiaomei Liu, Nanyang Xu, and Jiangfeng Du. Experimental realization of a quantum support vector machine. *Phys. Rev. Lett.*, 114:140504, April 2015.

[109] Marisa Pons, Veronica Ahufinger, Christof Wunderlich, Anna Sanpera, Sibylle Braungardt, Aditi Sen(De), Ujjwal Sen, and Maciej Lewenstein. Trapped ion chain as a neural network: Error resistant quantum computation. *Phys. Rev. Lett.*, 98:023003, January 2007.

[110] Hartmut Neven, Vasil S Denchev, Marshall Drew-Brook, Jiayong Zhang, William G Macready, and Geordie Rose. Binary classification using hardware implementation of quantum annealing. In *Demonstrations at NIPS-09, 24th Annual Conference on Neural Information Processing Systems*, pages 1–17, December 2009.

[111] Kamran Karimi, Neil G Dickson, Firas Hamze, M HS Amin, Marshall Drew-Brook, Fabian A Chudak, Paul I Bunyk, William G Macready, and Geordie Rose. Investigating the performance of an adiabatic quantum optimization processor. *Quantum Inf. Process.*, 11(1):77–88, 2012.

[112] Esma Aïmeur, Gilles Brassard, and Sébastien Gambs. *Machine Learning in a Quantum World*, pages 431–442. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[113] Rocco A. Servedio and Steven J. Gortler. Quantum versus classical learnability. In *Proceedings of CCC-01, 16th Annual IEEE Conference on Computational Complexity*, pages 138–148, June 2001.

[114] Alex Monràs, Gael Sentís, and Peter Wittek. Inductive quantum learning: Why you are doing it almost right. *arXiv:1605.07541*, May 2016.

[115] Srinivasan Arunachalam and Ronald de Wolf. Optimal quantum sample complexity of learning algorithms. *arXiv:1607.00932*, July 2016.

[116] Nathan Wiebe and Christopher Granade. Can small quantum systems learn? *arXiv:1512.03145*, December 2015.

[117] Ehsan Zahedinejad, Sophie Schirmer, and Barry C. Sanders. Evolutionary algorithms for hard quantum control. *Phys. Rev. A*, 90:032310, September 2014.

[118] Paul B. Wigley, Patrick J. Everitt, Anton van den Hengel, J. W. Bastian, Mahasen A. Sooriyabandara, Gordon D. McDonald, Kyle S. Hardman, C. D. Quinlivan, P. Manju, Carlos C. N. Kuhn, Ian R. Petersen, Andre N. Luiten, Joseph J. Hope, Nicholas P. Robins, and Michael R. Hush. Fast machine-learning online optimization of ultra-cold-atom experiments. *Sci. Rep.*, 6:25890, May 2016. Article.

[119] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. The quest for a quantum neural network. *Quantum Inf. Process.*, 13(11):2567–2586, November 2014.