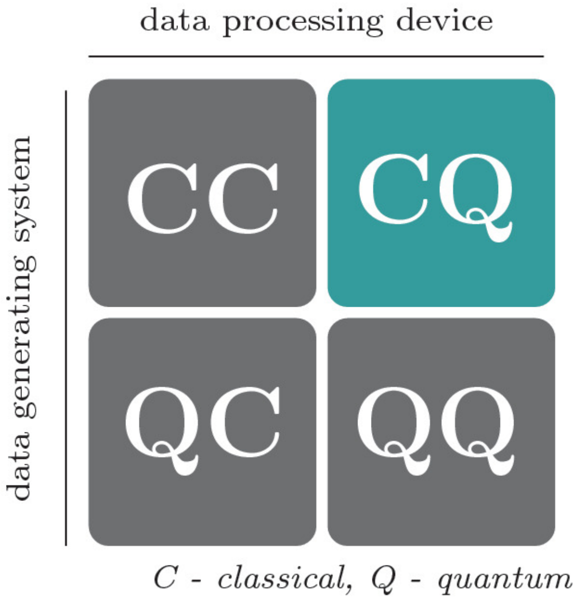
Quantum Machine Learning: A Computer Science and Engineering Perspective

Elías F. Combarro
combarro@gmail.com
University of Oviedo (Spain)

Workshop on Tensor Networks and (Quantum) Machine Learning for High-Energy Physics
CERN, Geneva - 4th November 2024

Universidad de Oviedo

What we talk about when we talk about QML



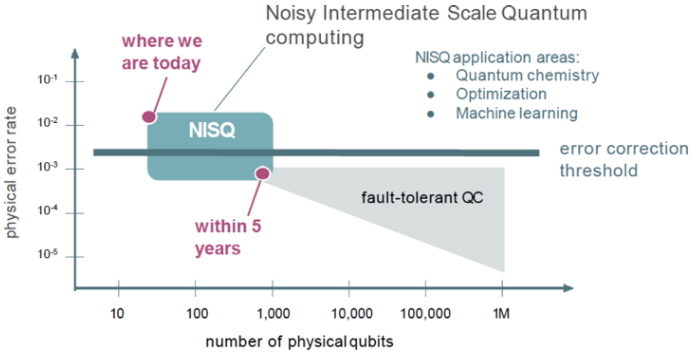
Why QML?

Method	Speedup	Amplitude amplification	HHL	Adiabatic	qRAM
Bayesian inference[106,107]	$O(\sqrt{N})$	Yes	Yes	No	No
Online perceptron[108]	$O(\sqrt{N})$	Yes	No	No	Optional
Least-squares fitting[9]	$O(\log N)$*	Yes	Yes	No	Yes
Classical Boltzmann machine[20]	$O(\sqrt{N})$	Yes/No	Optional/No	No/Yes	Optional
Quantum Boltzmann machine[22,61]	$O(\log N)$*	Optional/No	No	No/Yes	No
Quantum PCA[11]	$O(\log N)$*	No	Yes	No	Optional
Quantum support vector machine[13]	$O(\log N)$*	No	Yes	No	Yes
Quantum reinforcement learning[30]	$O(\sqrt{N})$	Yes	No	No	No

*There exist important caveats that can limit the applicability of the method[51].

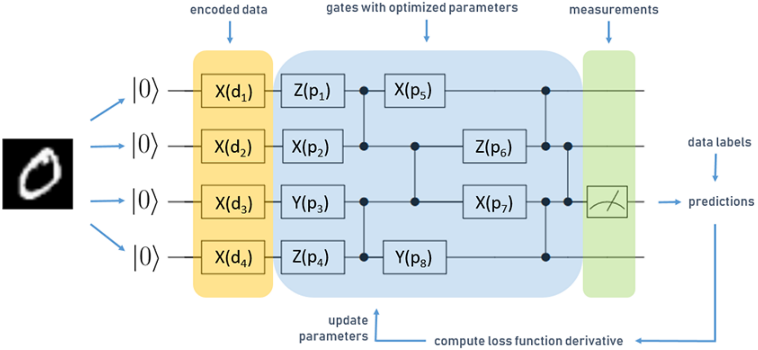
Biamonte, J., Wittek, P., Pancotti, N. *et al.* Quantum machine learning. *Nature* **549**, 195–202 (2017). https://doi.org/10.1038/nature23474

The NISQ era

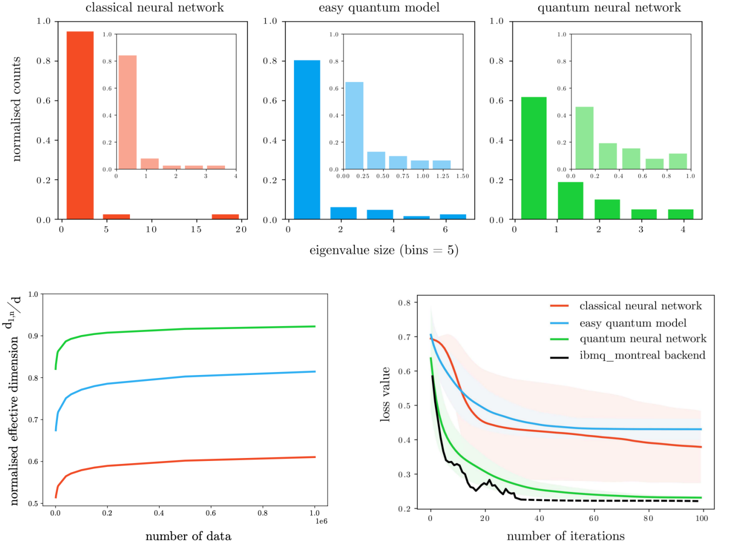


"Quantum computing in the NISQ era and beyond" Preskill, 2018 https://arxiv.org/abs/1801.00862

Variational QML



Why QML (again)?



Abbas, A., Sutter, D., Zoufal, C. *et al.* The power of quantum neural networks. *Nat Comput Sci* **1**, 403–409 (2021). https://doi.org/10.1038/s43588-021-00084-1

Better than classical? The subtle art of benchmarking quantum machine learning models

Joseph Bowles,[1, *] Shahnawaz Ahmed,[1, 2, †] and Maria Schuld[1, ‡]
[1] *Xanadu, Toronto, ON, M5G 2C8, Canada*
[2] *Chalmers University of Technology*

(Dated: March 15, 2024)

Benchmarking models via classical simulations is one of the main ways to judge ideas in quantum machine learning before noise-free hardware is available. However, the huge impact of the experimental design on the results, the small scales within reach today, as well as narratives influenced by the commercialisation of quantum technologies make it difficult to gain robust insights. To facilitate better decision-making we develop an open-source package based on the PennyLane software framework and use it to conduct a large-scale study that systematically tests 12 popular quantum machine learning models on 6 binary classification tasks used to create 160 individual datasets. We find that overall, out-of-the-box classical machine learning models outperform the quantum classifiers. Moreover, removing entanglement from a quantum model often results in as good or better performance, suggesting that "quantumness" may not be the crucial ingredient for the small learning tasks considered here. Our benchmarks also unlock investigations beyond simplistic leaderboard comparisons, and we identify five important questions for quantum model design that follow from our results.

Much has been written about the "potential" of quantum machine learning, a discipline that asks how quantum computers fundamentally change what we can learn from data [1, 2]. While we have no means of running quantum algorithms on noise-free hardware yet, there are only a limited number of tools available to assess this potential. Besides proving advantages for artificial problem settings on paper, certain ideas – most prominently, variational models designed for near-term quantum technologies – can be tested in classical simulations using small datasets. Such benchmarks have in fact become a standard practice in the quantum machine learning literature and are found in almost every paper.

A taste for the results derived from small-scale benchmarks can be obtained through a simple literature review exercise. Out of 55 relevant papers published on the preprint server *arXiv*[1] until December 2023 that con-

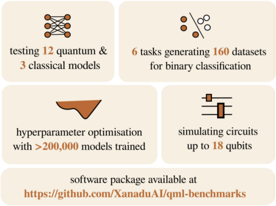


FIG. 1. The scope of the benchmark study at a glance.

Backpropagation

On quantum backpropagation, information reuse, and cheating measurement collapse

Amira Abbas[1,2,3,4], Robbie King[5], Hsin-Yuan Huang[5,6], William J. Huggins[1], Ramis Movassagh[1], Dar Gilboa[1], and Jarrod R. McClean[1*]

[1] *Google Quantum AI, Venice, California 90291, USA*
[2] *University of KwaZulu-Natal, Durban, South Africa*
[3] *Institute of Physics, University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands*
[4] *QuSoft, CWI, Science Park 123, 1098 XG Amsterdam, The Netherlands*
[5] *Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA 91125, USA*
[6] *Institute for Quantum Information and Matter, Caltech, Pasadena, CA 91125, USA*

Abstract

The success of modern deep learning hinges on the ability to train neural networks at scale. Through clever reuse of intermediate information, backpropagation facilitates training through gradient computation at a total cost roughly proportional to running the function, rather than incurring an additional factor proportional to the number of parameters – which can now be in the trillions. Naively, one expects that quantum measurement collapse entirely rules out the reuse of quantum information as in backpropagation. But recent developments in shadow tomography, which assumes access to multiple copies of a quantum state, have challenged that notion. Here, we investigate whether parameterized quantum models can train as efficiently as classical neural networks. We show that achieving backpropagation scaling is impossible without access to multiple copies of a state. With this added ability, we introduce an algorithm with foundations in shadow tomography that matches backpropagation scaling in quantum resources while reducing classical auxiliary computational costs to open problems in shadow tomography. These results highlight the nuance of reusing quantum information for practical purposes and clarify the unique difficulties in training large quantum models, which could alter the course of quantum machine learning.

Barren plateaus

ARTICLE

DOI: 10.1038/s41467-018-07090-4 OPEN

Barren plateaus in quantum neural network training landscapes

Jarrod R. McClean[1], Sergio Boixo[1], Vadim N. Smelyanskiy[1], Ryan Babbush[1] & Hartmut Neven[1]

Many experimental proposals for noisy intermediate scale quantum devices involve training a parameterized quantum circuit with a classical optimization loop. Such hybrid quantum-classical algorithms are popular for applications in quantum simulation, optimization, and machine learning. Due to its simplicity and hardware efficiency, random circuits are often proposed as initial guesses for exploring the space of quantum states. We show that the exponential dimension of Hilbert space and the gradient estimation complexity make this choice unsuitable for hybrid quantum-classical algorithms run on more than a few qubits. Specifically, we show that for a wide class of reasonable parameterized quantum circuits, the probability that the gradient along any reasonable direction is non-zero to some fixed precision is exponentially small as a function of the number of qubits. We argue that this is related to the 2-design characteristic of random circuits, and that solutions to this problem must be studied.

ARTICLE

Check for updates

https://doi.org/10.1038/s41467-021-21728-w OPEN

Cost function dependent barren plateaus in shallow parametrized quantum circuits

M. Cerezo[1,2]✉, Akira Sone[1,2], Tyler Volkoff[1], Lukasz Cincio[1] & Patrick J. Coles[1]✉

Variational quantum algorithms (VQAs) optimize the parameters $\boldsymbol{\theta}$ of a parametrized quantum circuit $V(\boldsymbol{\theta})$ to minimize a cost function C. While VQAs may enable practical applications of noisy quantum computers, they are nevertheless heuristic methods with unproven scaling. Here, we rigorously prove two results, assuming $V(\boldsymbol{\theta})$ is an alternating layered ansatz composed of blocks forming local 2-designs. Our first result states that defining C in terms of global observables leads to exponentially vanishing gradients (i.e., barren plateaus) even when $V(\boldsymbol{\theta})$ is shallow. Hence, several VQAs in the literature must revise their proposed costs. On the other hand, our second result states that defining C with local observables leads to at worst a polynomially vanishing gradient, so long as the depth of $V(\boldsymbol{\theta})$ is $\mathcal{O}(\log n)$. Our results establish a connection between locality and trainability. We illustrate these ideas with large-scale simulations, up to 100 qubits, of a quantum autoencoder implementation.

Defining new architectures

PHYSICAL REVIEW X **11,** 041011 (2021)

Absence of Barren Plateaus in Quantum Convolutional Neural Networks

Arthur Pesah,[1,2] M. Cerezo,[1,3] Samson Wang,[1,4] Tyler Volkoff,[1] Andrew T. Sornborger,[5] and Patrick J. Coles[1]

[1] *Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[2] *Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom*
[3] *Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87544*
[4] *Imperial College London, London, United Kingdom*
[5] *Information Sciences, Los Alamos National Laboratory, Los Alamos, New Mexico 87544, USA*

(Received 12 March 2021; revised 13 July 2021; accepted 2 August 2021; published 15 October 2021)

Quantum neural networks (QNNs) have generated excitement around the possibility of efficiently analyzing quantum data. But this excitement has been tempered by the existence of exponentially vanishing gradients, known as barren plateau landscapes, for many QNN architectures. Recently, quantum convolutional neural networks (QCNNs) have been proposed, involving a sequence of convolutional and pooling layers that reduce the number of qubits while preserving information about relevant data features. In this work, we rigorously analyze the gradient scaling for the parameters in the QCNN architecture. We find that the variance of the gradient vanishes no faster than polynomially, implying that QCNNs do not exhibit barren plateaus. This result provides an analytical guarantee for the trainability of randomly initialized QCNNs, which highlights QCNNs as being trainable under random initialization unlike many other QNN architectures. To derive our results, we introduce a novel graph-based method to analyze expectation values over Haar-distributed unitaries, which will likely be useful in other contexts. Finally, we perform numerical simulations to verify our analytical results.

DOI: 10.1103/PhysRevX.11.041011 Subject Areas: Quantum Information

Simulability

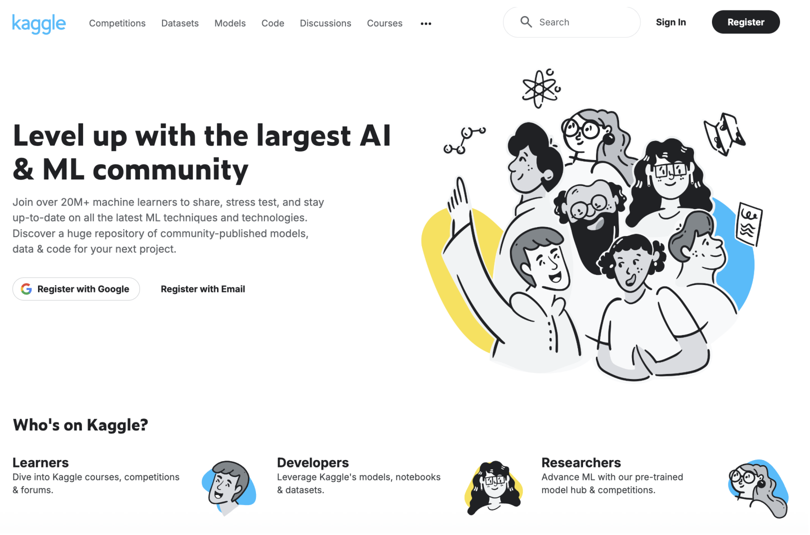
Does provable absence of barren plateaus imply classical simulability? Or, why we need to rethink variational quantum computing

M. Cerezo,[1,2,*] Martin Larocca,[3,4] Diego García-Martín,[1] N. L. Diaz,[1,5] Paolo Braccia,[3] Enrico Fontana,[6] Manuel S. Rudolph,[7] Pablo Bermejo,[8,1] Aroosa Ijaz,[3,9,10] Supanut Thanasilp,[7,11] Eric R. Anschuetz,[12,13] and Zoë Holmes[7]

[1]*Information Sciences, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2]*Quantum Science Center, Oak Ridge, TN 37931, USA*
[3]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[4]*Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA*
[5]*Departamento de Física-IFLP/CONICET, Universidad Nacional de La Plata, C.C. 67, La Plata 1900, Argentina*
[6]*University of Strathclyde, Glasgow G1 1XQ, UK*
[7]*Institute of Physics, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne CH-1015, Switzerland*
[8]*Donostia International Physics Center, Paseo Manuel de Lardizabal 4, San Sebastián E-20018, Spain*
[9]*Department of Physics and Astronomy, University of Waterloo, Ontario, N2L 3G1, Canada*
[10]*Vector Institute, MaRS Centre, Toronto, Ontario, M5G 1M1, Canada*
[11]*Chula Intelligent and Complex Systems, Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, 10330*
[12]*Institute for Quantum Information and Matter, Caltech, Pasadena 91125, USA*
[13]*Walter Burke Institute for Theoretical Physics, Caltech, Pasadena 91125, USA*

A large amount of effort has recently been put into understanding the barren plateau phenomenon. In this perspective article, we face the increasingly loud elephant in the room and ask a question that has been hinted at by many but not explicitly addressed: *Can the structure that allows one to avoid barren plateaus also be leveraged to efficiently simulate the loss classically?* We present strong evidence that commonly used models with provable absence of barren plateaus are also classically simulable, provided that one can collect some classical data from quantum devices during an initial data acquisition phase. This follows from the observation that barren plateaus result from a curse of dimensionality, and that current approaches for solving them end up encoding the problem into some small, classically simulable, subspaces. Thus, while stressing quantum computers can be essential for collecting data, our analysis sheds serious doubt on the non-classicality of the information processing capabilities of parametrized quantum circuits for barren plateau-free landscapes. We end by discussing caveats in our arguments, the role of smart initializations and the possibility of provably superpolynomial, or simply practical, advantages from running parametrized quantum circuits.

Datasets


kaggle
Competitions
Datasets
Models
Code
Discussions
Courses
Search
Sign In
Register
Level up with the largest AI & ML community
Join over 20M+ machine learners to share, stress test, and stay up-to-date on all the latest ML techniques and technologies. Discover a huge repository of community-published models, data & code for your next project.
Register with Google
Register with Email
Who's on Kaggle?
Learners
Dive into Kaggle courses, competitions & forums.
Developers
Leverage Kaggle's models, notebooks & datasets.
Researchers
Advance ML with our pre-trained model hub & competitions.

The Quantum Fourier Transform (again)

Inference, interference and invariance: How the Quantum Fourier Transform can help to learn from data

David Wakeham and Maria Schuld
Xanadu, Toronto, ON, M5G 2C8, Canada

How can we take inspiration from a typical quantum algorithm to design heuristics for machine learning? A common blueprint, used from Deutsch-Josza to Shor's algorithm, is to place labeled information in superposition via an oracle, interfere in Fourier space, and measure. In this paper, we want to understand how this interference strategy can be used for *inference*, i.e. to generalize from finite data samples to a ground truth. Our investigative framework is built around the Hidden Subgroup Problem (HSP), which we transform into a learning task by replacing the oracle with classical training data. The standard quantum algorithm for solving the HSP uses the Quantum Fourier Transform to expose an invariant subspace, i.e., a subset of Hilbert space in which the hidden symmetry is manifest. Based on this insight, we propose an inference principle that "compares" the data to this invariant subspace, and suggest a concrete implementation via overlaps of quantum states. We hope that this leads to well-motivated quantum heuristics that can leverage symmetries for machine learning applications.

I. INTRODUCTION

The Hidden Subgroup Problem (HSP) [1, 2] is the task of discovering a subgroup from information about the way it partitions the parent group. While abstract, it neatly generalizes many problems solved by quantum algorithms, from Deutsch-Jozsa [3] to Simon's problem [4] to Shor's algorithms for period-finding and discrete logarithms [5]. The standard quantum routine for the HSP [6] has a common and embarrassingly simple blueprint: label all inputs, uniformly superpose, apply the Fourier transform, and measure. Formally, this samples from a

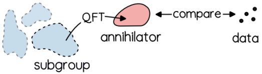


FIG. 1: Which hidden subgroup gave rise to the data? We propose to compare the data with the annihilator of a given subgroup. The annihilator is computed via a group Quantum Fourier Transform executed by a quantum computer.

Libraries

Qiskit
Elements for building a quantum future

Cirq

rigetti

pyQuil Referen

PENNYLANE

Other paradigms

Large-scale quantum reservoir learning with an analog quantum computer

Milan Kornjača[†,1,*] Hong-Ye Hu,[1,2,*] Chen Zhao,[1] Jonathan Wurtz,[1] Phillip Weinberg,[1] Majd Hamdan,[1] Andrii Zhdanov,[1] Sergio H. Cantu,[1] Hengyun Zhou,[1] Rodrigo Araiza Bravo,[2] Kevin Bagnall,[1] James I. Basham,[1] Joseph Campo,[1] Adam Choukri,[1] Robert DeAngelo,[1] Paige Frederick,[1] David Haines,[1] Julian Hammett,[1] Ning Hsu,[1] Ming-Guang Hu,[1] Florian Huber,[1] Paul Niklas Jepsen,[1] Ningyuan Jia,[1] Thomas Karolyshyn,[1] Minho Kwon,[1] John Long,[1] Jonathan Lopatin,[1] Alexander Lukin,[1] Tommaso Macrì,[1] Ognjen Marković,[1] Luis A. Martínez-Martínez,[1] Xianmei Meng,[1] Evgeny Ostroumov,[1] David Paquette,[1] John Robinson,[1] Pedro Sales Rodriguez,[1] Anshuman Singh,[1] Nandan Sinha,[1] Henry Thoreen,[1] Noel Wan,[1] Daniel Waxman-Lenz,[1] Tak Wong,[1] Kai-Hsin Wu,[1] Pedro L. S. Lopes,[1] Yuval Boger,[1] Nathan Gemelke,[1] Takuya Kitagawa,[1] Alexander Keesling,[1] Xun Gao,[3] Alexei Bylinskii,[1] Susanne F. Yelin,[2] Fangli Liu[†,1] and Sheng-Tao Wang[†1]

[1]*QuEra Computing Inc., 1284 Soldiers Field Road, Boston, MA, 02135, USA*

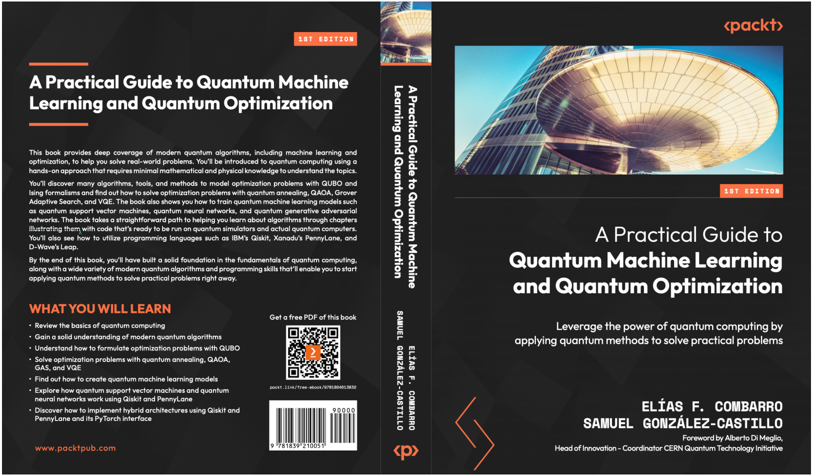
[2]*Department of Physics, Harvard University, Cambridge, MA 02138, USA*

[3]*JILA and Department of Physics, University of Colorado, Boulder, Colorado 80309, USA*

(Dated: July 4, 2024)

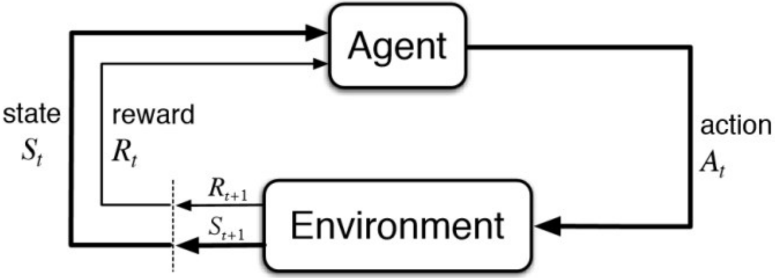
Quantum machine learning has gained considerable attention as quantum technology advances, presenting a promising approach for efficiently learning complex data patterns. Despite this promise, most contemporary quantum methods require significant resources for variational parameter optimization and face issues with vanishing gradients, leading to experiments that are either limited in scale or lack potential for quantum advantage. To address this, we develop a general-purpose, gradient-free, and scalable quantum reservoir learning algorithm that harnesses the quantum dynamics of neutral-atom analog quantum computers to process data. We experimentally implement the algorithm, achieving competitive performance across various categories of machine learning tasks, including binary and multi-class classification, as well as timeseries prediction. Effective and improving learning is observed with increasing system sizes of up to 108 qubits, demonstrating the largest quantum machine learning experiment to date. We further observe comparative quantum kernel advantage in learning tasks by constructing synthetic datasets based on the geometric differences between generated quantum and classical data kernels. Our findings demonstrate the potential of utilizing classically intractable quantum correlations for effective machine learning. We expect these results to stimulate further extensions to different quantum hardware and machine learning paradigms, including early fault-tolerant hardware and generative machine learning tasks.

Quantum Computing Education

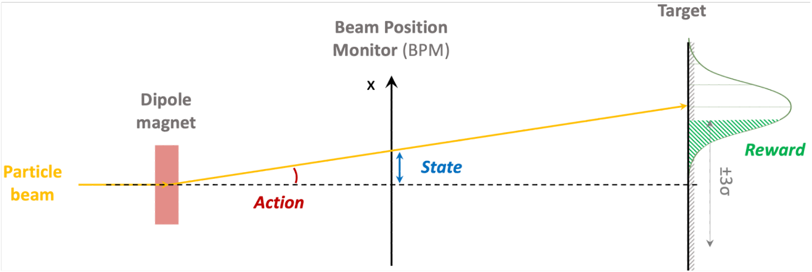


Quantum Computing Education

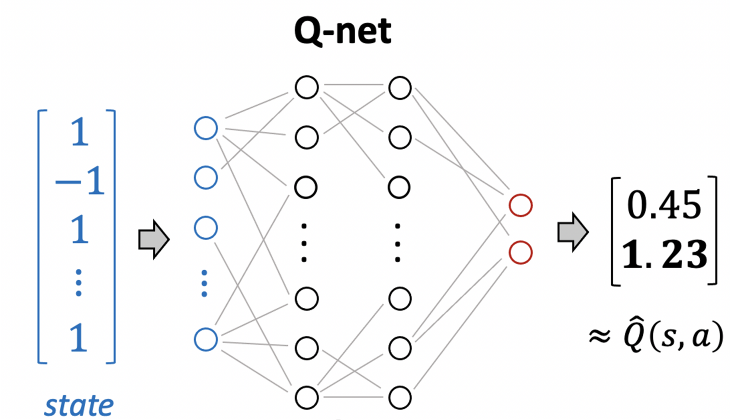
Reinforcement Learning



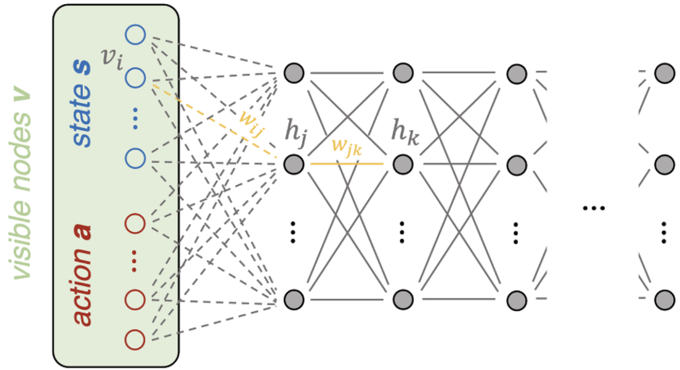
A (toy) problem in particle accelerator control



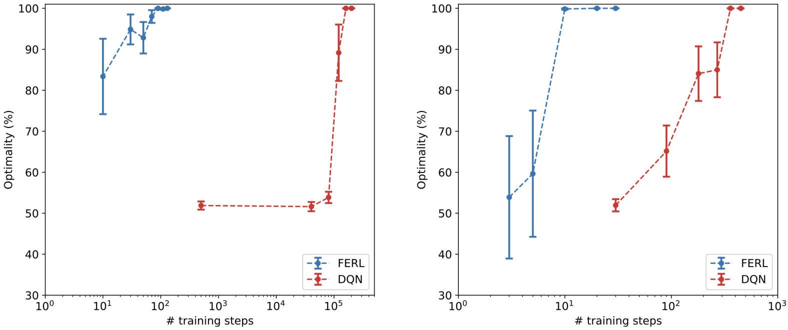
Approximation with a neural network



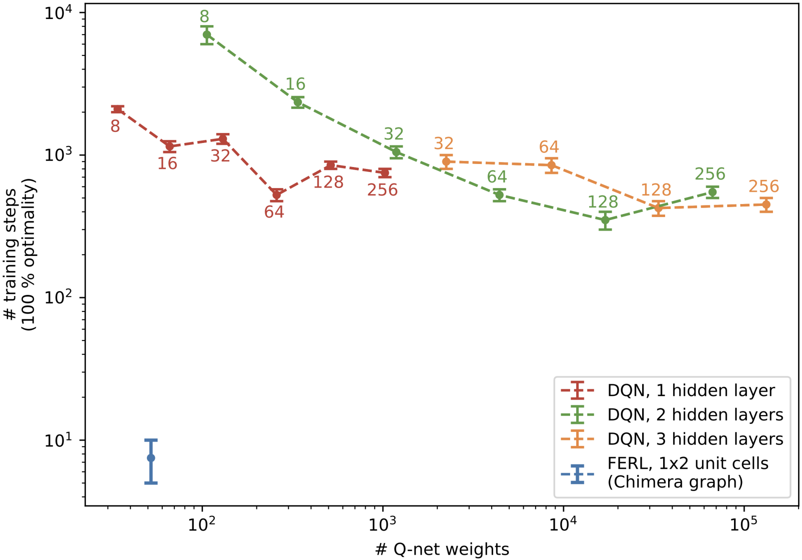
Approximation with a Quantum Boltzmann Machine



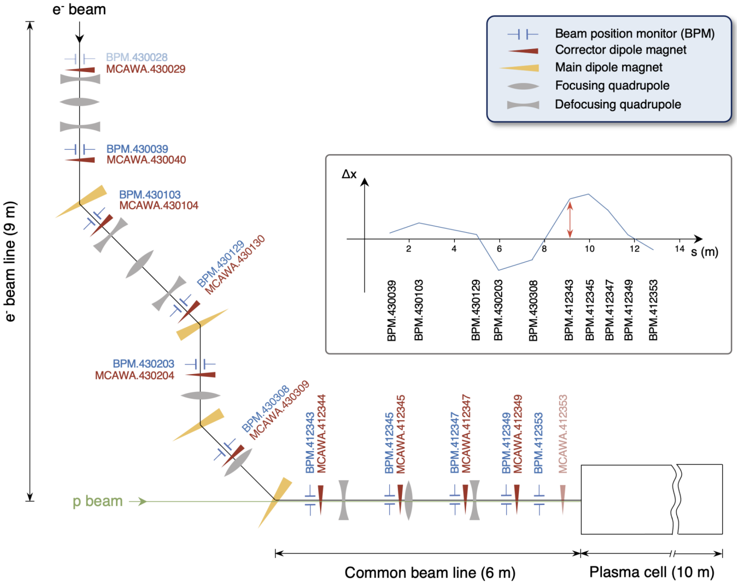
Results with and without experience replay



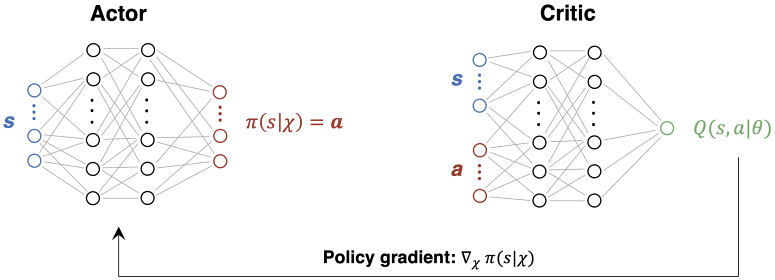
QBM vs neural networks



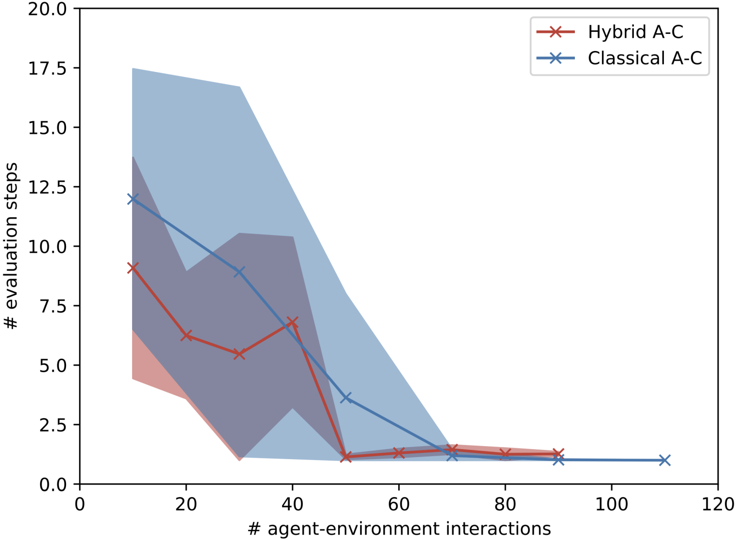
Another problem



Actor-critic architecture



Results



Hybrid A-C
Classical A-C
evaluation steps
agent-environment interactions
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
0
20
40
60
80
100
120

More details in...

IOP Publishing Quantum Sci. Technol. 9 (2024) 025012 https://doi.org/10.1088/2058-9565/ad261b

Quantum Science and Technology

RECEIVED
18 January 2023

REVISED
31 August 2023

ACCEPTED FOR PUBLICATION
5 February 2024

PUBLISHED
21 February 2024

PAPER

Hybrid actor-critic algorithm for quantum reinforcement learning at CERN beam lines

Michael Schenk[1,*], Elías F Combarro[2], Michele Grossi[1], Verena Kain[1], Kevin Shing Bruce Li[1], Mircea-Marian Popa[3] and Sofia Vallecorsa[1]

1 European Organisation for Nuclear Research, Espl. des Particules 1, 1211 Meyrin, Switzerland
2 Computer Science Department, University of Oviedo, C. San Francisco 3, 33003 Oviedo, Asturias, Spain
3 Politehnica University of Bucharest, Splaiul Independentei 313, 060042 Bucharest, Romania
* Author to whom any correspondence should be addressed.

E-mail: michael.schenk@cern.ch

Keywords: energy-based reinforcement learning, quantum machine learning, accelerator physics

Abstract

Free energy-based reinforcement learning (FERL) with clamped quantum Boltzmann machines (QBM) was shown to significantly improve the learning efficiency compared to classical Q-learning with the restriction, however, to discrete state-action space environments. In this paper, the FERL approach is extended to multi-dimensional continuous state-action space environments to open the doors for a broader range of real-world applications. First, free energy-based Q-learning is studied for discrete action spaces, but continuous state spaces and the impact of experience replay on sample efficiency is assessed. In a second step, a hybrid actor-critic (A-C) scheme for continuous state-action spaces is developed based on the deep deterministic policy gradient algorithm combining a classical actor network with a QBM-based critic. The results obtained with quantum annealing (QA), both simulated and with D-Wave QA hardware, are discussed, and the performance is compared to classical reinforcement learning methods. The environments used throughout represent existing particle accelerator beam lines at the European Organisation for Nuclear Research. Among others, the hybrid A-C agent is evaluated on the actual

Thank you for your attention!

QBM free energy

- The system's energy states are described by the Hamiltonian of the transverse-field Ising model

$$\mathcal{H}(\mathbf{v}) = -\sum_{\substack{i \in V, \\ j \in H}} w_{ij} v_i \sigma^z_{h_j} - \sum_{j,k \in H} w_{jk} \sigma^z_{h_j} \sigma^z_{h_k} \quad -\Gamma \sum_{j \in H} \sigma^x_{h_j},$$

- The negative free energy $F(\mathbf{v})$ of the clamped QBM is used to approximate the Q-function

$$\begin{aligned} Q(\mathbf{v}) &\approx -F(\mathbf{v}) \\ &= -\langle \mathcal{H}(\mathbf{v}) \rangle - \frac{1}{\beta} tr(\rho_{\mathbf{v}} \ln(\rho_{\mathbf{v}}))), \end{aligned}$$

with $\rho_{\mathbf{v}} = \frac{e^{-\beta H(\mathbf{v})}}{tr(e^{-\beta H(\mathbf{v})})}$.

QBM free energy (2)

- Effective Hamiltonian given by

$$\mathcal{H}^{\text{eff}}(\mathbf{v}) = -\frac{1}{N_r}\sum_{l=1}^{N_r}\left(\sum_{j,k\in H} w_{jk} h_{j,l} h_{k,l} + \sum_{\substack{i\in V,\\ j\in H}} w_{ij} v_i h_{j,l}\right)$$
$$-w^+\left(\sum_{j\in H}\sum_{l=1}^{N_r} h_{j,l} h_{j,l+1} + \sum_{j\in H} h_{j,1} h_{j,N_r}\right),$$

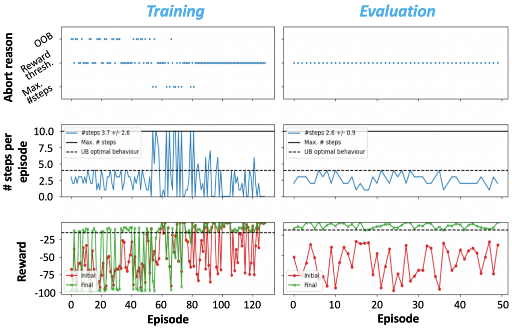
with $w^+ = \frac{1}{2\beta}\log\left[\coth\frac{\Gamma\beta}{N_r}\right]$

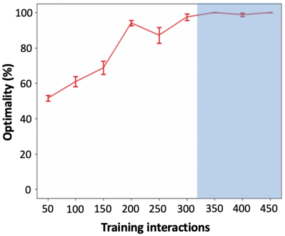
- Negative free energy $F(\mathbf{v})$ given by

$$Q(\mathbf{v}) \approx -F(\mathbf{v})$$
$$= -\langle\mathcal{H}^{\text{eff}}(\mathbf{v})\rangle - \frac{1}{\beta}\sum_c \mathbb{P}(c|\mathbf{v})\,\log\mathbb{P}(c|\mathbf{v}),$$

Results with Q-net

- Stable-baselines3 implementation
- Near-optimal behaviour after 300+ training interactions with environment
- Without experience replay: need roughly 10^4 interactions





Results with QBMs

- SQA and QBM with experience replay: 100 - 120 training interactions sufficient
- Relatively high variability depending on random seed and the states visited
- After hyperparameter tuning (SQA), train on D-Wave 2000Q annealer
- As for SQA: 120 training interactions sufficient
- Save trained QBM weights and evaluate agent with SQA