QA4EO

Quantum Advantage for Earth Observation

Soronzonbold Otgonbaatar[1,2], Olli Nurmi[3], Mikael Johansson[4], Jarmo Mäkelä[4], Tadeusz Kocman[5], Piotr Gawron[6,7], Zbigniew Puchała[7], Łukasz Pawela[7], Jakub Mielczarek[8], and Artur Miroszewski[8]

[1]German Aerospace Center (DLR) Oberpfaffenhofen
[2]Ludwig-Maximilians-Universitaet Muenchen
[3]VTT-Technical Research Centre of Finland Ltd.
[4]CSC–IT Center for Science Ltd.
[5]SYDERAL Polska sp. z o.o.
[6]AstroCeNT, Nicolaus Copernicus Astronomical Center of the Polish Academy of Sciences
[7]Etos Centrum Edukacji i Doradztwa Służby Zdrowia sp. z o.o.
[8]Jagiellonian University

Preprint.

ESA Contract No. 4000140122/22/I-DT

Contents

Quantum and Classical Computing Paradigm

1 Computational Complexity, Quantum, and Classical Computing:

Computational complexity lies at the heart of computer science, serving as a fundamental concept in understanding the efficiency and limits of problem-solving algorithms. It refers to the study of the resources required by algorithms to solve computational problems, typically measured in terms of time and space. Classical computing operates within the realm of classical physics, where computations are processed using classical bits that can represent either a 0 or a 1. Classical algorithms are designed to tackle problems efficiently based on their time and space complexity. However, as the scale and complexity of problems continue to grow, classical computers face inherent limitations in solving certain types of problems within a reasonable timeframe.

Quantum computing, on the other hand, takes advantage of the principles of quantum mechanics to offer potential solutions to computationally challenging problems. Instead of classical bits, quantum computers use quantum bits or qubits, which can represent both 0 and 1 simultaneously thanks to the concept of superposition. This property allows quantum computers to perform parallel computations, potentially leading to exponential speed-ups for specific problems. Quantum algorithms, such as Shor's algorithm for factoring large numbers, demonstrate the potential of quantum computing to solve problems that are currently intractable for classical computers. However, harnessing the power of quantum computing remains a significant technical and engineering challenge, as qubits are highly sensitive to noise and require careful control to maintain coherence (see Figure 1).

Important Cases:

- **Classical Computing:**
  - NP-complete problems: Examples include the Traveling Salesman Problem, Knapsack Problem, and Boolean satisfiability problem.
  - Sorting algorithms: Notable examples include Quicksort, Mergesort, and Heapsort.
  - Graph algorithms: Important cases include Dijkstra's algorithm, Prim's algorithm, and Kruskal's algorithm.
- **Quantum Computing:**
  - Shor's algorithm: A quantum algorithm for factoring large numbers, which has implications for breaking cryptographic protocols based on integer factorization.
  - Grover's algorithm: A quantum algorithm for searching an unsorted database with a quadratic speedup compared to classical algorithms.
  - Quantum simulation: Quantum computers can simulate quantum systems, enabling the study of molecular interactions, material properties, and quantum physics phenomena.

2 Taxonomy for optimization problems

A taxonomy for optimization problems is a classification framework that categorizes different types of optimization problems based on their characteristics and properties. It provides a structured way to analyze and compare various problem domains, algorithms, and solution approaches. A taxonomy typically takes into account factors such as problem structure, objective function type, constraints, and search space characteristics.

One common way to classify optimization problems is based on their problem structure. This classification includes categories such as linear programming, integer programming, nonlinear programming, combinatorial

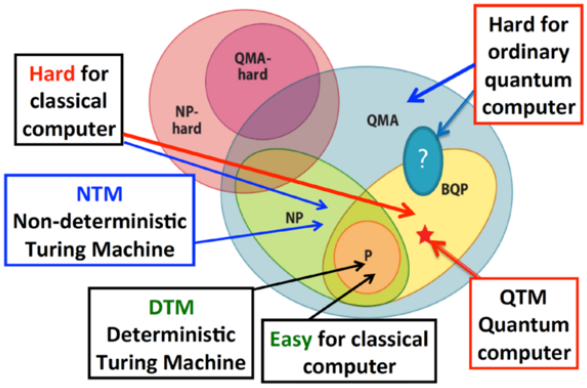


Figure 1: Computational complexity for easy and hard computational problems. Namely, the red star denotes a class of computational problems which is hard for a classical computer but easy for a quantum computer, that is, a polynomial depth quantum algorithm exists. Here, **NP** stands for non-deterministic polynomial time problems, **QMA** stands for quantum Merlin-Arthur problems, the quantum version of the **NP** problems, **P** stands for polynomial time problems, and BQP stands bounded-error quantum polynomial problems.

optimization, and multi-objective optimization. Each category represents a distinct problem structure and often requires specialized algorithms and techniques for efficient solution finding. Another dimension of classification is based on the type of objective function, such as minimization or maximization, and whether the objective function is continuous or discrete.

Additionally, taxonomy for optimization problems may consider the presence of constraints, such as equality constraints, inequality constraints, or both. Problems with constraints are often referred to as constrained optimization problems, and they require specialized methods, such as constrained optimization algorithms or mathematical programming techniques.

Overall, a taxonomy for optimization problems helps researchers and practitioners understand the characteristics and complexity of different problem domains, and it provides a foundation for developing efficient algorithms and solution approaches tailored to specific types of optimization problems.

- **Linear Programming:**
  - Resource allocation: Determining the optimal allocation of limited resources to maximize or minimize a certain objective, subject to linear constraints.
  - Production planning: Optimizing production levels and resource allocation to meet demand while minimizing costs.
  - Portfolio optimization: Finding the optimal allocation of investments to maximize returns while considering risk.
- **Integer Programming:**
  - Traveling Salesman Problem: Determining the shortest possible route that visits a given set of cities and returns to the starting point.
  - Knapsack Problem: Selecting a subset of items with maximum value, considering a limited capacity constraint.
  - Facility Location Problem: Deciding the optimal location for facilities to minimize costs and meet demand requirements.

- **Nonlinear Programming:**
  - Function Optimization: Finding the minimum or maximum of a nonlinear objective function, possibly subject to constraints.
  - Parameter Estimation: Adjusting the parameters of a mathematical model to best fit observed data.
  - Optimal Control: Determining the optimal control inputs over time to optimize a dynamic system's performance.
- **Combinatorial Optimization:**
  - Graph Coloring: Assigning colors to the vertices of a graph, such that no adjacent vertices share the same color.
- **Multi-objective Optimization:**
  - Pareto Front Exploration: Identifying a set of solutions that represents the trade-off between multiple conflicting objectives.
  - Multi-objective Resource Allocation in Project Management: Optimizing the allocation of resources to multiple projects while considering time, cost, and resource constraints.

Quantum Machines Assessment

3 The Assessment of Quantum Technology

The development of quantum computing encompasses a wide range of technologies from hardware systems to software tools depicted in Figure 2. The quantum computing industry is still in its infancy and like the early days of classical computing without well-defined interfaces between the various parts of the quantum computer. The quality of a quantum algorithm is affected not only by the quality of the individual constituent components (qubits, gates, measurements), but also by the interplay of global device and algorithmic properties such as device topology, multi-qubit noise correlations, and circuit structures. Also, the quantum compilers and middleware affect the algorithm performance to be run on certain hardware. Typically, the machine instructions are optimised for execution on all hardware platforms. After the execution additional postprocessing may also be employed to improve readout efficiency. These optimisations typically include:

1. depth reduction and logical transpilation: A sequence of compiler passes is used to mathematically reduce the gate depth (e.g., T-gate count) of the quantum circuit and the logical operations in the circuit are mapped to the native gates available on the hardware.
2. error-aware hardware mapping: Error-aware compilation is used to best select the appropriate subset and logical assignment of qubits on a device.
3. elimination of circuit crosstalk: Dynamical decoupling sequences are incorporated to mitigate various idling errors including dephasing and ZZ crosstalk at the algorithmic level.
4. optimised gate replacement: The process involves automated parsing of the device topology to ensure parallel gate optimizations do not share qubits; and relevant single and or multi-qubit gates are optimised.

QC's usefulness is heavily dependent upon the achievable fidelities, and the number of qubits of the Quantum Processing Unit (QPU). Scaling the quality and number of qubits will require advanced 3D architectures and assembly techniques. Some estimates say that to achieve practical quantum advantage requires running millions of parallel high-fidelity gates at high speed, as well as reading out millions of qubits in parallel. With current error-correction overheads, practical quantum advantage will be achieved, albeit only for algorithms with small I/O requirements and superquadratic (ideally exponential or quartic) speedups over their classical counterparts Beverland et al. [2022].

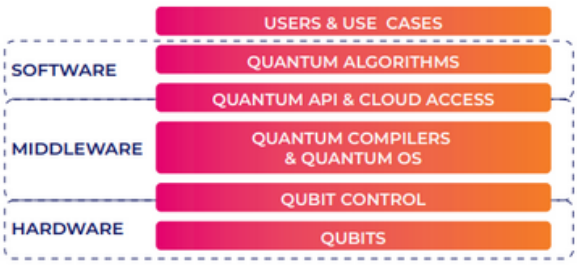


Figure 2: Quantum stack by QuiC showing the software, middleware and hardware layers which have direct impact on the use cases and their prospects.

In the current noisy intermediate-scale quantum (NISQ) era, the computation results are mostly limited by errors in single- and two-qubit quantum gates. To succeed roughly half of the time in a 100-qubit circuit

of depth five, one needs at least 99.9% gate fidelity. In practice, the number of qubits and especially the gate depth required for useful NISQ advantage is likely higher, leading to a fidelity target of 99.99% for all quantum gates, not yet demonstrated. Producing commercially viable QC's requires technologies that facilitate scalable manufacturing requiring manufacturing process efficiency and reliability, integration, and packaging. Due to manufacturing variability some of the qubits may not be functional and available for use; the exact number of qubits yielded will vary with each specific processor manufacturer. The enabling hardware that connects to the QPUs, such as cryogenic coolers, electronic systems, and cabling, need also to be matured. The widely accepted approach to remedy effects of noise and decoherence in quantum computers is the use of quantum error correction (QEC) Fowler and Gidney [2019]. While the hardware requirements to implement fault-tolerant (FT) quantum algorithms have not been met yet, the steady progress in the development of quantum hardware has initiated the introduction of a set of techniques that we refer to broadly as quantum error mitigation. These techniques immediately translate advances in qubit coherence, gate fidelities, readout precision, and speed to measurable advantage in computation. Quantum error mitigation offers the continuous path that will take us from today's quantum hardware to tomorrow's FT quantum computers. They might even be applicable to enable near-term practical quantum advantage without the use of QEC for certain use cases. A major use-case for near- to medium-term quantum computers is to accelerate existing HPC workflows. For this, a tight integration between HPC and QC, beyond cloud-access or the operation as separate compute systems, has been shown to be critical to avoid idle time either due to resource allocation or communications latency. Following three current trends can be identified: (1) stay at "small" scales (below 100 qubits) and try to solve coherence problems and create useful applications before scaling up; (2) go for large scales (over $1,000$ qubits) and try to implement quantum error correction for quantum advantage or superiority while scaling up; (3) scale up and solve large-scale hardware (HW) and software (SW) integration at systems levels.

We mentioned in the previous chapter that QC hardware can be characterised by the kinds of computation they can run into three categories:

1. **Annealers.** Quantum annealers are a kind of an analog quantum simulator relying on the adiabatic theorem and mimicking an Ising Hamiltonian to solve quadratic unconstrained binary optimization (QUBO) problems such as satisfiability problems and combinatorial search problems. QUBO problems are solved by finding their global minimum over a given set of their candidate solutions (candidate states), by a process using quantum fluctuations. In adiabatic computing noise- and error-tolerance are higher, and so hard to create entangled states, the main resource for quantum computational advantage over a conventional classical computer.
2. **Analog Quantum Simulators.** Analog quantum simulators are special purpose devices designed to study quantum systems in a programmable fashion. They exploit superposition and entanglement to provide insight about specific physics problems mimicking the Hamiltonian evolution of the system. Analog quantum simulators are especially suited for simulating quantum physical systems, also more general optimisation is possible. As the quantum interactions between quantum particles is a built-in feature of quantum simulators, near-term quantum advantage is expected for the specific class of problems that they can describe.
3. **Digital Universal Quantum Computers, that is, fault-tolerant universal quantum computers.** The most powerful class of quantum machines, that directly exploit superposition, entanglement, and wave-function interference and run quantum algorithms in a step-by-step procedure. A digital universal quantum computer can, in principle, solve some computable problems, with the additional advantage of up to exponential speed-up over classical computers. Digital quantum computers operate using quantum gates, logical operations on the basic quantum information primitives. These units are usually two-state quantum bits, qubits, but also continuous-variable (CV) approaches are under development. Qubits can be implemented using several different technologies, e.g., superconducting, trapped ions, neutral atoms or using photonics, which all come with their unique strengths and weaknesses. There are some differences in algorithms between discrete and continuous quantum states, with CV approaches especially suited for, e.g., sampling and regression tasks.

4 The qubit implementation techniques

There are plenty of approaches to develop scalable qubits with acceptable coherence time and error rate. Some of the approaches are on a very low TRL level and it's difficult to estimate their potential. In this chapter we describe six most promising approaches based on published information Cheng et al. [2023a]. The connectivity of a quantum gate processor impacts on the depth of actual quantum circuits. During transpilation, an input quantum circuit is compiled to a sequence of native gates or universal gate set such that all operations agree with the qubit topology and noise properties of a specific quantum processor. The signal-to-noise ratio impacts on the number of shots required to get a correct answer by recovering the signal. By Increasing the gate fidelity by a little bit, the number of shots and runtime of a given algorithm may go down drastically. Even a relatively modest 0.16 percentage point improvement in fidelity, could mean that it runs in less than half the time. Building large circuits requires long coherent times of the qubit, strong interqubit interaction for fast and high-fidelity two-qubit gates, and small to zero coupling between qubits when no interaction is needed. Transmon qubits allow for a variety of coupling concepts, with various pros and cons. Two of the most promising technologies are superconductors and ion traps. At the time of writing, at most 433 and 20 qubits are available for superconducting and ion trap devices, respectively, that is, the IBM Odsprey processor, USA and the AQT PINE processor, Austria. And at most 5627 qubits for quantum annealing devices, i.e., D-Wave Advantage. According to the roadmap in 2023 Advantage 2™ quantum system will incorporate a new qubit design that enables 20-way connectivity in a new topology containing 7000+ qubits and make use of the latest improvements in quantum coherence in a multi-layer fabrication stack.

1. **Superconducting circuits.** Physical implementations of superconducting qubits reside on the chip at fixed locations and are connected via a well-defined pattern, the so-called connectivity structure. Structures are designed to minimise the possibility of frequency collisions and optimise the hardware performances. The larger the number of neighbours of a qubit, the more frequencies are required to realise two qubit gates using cross-resonance interaction. Current technology can turn off the coupling of transmon qubits with close frequencies, but this is prone to crosstalk errors. With tunable couplers, a more efficient pulse shape could be optimised to achieve a CZ gate with a higher fidelity and lower unwanted leakage. Until recently, the mainstay devices have been fixed couplers with a constant coupling strength, but attention is now turning to tunable couplers, which are seen as offering the adjustable coupling strength necessary to improve performance. Roadmaps aim for increased coherence, yield and reproducibility, enabling higher gate fidelity and consequently larger circuit depth, on an equal footing with increased qubit number. Three-dimensional multi-chip allow massive scaling of QPUs. It is also necessary to reduce variation of all critical parameters and tolerances for all steps of chip fabrication and 3D integration. Chip engineering needs to consider signal routing, the electromagnetic environment, quantum coherence, and robustness against variations in device parameters. To advance the state of the art in quantum-processor performance require development of novel components for fast and highly selective multiplexed readout, elements for mid-circuit leakage detection, coupling schemes to accelerate parity measurements, conditional and unconditional reset capabilities, and highly parallelizable two-qubit gates. To ramp-up and operate large scale QPU requires also advancing the room-temperature electronic (RTE) systems with sufficient number of control and readout channels and capability for real-time quantum error correction.
2. **Trapped ions.** Ion traps use ions, single-charged atoms, as qubits. Information is encoded in the electronic state of ions that are confined using electric fields. Operations are performed with tailored laser pulses that modify the state of the ions. Ion-trap quantum computers provide optical interfaces and high-fidelity local operations. Multiple ion-trapping potentials can be connected deterministically by physically transporting ions across micro-scale segmented ion traps, which forms an architecture for a scalable quantum information processor. Realising trapped-ion qubits

requires the orchestration of several devices, including the ion source, dedicated lasers, several optical components and sensors, a vacuum, cooling mechanisms, and control and measurement electronics. The respective systems routinely operate with about 20-30 qubits but can be pushed (at reduced levels of control) up to 50 qubits. The devices hold fully connected quantum registers, which facilitate the implementation of quantum algorithms. For trapped ion qubits, the main noise is not relaxation with time T_1 but instead dephasing with time T_2 induced by fluctuation of magnetic fields. Also, the state-detection efficiency decreases with the motional heating of the ion without laser cooling.

3. **Photonic.** Qubits are realised by processing states of different modes of light through both linear and nonlinear elements. The fundamental building blocks include deterministic single photon sources, integrated photonic circuits, and efficient single-photon detectors. Photonic systems have the unique property that they can operate at room temperature and allow for easy transfer of quantum information. The main disadvantage of photonic systems is that performing a precise interaction between photons is a difficult task to achieve. In recent years a couple of programmable and scalable architectures for photonic quantum computing were introduced and specific quantum algorithms such as Gaussian boson sampling, molecular vibronic spectra and graph similarity were executed in laboratories. Photonic circuits, due to properties of photons, have different features from qubit-based systems from the point of view of computing and operations.

4. **Neutral atoms.** Qubits are realised by internal states of neutral atoms trapped in an optical lattice. Like ion-trap systems, qubits can be programmed using the energy levels of the atoms. Light, or electromagnetic radiation, can be used to trap and manipulate the quantum states of uncharged (neutral) atoms. Multiple qubits that are nearby in space can be programmed to interact with one another via two-qubit gates. This opens new possibilities for exotic quantum-computing circuit topologies. Neutral atom platforms for quantum processing have a unique potential for scalability: the size of the quantum register is only limited by the amount of trapping laser power and by the performance of the optical system generating the optical tweezers.

5. **Silicon spin.** QPU integrates both qubits and control electronics and operates at a liquid helium temperature (4K), which is higher than the usual millikelvin temperatures of superconducting qubit systems. The higher operating temperatures result in lower quality qubits, but extensive and efficient control electronics.

6. **NV diamond.** Qubits are realised by the electronic or nuclear spin of nitrogen-vacancy centres in diamond. In these artificial diamond structures, a carbon atom has been replaced by a nitrogen atom near a carbon atom gap. Qubit gates are implemented with microwaves, a magnetic field and an electric field. Qubit readout is using a laser and fluorescence detection.

4.1 QPU performance consideration

To implement a functional quantum computer requires an integrated system consisting of a quantum processor, its fabrication, packaging and wiring, room temperature electronics, enabling software, system integration, application development and testing system. Increasing QPU performance means improving all the subsystems and subcomponents of the machine individually and simultaneously, while ensuring all the systems continue to work well together. Here we focus on the Quantum Processing Unit, QPU. There is not yet a standard to assess the performance levels of the processor. Some of the approaches include benchmarking metrics such as Quantum volume, Algorithmic volume and Randomised benchmarking. To keep the qubit error rates below a certain threshold for fault-tolerant computation, it is crucial to extend the coherence time of qubits. Here we list some current critical areas in Qubit implementation, Qubit control, Qubit calibration and Code running. (operation times in the microsecond range, and qubits are more relevant for superconducting transmon qubits or spin qubits.

Table 1: Some dominating quantum machines in the global market offered by large organizations. See Figure 3 for the projection of the roadmap of some quantum machines and Table 2 for parameters of quantum machines QRM.

Organizations	Locations	Technology	Current qubits	Projected qubits (3-5 years)
IBM	USA	superconducting	433	$4,158$
Google	USA	superconducting	73	100
IQM	FI	superconducting	5	54
USTC	CN	superconducting	66	100
AQT	AT	trapped ions	20	200
IONQ	USA	trapped ions	29	256
Xanadu	CA	photonic	216	216
USTC	CN	photonic	113	300
D-Wave	CA	superconducting-annealing	$5,000$	$10,000$
QuEra	USA	neutral atoms	256	$1,000$

Currently, only trapped ions and superconducting qubits satisfy the five required criteria for quantum computing defined by DiVincenzo:

1. A scalable physical system with well characterised qubits;
2. The ability to initialise the state of the qubits to a simple fiducial state;
3. Long relevant decoherence times;
4. A "universal" set of quantum gates;
5. A qubit-specific measurement capability DiVincenzo [2000].

Typical physical indicators of quantum computers include T_1, T_2, single qubit gate fidelity, two qubit gate fidelity, and readout fidelity. The aggregated benchmarks can help the user to determine the performance of a quantum processor with only one or several parameters. The aggregated metrics can be calculated with randomly generated quantum circuits or estimated based on the basic physical properties of a quantum processor. Typical aggregated benchmarks include quantum volume (QV) and algorithmic qubits (AQ). Specific attributes (e.g., faster gate speeds, higher fidelities, denser connectivity) can make certain machines better at particular tasks than others, but not superior in absolute terms for the time being. Also, certain QPU may then fit better on certain QC4EO use cases based on the problem they need to solve instead of an arbitrary rating. For example, ion trap devices are able to make up for slower operation speeds with better connectivity, while the superconducting systems with sparser connectivities are still competitive due to their much faster operation times. Another point is that trapped ion qubits have very long coherence times, making them more robust to mid-circuit measurement — a key requirement for error correction. However, the 1000x faster gate speeds of superconducting are preferable for variational benchmarks like QAOA, which require millions of sequential iterations Cheng et al. [2023a]. We presented example quantum machines in the Table 1.

4.2 Sizing quantum machines

Modern classical central processing units (CPUs) operate at around 3GHz clock cycle speed, i.e., that is, around 0.30ns clock cycle time. Nowadays, computationally hard problems are even tackled on several

Table 2: Sizing quantum machines: SC−superconducting QCs Acharya et al. [2023], T.ions−trapped ions QCs Pogorelov et al. [2021], N.atoms−neutral atoms QCs Henriet et al. [2020], Photonic−photonic QCs Bartolucci et al. [2023], S.spin−silicon spin QCs Stano and Loss [2022], NV−nitrogen vacancy in diamond QCs Cheng et al. [2023a], CPUs− conventional central processing units. See also the Table 1.

Parameters	SC	T.ions	Photonic	N.atoms	S.spin	NV	CPUs
Clock cycle	1MHz	1KHz	10Hz	1MHz	0.76MHz	1MHz	3GHz
Measurement	$660ns$	$300\mu s$	x	$200ms$	$1.3\mu s$	x	x
2-qubit gate	$34ns$	$200\mu s$	x	$< 100\mu s$	x	$700ns$	x
1-qubit gate	$25ns$	$15\mu s$	x	x	x	$9ns$	x
Readout fidelity	99.4%	97.3%	50.0%	99.1%	99%	98%	x
1Q fidelity	99.99%	99.99%	99.84%	99.83%	99.99%	99.99%	x
2Q fidelity	99.97%	99.9%	99.69%	99.4%	99.5%	99.2%	x

hundreds of parallel CPUs as well as general processing units (GPUs). The fastest QPU is currently a superconducting-based QPU (see below table) in terms of the qubit and quantum gate operation time, that is, clock cycle time. However, I/O speed is 10,000 slower in QC compared to CPU. Logical qubit/Magic state distillation (creating more accurate quantum states from multiple noisy ones) are other restrictions, and another restriction is high-bandwidth low-noise classical electronics. Hence, to beat CPUs, there is the need to improve the speed of the whole I/O system in QPUs from register preparation to read-out. More than exponential speedup is also required in the quantum algorithm Beverland et al. [2022], and only some of the problems are meaningful to compare depending on their parallelizability on CPUs and GPUs (see Table 2 and Figure 3). Regardless of the qubit technology, there is the persisting challenge to scale logical error-free qubits due to the quantum state generation having a high fidelity and classical electronics controls, to name a few Cheng et al. [2023a].

4.3 Error mitigation and correction

Errors are generated by various interactions, electromagnetic or mechanical, between qubits and their immediate environment and are associated with the phenomenon of quantum decoherence. Error removal is progressing steadily but barely managing to gain one or two orders of magnitude in error whereas in an ideal world, we would need ten orders of magnitude improvements. It is possible to correct errors, even by using noisy gates, provided that the noise level remains below a certain threshold. The drawback is that it requires huge overhead of physical qubits and classical information processing (see Figure 4 and 5) Fowler and Gidney [2019]. There is an optimal "code size", i.e. number of physical qubits per logical qubit, that maximises the metric of performance— and beyond which more error correction degrades the computation accuracy rates. Also, less noise mandates a bigger code and more physical qubits, but more physical qubits give rise to more heat generation, hence more noise. To execute a quantum application successfully, QEC must be used to build logical qubits that can be used to store and manipulate quantum information better than raw physical qubits. This QEC capability is central to scalable quantum computers, but the costs are formidable, often multiplying the number of qubits needed by a factor of thousands, and runtimes by a factor of hundreds. One of the trends for improving the error correction rates characteristics is employing AI models for this process. This would in turn allow us to reduce the number of quantum computation instances needed before obtaining a reliable result, or decrease the number of physical qubits in QC systems. In Europe there exists a start-up which develops a toolkit for providing this form of QC improvement. An important metric for a QEC approach is its threshold, which specifies the maximum error rate that it can tolerate. To

avoid prohibitive QEC overheads, physical error rates on Clifford operations below 0.1% (including qubit preparations, measurements, and gates) are typically required. These values are possible to obtain only in the setting where operations can be applied in parallel, which may pose a significant hardware challenge for some platforms such as trapped ions. In many QEC schemes the non-Clifford gates (typically T gate) are quite costly when requiring fault tolerance Gidney and Fowler [2019]. The required low error rate T states are produced using a T state distillation factory involving a sequence of rounds of distillation, where each round takes in many noisy T states encoded in a smaller distance code, processes them using a distillation unit, and outputs fewer less noisy T states encoded in a larger distance code, with the number of rounds, distillation units, and distances all being parameters which can be varied. This procedure is iterated, where the output T states of one round are fed into the next round as inputs. T factories incur significant physical overheads, requiring several thousand physical qubits and only producing new T states once every 10 to 15 logical time steps Litinski [2019].

Microsoft (MS) has evaluated three use cases concluding that to achieve practical quantum advantage QC's need to be able both to control millions of parallel operations with low error rates, and to readout out those millions of qubits in parallel to enable decoding of the errors at speed; all while ensuring the overarching logical clock time is fast enough to complete the computation within a month runtime or less Reiher et al. [2017]. MS concluded that logical gate times under $10\mu s$, in turn requiring physical gate times around $100ns$, would be needed to complete the quantum chemistry algorithm within a month, using a few million physical qubits. To execute syndrome measurements on these qubits and communicate the quantum measurements to the decoder requires large quantum-classical bandwidth and processing power for decoding. The exact estimates of bandwidth requirements depend on the choice of QEC code, system size and physical operation times, but roughly, with a few million qubits, the estimation is that several terabytes per second of bandwidth will be required between the quantum and classical plane. Furthermore, processing these measurements at a rate that is sufficient to effectively correct errors demand petascale classical computing resources that are tightly integrated with the quantum machine.

5 Investment in quantum computing

Across Europe and the World quantum computing is gathering investment from either states and organizations, as well as private investors. In 2022, the investment in a quantum technology was globally around 30 billion euro, and in 2023, the investment amounts to 36 billion euro. By 2028, the overall investment in a quantum technology is projected to reach globally 53.2 billion euro, and quantum computing investment alone is estimated to be around 17.6 billion euro RM. Several major players are Qinsider:

1. **European Union** - The EU Chips Act with a total budget of around 43 billion euro has a quantum component included, and the European Quantum Flagship program invests around 1 billion euro in quantum computing excluding other quantum technologies like quantum sensing.
2. **USA**-The USA Chips Act with a total budget of around 50 billion euro has a quantum component included, and the US National Quantum initiative invests around 3.75 billion euro in quantum computing alone.
3. **China** - The one of leading players in quantum computing alongside the USA. Its quantum initiative invests around 15 billion euro in quantum computing.

6 When can we expect quantum advantage in Earth observation use-cases?

The United States is perceived as the leading player in quantum technology even though Europe has made clearly the most public investments in the industry. In the United States, big technology enterprises such as Microsoft, Google, Intel and IBM have driven commercial development efforts. In Europe, development has

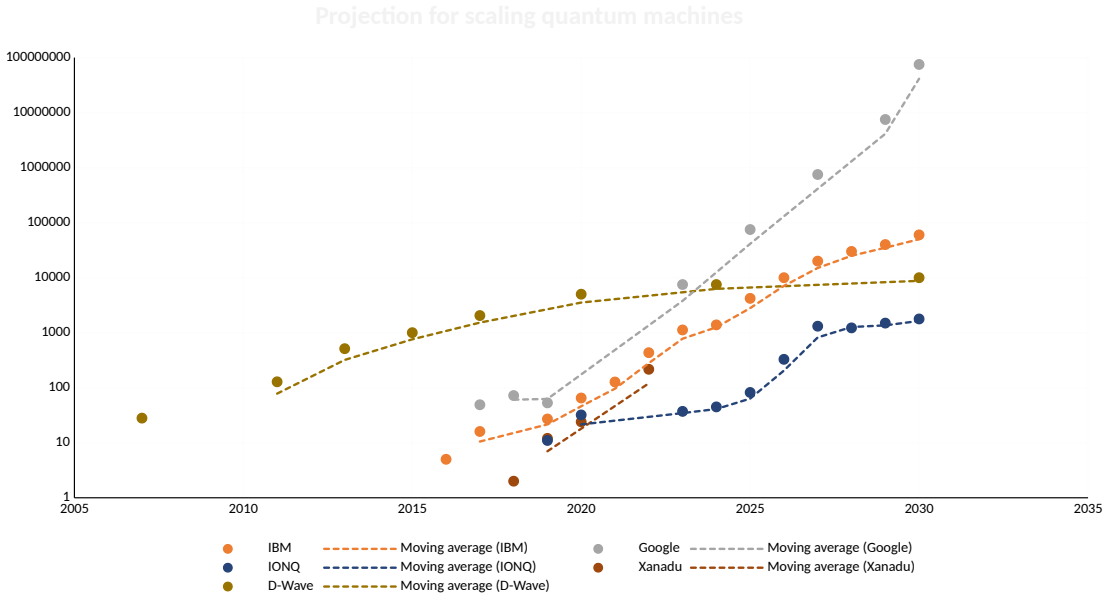


Figure 3: Quantum machines roadmap of some organizations which provide the open-data for their quantum development projection.

been slowed down by fragmentation. Currently there are about 140 projects, less than half commercial. Many of the groups listed are university or government labs, or departments within larger tech companies. Here we can make distinction between two approaches:

- components provided addressing parts of the HW stack which then may be integrated using so-called open architecture,
- a system integrator capable of bringing together and coordinating all the needed competences and components that will make up a commercially viable quantum computer.

Superconducting qubit-based approaches are the most researched (and have received the most development resources). Almost all the startups in this space are based on technology from university labs. To manufacture a stable QC requires more than an exploratory chip. As of early 2023, there have been around a few dozen successful attempts to build quantum computers around the world. There are some specialised companies that are developing middleware for the calibration, management and optimization of quantum computers to overcome some of the problems caused by errors.

Estimating if and when scalable and useful quantum computers will be available is a difficult art and science. The opinion spread between optimists and pessimists is quite large. As published in their 2020 roadmaps, Google, IBM and Amazon expect to achieve true quantum supremacy relatively quickly and create a quantum computer with 100 logical qubits in less than a decade. On the other end there are some pessimistic views saying that there is no hope to reach quantum speedup ever. There is not really any strong scientific obstacle preventing the creation of reliable quantum computers. In the scientific community there is a belief that the uncertainty is mostly a technological and engineering one and the pace to quantum usability is accelerating. However, there is some pessimism about the ability to fix the noise that affects qubits, whatever their type. A potentially exponential quantum speedup provided by quantum computers may vanish when there is big data that needs to be loaded from classical data, or when the full solution vector should be read out. Generally,

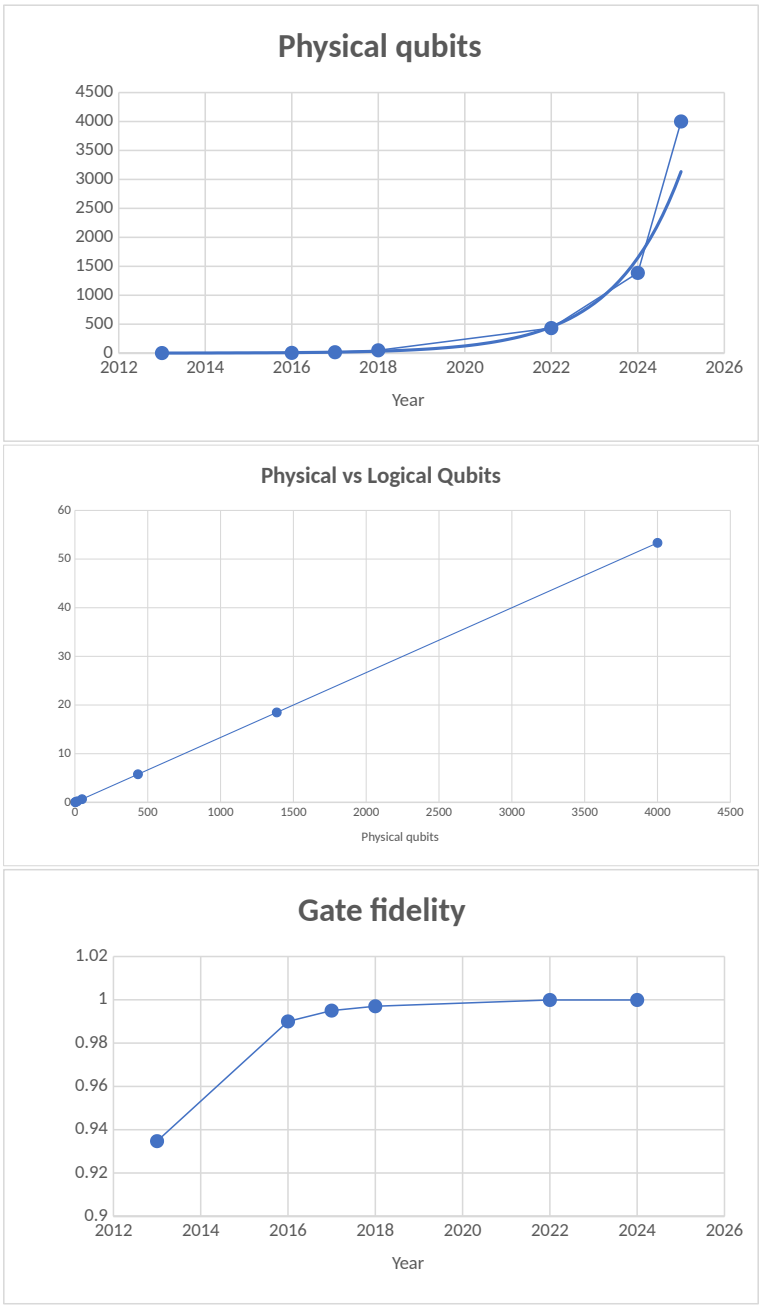


Figure 4: [Top] the trend line for scaling superconducting-based error-prone qubits, that is, physical qubits, [Middle] their corresponding logical qubits, when we assume that a single logical qubit is represented by 75 physical qubits based on the error-correction, [Bottom] the gate fidelity over the years.

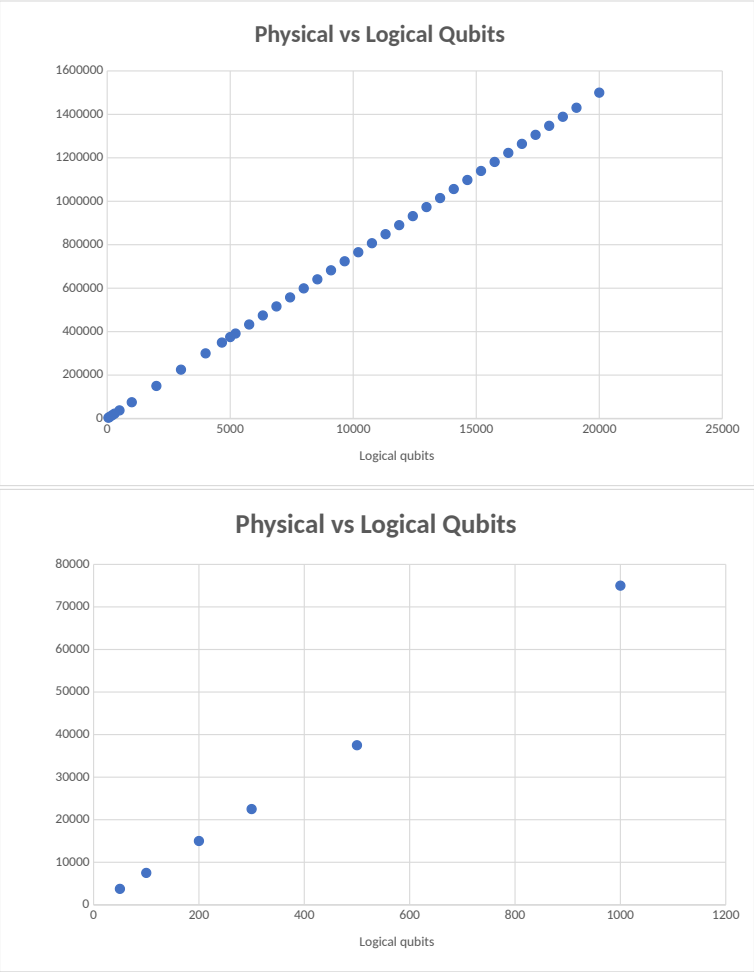


Figure 5: The trend line for physical and logical qubits in the future (beyond 2024) such that a single logical qubit is represented by 75 physical qubits.

quantum computers are considered as practical for "big compute" problems on small data, not big data problems. There is a growing number of informative end-to-end resource analyses, but typically these single out very specific algorithms and hardware and make very different assumptions across the stack. Different choices can result in different resource requirements. One can, for example, trade off more qubits against shorter run times, or trade off faster qubit gate operations against lower fidelities. It is obvious that the number of physical qubits and the duration of a logical time step reduce as physical error rates improve. Entanglement has long been considered to play an essential role in quantum computing and promise for exponential speedup of various quantum algorithms that require asymptotically fewer operations than their classical counterpart. Specific examples where this is the case are quantum problems in chemistry and materials science. Entanglement can be seen as the key feature that sets quantum computing apart from classically simulable processes. Thus, the key metrics to follow the development should include the number and quality of entangling gates provided. The GHZ states provide the strongest non-local correlations possible for an n-particle entangled state. These GHZ states are very fragile, as loss of a single particle completely

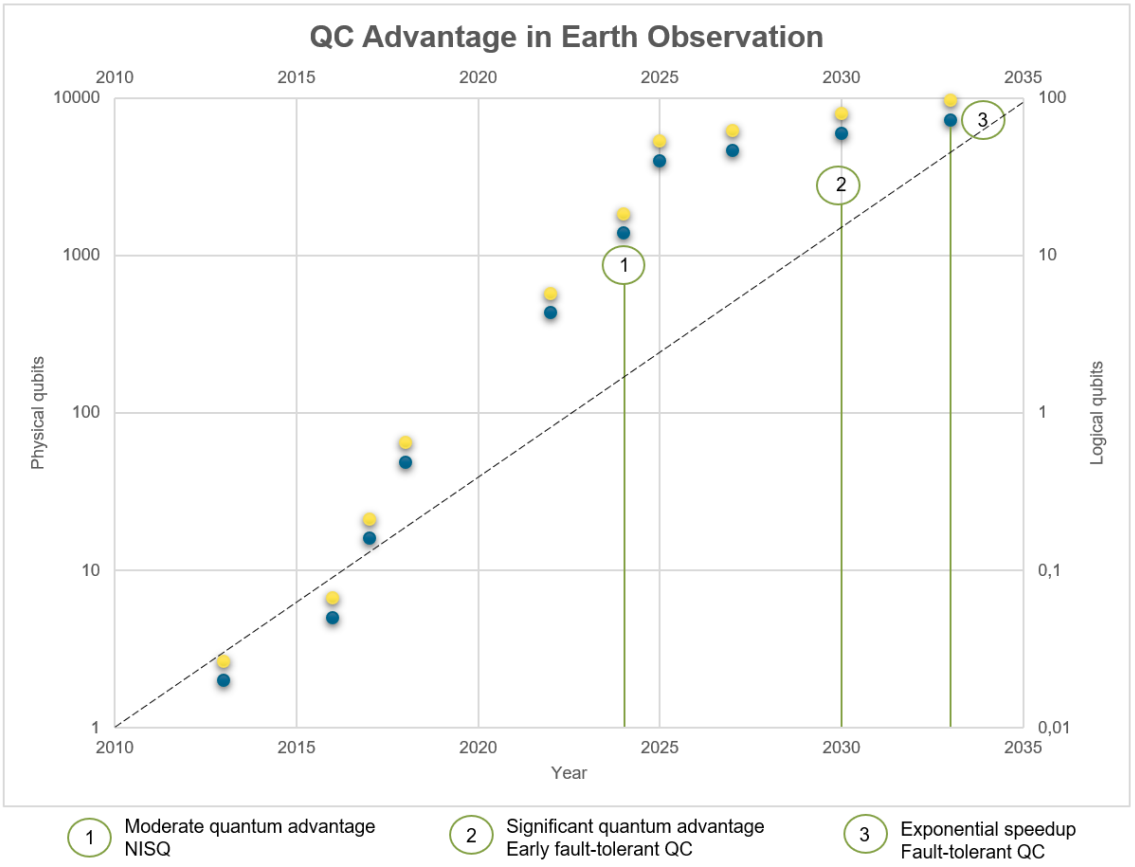


Figure 6: Estimation of quantum advantage in the Earth Observation using gate based computers. In the left vertical axis is the number of physical qubits denoted by blue which can be in superposition. On the right vertical axis is the number of logical qubits related to their physical qubits denoted by yellow over the years.

destroys the entanglement. Also, because all particles contribute to the phase evolution, the dephasing time decreases with the particle number. Such states are challenging to create, requiring either many particles to interact with each other or a series of two-particle interactions performed in sequence. Some of the recent approaches to improve the SC qubit fidelities include

- redesigning the qubit geometries,
- use of new low-loss materials and
- optimising the control pulse that drives the quantum system.

Based on the current quantum volume indicator, which is a marketing simplification tool from IBM, current quantum computers are more than easily emulable in a simple classical computer. In current NISQ-computers the fully manageable number of qubits is somewhere below 25, way below 50 which is the limit for emulation. Quantum volume sets limit to Hilbert's vector space, i.e. the number of different superposed states that is manageable from a practical point of view with a depth of computation equal to the number of corresponding qubits. Based on the expert estimation we believe that starting in 2025, we see some relevant quantum

advantages with actual data and useful algorithms running on NISQ hardware in the Earth Observation domain. A quantum advantage can come from the computing time, system energetic footprint and/or the precision of the outcome (metrics: time to solution, energy consumed to reach the solution and precision of the solution). We estimate that the threshold of 150-ish high quality qubits, with a low error rate and a long coherence time, will be needed to achieve any real quantum advantage. With these qubits it may be possible to form about 10 logical qubits. However, entangled qubits are required for exponential speedup and significant quantum advantage. We estimate that the number of maximally entangled logical qubits will start growing exponentially around 2030 with advancement in the qubit engineering. We summarise this development in the following graph in three phases (see Figure 6).

1. Late NISQ era: ($100 - 200+$ physical qubits; $99.99\%+$ fidelities, especially 2Q gate fidelity; high qubit connectivity) ($3 - 5$ years from now).
2. Early Fault Tolerant QC era delivering significant advantage (< 10 maximally entangled logical qubits) ($5 - 10$ years from now).
3. Fault Tolerant QC era delivering exponential advantage (> 50 maximally entangled logical qubits) ($10 - 20$ years from now).

Climate Adaptation Digital Twin HPC+QC Workflow

Abstract

A Digital Twin is a virtual model of a given system or process, where one as accurately as possible attempts to reproduce its behaviour. The goal is to model the time evolution based on initial values and possible perturbations. Some of the most ambitious projects of this type are found within the Destination Earth initiative, where digital twins of the Earth are constructed. Modelling processes at this scale requires massive amount of computational resources. Only now, with the advent of pre-exascale and exascale supercomputers, has this become a realistic prospect. Here, we scrutinise the present state-of-the-art in climate modelling, the Climate Adaptation Digital Twin (ClimateDT), which is currently being prepared for running on the most powerful supercomputing infrastructures in the world. We discuss the potential of quantum computers for further improving both the efficiency and accuracy of the models, by combining high-performance computing (HPC) and quantum computing (QC) in a hybrid HPC+QC manner. The biggest promise of quantum acceleration comes from aiding machine learning models and from improving the climate models by including processes that presently are computationally intractable due to their complexity. These include, for example, chemical processes in the atmosphere.

7 Introduction

The Climate Adaptation Digital Twin (ClimateDT) is a project issued by the European Centre for Medium-Range Weather Forecasts (ECMWF) in the Destination Earth initiative, where the goal is to develop a highly accurate digital model of the Earth see Figure 7. The aim is to develop an accurate model of the Earth in order to monitor and simulate the interactions between the natural environment and human activities with as high precision as possible. Through this, the effects of various natural phenomena and human actions on the climate can be studied. The underlying goal is to move from plausibility assessments of local and regional climate to fully-developed risk assessments. The ClimateDT is being developed in response to the European Commission's Green Deal and Digital Strategy and it will make it possible to predict the effect of specific climate actions, which will aid policy makers to make informed decisions on how to best mitigate the effects of climate change.

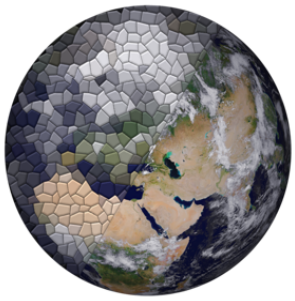


Figure 7: Digital twins of the Earth attempt to replicate the behaviour of certain aspects of the planet based on Earth Observation data and physical models.

The ClimateDT is being developed around two Earth System Models (ESMs), the Max Planck Institute for Meteorology's (MPI-M) Icosahedral Nonhydrostatic Weather and Climate Model (ICON) and ECMWF's Integrated Forecast System (IFS), the latter coupled with two different ocean models: FESOM and NEMO. Both of these ESMs have demonstrated the ability to run at grid scales finer than 5 km globally, coupled to an ocean model. The Climate DT introduces the idea of a generic state vector (GSV), which is evolved by the ESMs, quality controlled and interpolated to a common grid (5km or finer global mesh), and "streamed" to applications. This creates an information system that can scale across an unlimited number of applications that have access to all necessary data and achieve the long-sought goal of interactivity and new ways of co-design.

The ESMs in ClimateDT are being developed in three configuration: a coarse 10 km grid resolution for development, 5 km grid for production, and 2.5 km grid for prototyping subsequent Destination Earth phases. The ClimateDT focuses on five use cases (with a couple of keywords) drawn from climate impact sectors: forestry (wildfires, forest management), urban environments (heatwaves and heat island effect), hydrology (river flows, fresh water availability), hydro-meteorology (extreme-events, flooding), and energy (changing patterns of wind and sunshine, storm vulnerability). The ClimateDT is also complemented by other projects, such as the biodiversity digital twin BioDT [https://biodt.eu/].

Accurate digital twins of the entire Earth have only become possible with the latest generation of supercomputers, that is, the pre-exascale and exascale systems. There are presently five supercomputers on the Top500 list (June 2023) that have a sustained performance of over 100 petaFLOPS, that is, the capacity to perform over 10^{17} floating point operations per second (two of these, LUMI and Leonardo, are European). Even with this impressive increase in data processing capacity, an increase of a factor of million over the last 25 years, digital twins are an ambitious undertaking, and the models necessarily include approximations that affect the accuracy and reliability of the predictions.

Here, we scrutinise the present state-of-the-art workflows for setting up digital twins for climate adaptation, with the intent of identifying areas where quantum computing has a potential for speeding up or increasing the accuracy of selected parts of the entire simulation. Further, we identify the sources of the largest uncertainties in the model in the form of missing parameters or physics in the model, again with the aim to identify areas where quantum computing could provide an advantage.

8 Climate Digital Twin Workflow Analysis

8.1 Present Classical Approach

The Climate Digital Twin workflow is presented in Figure 8. The workflow begins with the typical initialisation and preparatory steps required by a climate or Earth System Model (ESM). In the Climate DT project, the ESMs in use are ICON and IFS. In the workflow, the current model state, illustrated as a Model State Vector (MSV), is propagated forward in time to produce a new state and, simultaneously, the model output or Output State Vector (OSV). This output is streamed (not saved) through a processing pipeline – that introduces additional diagnostic variables and handles interpolation, meta-data conversion and simple operations on the fields – to generate a Generic State Vector (GSV). The GSV is saved directly to Fields DataBase, which is a domain-specific object store developed at the ECMWF; another streaming approach is also being developed with the use of Maestro (https://www.maestro-data.eu/). The GSV is then forwarded to the applications and quality assessment and uncertainty quantification (AQUA), all of which can also utilise external data sources, e.g., observations, climatologies and reanalysis.

The most resource-heavy and time consuming part of this workflow, i.e., the bottleneck, is the climate model itself. Here we note, that the amount of data in a climate model is large. With a typical grid resolution of 10 km, the total number of grid points representing the atmosphere is in the hundreds of millions. Each grip

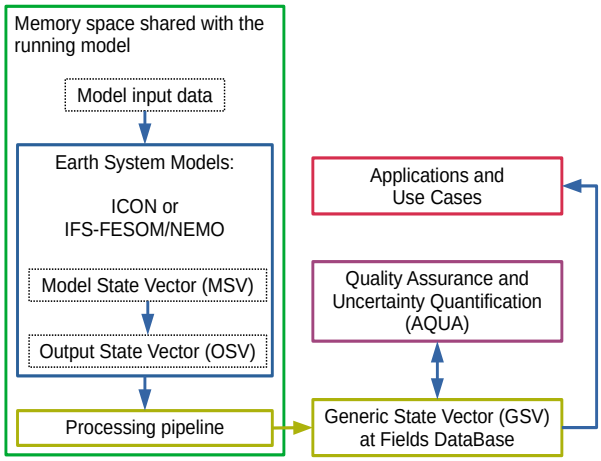


Figure 8: Operational version of the ClimateDT workflow that will be developed during 2023.

point has several variables associated with it, such as air density, temperature, wind speed, humidity, etc. The total parameter space is thus counted in the billions.

Figure 9 shows the relation between different processes in the ICON-Sapphire Earth system model Hohenegger et al. [2023]. What can be seen is that different processes are updated at different intervals, that is with different Δt. This is partly due to the varying computational complexity for propagating specific processes in time in the Earth and climate models. The shortest time steps are those of the dynamical core computations that solve the fluid dynamics equations of atmospheric motions, while the radiative transfer computations have the longest time steps. There is roughly a 1:30 ratio between the shortest and longest time steps. In the latest climate models within ClimateDT, with a resolution of 10 km, the time steps for dynamics and radiation are typically 60 s and 30 min, respectively. Presently, the wall-time for computing the individual time-steps range from the subsecond regime to around 10 s on the LUMI supercomputer. We note that doubling the resolution of the model typically requires halving of the time steps, following the Courant-Friedrichs-Lewy condition Courant et al. [1928]. Thus, doubling the resolution, e.g., going from 10 km to 5 km increases the computational complexity roughly by a factor of 8.

8.2 Quantum Perspective

From the previous section, we can identify two main challenges that hamper direct adoption of quantum computing to climate modelling problems within ClimateDT:

1. ”big data” problem
2. short wall-time for individual calculations

First, the climate models work on a large amount of data, both as input and output. These ”Big data” problems are, however, not directly suitable for quantum computers. The strength of quantum computers lies in being able to solve problems with a *moderate* amount of both input and output variables, where the relation between input and output variables is a highly complex equation that can be solved efficiently by some quantum algorithm, exploiting quantum parallelism Hoefler et al. [2023]. In other words, quantum computing typically requires problems that have a large potential solution space, but only a small set or even a single solution, with the additional provision that the input parameters need to be of the same order of magnitude as the number of qubits in the system.

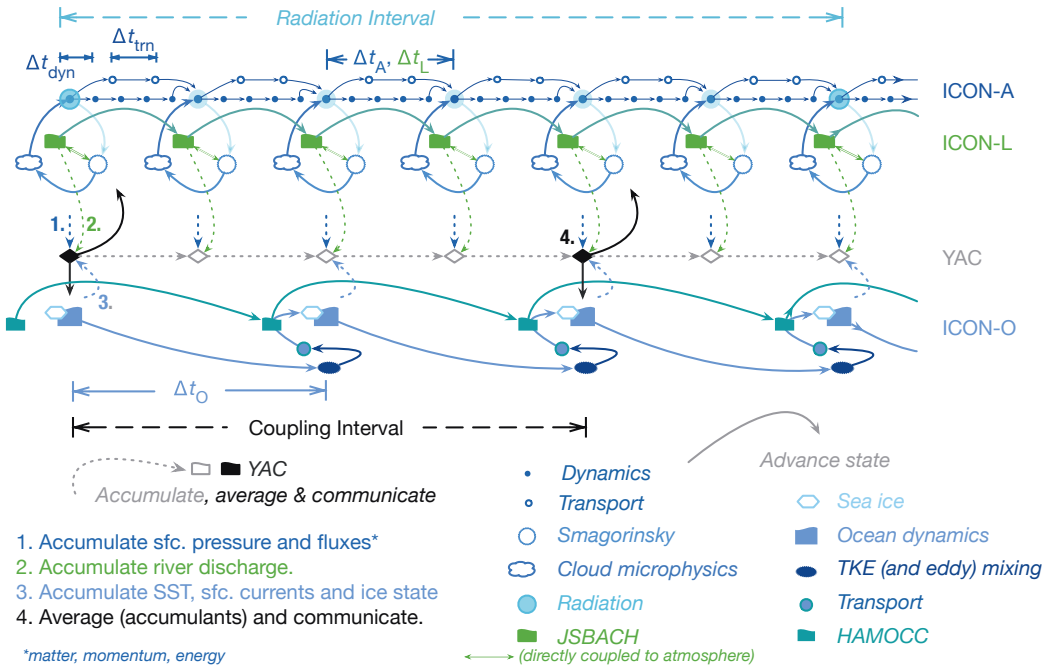


Figure 9: Time stepping in the ICON-Sapphire Earth system model. ICON-A: atmosphere component; ICON-L: land component; YAC: atmosphere-ocean coupler; ICON-O: ocean component. Reproduced from Hohenegger et al. [2023] under the Creative Commons Attribution 4.0 license.

Second, for quantum computers to be able to show a wall-time advantage over classical computers, they need to solve sufficiently complex algorithms. This means that the algorithms have to be sufficiently deep, that is, the number of basic operations has to be high. In practice, this means that single useful quantum computing calculations will take at least seconds to complete Humble et al. [2021]. Individual variational circuits can and do take shorter time, but the wall-time to solution is of course much longer, as several iterations need to be performed. On the other hand, already now, the shortest individual time-steps in the climate digital twins take less than a second, and even the longest around 10 seconds. Further, the aim of the ClimateDT initiative is to speed up the individual time steps significantly, with up to a factor of one hundred. This would push *all* of the individual propagation calculations into the sub-second regime. Thus, quantum computers cannot speed up these calculations further, as they already are faster than the fastest useful quantum computer calculations.

Climate models would thus, at a first glance, seem to be rather unsuitable for quantum acceleration. In order to gain some quantum advantage, we need to consider the problem at hand from a broader perspective. Simply taking present classical algorithms and the approximations they include and rely on, and transforming these to quantum versions of the same will not work. Instead, quantum advantage will be found by approaching the problem from different, new angles, utilising the unique features of quantum machines.

A large part of the calculations in the current workflows are in effect Computational Fluid Dynamics (CFD). Here, we have a direct connection to solving linear systems of equations. The HHL quantum algorithm for linear systems of equations, named after its authors Harrow, Hassidim, and Lloyd Harrow et al. [2009], and variations thereof, thus have the potential to speed up also CFD simulations. As noted by Lapworth Lapworth [2022], classical algorithms running on supercomputers are highly efficient at solving matrix equations by,

for example, side-stepping the need for matrix inversions. Quantum algorithms do not need to, even *should* not rely on the same approximations as classical algorithms, however. Quantum algorithms like HHL and the Quantum Singular Value Transformation (QSVT) Gilyén et al. [2019] can efficiently perform direct matrix inversions, and should therefore be utilised for quantum advantage. The approach presented by Lapworth Lapworth [2022] relies on fault-tolerant quantum computers, but also hybrid classical/quantum algorithms for the NISQ era have been proposed and discussed Kyriienko et al. [2021].

8.3 Enhancing Machine Learning Approaches

The use of Machine Learning (ML) and Artificial Intelligence (AI) in climate modelling and related fields is a hot topic of research Tuia et al. [2023]. ML and AI show promise for accelerating the resource heavy calculations involved also in digital twins of the Earth Chantry et al. [2021]; Watson-Parris [2021]. Computational Fluid Dynamics (CFD), as discussed above, is central to climate models. A promising approach is, e.g., to decrease the grid size without losing accuracy, by using ML for interpolation Kochkov et al. [2021]. As the atmospheric events to a large part are rather smoothly changing, the speed-up from machine learning can be expected to be significant.

The connection to quantum computing here comes at a general level. As discussed in other sections of this report, quantum machine learning has potential advantages over purely classical machine learning. Also here, we need to keep in mind that present-day classical machine learning approaches are immensely powerful. Therefore, quantum algorithms will have a hard time directly competing with classical algorithms from a pure speed-up perspective Schuld and Killoran [2022]. Instead of speed-up, quantum computers can possibly increase the accuracy of the models or decrease the amount of training data required for building the models. Here, the advantage arises from doing the training *differently*, not necessarily faster. For example, hybrid quantum/classical neural networks, where a neural network consists of both classical and quantum layers, has the potential to outperform purely classical and purely quantum approaches Xia and Kais [2020]; Arthur and Date [2022]. As another example, Quantum Support Vector Machines (QSVM) have the potential to perform classification tasks more efficiently than their classical counterparts Havlíček et al. [2019]; Schuld and Killoran [2019]; Jäger and Krems [2023].

The understanding of where classical ML and AI can be utilised in climate modelling is thus being established through significant global efforts. The next step is to identify those machine learning tasks that can benefit from a quantum ingredient. This task is, however, highly empirical by nature. Only by testing, trial, and error, can the most successful quantum machine learning approaches be identified and refined. At this stage, it is too early to predict the future impact of QML on climate modelling. There is cause for careful optimism however, especially for improving the quality of the models, if not directly the time required for establishing them.

9 Missing Physics in the Models

In this section, we discuss two of the major approximations in the present climate DT models: clouds and atmospheric chemistry. Their inclusion is presently prohibitively expensive from a computational resource point-of-view. Quantum algorithms and quantum computing could bring about the necessary reduction in required computational resources in order to enable the inclusion of more parameters and additional physics into the climate digital twins, also beyond these two examples.

9.1 Clouds

Cloud feedbacks and cloud-aerosol interactions are the most likely contributors to the high values and increased range of equilibrium climate sensitivity in CMIP6 Meehl et al. [2020]. In the past, clouds have been poorly represented in Earth System Models (ESMs) due to the complex cloud formation process and because the models could not be run on the scales at which clouds form. Additionally, numerical cloud

modelling has relied on the Eulerian continuous medium approach for all cloud thermodynamic variables. However, recently modelling has shifted towards Lagrangian particle-based probabilistic approaches in small and cloud-scale simulations. Clouds are being taken seriously – the World Climate Research Programme has launched a Grand Challenge on Clouds, Circulation and Climate Sensitivity and NASA has a Grand Challenge "Uncertainty Project" Fridlind et al. [2021] tackling cloud physics knowledge on ESMs.

Clouds are also a focus point for the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) initiative, where a relatively recent review Stevens et al. [2019] proposed a protocol for the first intercomparison project of global storm-resolving models. The review presents 40-day global model simulations (these include ICON and IFS) with a grid resolution uniformly lower than 5km and addresses both scientific aspects and computational performance analysis. The outlook is optimistic even though the authors note that fully resolving shallow cloud systems, whose vertical (and hence horizontal) scale may be only a few kilometers, requires substantially smaller grid distances. This ties in with machine learning efforts for cloud cover modelling Grundner et al. [2022], and consequently with quantum machine learning efforts discussed above. We expect cloud representation to improve in all ESMs, including ICON and IFS; In the first phase, using purely classical supercomputing, and subsequently, by quantum-accelerated HPC.

9.2 Atmospheric Chemistry

Li et al. report that "climate models indicate at least a 30% uncertainty in aerosol direct forcing and 100% uncertainty in indirect forcing due to aerosol–cloud interactions" Li et al. [2020]. The accurate modelling of atmospheric chemistry would thus be of crucial importance for increasing the reliability of the climate digital twins. Atmospheric chemistry is highly challenging from a modelling perspective. Many of the reactions involve radicals, and several are photochemical in nature. This necessitates the use of highly sophisticated electron-correlation methods for describing the electronic structure of the molecular species. Highly correlated quantum chemical wave-function methods are notoriously difficult for classical computers. This is due to the expensive scaling of the so-called multi-reference methods that are required. These scale exponentially with the size of the problem (effectively, the number of electrons) at the limit of sufficient accuracy.

On the other hand, the electronic structure problem is naturally suited for quantum computers Cao et al. [2019]. Quantum chemical simulation is one of the major areas of research and development on all platforms of quantum computing, from quantum annealers, via quantum simulators, to general-purpose quantum computers. Quantum phase estimation (QPE) can provide sub-exponential solutions to electronic structure problems, but requires fault-tolerant quantum (FTQ) computers, due to massive circuit depth requirements. For NISQ devices, several iterative, variational algorithms have been devised, and progress is rapid. The importance of error-mitigation is recognised Glos et al. [2022]; Cai et al. [2023]; Kim et al. [2023], which gives hope for notable quantum advantage in the simulation of highly correlated electronic structure problems already before FTQ computing becomes a reality.

Having fast, highly accurate methods for simulating atmospheric chemistry is crucial, as the number of possible reaction pathways also grows rapidly with the size of the molecules involved in the reactions. With present-day supercomputers and modern classical quantum-chemical approximations, it is feasible to model a limited set of possible reaction pathways for smaller individual molecules. This is sufficient for demonstrating the importance of including chemical reactions and their interactions with the rest of the climate system in the models. In order to be of predictive accuracy and for decreasing the related uncertainties, massive improvements in modelling methods and capacity is, however, needed. Here, quantum algorithms and quantum computers can play a decisive role.

On-the-fly calculations of chemical reactions in the atmosphere within the climate models will remain out of reach for a long time, even with powerful quantum computers of the future. The run-time of individual

quantum computing subroutines, including the necessary pre- and post-processing of data, will be significant, especially for cases where any quantum-advantage can be expected.

For climate modelling, where the average wall-time spent on each time step in the state evolution needs to be sufficiently short (below 1 second), atmospheric chemistry can be included in a parameterised manner instead. For creating accurate ML approaches for atmospheric chemistry, huge amounts of training data is required, and this can often only be obtained by performing quantum chemical calculations Kubečka et al. [2023]. Then, the advantage of quantum computers would come from enabling unprecedented accuracy for the data that forms the basis for the models.

10 Combining High-Performance Computing and Quantum Computing: HPC+QC

There are presently major ongoing efforts around the globe for connecting HPC infrastructure with quantum computers. This is perhaps even somewhat surprising, considering that quantum computers presently cannot solve any useful real-world modelling problem more efficiently than a single node of a supercomputer. At the same time, it is testament to the potential, and the *belief* in the potential of quantum computing for scientific modelling.

In Europe, the plans for making quantum computing relevant for research and development in academia and industry alike have been outlined, with the goal of having a European quantum computing infrastructure exhibiting quantum advantage by 2030. The first quantum simulators are already being integrated with HPC infrastructure in the HPCQS project [https://www.hpcqs.eu/]. In June 2023, the EuroHPC Joint Undertaking has signed hosting agreements for six different quantum computers to be placed in HPC centres around Europe, with the plan to make these available to European users in 2024. These first quantum computers are only the beginning, several updates and new procurements are already planned.

The actual (future) HPC infrastructure and its implementation needs to be accounted for. Already in the near-term, it is expected that individual supercomputers will be connected to several quantum machines of various types and implementations Johansson et al. [2021]. The initial setups, with individual QPUs distributed throughout the continent, connected to an HPC system, can be seen as precursors to a future where QPUs will be connected in parallel, either entangled or not. Plans for even tighter, on-chip integration of QPUs with classical processing units already exist, and may well be the way for reaching fault-tolerant quantum computing. With this in mind, more emphasis on developing parallel quantum algorithms, which simultaneously utilise several QPUs, in an HPC+nQC manner, would seem appropriate. For time-evolution problems like climate modelling, this can be a necessary development at a relatively early stage, in order to enable the quantum processing part to keep up with the classical computing tasks at each time step.

Reassuringly, the importance of investing in software development for hybrid HPC+QC applications has been recognised. These developments complement the efforts for developing purely classical software for exascale supercomputers and beyond, exemplified by the Destination Earth initiative.

Here, it is apt to note that there is a need for significant classical software development alongside the quantum algorithm research. Presently, pre- and post-processing tasks take up a significant portion of the total wall-time of executing a quantum algorithm. As an example, the recent experiment on spin dynamics using IBM's 127 qubit QPU, the actual time spent on the QPU was 5 minutes, while the wall-time of the experiment was a hundred times longer, over 9 hours Kim et al. [2023]. These overheads will decrease in the future, but at the same time, increasing qubit count will again increase the complexity of pre- and post-processing. Part of this overhead lies within the domain of hardware development, e.g., qubit reset and readout. Much of this is, however, classical computing routines, such as compiling, transpiling, qubit routing optimisation, error mitigation, noise cancelling, to name a few. All of these will become computationally more demanding with increasing qubit count, and will therefore require increasing amounts of classical computing power. Thus,

efficiently operating the quantum machines of the future will require an HPC infrastructure in itself, as well as the classical software to run on it.

For reaching quantum advantage as soon as possible, both in general and especially within climate modelling, it is important to develop quantum algorithms keeping the immense, existing classical supercomputing power in mind. This means for example taking full advantage of the available HPC infrastructure for performing the necessary pre- and post-processing of data to and from the quantum machines. For electronic structure problems, as in the case of modelling atmospheric reactions discussed above, HPC resources are needed for providing an initial guess for the quantum computer; in other words, provide the best approximation to the true electronic structure that classical methods can provide, and refine it further on the quantum computer. This exemplifies the need for a broad, multidisciplinary approach to quantum advantage. We need to combine expertise in quantum algorithms, classical HPC algorithms, computer science, AI/ML, and specific domain expertise, also from the end-user side.

11 Sizing Quantum Machines for Climate Modelling

As discussed above, there are several means to achieve quantum advantage for the climate digital twins. Different problems are suitable for different quantum machines and implementations. Efficient quantum solutions for linear systems of equations and computational fluid dynamics have mostly been proposed for gate-based quantum computers. (Hybrid) quantum algorithms for machine learning and quantum chemistry have been proposed for all three major quantum machines classes, quantum annealers, quantum simulations, and gate-based quantum computers. Thus, for climate digital twins, all three classes can potentially be useful. It is important not to focus efforts too narrowly, say, only on gate-based algorithm development. Various approaches should be explored for the different quantum machines, and also combination of approaches, like digital-analog quantum computing and simulation.

We want to emphasise the difficulty in predicting developments of both quantum hardware and software, quantum algorithms. The expectations on hardware are reasonable, but naturally come with large error bars. Still, the progress can be expected to be rather smooth. A much larger uncertainty still comes from the quantum software side, as new discoveries can truly revolutionise the utility of quantum machines. It is completely possible that a novel "Shor's algorithm" for climate modelling will be invented, or that new algorithmic breakthroughs for relaxing the requirements on, say, coherence times for the qubits will be developed. In the case of software, the progress *can* be smooth, but the possibility for (quantum) leaps in efficiency is ever-present.

11.1 Present Day

We are still at the development phase of what is to become a mature supercomputing infrastructure incorporating quantum machines for climate modelling. The quantum hardware is becoming sufficiently stable for performing real-world testing at small scale. Quantum machines are being incorporated with HPC infrastructure in various manners, from cloud access models, via co-located installations, to truly distributed approaches Johansson et al. [2021]. Standards are under development, but not yet in place, for the various components of the full HPC+QC software stack. Different programming models are still developed in parallel. This is a necessary step in the evolution of a fully mature quantum computing infrastructure: we need to try out several different approaches, even through blunt trial-and-error, in order to learn which methods, protocols, and models work best. We have to resist the temptation of unifying different models at a too early stage, even if this complicates the life for hybrid HPC+QC software developers somewhat.

At the time of writing, none of the three platforms, annealers, simulators, or gate-based "universal" quantum computers provide any computational quantum advantage over purely classical methods. All of them have, however, demonstrated the potential for inching closer to the point of quantum advantage for problems

resembling real-world use cases. The D-Wave "Advantage" quantum annealer has over 5600 spin qubits. The largest available Pasqal neutral atom simulator presently has 100 qubits, with 200 planned for 2023 and 1000 for 2024. IBM's "Osprey" gate-based QPU has 433 qubits, with a thousand-qubit QPU announced still for 2023.

In recent years, several demonstrations of experiments on quantum machines, performing tasks that would be very hard for classical computers to simulate, have been reported. Keeping in mind that full simulation of sixty fully entangled qubits is already far beyond the capacity of the worlds largest supercomputers, we *should* be entering the regime of potential quantum advantage right now.

Naturally, qubit count is only one figure of merit, as discussed previously. Coherence times, operation fidelity, and qubit connectivity are at least as important; arguably, already around a fifty qubits, these already become more important than the physical qubit count. Presently, the state-of-the-art quantum computers feature roughly a 99.9% fidelity on their operations. This means that on average, one in a thousand operations fail. In order to be able to efficiently suppress errors, a minimum fidelity of 99.9999% is required Google Quantum AI [2023]. We are thus presently roughly three orders of magnitude below target on fidelity, one of the main reasons why useful quantum advantage has not yet been demonstrated.

11.2 3-5 years

Incremental advantage, or at least a convincing prospect of advantage over purely classical methods relevant to ClimateDT should by this time be exhibited. The qubit count and quality of all three classes of quantum machines should be such that reasonably reliable estimates of what types of algorithms can be expected to show true quantum advantage can be made.

The availability of quantum computers will still be so low, that for production-scale climate modelling, purely classical HPC infrastructure will be used. Benchmarking of hybrid HPC+QC methods for relevant machine learning and electronic structure tasks will be underway, paving the way for models based at least partly on data produced by quantum machines.

11.3 15 years

By this time, all surviving quantum machine classes would have sizes in the hundred-thousand to a million physical qubit regime. This implies a logical qubit count of at least a hundred, possibly thousands, and would be sufficient for executing sufficiently complex simulations or circuits for being directly relevant for climate modelling.

The manner in which the quantum machines will be utilised depends heavily on the progress in "clock speed": how fast can the Hamiltonian evolution be driven, how short are the gate-operation times, and so on. We estimate that execution speed will still be much too slow for incorporating quantum operations directly inside the workflow of the digital twin, due to the strict and short wall-time requirements imposed on each time step, which at this time would be counted in milliseconds.

It is possible that on-chip QPU technology will be implemented within 15 years, although we estimate that sufficiently mature solutions would be some time away still. With quantum and classical processing on the same chip, it may be possible to seamlessly execute short quantum-accelerated subroutines as part of a more complex calculation. This would require efficiently integrated and error-corrected/mitigated quantum coprocessing technology, in an analogous manner to floating-point units (FPUs) of standard CPUs. Tightly integrated QPUs would have a major impact on time-step evolution problems, like ClimateDT, as they would enable quantum acceleration within the individual time steps.

With a million physical qubits, or a thousand logical ones, highly accurate modelling of atmospheric reactions of small molecules exhibiting complex electronic structures will be possible, which will increase the accuracy

of the digital twins notably. We note that atmospheric chemistry is a prime candidate for early advantage from quantum-accelerated electronic structure solutions, due to the relatively small size of the molecules involved; the number of electrons that need to be described is very moderate compared to, say, enzymatic reactions. The atmospheric reaction flows can then be included in the digital twins in a parameterised manner. In the same vein, quantum-assisted machine learning will be used for training models relevant for various parts of the digital twins that now rely on machine learning for speed and/or accuracy.

12 SWOT analysis

12.1 Strengths

- The climate modelling community has a deep understanding of the problem at hand, and the bottlenecks present, both from the efficiency and accuracy points-of-view
- A recognised high-priority problem: resources available for finding solutions

12.2 Weaknesses

- Understanding of the applicability of quantum computing to climate modelling limited
- Quantum-acceleration presently not seen as a viable route, due to the "big data" nature of digital twins

12.3 Opportunities

- Progress in QC hardware and software capacity can enable more accurate models
- Global drive for supporting hybrid HPC+QC software development

12.4 Threats

- Development of sufficiently powerful QC hardware/software delayed
- Lack of long-term funding commitment to development, in case near-term gains do not live up to (inflated) expectations.

13 Conclusions

In their present form, digital twins of the climate are largely not amenable to quantum acceleration, due to their reliance on large amounts of both input and output data, and very short wall-time of the individual time steps. Despite this, quantum computers have the potential to both speed up current climate digital twins, as well as increase their accuracy. The increase in accuracy follows both from enabling higher resolution of the digital twins, and from the possibility of setting up a more complete physical model of the Earth and the dynamic processes that govern the time-evolution of the climate. The build-up towards computational quantum advantage is steady. In the medium term, within the next ten years, quantum computing can provide incremental but notable improvements to the accuracy of the models.

In the longer term, with fault-tolerant, or at least near-FT quantum machines, the gains can be of decisive importance. For computational quantum advantage in climate modelling, we estimate that this will require at least 15-20 years of further hardware development. Digital twins of the Earth and climate will thus not be among the first applications where notable quantum advantage will be found, but in time, also climate modelling will experience a quantum revolution.

Alongside the hardware, it is crucial to actively develop a diverse software ecosystem around the maturing quantum machines. The problems and subtasks that make up ClimateDT will need to be reformulated in a suitable manner in order to be amenable to quantum computing. The new quantum software will need to be seamlessly integrated with existing workflows based on classical processing. Supporting software, such as compilers, error mitigation and error-correction routines, even programming languages themselves need to be created.

Even if we would have sufficiently mature and powerful quantum machines right now, they would be practically useless for climate modelling, due to the lack of software. Just preparing the software framework for efficiently running a climate digital twin partly on quantum machines can take a decade. The Destination Earth initiative itself is about making existing classical software run efficiently on the latest supercomputers. In essence, this largely means switching the code base from CPUs to GPUs, from one classical computing platform to another. Switching from classical to quantum is a whole different level of hard. Therefore, we need to start the transition now, in order for the software to be ready when the hardware is. If and when this happens, it will boost information-based climate adaptation efforts significantly.

The software development will most likely require dedicated non-commercially motivated funding. Long-term commitment is needed, and the problem to be solved, creating a highly reliable and efficient long-term climate model, has little direct economic impact; much of the missing work is still fundamentally basic research. The indirect impact on economy and society as a whole is, of course, immense.

It is crucial to recognise the massive computational power of already existing high-performance computing (HPC) infrastructure. Direct replacement of classical methods by analogous quantum methods will not bring significant speedup. We cannot just recompile existing classical subroutines in the ESMs to run on quantum machines and expect any advantage. Instead, quantum computing needs to augment and enhance present modelling procedures. By approaching the actual problem from new angles, presently used classical subtasks and approximations can also be rendered obsolete. Quantum computing is *different*, and therein lies its strength. By solving old problems in a new way, or enabling solutions to previously intractable problems, quantum machines can accelerate HPC infrastructure in a meaningful way, complementing binary supercomputers.

Here, we have identified two main approaches for gaining quantum advantage for the climate DTs: quantum-enhanced machine learning approaches and the extension of the models to include presently missing physics and chemistry, such as atmospheric reactions. These are naturally just two examples of latent quantum advantage. Many more are expected to be uncovered by dedicated efforts. The futures of HPC and climate digital twins are quantum-accelerated; how far ahead that future lies, depends on the combined efforts of several fields of science and technology.

Uncertainty Quantification for Remotely-Sensed Datasets

Abstract

Deep learning (DL) models are extensively used to analyze and monitor the Earth's surface due to their scalability on large-scale EO datasets and their computational efficiency compared to conventional statistical approaches such as Bayesian analysis. However, they are not capable of explaining their predictions; namely, their outputs, given large-scale datasets as input, are not trustworthy, reliable, and robust which can be measured using uncertainty quantification. In fact, DL models are often considered as uninterpretable black boxes with unknown uncertainties. In contrast, Bayesian analysis is a gold standard technique for uncertainty quantification in order to obtain trustworthy and reliable predictions generated by models fitted small- or moderate-scale datasets (observations) due to its high computational cost. Hence, DL models integrated with Bayesian analysis, that is, Bayesian Neural Networks (BNNs), are slowly gaining great interest, since they allow make their outputs interpretable together with trustworthy and reliable uncertainties. However, BNN inference on large-scale datasets persists high computational cost even on the HPC system, and commonly used methodologies to overcome this challenge are Monte Carlo Markov Chain (MCMC) and variational inference (VI) approaches. Moreover, the VI approach, returning approximate samples, can be scaled on big datasets in contrast to the exact sampling MCMC. Therefore, this study assesses and examines quantum VI paradigm for processing BNN inference on small-scale EO datasets (in our case, hyperspectral images (HSIs)) to improve the sampling power of a conventional VI method. More importantly, we estimate quantum resource required for some example quantum VI models in terms of T-gates but not the implementation of quantum VI models on small-scale HSIs.

14 Introduction

Deep Learning (DL) models are employed for distinct tasks such as recognizing informative patterns in large-scale, high-dimensional datasets (in our case, satellite images or remotely-sensed datasets) and discovering their underlying distributions; here, large-scale, high-dimensional datasets can be denoted by either $\mathcal{S} = \{y_i, \mathbf{x}_i\}_{i=1}^N$ or $\mathcal{S} = \{\mathbf{x}_i\}_{i=1}^N$ depending on whether or not a learning task is to discover underlying distributions generating datasets, where y_i is their true labels, $\mathbf{x}_i$ is their high-dimensional elements, and N refers to their large-scale size. In general, these different tasks can be divided into two categories so-called supervised and unsupervised learning paradigm Murphy [2012]; Goodfellow et al. [2016]. Supervised learning paradigm refers to DL tasks for recognizing informative patterns in big datasets $\mathcal{S} = \{y_i, \mathbf{x}_i\}_{i=1}^N$ in order to predict labels $\hat{y}_i$ with the highest probability $p(\hat{y}_i|\hat{\mathbf{x}}_i, \boldsymbol{\theta})$ given $\hat{\mathbf{x}}_i$ such that the loss function $\mathcal{L}_{\boldsymbol{\theta}}(y, \hat{y})$ between true and predicted labels optimized over trainable parameters $\boldsymbol{\theta}$ is at its minimum value, while unsupervised learning paradigm is to approximately obtain the underlying distribution $p(\mathbf{x})$ of $\mathcal{S} = \{\mathbf{x}_i\}_{i=1}^N$, where labels y_i are not provided as input, such that the approximated distribution $q(\mathbf{x}|\boldsymbol{\theta})$ is closer to the true distribution $p(\mathbf{x})$ measured by optimizing some metrics over trainable parameters $\boldsymbol{\theta}$ such as Kullback-Leibler divergence (KL-divergence). Moreover, DL models are widely used to find solutions to real-world, data-driven and model-driven problems in industry and science − even in Earth observation (EO) Cheng et al. [2020] − due to their scalability on large-scale datasets and their computational efficiency on powerful computing resources (i.e., GPU tensor cores) compared with conventional statistical methods such as intractable Bayesian analysis Pandey et al. [2022]; Willard et al. [2020]. However, DL models are often perceived as uninterpretable

black-boxes generating untrustworthy and unreliable predictions, while intractable Bayesian analysis outputs predictions with trustworthy and reliable uncertainties Gal et al. [2022]. Hence, DL models with the help of intractable Bayesian analysis are gaining great attention for designing novel learning models, that is, Bayesian Neural Networks (BNNs). In fact, BNNs combining the best of both DNN and Bayesian analysis can be scaled on large-scale datasets and computed cheaply on GPU tensor cores if an efficient sampling technique for them is available while generating predictions with trustworthy and reliable errors/uncertainty estimates at the same time Jospin et al. [2022]; Olivier et al. [2021].

In contrast to DNNs, BNNs still persist high computational cost for computing a posterior distribution $p(\theta|\mathcal{S})$, directly proportional to the product of a likelihood $p(\mathcal{S}|\theta)$ and a prior $p(\theta)$, and inversely proportional to an evidence $p(\mathcal{S})$, which is a probability distribution function integrating out parameter space Olivier et al. [2021]; Zhang et al. [2017]. Moreover, the evidence is an intractable function due to its high dimensional parameter space, and so the posterior (i.e., BNNs). To weaken this intractable BNN problem, the authors of the articles MacKay [1992]; Neal [1995]; Blei et al. [2017] proposed a machinery so-called variational inference (VI) which approximates the posterior by a tractable parametrized distribution. Another method for tackling an intractable BNN is a Monte Carlo Markov Chain (MCMC) technique Brooks et al. [2011]; Hoffman and Gelman [2011] which, however, does not scale on big datasets as a VI method does. In fact, both MCMC and VI are indispensable tools for generating samples efficiently from intractable posteriors and for quantifying parameter uncertainty for safety-critical and human-centered EO tasks regardless of their respective imperfection, i.e., MCMC does not scale well on large scale datasets but generates exact samples from a posterior, while VI scales on large scale datasets but generates approximate samples from a posterior.

The emergence of quantum algorithms for accelerating some conventional algorithms attracts engineers and scientists alike who persistently attempt to find solutions to intractable problems efficiently by inventing and designing classical algorithms An et al. [2021]; Harrow et al. [2009]. Moreover, there exists the quantum versions of MCMC and VI methods (for short, quantum MCMC and VI) which promise theoretical quantum advantage for some computational problems over their classical counterparts due to the inherent probabilistic nature of quantum machines, that is, a quantum annealer, a quantum simulator, or a universal quantum computer Montanaro [2015]; Layden et al. [2022]; Benedetti et al. [2021]. However, no quantum advantage is demonstrated for finding solutions to practically relevant problems since currently existing quantum machines, that is, noisy intermediate-scale quantum (NISQ) computers, comprise a limited number of error-prone quantum bits (qubits) ≤ 100 and quantum gates Preskill [2018a], while there is a theoretical guarantee to build fault-tolerant quantum (FTQ) computers having error-free qubits > 100 and quantum gates for demonstrating quantum advantage for real-world problems Preskill [1997].

Therefore, in this use-case study, we survey and examine theoretically a quantum/classical VI method due to its scalability on large-scale datasets as opposed to a MCMC technique from the perspective of computational complexity theoretic conjectures. More importantly, a quantum/classical VI method returns solutions to BNNs with trustworthy and reliable uncertainty estimates more efficient than their classical counterparts, while we apply BNNs to safety-critical and human-centered EO tasks specifically involving small-scale real-world datasets, i.e., EO Use-Cases using small-scale hyperspectral image datasets DLR. In addition, we provide the quantum resources required for computing BNNs on proof-of-concept-small (e.g., small-scale) and operational-size-big (e.g., large-scale) EO datasets as we critically stick to the scalability and development roadmap of quantum machines provided by industry and academia.

15 Problem Definition: Earth Observation Use-Case

Sensors on Earth observation satellites detect spectral signals reflected on natural and human-made objects on Earth's surface, and huge amounts of spectral signals in distinct wavelength ranges (Terabytes of data per day) are archived in data storage devices day and night ESA. A hyperspectral imaging satellite, e.g., an EnMAP satellite DLR, is imaging sensors mounted on a satellite for sensing spectral wavelengths in ranges

of 420 nm to 1000 nm (VNIR) and from 900 nm to 2450 nm (SWIR). Its mission is to collect hyperspectral imaging data in order to provide vital information for scientific inquiries, societal grand challenges, and key stakeholders and decision makers relating to DLR

- climate change impact and interventions,
- hazard and risk assessment,
- biodiversity and ecosystem processes, and
- land cover changes and surface processes.

DNNs for data-driven tasks require big labeled datasets (i.e., a data-driven approach), while hyperspectral images (HSIs) for e.g., hazard and risk assessment, are images with limited label information, namely, the limited availability of training (benchmark) HSIs, compared to conventional benchmark remote-sensing datasets like multispectral images Cheng et al. [2017]; Paoletti et al. [2019]. There is also the commonly known limitation that DNNs do not yield their confidence level for making high stake decisions for hazard and risk assessment. Hence, the persisting challenge is to invent and design inherently interpretable data-driven models for HSIs together with error/uncertainty estimates due to both uninterpretable black-box DNN models and (almost) lack of benchmark labeled-HSI datasets, since the answers to the above-questions are already utilized to make high stake decisions − safety-critical and human-centered EO decisions Rudin [2018].

BNNs combining both DNN and Bayesian model are widely believed to be inherently interpretable data-driven models for both small- and large-scale datasets only if the efficient sampling algorithms from them are available due to their infeasible evidence Jospin et al. [2022]; Olivier et al. [2021]. More importantly, BNNs are data-efficient models which can be trained on limited label datasets, since they provide uncertainty information in their predictions and weights. The authors of the articles Alcolea and Resano [2022]; Joshaghani et al. [2022] utilized and assessed BNNs for limited benchmark labeled-HSI datasets to generate predictions with trustworthy and reliable uncertainties as they used classical MCMC and VI techniques to generate samples from the posterior of BNNs. Both MCMC and VI sampling tools are far from perfect, and their imperfection is inspected and analyzed numerically on limited labeled-HSIs, e.g., their scalability and precision, by the authors of the article Ries et al. [2022]. Regardless of their imperfection, these sampling methods are base algorithms to invent and benchmark novel sampling methods like Generative Quantum Machine Learning Zoufal [2021], Evidential Deep Learning Sensoy et al. [2018], and Dempster–Shafer Theory of Evidence Deng [2015].

Furthermore, the authors of the articles Layden et al. [2022]; Montanaro [2015]; An et al. [2021] conducted a research on quantum MCMC and quantum-enhanced MCMC methods through the lens of theoretical computational complexity in order to speed-up the conventional MCMC algorithm, while the authors of the article Benedetti et al. [2021] focused on a quantum VI method to show quantum advantage over its classical counterpart because of classically hard-to-simulate quantum circuits such as Instantaneous Quantum Polynomial (IQP) circuits Bremner et al. [2010] and Quantum Approximate Optimization Algorithm (QAOA) sampling Farhi and Harrow [2016]. IQP circuits are quantum circuits equivalent to a so-called partition function, not efficiently simulable on conventional computers. More importantly, the impact of these quantum algorithms will be enormous for processing BNNs on limited benchmark labeled-HSI datasets for making high stake decisions − safety-critical and human-centered EO decisions when we have an access to reasonable noisy intermediate-scale and fault-tolerant quantum computers (QCs) integrated with supercomputers, high performance computing (HPC): That is, HPC+QCs for computational problems of practical significance.

This Earth Observation Use-Case (EO UC) study surveys and examines a quantum VI tool together with a hybrid approach (i.e., HPC+QCs) for the limited HSIs while assessing distinct quantum computers including a quantum annealer, a quantum simulator, or a universal quantum computer by critically sticking to their scalability and development roadmap provided by industry and academia. In addition, we provide our pseudo-algorithms for processing BNNs via the quantum VI technique on high stake problems listed above.

16 Classical Bayesian Neural Networks

Classical Bayesian Neural Networks, for short, Bayesian Neural Networks (BNNs), are referred to as stochastic Deep Neural networks (DNNs) trained using Bayesian analysis on datasets. BNNs generating probability distributions of predictions and parameters (i.e., weights) are natural data-efficient and inherently interpretable models thanks to their respective uncertainties, that is, uncertainties in predictions and weights Jospin et al. [2022]; Koller et al. [2022]. In contrast, conventional DNNs considered as uninterpretable black-box models require big labeled datasets, and even they are needed to be trained and tested on sub-datasets including training, test, and validation sets, while one does not need to divide datasets into training, test, and validation sets for training BNNs. For limited labeled datasets, this division of a dataset raises a challenge for training DNNs but not for BNNs Olivier et al. [2021]. Moreover, DNNs yield also point estimates of predictions with point weights lacking their uncertainty, i.e., lacking explainability due to the uninterpretable black-box paradigm Rudin [2018].

Furthermore, BNNs combine DNNs and Bayesian analysis in order to quantify uncertainties in their predictions and weights, since they better utilize the available dataset, either small or big datasets. Namely, they are DNN models analyzed using Bayesian analysis while their weights and predictions follow certain probability distributions. To design BNNs, we first choose an appropriate DNN model $F_{\boldsymbol{\theta}} = F_{\boldsymbol{\theta}}(\cdot)$ for a given dataset $\mathcal{S} = \{y_i, \mathbf{x}_i\}_{i=1}^N$. Secondly, its weights and predictions are needed to be defined according to some prior $p(\boldsymbol{\theta})$ and likelihood $p(\mathcal{S}|F_{\boldsymbol{\theta}})$ distributions:

$$\begin{aligned} &\boldsymbol{\theta} \sim p(\boldsymbol{\theta}) = \mathcal{N}(0, \sigma^2\mathbf{I}), \\ &p(\mathcal{S}|F_{\boldsymbol{\theta}}) = p(\mathcal{S}_y|\mathcal{S}_x, F_{\boldsymbol{\theta}}) = \mathcal{N}(\mathcal{S}_y; F_{\boldsymbol{\theta}}(\mathcal{S}_x), \sigma^2\mathbf{I}); \end{aligned} \tag{1}$$

where weights $\boldsymbol{\theta}$ are sampled from a normal distribution $\mathcal{N}(0, \sigma^2)$ with zero mean and known uncertainty σ^2. $\mathcal{S}_y$ and $\mathcal{S}_x$ denote labels $\{y_i\}_{i=1}^N$ and input data points $\{\mathbf{x}_i\}_{i=1}^N$ for BNNs, e.g., $F_{\boldsymbol{\theta}}(\mathcal{S}_x)$. We note that one can represent a prior and likelihood by any probability distribution function instead of a normal distribution. For simplicity, we utilized a normal distribution $\mathcal{N}(\cdot)$. To quantify uncertainties in predictions and weights, BNNs utilize the Bayes' theorem:

$$p(F_{\boldsymbol{\theta}}|\mathcal{S}) = \frac{p(\mathcal{S}|F_{\boldsymbol{\theta}})p(\boldsymbol{\theta})}{p(\mathcal{S})} \quad \longleftrightarrow \quad p(\boldsymbol{\theta}|\mathcal{S}) = \frac{p(\mathcal{S}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathcal{S})}, \quad \textit{given} \quad p(\mathcal{S}) = \int_{\Omega_\theta} p(\mathcal{S}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}; \tag{2}$$

here $p(\boldsymbol{\theta}|\mathcal{S})$ is the posterior, and $p(\mathcal{S})$ is the evidence integrating over parameter space Ω_θ. Finally, after computing the posterior distribution expressed by Eq. (2), we can calculate a probability to predict a label $\hat{y}$ given a test data point $\hat{\mathbf{x}}$ and dataset $\mathcal{S}$, that is, a predictive posterior:

$$p(\hat{y}|\hat{\mathbf{x}}, \mathcal{S}) = \int_{\Omega_{\boldsymbol{\theta}}} p(\hat{y}|\hat{\mathbf{x}}, \boldsymbol{\theta})p(\boldsymbol{\theta}|\mathcal{S})d\boldsymbol{\theta}. \tag{3}$$

In particular, the posterior $p(\boldsymbol{\theta}|\mathcal{S})$ gives uncertainties in weights $-$ this uncertainty is called an epistemic uncertainty, while the predictive likelihood $p(\hat{y}|\hat{\mathbf{x}}, \boldsymbol{\theta})$ yields uncertainties in predictions $-$ this uncertainty is called an aleatoric uncertainty. Therefore, the predictive posterior $p(\hat{y}|\hat{\mathbf{x}}, \mathcal{S})$ generates total uncertainties in predictions by leveraging both epistemic and aleatoric uncertainties Hüllermeier and Waegeman [2021]; Gawlikowski et al. [2022]. By convention, the epistemic uncertainty related to the random noise (randomness) in a dataset can be reduced by increasing the size of a dataset, while the aleatoric uncertainty associated with a lack of knowledge in a model $\boldsymbol{\theta}$ is an irreducible uncertainty even by increasing the size of a dataset.

The parameter space $\Omega_{\boldsymbol{\theta}}$ of modern DNNs includes several thousands to millions of tuneable parameters $\boldsymbol{\theta}$. This high dimensional space of parameters raises a challenge to integrate the evidence $p(\mathcal{S})$ as well as predictive posterior $p(\hat{y}|\hat{\mathbf{x}}, \mathcal{S})$ over Ω_θ; namely, computing the evidence and predictive posterior is an

intractable problem Arora and Barak [2009]. Thus, the posterior $p(\boldsymbol{\theta}|\mathcal{S})$ is a hard-to-compute function on conventional computers due to the intractable evidence.

In order to tackle these intractability challenges, studies proposed several different techniques including so-called variational inference (VI) MacKay [1992]; Neal [1995]; Blei et al. [2017] and Monte Carlo Markov Chain (MCMC) Brooks et al. [2011]; Hoffman and Gelman [2011]. VI is a method to approximate the posterior $p(\boldsymbol{\theta}|\mathcal{S})$ by a tractable variational distribution $q_c(\boldsymbol{\theta};\boldsymbol{\lambda})$ via some distance metric over a variational parameter $\boldsymbol{\lambda}$, and to collect samples from this tractable distribution, while MCMC returns (almost) exact samples directly from the posterior $p(\boldsymbol{\theta}|\mathcal{S})$ via like the No-U-Turn Sampler Hoffman and Gelman [2011]. Note that these techniques have their own advantages and imperfections for approximate sampling and scalability on big datasets Ries et al. [2022]. In particular, the VI method returns approximate samples and scales well on big datasets (i.e., computationally cheap), while the MCMC generates almost exact samples and poorly scales on big datasets (i.e., computationally expensive). Thus, it is of great importance to design and assess the quantum VI instead of the quantum MCMC due to its scalability on big datasets in order to make it better on approximate samples.

16.1 Classical Variational Inference

Variational inference (VI) is a machinery to approximate the posterior written in Eq. (2) by some easy-to-sample distribution $q_c(\boldsymbol{\theta};\boldsymbol{\lambda})$. To define a easy-to-sample distribution, we optimize a reverse Kullback-Leibler divergence (KL-divergence) as done for training conventional DNN models Jospin et al. [2022]:

$$\begin{aligned}&\text{argmin}_{\boldsymbol{\lambda}} KL(q_c(\boldsymbol{\theta};\boldsymbol{\lambda})||p(\boldsymbol{\theta}|\mathcal{S}))\\ &\quad= \sum_{\boldsymbol{\lambda}} q_c(\boldsymbol{\theta};\boldsymbol{\lambda}) \log\left(\frac{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}{p(\boldsymbol{\theta}|\mathcal{S})}\right) = \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log\left(\frac{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}{p(\boldsymbol{\theta}|\mathcal{S})}\right)\right],\end{aligned} \tag{4}$$

where $KL(q_c(\boldsymbol{\theta};\boldsymbol{\lambda})||p(\boldsymbol{\theta}|\mathcal{S}))$ is equal to zero or minimized if and only if $q_c(\boldsymbol{\theta};\boldsymbol{\lambda}) \approx p(\boldsymbol{\theta}|\mathcal{S})$. If we expand the above KL-divergence by using the posterior expressed in Eq. (2) and rearrange it then we have:

$$KL(q_c(\boldsymbol{\theta};\boldsymbol{\lambda})||p(\boldsymbol{\theta}|\mathcal{S})) = -\left(\mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log(p(\mathcal{S}|\boldsymbol{\theta})p(\boldsymbol{\theta}))\right] - \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log q_c(\boldsymbol{\theta};\boldsymbol{\lambda})\right]\right) + \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log p(\mathcal{S})\right]. \tag{5}$$

We easily notice that the KL-divergence is still a hard-to-optimize function, since it includes a log evidence $\log p(\mathcal{S})$ where the evidence $p(\mathcal{S})$ is an intractable distribution function. To overcome the hardness of computing the KL-divergence, we utilize the fact that $KL(q_c(\boldsymbol{\theta};\boldsymbol{\lambda})||p(\boldsymbol{\theta}|\mathcal{S})) \geq 0$, and so:

$$\mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log p(\mathcal{S})\right] \geq \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log(p(\mathcal{S}|\boldsymbol{\theta})p(\boldsymbol{\theta}))\right] - \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log q_c(\boldsymbol{\theta};\boldsymbol{\lambda})\right] \tag{6}$$

where the expression on right hand side is called an evidence lower bound (in short, ELBO):

$$\begin{aligned}ELBO(p(\mathcal{S},\boldsymbol{\theta})||q_c(\boldsymbol{\theta};\boldsymbol{\lambda})) &= \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log(p(\mathcal{S}|\boldsymbol{\theta})p(\boldsymbol{\theta}))\right] - \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log q_c(\boldsymbol{\theta};\boldsymbol{\lambda})\right]\\ &= \mathbb{E}_{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log\frac{p(\mathcal{S},\boldsymbol{\theta})}{q_c(\boldsymbol{\theta};\boldsymbol{\lambda})}\right] = \mathcal{L}(\boldsymbol{\lambda},\boldsymbol{\theta};\mathcal{S}).\end{aligned} \tag{7}$$

More importantly, the ELBO is a tractable metric function compared with the KL-divergence, and the following condition is satisfied:

$$\text{argmax}_{\boldsymbol{\lambda}}\mathcal{L}(\boldsymbol{\lambda},\boldsymbol{\theta};\mathcal{S}) = \text{argmin}_{\boldsymbol{\lambda}} KL(q_c(\boldsymbol{\theta};\boldsymbol{\lambda})||p(\boldsymbol{\theta}|\mathcal{S})). \tag{8}$$

For scaling and optimizing the ELBO on small- and large-scale datasets $\mathcal{S} = \{y_i, \mathbf{x}_i\}_{i=1}^N$, stochastic variational inference (SVI) is extensively utilized for scaling BNNs on datasets, and it is a machinery to randomly generate M mini-batches and to optimize the ELBO on those batches:

$$\mathcal{L}(\boldsymbol{\lambda}) = \frac{N}{M} \sum_{i=1}^{M} \mathcal{L}(\lambda_i, \boldsymbol{\theta}; \mathcal{X}_i), \quad \text{argmax}_{\boldsymbol{\theta},\boldsymbol{\lambda}}\, \mathcal{L}(\boldsymbol{\lambda}). \tag{9}$$

We refer to the article Blei et al. [2017] for the interested readers for detailed discussions on the SVI optimizing algorithm. In general, the SVI algorithm returns approximate samples of variational parameters $\{\lambda_i\}_{i=1}^N$ and model weights $\boldsymbol{\theta}$.

17 Quantum Bayesian Neural Networks

Quantum Bayesian Neural Networks (QBNNs) are BNNs boosted by quantum algorithms which are designed to solve efficiently some hard computational problems on quantum computers Benedetti et al. [2021]. Moreover, they promise to generate solutions to a class of computational problems much faster than conventional computing resources, and quantum computers (i.e., quantum circuits) are even able to represent classically intractable probability distributions due to their inherently probabilistic nature and non-classical correlation property, that is, quantum circuits with large entanglement Bremner et al. [2010].

There are some proposals to design QBNNs based on quantum DL techniques Allcock et al. [2020] combined with BNNs Berner et al. [2021], and computing the VI approach on a quantum computer Benedetti et al. [2021]. Namely, it is a machinery to approximate the posterior by a tractable distribution $q_c(\boldsymbol{\theta}; \boldsymbol{\lambda})$ by optimizing the ELBO measure expressed in Eq. (7). Indeed, this approximation makes BNNs scalable on large-scale datasets but generates approximate samples − not exact samples. Therefore, we propose to utilize a parametrized quantum circuit (PQC) which represents a family of probability distributions $q_Q(\boldsymbol{\theta}; \boldsymbol{\lambda})$ in order to make the VI technique better on generating good approximate samples. Some PQCs used to generate samples even are known to be not simulable on a conventional computer Bremner et al. [2010]; Farhi and Harrow [2016]. Hence, we approximate the posterior by the quantum variational distribution $q_Q(\boldsymbol{\theta}; \boldsymbol{\lambda})$ due to its representational power over the classical variational distribution $q_c(\boldsymbol{\theta}; \boldsymbol{\lambda})$ Benedetti et al. [2021]. This quantum approximation can be dubbed "quantum BNNs" or "quantum variational inference".

17.1 Quantum Variational Inference

Quantum variational inference (QVI) is referred to as representing a variational distribution $q_Q(\boldsymbol{\theta}; \boldsymbol{\lambda})$ and sampling it on quantum machines:

$$q_Q(\boldsymbol{\theta}; \boldsymbol{\lambda}) = |\langle \boldsymbol{\theta} | \psi(\boldsymbol{\lambda}, \mathbf{x}_i) \rangle|^2, \tag{10}$$

where $\mathbf{x}_i$ and $\boldsymbol{\lambda}$ denote an input data point and trainable parameters, respectively. Namely, given a quantum learning model, $\hat{\mathcal{O}}(\mathbf{x}_i) H^{\otimes n} \hat{\mathcal{O}}(\boldsymbol{\lambda})$, where data encoding quantum gate $\hat{\mathcal{O}}(\mathbf{x}_i)$ and an initial quantum state $|0\rangle^{\otimes n}$, the final quantum state is prepared in the state $|\psi(\boldsymbol{\lambda}, \mathbf{x}_i)\rangle$ by an evolution $\hat{\mathcal{O}}(\boldsymbol{\lambda}) H^{\otimes n} \hat{\mathcal{O}}(\mathbf{x}_i) |0\rangle^{\otimes n}$, and then we measure it in the basis $|\boldsymbol{\theta}\rangle$; here, $\hat{\mathcal{O}}(\boldsymbol{\lambda})$ is a PQC model, and $H^{\otimes n}$ are Hadamard gates. In fact, this basis measurement outputs $\boldsymbol{\theta}$ samples with their corresponding probabilities $q_Q(\boldsymbol{\theta}; \boldsymbol{\lambda})$ Bremner et al. [2010]; Benedetti et al. [2021].

To obtain the variational parameters $\boldsymbol{\lambda}$ of the PQC, we can optimize, e.g., the ELBO expressed in Eq. (7) while replacing its classical variational distribution $q_c(\boldsymbol{\theta}; \boldsymbol{\lambda})$ by the quantum variational distribution $q_Q(\boldsymbol{\theta}; \boldsymbol{\lambda})$:

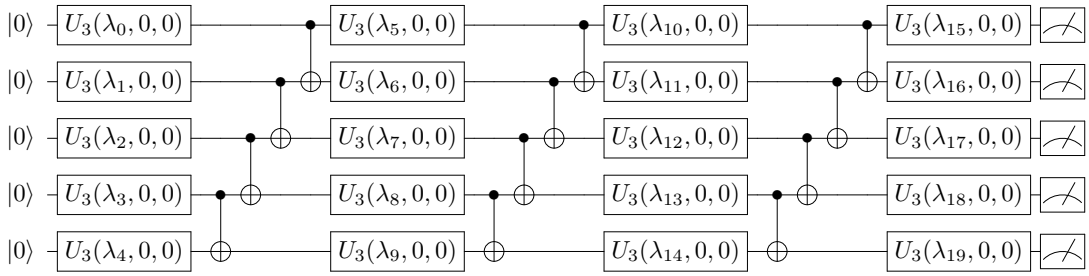


Figure 10: A real-amplitude quantum circuit having depth-one is transpiled into the Clifford+T gate set. It is used to demonstrate the power of QML models by the authros of the article Abbas et al. [2021].

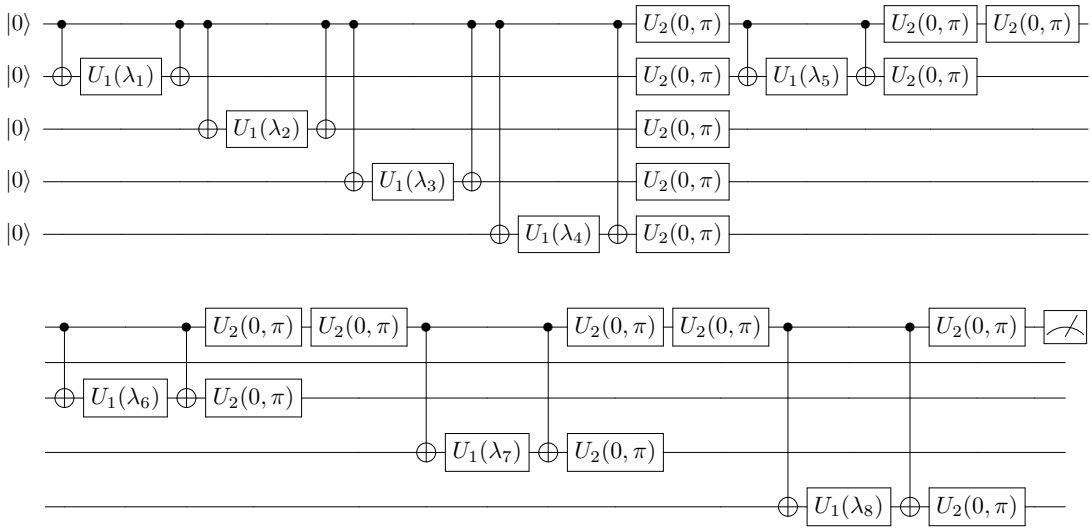


Figure 11: An energy-based quantum circuit having depth-one is transpiled into the Clifford+T gate set. This QML model is proposed for the NISQ device by the authors of the article Farhi and Neven [2018].

$$
\begin{aligned}
ELBO(p(\mathcal{S},\boldsymbol{\theta})||q_Q(\boldsymbol{\theta};\boldsymbol{\lambda})) &= \mathbb{E}_{q_Q(\boldsymbol{\theta};\boldsymbol{\lambda})}\left[\log \frac{p(\mathcal{S},\boldsymbol{\theta})}{q_Q(\boldsymbol{\theta};\boldsymbol{\lambda})}\right] \\
&= \mathcal{L}_Q(\boldsymbol{\lambda},\boldsymbol{\theta};\mathcal{S}).
\end{aligned} \tag{11}
$$

The ELBO function $\mathcal{L}_Q(\boldsymbol{\lambda},\boldsymbol{\theta};\mathcal{S})$ can be maximized on quantum computers thanks to quantum differentiable programming paradigm identical to classical differentiable programming one, namely, automatic differentiation for AI models Bergholm et al. [2022].

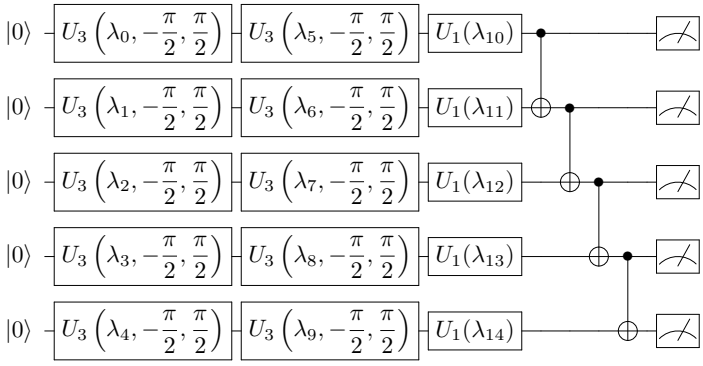


Figure 12: A strongly-entangling quantum circuit having depth-one is transpiled into the Clifford+T gate set. This QML model is proposed to build a powerful quantum learning model in the article Schuld et al. [2020].

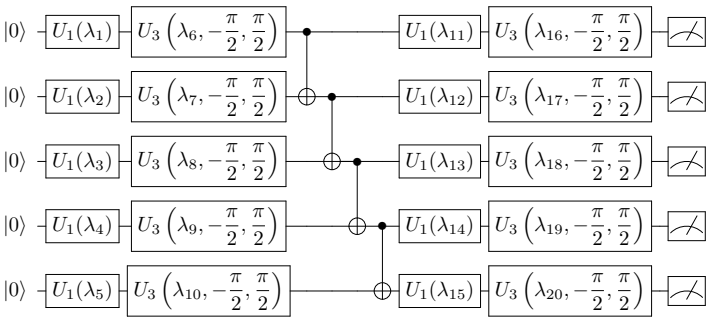


Figure 13: A hardware efficient quantum circuit having depth-one is transpiled into the Clifford+T gate set. This PQC is used for quantum variational inference in the article Benedetti et al. [2021].

18 Sizing Quantum Machines

For training PQC models on limited benchmark labeled-HSI datasets, we utilize a classical layer for reducing the dimensionality of the features of the HSI datasets due to a limited number of input qubits. However, how much one needs to reduce the dimensionality of the given HSI dataset depends on quantum computers utilizing, that is, whether we have an access to a NISQ device having error-prone qubits ≤ 100 or a fault-tolerant quantum (FTQ) computer having error-free qubits > 100. In particular, the classical machine plays a less role for pre-processing the HSI dataset, and we can feed many informative features to quantum computers (less dimensionality-reduction) as the number of the error-free qubits of quantum machines increases. In particular, we assume that we use EnMAP HSIs with 230 spectral bands and 145×145 spatial dimensions, that is, a size of the dataset. Moreover, EnMAP HSIs having $21,205$ data points and 230 features are a small-scale image dataset compared with conventional multispectral images for training DL models. To execute the PQC model on NISQ machines having ≤ 100 input qubits, we either reduce the spectral bands of the EnMAP HSIs from 230 to at most 100 or select the most of informative 100 bands to be compatible with the input qubits by utilizing a classical machine. Instead, for FTQ machines having more than 100 input

qubits, we persevere more spectral bands of EnMAP HSIs when performing the dimensionality-reduction or feature selection technique in their spectral bands by using a classical machine.

Towards quantum resource estimation, we assess four different PQC models expressed by the Clifford+T gate set (see Figures 10-13). The Clifford-T gate set is defined by U_1, U_2, U_3, and CNOT gates:

$$\begin{aligned} U_1(\lambda) &= \begin{pmatrix} 1 & 0 \\ 0 & e^{i\lambda} \end{pmatrix}, \quad U_2(\lambda, \phi) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -e^{i\phi} \\ e^{i\lambda} & e^{i(\lambda+\phi)} \end{pmatrix}, \\ U_3(\lambda, \phi, \gamma) &= \begin{pmatrix} \cos(\lambda/2) & -e^{i\gamma}\sin(\lambda/2) \\ -e^{i\phi}\sin(\lambda/2) & e^{i(\phi+\gamma)}\cos(\lambda/2) \end{pmatrix}, \end{aligned} \tag{12}$$

where, for example, $U_1(\pi/4) = T, U_1(\pi/2) = S, U_2(0, \pi) = H$. Hence, the Clifford-T gate set is $\{U_1(\pi/2)$, $U_2(0, \pi)$, CNOT, $U_1(\pi/4)\}$. Given a HPC+QC system, the four PQC models shown in the Figure 10-13 comprise several parametrized $U_1(\lambda)$ gates. We can execute them on the HPC instead of QCs, and quantum resource required for executing them on QCs is then $\mathcal{O}(1)$ (constant time) if there is either no sign of T-gates or a small number of T-gates. In particular, we deploy them on either HPC or quantum computers depending on the existence and a number of T-gates in their configuration during the training phase via stochastic variational inference (SVI) expressed by the equation (8). Furthermore, a number of T-gates defines quantum resource required for deploying QML models on NISQ and FTQ computers. To determine the number of T gates, we use the concept of symmetry breaking of conventional neural networks Fok et al. [2017]. We strongly emphasize that QML models also breaks the symmetry in their weights in order to decrease their redundant parametrized quantum gates and to generalize better on unseen data points than conventional neural networks. In particular, each weights within a quantum layer must have different digital values for capturing particular features. Hence, we assume that each layer of QML models must have at most a single T-gate at each learning iteration. Hence, our QML models having depth-one can have only one T-gate. Towards quantum resource required for executing them on digital quantum computers Fowler and Gidney [2019]:

1. If our PQCs have 10^8 T-gates and five logical qubits then we need $158,431$ physical qubits (i.e., $9,375$ state distillation qubits and $149,056$ physical qubits) on the surface code distance of $d = 25$, and our QML models then take around 5 hours.
2. If our PQCs have three T-gates and five logical qubits then we need $50,700$ physical qubits (i.e., $14,400$ state distillation qubits and $36,300$ physical qubits) on the surface code distance of $d = 11$, and our QML models then take around 8 hours.
3. If our PQCs have one T-gates and five logical qubits then we need $15,135$ physical qubits (i.e., $14,400$ state distillation qubits and 735 physical qubits) on the surface code distance of $d = 7$, and our QML models then take around 2 hours.

Quantum resource estimation demonstrates that some QML models can not be simulated on the HPC system if a number of T-gates is sufficiently high at the quantum ISA level, and otherwise, we deploy them on quantum computers Beverland et al. [2022]; Reiher et al. [2017]. In addition, we present the scaling of physical qubits and surface (code) distance with respect to the gate error rate in the Figure 14, since our PQC models require the logical error rate denoted by P_L of around 10^{-15} and the gate error p of 10^{-3} given the threshold error rate p_th of 0.57 (the green line in the Figure 14). See also the chapter 3 for the detailed discussion on the assessment and quantum development roadmap of quantum machines.

18.1 Present Day

Superconducting-based quantum machines in the current market comprise around 100 error-prone qubits and depth-5 faulty quantum gates, while quantum learning models require more than depth-5 quantum gates. Hence, quantum variational models can be only implemented as a proof-of-the-concept, when the elements in HSIs have no more than 5-10 percent overlap.

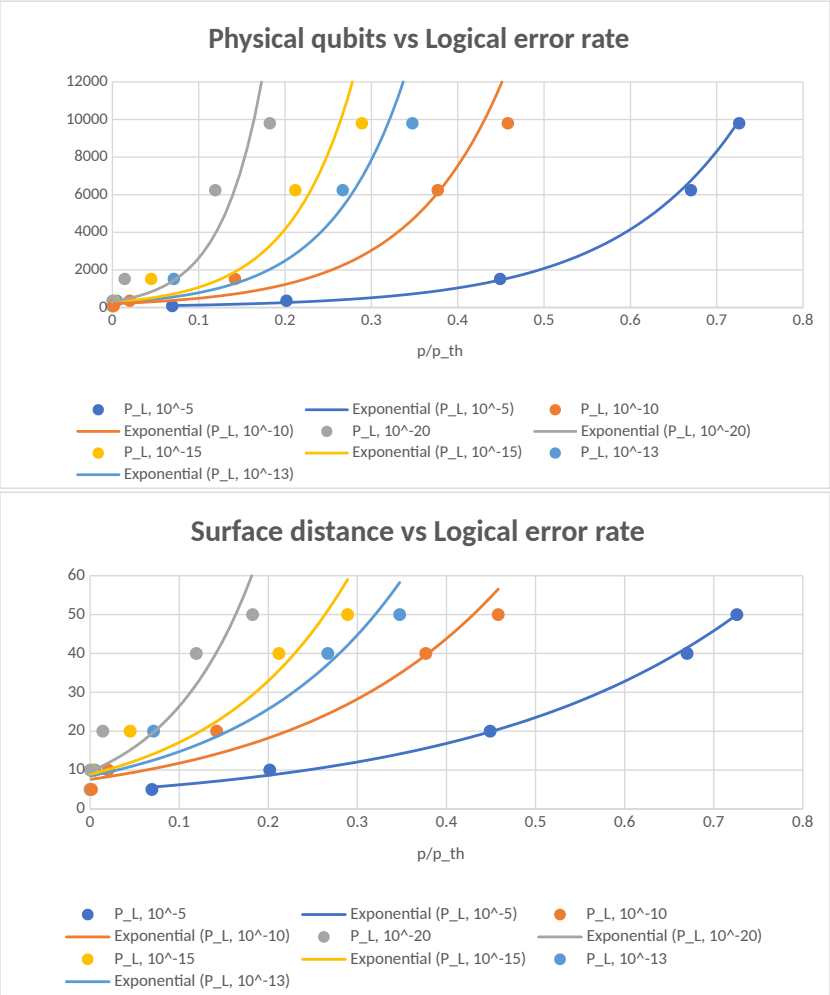


Figure 14: Quantum resource estimation for different logical error rates P_L.

18.2 3-5 years

In 3-5 years, quantum machines begin to have 100 error-prone input qubits and depth-100 faulty quantum gates. Quantum variational inference models can be executed on those quantum machines, while the elements in HSIs could be overlapping up to 10-30 percentage.

18.3 15 years

By this time, quantum machines will have around thousands of error-corrected input qubits and more than depth-100 quantum gates. Quantum variational inference models then can be implemented for operational-sized HSIs having more than 30 percentage overlap in their elements on fault-tolerant quantum machines.

19 SWOT analysis

19.1 Strengths

- Quantum machines could be applied to generate data samples from classically difficult distributions Coyle et al. [2020].
- Proved exponential speed-up in at least one scenario Liu et al. [2021].

19.2 Weaknesses

- Data loading is a major obstacle for achieving exponential speed-up of some QML algorithms Tang [2021b].
- Measurement error mitigation is limited very strongly by the number of qubits and the circuit depth. Quek et al. [2022].
- Quantum machines can be difficult to train due to the error-correction scheme Stilck França and García-Patrón [2021].

19.3 Opportunities

- Major shift in the quality of quantum computers. NISQ machines may be available with less 100 high-quality error-prone qubits in the reference time-frame of 3-5 years, and FTQ machines available with more than 100 fully error corrected qubits in the reference time-frame of 15 years.
- New applications of classical machine learning for quantum computing: compiling, mapping, control, error correction.

19.4 Threats

- Fundamental lack of ability to control, mitigate and correct sources of noise in the quantum machines.
- Novel classical algorithms inspired by quantum computing may outperform some pure quantum algorithms.

20 Conclusion

Deep Learning (DL) models are extensively applied to process big EO datasets due to the powerful computing machines like GPU tensor cores and availability of benchmark labeled-datasets. They are often considered as uninterpretable black-box models with dubious uncertainties: their outputs are not trustworthy and reliable estimates for making high stake decisions involving EO datasets. As opposed to DL models, classical Bayesian statistical approaches are inherently interpretable models generating trustworthy and reliable predictions with error/uncertainty estimates but there is the challenge that they do not scale well as the size of datasets increases or computationally expensive. This challenge can be tackled by combining the best of both DL model and Bayesian analysis, that is, Bayesian Neural Networks (BNNs); namely, DL models scale well on increasing the size of benchmark labeled-datasets, while Bayesian approaches generate the trustworthy and reliable predictions with their confidence level. However, BNNs are still computationally expensive due to their intractable posterior distributions. To weaken BNNs, variational inference (VI) paradigm approximates the intractable posterior by a tractable variational distribution function by optimizing the ELBO metric. Hence, BNNs become scalable interpretable models as the size of benchmark label-datasets increases. More importantly, they generalize well on small-scale datasets compared with DL models. There persists, however, the imperfection that the tractable variational distribution returns approximate samples for uncertainty quantification $-$ not exact samples.

Hyperspectral image (HSI) datasets obtained by hyperspectral imagery satellites are used to make safety-critical and human-centered EO decisions such as hazard and risk assessment (that is, EO Use-Case). In contrast to conventional benchmark labeled-multispectral satellite images, there is the limited availability of benchmark HSIs, namely, there is either the lack of labeled-HSI datasets or small-scale benchmark labeled-HSIs, while DL models require large-scale datasets as input. Hence, we propose to apply BNNs to small-scale benchmark labeled-HSIs for making high stake decisions, since they provide the confidence in their predictions measured by error/uncertainty estimates. In addition, to estimate their uncertainties with high precision, we utilize a quantum variational inference instead of its classical counterpart. For quantum variational inference paradigm, we replace a classical variational distribution function by a parameterized quantum circuit (PQC). According to computational complexity theoretic conjectures, PQCs can not be sampled on a conventional computer. This fact proves that quantum variational inference exhibits so-called quantum advantage over its classical counterpart. The quantum variational distribution approximates the intractable posterior better and generates more superior samples for uncertainty quantification than ones generated by the classical variational distribution $-$ closer to exact samples. In particular, the PQCs can be executed on *superconducting- and photonic-technology machines* integrated with a classical HPC workflow; HPC+QCs paradigm. The classical part selects informative features from limited labeled-HSI images and performs the dimensionality-reduction on them depending on NISQ machine (3-5 years) or FTQ machine (15 years). The larger and more error-free the qubits, the less the classical resource usage for pre-processing HSI datasets.

Quantum Algorithms for Earth Observation Image Processing

Abstract

Satellite-based Earth observations have a broad range of applications, such as natural disaster warnings, analysis of global temperature impacts, weather conditions analysis, and land-use classification. However, current machine learning techniques for land-use classification are costly in terms of time and energy. There two possible approaches to solving this problem. The first one are Variational Quantum Algorithms. They are a class of quantum algorithms that is aimed at the application in the Near Intermediate-Scale Quantum computing era. These algorithms employ jointly parametrized quantum circuits and classical optimization techniques for finding quantum circuits or states that have desirable properties from the point of a given application. VQAs find applications typically in finding low energy states of quantum Hamiltonians, solving approximately Quadratic Unconstrained Binary Optimization problems and training Quantum Neural Networks. In the area of Earth observations, the most promising area of applications lies with QNNs since the application of VQAs allows for the creation of new classification methods that employ quantum information processing tools. The second approach is to use quantum computers for a hybrid machine-learning approach utilizing an autoencoder for dimensionality reduction and a quantum algorithm powered by quantum annealer to reduce training costs. The autoencoder, using conventional deep learning techniques, is executed on GPUs, while the Deep Belief Network is run on a D-Wave quantum annealer. This hybrid approach allows for independent training of both modules, partially reducing the time and energy required to retrain the model.

21 EO use-case problem description

Distinct sensors on the satellite platforms and aircraft monitor Earth's surface day and night. They produce and transfer several terabytes of raw EO data to data storage on the ground. The stored data are only relevant when processed. Currently, deep learning becomes an indispensable tool for extracting informative information from raw EO datasets. Unfortunately, training large deep-learning neural networks is costly and consumes a significant amount of energy. Therefore, it is desirable to assess the possibility of the application of quantum computers for tasks related to processing EO data. Typical tasks related to EO data processing are mostly related to image classification or segmentation. It was shown that quantum algorithms can perform these tasks on EO data Gawron and Lewiński [2020]; Gupta et al. [2022, 2023]. In particular, Variational Quantum Algorithms are suitable and applicable in the area of EO data classification. It was shown rigorously Gyurik and Dunjko [2022] that quantum machine learning methods can have advantages over their classical counterparts. Therefore, there may exist substantial advantages for their application for EO data processing.

Satellite-based Earth observations have a broad spectrum of use cases Kansakar and Hossain [2016], Zhao et al. [2022]. This naturally leads to a wide range of potential real-life applications. Those include tasks related to various important "how to" questions such as how to

1. warn people against natural disasters (e.g. floods, fires),
2. analyse the impact of rising global temperatures on ocean levels and rate of glacier melting.

3. analyse weather conditions (e.g., for optimizing the localization of green energy installations)
4. analyze the structure of crops,
5. perform assessment of forest area in different countries and identification of areas of heavy defrostation.

Potential answers/solutions, especially those leading to sustainable development, require a constant flow of information regarding environmental changes. Currently, machine learning techniques for land-use classification of the Earth's surface often employ (deep) neural networks. While those networks are very efficient in a variety of human-like tasks, they are very costly to train (in terms of time and energy). This naturally calls for an investigation to what extent it is possible to offload the training process to a quantum computer to reduce the training costs.

22 Vartiational quantum algorithms

22.1 Introduction

Variational quantum algorithms (VQA) are a class of quantum-classical heuristics that employ both classical computers and quantum computers to perform optimization of a function computed using a quantum computer. The main idea behind VQA is as follows. The algorithm designer chooses a class of parametrized quantum circuits that can be executed on a quantum computer and an observable that can be measured on this computer. The goal of computation is to find such parameters of the quantum circuits that generate the state for which the expectation value of this observable is optimized. The optimization procedure is performed by iterative varying of the parameters, executing the quantum circuit, and measuring the expectation value of the observable. This procedure is controlled by the classical computer and repeated until a given stop criterion is reached.

Variational quantum algorithms can be applied to:

- solving (approximately) Quadratic Unconstrained Binary Optimization (QUBO) Farhi and Harrow [2016] problems using Quantum Adiabatic Optimization Algorithm (QAOA) Farhi and Harrow [2016],
- minimizing the energy of quantum Hamiltonian — Variational Quantum Eigensolver (VQE) Peruzzo et al. [2014], and
- training Quantum Neural Networks (QNNs).

22.2 Technical description

The mathematical formulation of a VQA is the following. Given an initial state $|\psi\rangle \in \mathbb{C}^n$, parametrized quantum circuit $U(\theta) \in \mathbb{C}^{n\times n}$, an observable $O \in \mathbb{C}^{n\times n}$ the goal is to find such a set of parameters θ that minimizes the expectation value of observable O in the state $U(\theta)|\psi\rangle$ *i.e.* $\langle O\rangle_{U(\theta)|\psi\rangle} = \langle\psi| U^\dagger(\theta) O U(\theta) |\psi\rangle$. The optimization is performed using a classical computer by varying the parameters θ and minimizing the function $f(\theta) := \langle O\rangle_{U(\theta)|\psi\rangle}$.

There exist variations of the abovementioned algorithm that can be adapted to a variety of tasks.

22.2.1 Parametrized quantum circuits

A common technique to implement Variational Quantum Algorithms is to employ Parametrized Quantum Circuits (PQC). PQC $U(\theta) = (U(\theta_i))_{i=1}^N$ is a sequence of N quantum gates that depend on one or more

real-valued parameters θ_i. Once the parameters are defined the PQC becomes an ordinary quantum circuit that can be—in principle—executed on a quantum machine.

22.2.2 Variational Quantum Eigensolver

The Variational Quantum Eigensolver is the most important example of VQAs because it provides an approximate solution to the problem of finding the ground state of a quantum Hamiltonian that can be expressed in the following way. Given Hamiltonian H find its minimal eigenvalue E_{min} (minimal energy) and minimizing eigenvector $|\psi_{\text{min}}\rangle$ (minimal energy state). It is known that for any quantum state $|\psi\rangle \geq E_{\text{min}}$. Therefore, one can find an approximate value of E_{min} by minimizing $\langle H \rangle_{U(\theta)|0\rangle}$ by varying parameters θ of a PQC $U(\theta)$. To perform this operation on a quantum the Hamiltonian H has to be decomposed into a linear combination $H = \sum_\alpha h_\alpha P_\alpha$ of Pauli strings $P_\alpha = \sigma_1^{\alpha_1} \otimes \sigma_2^{\alpha_2} \otimes \ldots \otimes \sigma_N^{\alpha_N}$ with $\sigma_i^j \in \{\mathbb{1}_i, \sigma_i^x, \sigma_i^y, \sigma_i^z\}$, so that $\langle H \rangle_{U(\theta)|0\rangle} = \sum_\alpha h_\alpha \langle P_\alpha \rangle_{U(\theta)|0\rangle}$. The quantum circuit $U(\theta)$ should ideally dependent on a few parameters θ, be able to explore the Hilbert space of the quantum system, and be efficiently implementable on the quantum computer.

22.2.3 Quantum Neural Networks

Quantum Neural Networks (QNNs) are a class of machine-learning models that can be evaluated on a quantum computer. QNNs, similarly to classical neural networks are parametrized functions that can be trained using data to perform common machine learning tasks such as classification, regression, sampling from a complex probability distribution or generating new data. QNNs can be composed—in the mathematical sense—with classical neural networks forming hybrid quantum-classical NNs and jointly trained using backpropagation.

The visual representation of a quantum variational algorithm that employs both data x and parameters θ is presented in Fig. 15. In the figure, the quantum computer is driven by the classical computer that is responsible for transferring the data x and parameters θ to the controller of the quantum computer that uses these pieces of information to generate quantum circuits $U_{\text{load}}^{(k)}(x)$ and $U_{\text{var}}^{(k)}(\theta)$ that encode data and model respectively. After those circuits are executed measurements of quantum observables O_i is performed and the outcomes of the measurements $f_i(x, \theta)$ are returned to the classical computer and combined jointly using—possibly parametrized—function $f_{\text{classical}}(f_1(x, \theta_1), \ldots, f_I(x, \theta_I), \theta_{\text{classical}})$.

A typical implementation of the forward part of the training of a quantum neural network is presented in Algorithm 1.

Algorithm 1 Forward algorithm for a quantum neural network.

```
procedure FORWARD(x, θ)
    for k = [K] do                                   ▷ Repeat for each QNN layer
        APPEND(Q_tape, U_var^(k)(θ))                 ▷ Append the variational circuit to the tape
        APPEND(Q_tape, U_load^(k)(x))                ▷ Append the data loading circuit to the tape
    APPEND(Q_tape, O)                                ▷ Append observable O to the tape
    f(x, θ) ← QCRUN(Q_tape)                          ▷ Execute quantum tape and measure results
    return f(x, θ)
```

Data encoding Classical data can be encoded on the computational basis of quantum states—as binary strings, amplitudes of the quantum states or observables.

Quantum neural network architectures Quantum neural networks can be implemented in a variety of architectures. The simplest case is when the quantum evolution is completely unitary e.g. Figure 16b presents

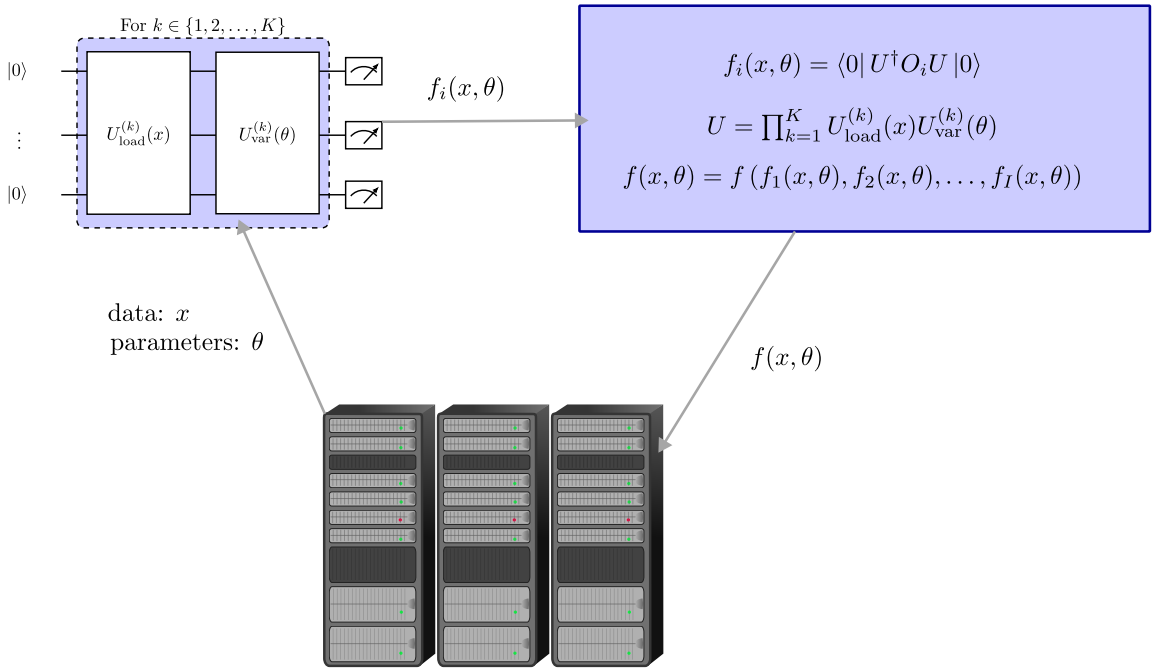
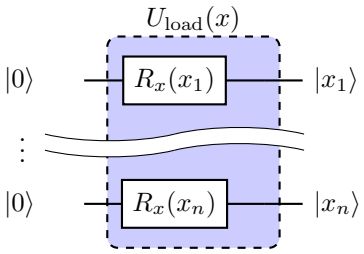
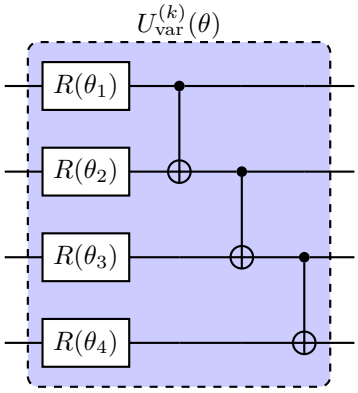


Figure 15: A simple depiction of QNN main quantum loop.



(a) QC for loading of classical data to a quantum computer.



(b) An example of a quantum neural network layer.

an example of a simple layer of a quantum neural network with general parametrized qubit rotations being controlled by the model parameters.

Other architectures can introduce mid-evolution measurements and classically controlled gates or extending the number of used qubits in subsequent layers of QNN.

22.2.4 Quantum Adiabatic Optimization Algorithm

Quantum Adiabatic Optimization Algorithm QAOA uses the VQA principle to solve, possibly approximately Quadratic Unconstrained Optimization (QUBO) problems. The algorithm does it by simulating the adiabatic quantum computing process.

There exist several issues related to each of the steps of the above algorithm. The architecture of a quantum computer and the structure of a particular problem have to be taken into account while designing and executing a VQA. Additionally, it is important to take into the account the interplay between the classical computing systems: storage, information transfer and compute units with the quantum computer.

22.3 Sizing quantum machines for VQAs

Authors of an in-depth overview Bharti et al. [2022] provide an overview of the current state of quantum computers concerning implementations of VQAs as well as an outlook for the future. They divide the future into two main eras: one of near intermediate-scale quantum computer NISQ and one of the fully error-corrected quantum (FEC) computers. Unfortunately, they do not attempt to provide any concrete timeline for the possible future development of VQAs. Since the efficiency of variational quantum algorithms depends on multiple factors, such as:

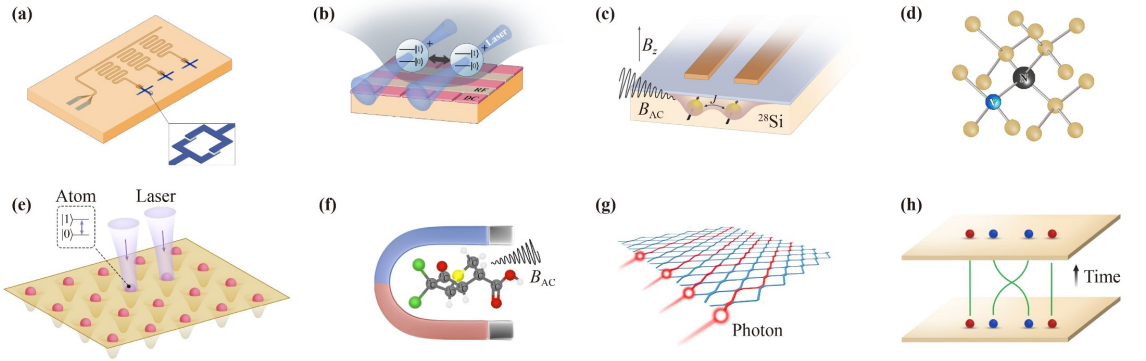
- number of qubits,
- qubits connectivity,
- single-qubit, two-qubit or multi-qubit gate fidelities,
- measurement errors,
- quantum system coherence time,
- execution time of operations reset, gate, and measurement,
- scalability of the quantum computing hardware platform,
- precision of control pulses,
- possibility to perform mid-quantum computing measurement and classical computing,
- classical optimization method,
- ansatze,

the following sizing assessment is an educated guess about the timeline for future VQAs applicability to real-life problems related to EOs.

We can use the method for defining practical quantum advantage and application readiness levels (ARLs) as presented recently by Herrmann et al. [2023]. The authors define quantum advantage using the notions of quantum utility or quantum dominance, where the former notion requires that a quantum (possibly hybrid) system "(i) requires less computing time, or (ii) requires less power, or (iii) yields more accurate results [...] to the best classical device of similar size, weight, and cost.", and the latter notion requires that points (i)–(iii) are "compared to any other classical device".

The authors of Herrmann et al. [2023] define five levels of application readiness levels:

- ARL 1: concept with unknown potential,
- ARL 2: beneficial in small idealized systems,
- ARL 3: utility indicated by theory or resource estimations,
- ARL 4: simulated utility demonstration,
- ARL 5: utility demonstration on quantum hardware.



(a) Superconducting quantum processor:
An array of artificial atoms made of Josephson junctions and capacitors, fabricated with lithography, controlled with microwave electronics
Number of qubits with individual control: ~100
Gate fidelities:
Systems with 53 qubits: $F_1 > 99\%$; $F_2 > 99\%$
Pros:
Fast gate speed, tailored qubits, high controllability and scalability
Cons:
Crosstalk between qubits, low temperature, supporting technology for scaling up

(b) Trapped ion quantum computing:
Laser-cooled atomic ions held by radio-frequency electric fields in ultra-high vacuum environment
Number of qubits with individual control: 20−30
Gate fidelities:
Systems with dozens of ions
$F_1 > 99.9\%$; $F_2 > 99\%$
Pros:
Extremely long coherence time and excellent gate operations on few ions;
Reconfigurable connectivity between ion qubits
Cons:
Technically challenging in integration

(c) Semiconductor spin-based quantum computing:
Electron or hole spins in semiconductor (silicon) quantum dot
Number of qubits with individual control: 6
Gate fidelities:
Donor spin qubit: $F_1 \sim 99.99\%$; $F_2 \sim 99.5\%$
Gate-defined qubit: $F_1 \sim 99.9\%$; $F_2 \sim 99.51\%$
Pros:
Semiconductor fabrication
Long coherence and fast high-fidelity gate
Work at temperature > 1 K
Small footprint
Cons:
Challenge of nanoscale fabrication

(d) NV center quantum computing:
Point defects in diamond; electron and nuclear spins with long coherence time; atom-like properties and solid-state host environment
Number of qubits with individual control: 10 (one electron + nine ^{13}C nuclear spins)
Gate fidelities:
$F_1 \sim 99.995\%$; $F_2 \sim 99.2\%$
Pros:
Work at room temperature
Excellent quantum sensor
Handy in quantum network
Cons:
Hard to scale up

(e) Neutral atom array quantum computing:
Neutral atom arrays trapped in optical tweezers with controlled interactions based on the Rydberg interactions
Number of qubits in digital quantum processors: 24
Number of qubits in analog quantum simulators: 289
Pros:
Both digital and analog quantum simulations
Scaling up beyond 100 qubits in programmable geometries
Cons:
Improving fidelities of two-qubit gates

(f) NMR quantum computing:
Nuclear spins in molecule
Number of qubits with individual control: 12
Gate fidelities:
$F_1 \sim 99.98\%$; $F_2 \sim 99.3\%$
Pros:
Work at room temperature
Long coherence time
Digital quantum simulator
Cons:
Hard to scale up

(g) Photonic quantum computing:
Coherently manipulating a large number of single photons or canonically conjugate pair of variables for electromagnetic modes to process quantum information
Number of qubits with individual control: 18
Number of photons with coherent control: 255 (Jiuzhang 3.0 (to be published))
Gate fidelities:
$F_1 \sim 99.84\%$; $F_2 \sim 99.69\%$
Pros: robust against decoherence, working at room temperature, compatible with CMOS fabrications, natural interface for distributed quantum computing
Cons: Big challenge to realize deterministic photon-photon gate

(h) Topological quantum computing:
Fault-tolerant quantum computation based on non-abelian braiding of anyons
Number of qubits with individual control: N/A
Gate fidelities:
N/A
Pros:
Intrinsic topological protection
Few physical qubits to construct a logic qubit
Promising to achieve large-scale, error-corrected computation
Cons:
The ideal materials or systems not found yet.
Zero topological qubit so far

Figure 17: Reproduction of Fig. 2 from Cheng et al. [2023b] presenting a selection of quantum computing hardware. (CC-BY 4.0)

22.3.1 Currently

Currently-existing quantum computers belong to the NISQ era and can not claim any applicable advantage over classical computers. Those are experimental devices testing the limits of the existing technology. Decoherence, limited qubit connectivity, gate errors and measurement errors. The training of a QNN on a quantum computer is currently very difficult and not efficient.

We have currently reached ARL level of one for most VQAs applications. There exists several proposals of applying VQAs for remote sensing data processing such as e.g. Gawron and Lewiński [2020]; Gupta et al. [2022, 2023]; Nalepa et al. [2022]; Miroszewski et al. [2023]; Otgonbaatar and Datcu [2021b,c], and also for EO mission planning Rainjonneau et al. [2023].

22.3.2 3–5 years

It is unlikely that any neither useful implementation of a Variational Quantum Eigensolver or Quantum Machine Learning algorithm will be demonstrated. The quality of the quantum computers will be too low. But we can observe steady progress in the quality of quantum hardware, the development of new practical algorithmic ideas, and the development of quantum software stack.

After the mark of five years, we should be able to achieve ARL of two or three for at least a couple of VQAs use cases if major efforts are put into R&D activities.

22.3.3 15 years

While it is very difficult to predict the future of disruptive technology, such as quantum computing, 15 years ahead one can hope for the existence of fully error-corrected quantum (FEC) computers with hundreds of logical qubits. Such computers would be able to tackle machine learning problems that are impossible to be solved today. Especially if supplied with coherent quantum information e.g. acquired from quantum sensors.

After the mark of 15 years, we can be hopeful to show ARL four or five for at least one or two use cases.

22.4 SWOT analysis

22.4.1 Strengths

- Quantum Neural Networks have larger effective dimensions than Deep learning models Abbas et al. [2021].
- Quantum kernel methods can provide better classification results by transforming quantum encoded features using projected quantum kernels than classical kernel methods Huang et al. [2021].
- Quantum Born machines could be applied to generate data samples from classically difficult distributions Coyle et al. [2020].
- Proved exponential speed-up in at least one scenario Liu et al. [2021].

22.4.2 Weaknesses

- Data loading is a major obstacle for achieving exponential speed-up of some QML algorithms Tang [2021b].
- Limited number of samples obtained from quantum devices leads to measure concentration Thanasilp et al. [2022] and difficulty to train quantum kernel methods.
- Measurement error mitigation is limited very strongly by the number of qubits and the circuit depth. Quek et al. [2022].
- VQAs can be difficult to train due to barren plateaus McClean et al. [2018].
- Noisy quantum devices have major limitations Stilck França and García-Patrón [2021].

22.4.3 Opportunities

- Major shift in the quality of quantum computers. Fully-error corrected quantum computer available with ≈ 100 fully error qubits.
- New applications of classical machine learning for quantum computing: compiling, mapping, control, error correction.

22.4.4 Threats

- Fundamental lack of ability to control, mitigate and correct sources of noise in the QCs.
- Unlikely collapse of the complexity hierarchy will likely lead to a lack of quantum advantage.
- Potential new "no-go theorems".

23 Hybrid approach

23.1 Introduction

We propose an autoencoder for the dimensionality reduction of input EO images, and a quantum algorithm powered by the quantum annealer (quantum machines) to reduce the training costs. Ideally, in such a hybrid machine-learning approach, one would want to combine at least two distinct modules. The first one is precisely an autoencoder, with latent binary representation, that essentially prepares data to be used by the (quantum) annealer. The second is a Deep Belief Network, a stack of Restricted Boltzmann Machines, which is used for classification purposes Hua et al. [2015], Dixit et al. [2020], Dixit et al. [2021]. The crucial part related to the training of every neural network is an update of all connections between neurons. This process involves calculating many average values of specific functions (which are problem-dependent). The computation can be accelerated if one can sample from a particular distribution, which again is problem-dependent. Then, and only then, all complicated expectations values involving that distribution simplify to weighted sums. This is precisely what a quantum annealer allows one to do. Independent samples can be drawn from the Boltzmann distribution quickly (even in microseconds), accelerating the training stage. However, to take advantage of such capabilities, one needs to (re)formulate the original classification problem using the Ising Hamiltonian (a model of interacting spin-1/2 particles). How to perform such a mapping effectively is an open problem, being part of this study. Interestingly, both modules of such hybrid architecture can be trained independently. The autoencoder, which uses conventional deep learning techniques, can be executed on the GPUs, and the second one — deep belief network — on the D-Wave quantum annealer. This separation allows us to partially reduce the amount of time and energy needed to retrain the model.

23.2 Technical description

As a proof-of-concept of our approach, we also provide pretrained models for a selected set of data. The solution consist of a machine learning system. The client will be able to request the training of its model on the Sentinel-2 multispectral data they select and demand the land-use labels for a particular set of land patches. The solution uses the D-Wave quantum annealer during the training process. The annealer is used in the most difficult part of the machine learning pipeline. Namely, in generation of the multispectral data representation and multilabel classification. We aim to use a two-stage data transformation process. In the first stage, the data will be transformed from its natural representation into a binary sequence, which is more natural for quantum machines. This binary string should encode most of the relevant information about a hyperspectral data cube patch.

The processing pipeline of the presented approach is as follows (see Figure 18).

First, the multispectral data is compressed into a binary representation using an autoencoder schematically presented in Figure 19. We considered two autoencoders, like latent Bernoulli autoencoder (LBAE) and Binary variational autoencoder (BVAE). The LBAE model is used as an example in the compiled version of our solution. As such, it will be discussed in greater detail in the following.

Then the compressed data are used to train a restricted Boltzmann machine (RBM). We consider three possible training backends for this process:

- Contrastive divergence (run on a classical computer).

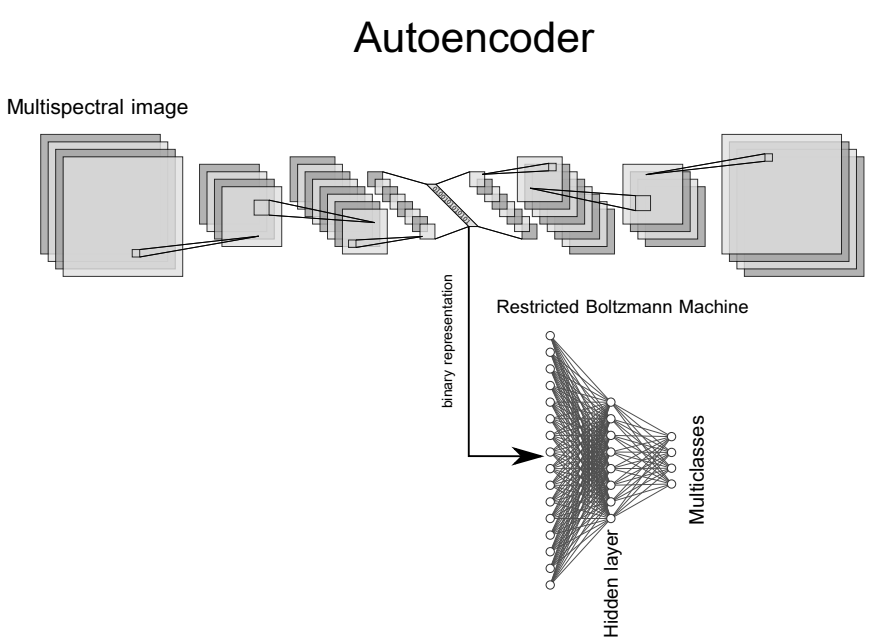


Figure 18: Processing pipeline of hybrid machine-learning approach based on autoencoder and RBM.

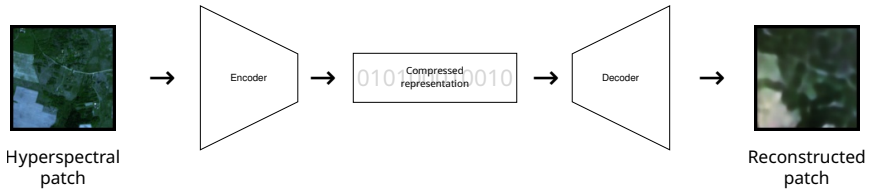


Figure 19: Processing pipeline of the LBAE training.

- Coherent Ising machine (CIM). This approach takes as an input an Ising optimization problem and solves it using a system of dynamical equations resulting in a simulation of the behavior of an actual quantum annealer. A general overview of CIM is presented in the following subsections.
- Quantum annealing backend. Currently, this backend uses the D-Wave annealer. A short note on utilization of quantum annealers is presented in the following subsections.

In the next step of the weights obtained form the training of the RBM are used as weights for layers of a neural network.

Lastly, we train the final layer of our network in a supervised manner. All other weights remain fixed. This is the only supervised part of our training pipelie.

23.2.1 Latent Bernoulli autoencoder

The need for differentiability of each layer represents a challenge if one desires to train stochastic neurons or other non-differentiable functions such as quantization Fajtl et al. [2020]. Sampling from and interpolating in the discrete latent space is equally challenging. Unlike multimodal, Gaussian and many other real-valued distributions, the multivariate Bernoulli distribution concentrates most of the information on the second and

higher moments, since the marginals are strictly unimodal and entirely described by the mean. Given that this model learns a distribution with unknown prior, and based on the aforementioned premise, the model parametrizes the learned distribution by its first two moments. The main advantage of the model is the fact that there is no sampling of pseudo-random numbers during the training step.

23.2.2 Coherent Ising machine

The coherent Ising machine is an iterative algorithm for sampling low-energy spin configurations in the classical Ising model Goto et al. [2021]. It treats each spin value as a continuous variable from the range $[-1, 1]$. Each iteration begins with calculating the mean field acting on each spin by all other spins. Then the gradients for the spin values are calculated. Then the spin values are updated according to the gradients and some chosen activation function. After multiple updates, the spins will tend to either -1 or +1 and the final discrete spin configuration is obtained by taking the sign of the continuous variables. CIM has been tested on a variety of problems. Implemented on a consumer graphic processor, this algorithm runs faster and generates higher quality samples than many analogue and digital annealing processes. Typical results from these simulations are presented below (see Figure 20).

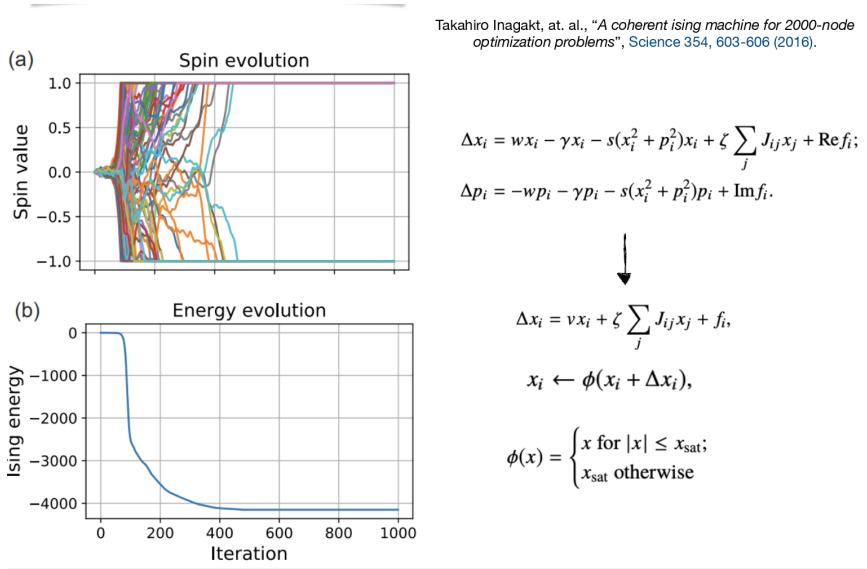


Figure 20: Results from simulations of training a restricted Boltzmann machine with training backend based on coherent Ising machine.

23.2.3 D-Wave annealer

The quantum annealer allows us to sample from the Boltzmann distribution, which is a crucial part of training an RBM Dixit et al. [2021]. With the rapid advancement of this technology (see Figure 21) this shows a great promise for acceleration of classical training.

The learning process is a hybrid of classical and quantum computation. The weights of the RBM are stored on a classical computer and are updated based on samples from the Boltzmann distribution obtained from the quantum annealer. This can be summarized as

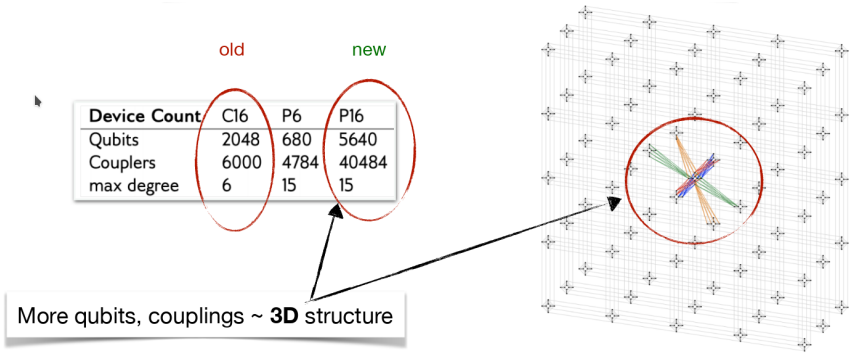


Figure 21: Figure show comparison of parameters D-wave architectures like Chimera and Pegasus.

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial J_{ij}} &= \sum_{\mathrm{S}} \left(s_i s_j - \langle s_i s_j \rangle_{\{\mathrm{S}\}} \right) \\ \frac{\partial \mathcal{L}}{\partial b_i} &= \sum_{\mathrm{S}} \left(s_i - \langle s_i \rangle_{\{\mathrm{S}\}} \right) \end{aligned} \quad \longrightarrow \quad \begin{aligned} \Delta \hat{J} &\approx \xi \left(\frac{1}{N} \sum_{n=1}^{N} \mathrm{s}_n \mathrm{s}_n^{\mathrm{T}} - \frac{1}{M} \sum_{m=1}^{M} \hat{\mathrm{s}}_m \hat{\mathrm{s}}_m^{\mathrm{T}} \right) \\ \Delta b &\approx \xi \left(\frac{1}{N} \sum_{n=1}^{N} \mathrm{s}_n - \frac{1}{M} \sum_{m=1}^{M} \hat{\mathrm{s}}_m \right). \end{aligned} \tag{13}$$

23.2.4 Hybrid approaches with simulated bifurcation machines

Hybrid quantum-classical optimization is an approach that combines classical optimization algorithms with quantum computing techniques to solve complex optimization problems more effectively. In the context of Simulated Bifurcation Machines (SBM), the idea is to use SBM as a quantum-inspired technique to complement classical optimization algorithms.

An (SBM) is a type of quantum-inspired computing technology designed to solve complex combinatorial optimization problems. It was developed by researchers at Toshiba, and it is based on the concept of bifurcation, which is a phenomenon in dynamical systems where a small change in a system's parameters can cause a sudden shift in its behaviour.

The SBM leverages a classical computer to simulate the behavior of a quantum system undergoing bifurcation. It uses this behavior to explore the solution space of the given optimization problem more efficiently than traditional classical methods. The key idea behind the SBM is to take advantage of the sudden transitions that occur in bifurcation to jump between possible solutions, allowing the algorithm to converge to an optimal or near-optimal solution quickly.

While the SBM is not a true quantum computer, it is inspired by and seeks to harness some of the benefits of quantum computing. This technology has shown promise in solving a variety of optimization problems, such as the traveling salesman problem, portfolio optimization, and drug discovery, among others. However, it is important to note that the SBM has its limitations and is not a universal solution for all optimization problems.

The technical aspects of a Simulated Bifurcation Machine (SBM) involve a combination of classical computing and concepts inspired by quantum mechanics. The underlying mechanism of SBM is based on the phenomenon of bifurcation, which is characterized by sudden changes in the behavior of dynamical systems. Here are some key technical aspects of SBM:

1. Hamiltonian dynamics: In SBM, the optimization problem is mapped to a continuous-time Hamiltonian system. A Hamiltonian function is used to describe the total energy of a system, which is the

sum of kinetic and potential energies. The Hamiltonian dynamics is used to navigate the solution space of the optimization problem Kalinin et al. [2021]; Goto [2021].

2. Simulated bifurcation algorithm: The SBM uses a simulated bifurcation algorithm that mimics the behavior of quantum systems undergoing bifurcation. The algorithm leverages the classical computer's ability to simulate these quantum-like transitions, allowing for an efficient exploration of the solution space Kalinin et al. [2021].
3. Adiabatic transitions: The SBM algorithm incorporates a concept similar to adiabatic quantum computing, where the system transitions slowly between different energy levels, staying close to the ground state. This allows the SBM to explore the solution landscape more efficiently Farhi et al. [2000].
4. Parameter tuning: In the SBM, the Hamiltonian system's parameters are carefully tuned to induce bifurcation points. These bifurcation points cause sudden transitions between different solutions, allowing the algorithm to jump from one solution to another and explore the solution space more effectively Kalinin et al. [2021].
5. Near-optimal solutions: The SBM is designed to find near-optimal solutions to combinatorial optimization problems. While it may not always find the absolute best solution, it can often find high-quality solutions in a relatively short amount of time compared to classical optimization algorithms Kalinin et al. [2021].

The Simulated Bifurcation Machine combines these technical aspects to provide an efficient, quantum-inspired approach to solving combinatorial optimization problems. However, it is important to note that the SBM has its limitations and is not a universal solution for all optimization problems Kalinin et al. [2021].

23.3 Sizing quantum machines for the hybrid approach

23.3.1 Currently

Currently-existing quantum annealers belong to the so-called NISQ era and can not claim any applicable advantage over classical computers. The current state of the art annealers has 5640 qubits and 40484 connections Systems [2023]. The device is susceptible to noise and for large families of instances, the device finds solutions far from the ground state.

23.3.2 3–5 years

D-Wave road map suggests that within 5 yeas the number of qubits will increase to around 8k and number of connection around 80k D-Wave Systems Inc. [2021]. The devices will still be prone to noise but it is likely that we will observe steady progress in the quality of the hardware. There will exist a hybrid (classical-quantum) solvers to solve large problems 1M variables. Those solvers will utilize sophisticated classical methods (simulated bifurcations) combined with quantum annealing.

23.3.3 15 years

In the forthcoming era, D-Wave quantum annealers will encompass 100,000 highly interconnected qubits, empowering them to address intricate optimization conundrums D-Wave Systems Inc. [2021]. Utilizing reversed annealing techniques will facilitate the generation of more sophisticated quantum states, while the hybrid solver methodology will amalgamate classical and quantum computing paradigms to ascertain optimal solutions. These cutting-edge developments will usher in unprecedented breakthroughs across domains such as optimization, machine learning, and logistics.

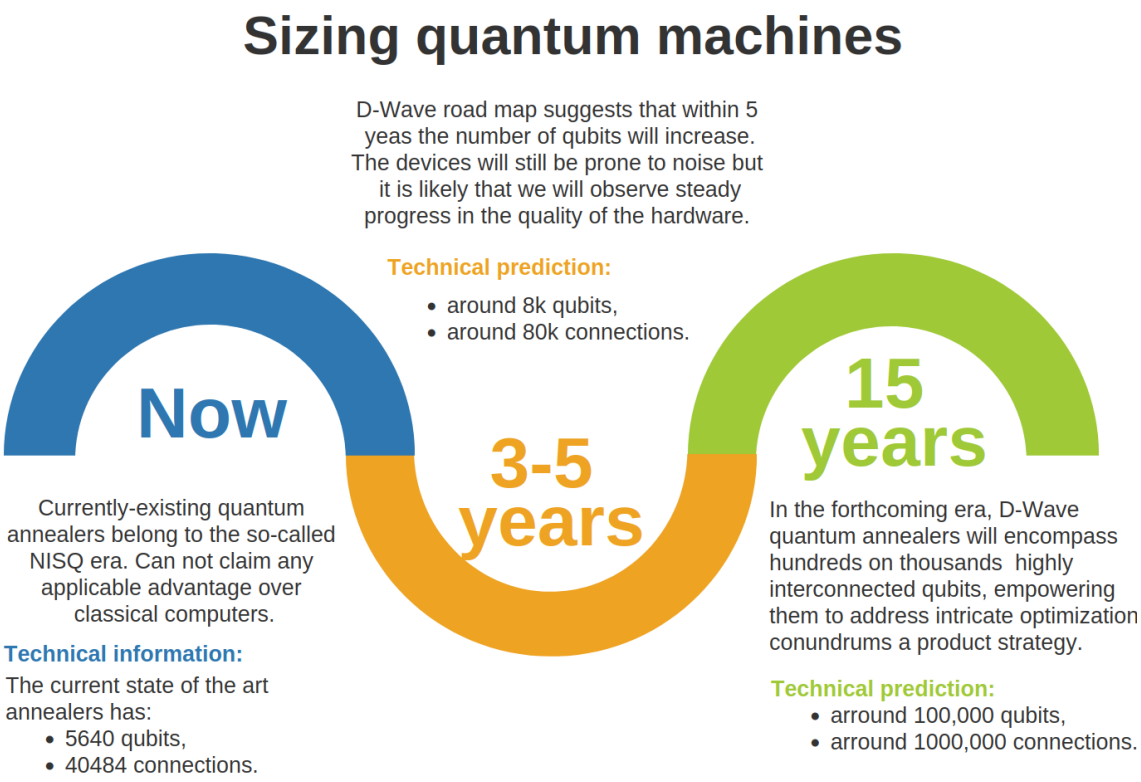


Figure 22: Figure shows sizing of quantum annealers

23.4 SWOT analysis

23.4.1 Strengths

- Quantum advantage: D-Wave annealers have the potential to process and analyze large amounts of data (utilizng hybrid approaches) significantly faster than classical computers Systems [2023].
- Hybrid solver approach: By combining classical and quantum computing methods, D-Wave's hybrid solvers can efficiently navigate the solution space and arrive at optimal or near-optimal solutions Systems [2023].
- Scalability: With 100,000 highly connected qubits, D-Wave quantum annealers can address increasingly complex Earth Land Cover Understanding problems as technology advances Systems [2023].

23.4.2 Weaknesses

- Noise sensitivity: Quantum systems, including D-Wave annealers, are susceptible to noise and errors, which may impact the accuracy of the results Preskill [2018b].
- Limited availability: D-Wave systems are not yet widely accessible, and their usage requires specialized knowledge and expertise Systems [2023].

- Problem-specific applicability: D-Wave annealers are primarily suited for optimization problems, which might limit their applicability in other aspects of Earth Land Cover Understanding Systems [2023].

23.4.3 Opportunities

- Enhanced remote sensing: D-Wave annealers can be applied to process and analyze remote sensing data, enabling more efficient and accurate land cover classification and monitoring Mallet and Bretar [2009].
- Climate change research: Quantum computing can potentially improve climate models and predictions, contributing to a better understanding of the impacts of land cover changes on the environment Biamonte et al. [2017].
- Interdisciplinary collaboration: The use of D-Wave annealers in Earth Land Cover Understanding can foster collaboration between researchers in quantum computing, remote sensing, and environmental sciences, leading to new insights and innovations.
- Hybrid approaches: Possibility to utiliza quantum-classical problems are split between a classical approach and the quantum annealer allows to tackle large problems.

23.4.4 Threats

- Competition: As more companies and researchers develop quantum computing technologies, there will be increased competition for D-Wave annealers in the Earth Land Cover Understanding domain Preskill [2018b].
- Technological obsolescence: The rapid pace of quantum computing advancements may render current D-Wave annealer technology obsolete or less competitive in the future Preskill [2018b].
- Funding constraints: The high costs associated with quantum computing research and infrastructure may limit the availability of funding for D-Wave annealer projects in Earth Land Cover Understanding.

24 Conclusions and recommendations

The only recommendation that can be provided is that further research is needed. It is important to identify bottlenecks where classical computers struggle to provide efficient solutions for the Earth Observations and try to pair them with the incoming development in the field of Quantum Machine Learning. Those bottlenecks are defined by the computation time, energy consumption and quality of obtained results. A wide variety of stakeholders such as computer scientists, machine learning experts, Earth observation experts, agricultural experts, climate scientists, the members of various communities should be involved in identifying mentioned bottlenecks and hard problems related to Earth observations, and later defining the road forward.

It would be advantageous for the research community if several projects were funded. It would allow the gathering of experts from a variety of domains and focus their work on identifying important specific problems in EO that might be solvable using quantum computing and VQAs in particular.

There exists a significant gap between the skills and knowledge of computer scientists of physicists working on quantum computing and practitioners working on e.g. climate change assessment. Therefore, common platforms for scientific discussions have to be organized in order to facilitate communication.

The proposed hybrid machine-learning approach amalgamates an autoencoder for dimensionality reduction of input Earth Observation (EO) images and a quantum algorithm powered by a quantum annealer for mitigating training costs. This innovative methodology capitalizes on the synergies between classical and

quantum computing, offering a propitious solution for proficient and accelerated training in Earth Land Cover Understanding. Notably, the independent training capabilities of the two modules facilitate a more adaptive and energy-efficient system.

Considering the potential of this hybrid architecture, it is recommended to investigate hybrid solvers proffered by multiple startups advancing cutting-edge hybrid technology. By establishing collaborations with these startups, researchers and practitioners can gain access to the latest breakthroughs and proficiency in both classical and quantum computing domains. This cooperative endeavor could culminate in the development of more efficacious and efficient hybrid machine-learning models for Earth Land Cover Understanding, ultimately yielding enhanced insights and decision-making capabilities in this field.

As the contemporary landscape of quantum computing continues to evolve rapidly, integrating cutting-edge hybrid technology into Earth Land Cover Understanding can potentially revolutionize the domain. By embracing the strengths of both classical and quantum computing, researchers and practitioners can unlock new possibilities for analyzing and interpreting complex data sets. Ultimately, the pursuit of such hybrid architectures can lead to unprecedented breakthroughs in understanding our planet's land cover, informing crucial decisions related to environmental conservation, climate change, and sustainable development.

We recommend the exploration of hybrid classical-quantum solvers and simulated bifurcation machines for Earth Land Cover Understanding problems. These bleeding-edge technologies offer a unique combination of computational capabilities, which can significantly improve the efficiency and effectiveness of solving complex optimization and classification tasks associated with land cover analysis.

Hybrid classical-quantum solvers, by leveraging the strengths of both classical and quantum computing, can efficiently navigate the solution space and arrive at optimal or near-optimal solutions for Earth Land Cover Understanding problems. Quantum annealers, in particular, can accelerate the training stage by quickly drawing independent samples from the Boltzmann distribution, while classical computing methods can handle the autoencoder and other preprocessing steps. This synergy can reduce the time and energy needed for training, ultimately leading to faster and more accurate results.

Simulated bifurcation machines, on the other hand, provide an alternative approach to quantum annealing that is based on classical computing resources. These machines offer a powerful means to solve combinatorial optimization problems by simulating the bifurcation dynamics of quantum systems, without the need for specialized quantum hardware. As such, simulated bifurcation machines can offer a more accessible and cost-effective solution to Earth Land Cover Understanding problems, while still providing significant performance gains compared to traditional classical computing methods.

In conclusion, adopting hybrid classical-quantum solvers and simulated bifurcation machines for Earth Land Cover Understanding problems can lead to enhanced insights and decision-making capabilities in the field. By exploring these bleeding-edge technologies, researchers and practitioners can unlock new possibilities for analyzing and interpreting complex data sets, ultimately contributing to a better understanding of our planet's land cover and informing critical decisions related to environmental conservation, climate change, and sustainable development.

Quantum computers will likely be only one component of many non-Von Neuman computational accelerators such as e.g. analog, photonic or neuromorphic computers Cavallaro et al. [2022] what makes the landscape of possible non-classical solutions for EO-related problems even more interesting and difficult to navigate in the near future.

Feature Selection and Feature Extraction for Satellite Hyperspectal Imagery Data

Abstract

This section explores the utility of quantum algorithms in the tasks of feature selection and feature extraction of hyperspectral satellite data. The feature selection approaches discussed include the quantum and hybrid extensions of the classification-based RFE (Recursive Feature Eliminati on) algorithm, as well as an approach formulated as the QUBO (Quantum Unconstrained Binary Optimization) problem. The application of the quantum variational method is explored in both feature selection and feature extraction. Additionally, the feature extraction techniques of quantum PCA (Principal Component Analysis) and quantum autoencoders are analyzed. The practical implementation and hardware requirements for each of these approaches are discussed.

25 Introduction

Feature selection and feature extraction are common methods for reducing the number of features in large, high-dimensional data sets. A basic distinction between these methods is that the first involves transforming the original features, while the second preserves the features.

The procedures have profound practical consequences, allowing for more effective data storage, transferring, reduction, and analysis. The hyperspectral data satellite data, with even hundreds of narrow spectral bands, provide an example of the area in which utilization of the methods seems virtually unavoidable. The rich spectral information may simply surpass the needs of certain applications.

On the other hand, since the number of possible selections (subsets) grows exponentially with the number of features, the application of the selection methods involves hard optimization tasks. Within the use case, the possibility of applying quantum algorithms to improve the selection methods is examined. The discussion focuses on hyperspectral imagery data as a promising but extremely challenging data type in the field of satellite Earth imagery.

25.1 About hyperspectral imaging

Hyperspectral satellite imaging is a remote sensing technology that captures images of the Earth's surface at high spectral resolution. Unlike traditional satellite imagery, which captures images in several discrete broad spectral bands (usually red, green, blue, and bands in the infrared wavelength spectrum), hyperspectral imaging captures images in hundreds of narrow and contiguous spectral bands, providing a detailed and comprehensive view of the Earth's surface Chang [2003]; Grahn and Geladi [2007]; Lv and Wang [2020].

Each pixel in a hyperspectral image contains information about the reflectance or emission of electromagnetic radiation at each of the spectral bands. This information can be used to identify and map the distribution of different materials and substances on the Earth's surface, such as vegetation, minerals, water, and pollutants.

Hyperspectral satellite imaging has numerous applications in fields such as environmental monitoring Moroni et al. [2013], agriculture Nalepa et al. [2022], mineral exploration Booysen et al. [2022], and urban planning Karoui et al. [2019]. It can provide valuable information for resource management, land-use planning, and environmental protection, among other purposes.

Some examples of hyperspectral satellite missions include:

- Hyperion: Launched in 2000, Hyperion was a hyperspectral imaging instrument carried aboard the Earth Observing-1 (EO-1) satellite. It provided high-resolution hyperspectral data for a variety of applications, including land use and land cover classification, mineral mapping, and more.
- EnMAP: The Environmental Mapping and Analysis Program (EnMAP) is a German hyperspectral satellite mission that was launched in 2021. Its primary goal is to monitor the Earth's environment, including vegetation, water bodies, and urban areas, to support a range of environmental applications.
- PRISMA: The Italian Space Agency's Precursore Iperspettrale della Missione Applicativa (PRISMA) satellite mission was launched in 2019. Its primary objective is to provide high-resolution hyperspectral data for a variety of applications, including agriculture, forestry, and natural resource management.
- HyspIRI: The Hyperspectral Infrared Imager (HyspIRI) is a proposed NASA mission that would carry a hyperspectral sensor suite to monitor the Earth's terrestrial ecosystems, carbon cycle, and natural hazards. The mission is currently in the planning phase and is expected to launch in the mid-2020s.

The rapid growth of hyperspectral imagery product availability translates also to the technical challenges connected with data processing and analysis:

The increased spectral resolution of hyperspectral images allows for more precise identification and analysis of different materials and features within an image, but also results in a much larger amount of data being generated. The size of a hyperspectral image can be several times larger than that of a multispectral image covering the same area, which can present challenges for data storage and processing.

The construction of contiguous spectral coverage between 400 and $2500nm$ with a spectral sampling interval (of several to tens of nm) Rast and Painter [2019], such that the adjacent channels overlap, makes the information carried by the spectral bands strongly correlated and hence partially redundant. This makes pattern recognition tools for hyperspectral data susceptible to the Hughes phenomenon Hughes [1968]; Theodoridis and Koutroumbas [2006]; Ma et al. [2013]. The Hughes phenomenon occurs when the classification accuracy increases gradually in the beginning as the number of spectral bands or dimensions increases but decreases dramatically when the band number reaches some value. It is a manifestation of the wider phenomenon, collectively called the 'curse of dimensionality' Bellman [1957]; Bishop and Nasrabadi [2006].

The above motivates the use of data reduction techniques for hyperspectral data. In the following, we investigate the possibility and feasibility of quantum-assisted feature selection and feature extraction methods.

26 Feature selection

Complex datasets, such as the introduced hyperspectral images, may contain much more information than is actually needed for a specific application. Furthermore, due to the spectrum's dense band coverage, significant data redundancy may be present. Consequently, data analysis is laborious and may exhibit undesirable consequences, such as the Hughes phenomenon or low interpretability.

Feature selection is a common strategy in data analysis to overcome these difficulties. The price to pay is, however, the high computational cost of the feature selection procedure.

In what follows, we review four possible paths of applying quantum computing methods which promise to improve the classical feature selection methods. We begin with the possible quantum extensions of the Recursive Feature Elimination (RFE). This will mainly concern the employed quantum and hybrid Support Vector Machines (SVM) algorithms. After that, a relation with the quantum kernel methods will be discussed. Another branch of the methods employs quantum optimization, considered one of the most promising applications of the NISQ-era quantum computers. Finally, the utility of quantum variational methods will be discussed.

26.1 Recursive Feature Elimination (RFE) based on quantum/hybrid SVM

Recursive Feature Elimination (RFE) Guyon et al. [2002] is a systematic feature selection algorithm that enables the identification of the most important variables within a given dataset. This technique iteratively eliminates the least significant features and re-calibrates the model with the remaining ones until the optimal number of features is achieved.

By systematically removing the least relevant features during each iteration, RFE facilitates the identification of the feature subset that is most adequate to the model's predictive performance. Consequently, this results in the development of more effective and precise models, as well as a simplified interpretation of the model's outcomes.

RFE can be employed with a diverse range of machine learning models, including linear and logistic regression, Support Vector Machines (SVM), and random forests, amongst others. This technique can be a powerful tool to improve the efficiency and effectiveness of predictive models in numerous applications. In what follows, we will focus on the SVM-based approach to the RFE. The reason for this is that, already in 2014, a promising quantum extension of the SVM algorithm was proposed Rebentrost et al. [2014] — the QSVM (Quantum Support Vector Machine).

The SVM algorithm consists of two main steps. The first is a calculation of the so-called *Kernel function* $K(\mathbf{x}_i, \mathbf{x}_j)$, which is a generalization of the scalar product between the data vectors, $\mathbf{x}_i \in \mathbb{R}^n$, where n is the dimension of the data under consideration. The second step uses the kernel function to solve a system of equations that determine the position of the hyperplane of the decision (separating the two classes of data). In the classical version of SVM, the computational complexity of the algorithm is polynomial in both the number of vectors of the training data (m) and their dimension (n): $\mathcal{O}(m^2(n+m))$. It is proven that for QSVM, the computational complexity reduces to logarithmic in both m and n: $\mathcal{O}(\log nm)$ Rebentrost et al. [2014], which is an example of the so-called *exponential speedup*.

The first step of the QSVM algorithm can be performed using a quantum register with a number of logical qubits equal to n (which corresponds to the dimension of the data). Therefore, in light of the current and expected in the near future quantum computers, the method is suitable for multispectral and even hyperspectral data with dozens of wavebands. Further improvement can be achieved under the dense embedding of the classical data vector on the quantum register. This may ultimately require $\log_2 n$ qubits for n hyperspectral bands. In this limiting case, only eight logical qubits are sufficient to encode 256 spectral bands. Ultimately, future quantum computers operating on hundreds of noisy physical qubits may allow the direct quantum kernel function determination for hyperspectral data satellites.

However, the second QSVM step, equivalent to finding the inverse of a matrix of dimension $(1+m)\times(1+m)$, requires using a number of qubits that scales linearly with the number of training data vectors. Given that in typical cases, the training of the dataset contains up to several hundred thousand data vectors, performing this step is not possible using NISQ technology. In other words, the NISQ technology will allow one to perform this step only for very small ($m \sim 10^1 - 10^2$) training data sets, which is uninteresting from a practical point of view. Therefore, the application of the full, original QSVM algorithm for the purpose of the RFE procedure is impractical from the perspective of the NISQ technology. The QSVM-based RFE for the satellite hyperspectral data may, however, be implemented in the more distant future, say in more than fifteen years.

Earlier, say in 5 years, hybrid methods, in which the first step of the SVM algorithm is quantum-aided, while the second step utilizes classical computing resources, have a chance to be implemented successfully. Furthermore, besides the expected reduction of the computational complexity, quantum evaluation of the kernel function opens up the possibility of improving the classification accuracy and, in consequence, the selection of the feature. This is because quantum algorithms give rise to a whole new family of kernel functions. Some of them may outperform the considered classical ones in classification accuracy. Furthermore, they may turn out to be hard to emulate on classical machines, justifying the use of quantum computers. We discuss the quantum kernel-based approach in the next subsection.

26.2 Feature selection based on quantum kernel methods

There are various possible maps from the original data space, say $\mathbb{R}^n$ to the Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n'}$ in which a state of n'-qubit quantum register is defined. Let the map, which we call a quantum feature map, be defined as

$$\mathbf{x} \to |\Phi(\mathbf{x})\rangle = \hat{\mathcal{U}}_{\Phi(\mathbf{x})}|0\rangle^{\otimes n'}, \tag{14}$$

where $|0\rangle^{\otimes n'}$ is an initial state of the quantum register and $\hat{\mathcal{U}}_{\Phi(\mathbf{x})}$ is the unitary quantum operator that defines the quantum map.

Therefore, the kernel function is determined as the following square of the amplitude modulus probability:

$$K(\mathbf{x}_i, \mathbf{x}_j) = |\langle \Phi(\mathbf{x}_i)|\Phi(\mathbf{x}_j)\rangle|^2. \tag{15}$$

In the 2019 article, Havlíček et al. [2019], a promising proposal for the kernel function has been studied. The corresponding quantum feature map was given by the operator $\hat{\mathcal{U}}_{\Phi(\mathbf{x})} = (\hat{U}_{\Phi(\mathbf{x})}\hat{H}^{\otimes n})^d$, where $\hat{H}$ is the Hadamard gate and $d \in \mathbb{N}_+$ is the number of repeats. The operator $\hat{U}_{\Phi(\mathbf{x})}$ is given by the following expression:

$$\hat{U}_{\Phi(\mathbf{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(\mathbf{x}) \Pi_{i \in S} \hat{Z}_i\right), \tag{16}$$

where $\hat{Z}_i$ is a Pauli-Z gate, and the functions $\phi_S(\mathbf{x}) \in \mathbb{R}$ introduce the data vectors $\mathbf{x} \in \mathbb{R}^n$ into the register. The map is the so-called ZZ map.

For the data representation given by the operator (16), determining the function of the kernel is a difficult task ($\#P$-hard) for classical computers.

This requires, in general, operating on 2^n-dimensional vectors in the feature space, which, for example, for hyperspectral data containing $n = 100$ bands correspond to composite vectors of dimension $2^{100} \sim 10^{30}$. This is beyond the capabilities of the current and imaginable future classical supercomputers. On the other hand, quantum computations require, in this case, a quantum register composed of $n = 100$ logical qubits and possibly a few thousand noisy physical qubits, which sounds realistic from the 15-year time perspective. This justifies approaching the problem from the side of quantum computing, especially in the case of hyperspectral data (with large n).

However, validating the utility of the map introduced for real-life hyperspectral is currently a challenging task. This is simply because of the high complexity of the map, which makes it difficult to emulate with classical HPC. Therefore, some lower-complexity maps have recently been investigated to make the application to real-world multispectral data possible Miroszewski et al. [2023].

26.3 Feature selection based on quantum optimization

Another approach is to perform feature selection using quantum optimization. Here, one encodes the tasks into a mathematical optimization problem that can be solved using a quantum computer. In Refs. Otgonbaatar and Datcu [2021a]; Mücke et al. [2023], a feature selection approach that utilizes the QUBO (Quantum Unconstrained Binary Optimization) problem and quantum annealing for optimization has been proposed for processing hyperspectral images (HSIs) and practical datasets. This approach involves mapping the feature selection problem to a binary optimization problem, where each feature corresponds to a binary variable, and the objective is to minimize or maximize a particular performance metric, such as precision or accuracy. The resulting optimization problem can then be solved using a quantum optimization algorithm, such as quantum annealing or the Quantum Approximate Optimization Algorithm (QAOA).

Following Ref. Mücke et al. [2023] let us consider labeled dataset $D := \{(\mathbf{x}_i, y_i), i \in [N]\}$, where $[N] = \{1, 2, ..., N\}$ is the set of data point indices and $\mathbf{x}_i \in \mathbb{R}^n$ – the quantum feature selection is even performed on HSIs on a D-Wave quantum annealer Otgonbaatar and Datcu [2021a]. Applying the feature selection procedure reduces the set D to $D_S := \{(\mathbf{x}_{S,i}, y_i), i \in [N]\}$, where the feature-selected data points $\mathbf{x}_{S,i}$ are defined so that they lead to a performance comparable to the original classification problem.

The selected features from the original data vector can be indicated by the binary vector $\mathbf{X} = \{X_1, ..., X_n\} \in \{0, 1\}^n$, so that $X_i = 1$ if a given feature is selected and $X_i = 0$ if it is not. Then the following cost function can be introduced:

$$Q(\mathbf{X}, \alpha) = \mathbf{X}^T \mathbf{Q}(\alpha) \mathbf{X} = -\alpha \sum_{i=1}^{n} I_i X_i + (1 - \alpha) \sum_{i,j=1}^{n} R_{ij} X_i X_j, \tag{17}$$

where the parameter $\alpha \in [0, 1]$ balances the contribution of the two terms. The first term, where $I_i := I(X_i, y) \geq 0$, quantifies the importance of a given measure. Here, $I(X_i, y)$ is the mutual information between the individual features X_i and the class label y. The second is the redundancy term where $R_{ij} := I(X_i; X_j) \geq 0$ is the mutual information among the individual features. In the matrix notation, the elements of $\mathbf{Q}$ can be written as $Q_{ij}(\alpha) = R_{ij} - \alpha(R_{ij} + \delta_{ij} I_i)$.

The task is now to find such a configuration of X_i, call it X_i^*, that both the importance and the redundancy terms are minimized. So, we are looking for:

$$\min_{\mathbf{X} \in \{0,1\}^n} \mathbf{X}^T \mathbf{Q}(\alpha) \mathbf{X}, \tag{18}$$

and in consequence, the set of the selected features is given by

$$\mathbf{X}^* = \arg \min_{\mathbf{X} \in \{0,1\}^n} \mathbf{X}^T \mathbf{Q}(\alpha) \mathbf{X}. \tag{19}$$

The task introduced above is the QUBO (Quantum Unconstrained Binary Optimization) problem, which typically is of the NP-hard type. This is because for every X_i, two values of $\{0, 1\}$ are allowed. In consequence, for n features, there are 2^n configurations to be explored. Therefore, finding the global minimum requires performing the number of steps that grow exponentially with n. So for the $n = 100$ hyperspectral bands, the number of configurations to explore is $2^{100} \sim 10^{30}$ and for $n = 100$ we have $2^{200} \sim 10^{60}$. Both numbers are beyond the scope of any current classical computing resources. Nevertheless, in some cases, not all configurations need to be explored, and classical ergodic algorithms, such as those based on the Markov chain, can be applied.

The QUBO is known to be related to the problem of finding the ground state of the Heisenberg-type quantum Hamiltonian:

$$\hat{H} = \sum_{<i,j>} a_{ij} \hat{Z}_i \hat{Z}_j + \sum_i b_i \hat{Z}_i + c, \tag{20}$$

where $\hat{Z}_i$ are the Pauli-Z operators and the summation $< i, j >$ is defined such that it does not repeat over pairs. The coefficients a_{ij}, b_i and c can be directly related to the parameters of the original problem defined by Eq. 17.

There are various approaches to attempting to determine the ground state of the Hamiltonian (20). On the one hand, the task can be formulated as the quantum annealing problem, which may be implemented on a quantum annealer, such as the one provided by the D-Wave company. On the other hand, optimization algorithms that operate on gate-based quantum computers can be utilized. This, in particular, concerns the QAOA algorithm, and the variational algorithms being a subject of the subsequent subsection.

26.4 Feature selection based on variational methods

Quantum variational methods play a significant role in the early stage of applications of quantum computing technologies. It turns out that the approach is NISQ-friendly and may provide an improvement in solving real-life problems. These algorithms use a quantum computer to find the optimal solution to a given problem by minimizing a cost function. The cost function is typically defined in terms of the parameters of a quantum circuit, and the algorithm iteratively updates these parameters until the optimal solution is found. Importantly, variational methods are typical of the hybrid type. Some of the applications of variational quantum methods include: quantum chemistry (determination of the ground stats of molecules), machine learning (optimization of the parameters of machine learningmodels), and quantum error correction (optimization of the parameters of the quantum error correction codes).

In the context of feature selection, the variational methods can be employed both in the quantum/hybrid RFE approach and the QUBO problem discussed in the previous subsection.

Concerning the RFE, a concrete proposal for a quantum variational classifier has been presented in Ref. Havlíček et al. [2019]. Its undoubted advantage is that, unlike the original QSVM method, the number of qubits in the quantum register is independent of the number of training data m. The approach is based on the so-called *Variational Quantum Classifier* (VQC), which uses a variational quantum system to classify a training set in direct analogy to conventional SVM. In this method, the first step is the same as in QSVM, but in the second step, instead of matrix inversion, a parametrized quantum circuit $\hat{W}(\theta)$ is used. Training consists of minimizing a certain cost function with respect to the parameters of the circuit (given by the vector θ).

27 Feature extraction

Feature extraction methods take the original input features and transform them into new, usually lower-dimensional target features while maintaining as much of the informational content of the data as possible Ding et al. [2012].

27.1 Quantum PCA

Principal Component Analysis (PCA) is an over-century-old technique of dimensionality reduction Pearson and Lines [1901]. It is typically performed by diagonalizing the covariance matrix of the given data and keeping the subset of the eigenvectors, which correspond to the largest eigenvalues.

The approach to quantum principal analysis consists of two steps: association of the covariance matrix K with the density matrices ρ and diagonalization of the resulting density matrix.

Suppose that we have n-dimensional random vectors $X = \{X_1, \ldots, X_n\}$. The entry $K_{\alpha\beta}$ of the covariance matrix is defined as

$$K_{\alpha\beta} = \mathrm{cov}(X_\alpha, X_\beta) = E\left((X_\alpha - E(X_\alpha))\,(X_\beta - E(X_\beta))\right), \tag{21}$$

where E denotes the expectation value. Now, for the $L2$ normalized dataset $\mathcal{D} = \{x_\alpha^{(1)}, x_\alpha^{(2)}, \ldots, x_\alpha^{(m)}\}$, $\sum_{\alpha=1}^{n}\left(x_\alpha^{(i)}\right)^2 = 1$ one can perform the amplitude encoding Grover [2000]; Grover and Rudolph [2002]; Plesch and Brukner [2011] of each of the m data points $x^{(i)} \mapsto |x^{(i)}\rangle\langle x^{(i)}| = \rho^{(i)}$. The authors of Gordon et al. [2022] have noticed that introducing the ensemble average density matrix $\bar{\rho} = E\left(\rho^{(i)}\right)$ leads to a useful equivalence

$$\bar{\rho}_{\alpha\beta} = K_{\alpha\beta} + \mu_\alpha\mu_\beta, \tag{22}$$

where $\mu_\alpha = \sum_{i=1}^{m} x_\alpha^{(i)}$ is the mean value of the feature α in the data set. Furthermore, in accordance with Weyl's theorem for the eigenvalues of Hermitian matrices, the eigenvalues of K and $\bar{\rho}$ are closely related. This

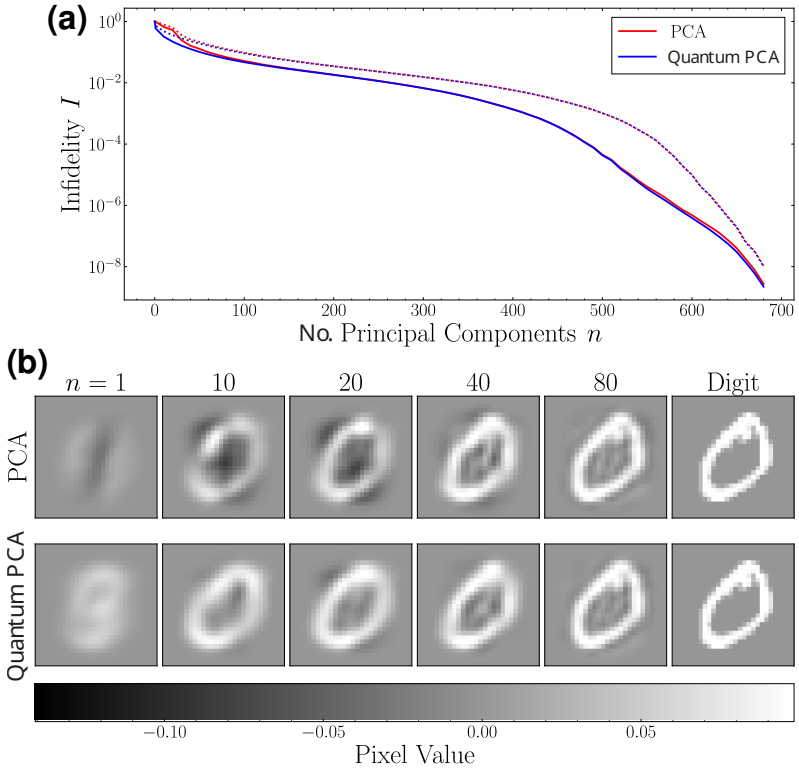


Figure 23: Projection onto n principal components for PCA and quantum PCA of the MNIST dataset. (a) The plot shows the median (solid) and upper 90% interval (dotted) infidelity (I) values between the projected state constructed using n principal components calculated from PCA and quantum PCA. The median and upper interval is calculated over all 50000 digits. (b) We show how the images are reproduced when using $n = 1$ to $n = 80$ principal components from PCA and quantum PCA in the upper and lower rows, respectively. From Gordon et al. [2022] (CC BY 4.0).

gives a concrete implementation of the covariance matrix preparation for quantum computers and guarantees the quality of the quantum principal component analysis (qPCA); see Fig. 23.

For the diagonalization of the covariance matrix on quantum computers, one long-term approach and one near-term approach are presented. In 2014 Lloyd et al. Lloyd et al. [2014] introduced a quantum principal component analysis (qPCA) algorithm which was supposed to provide an exponential speed-up over PCA for low-rank covariance matrices. The author of Tang [2021a] noted that the exponential speedup is possible only with the unjustified assumption of state preparation. With further discussions Cotler et al. [2021] it seems that one can hope for the speedup for qPCA on classical data, but it will not be exponential. The original qPCA algorithm was designed for large, fault-tolerant quantum computers, and hence is not suitable near-term devices. Recently, NISQ-era friendly algorithms, useful for qPCA implementation, were proposed LaRose et al. [2019]; Cerezo et al. [2022]; Verdon et al. [2019]; Ezzell et al. [2022]; Gordon et al. [2022]. Contrary to the original qPCA formulation, the variational quantum state diagonalization (VQSD) LaRose et al. [2019] and variational quantum state eigensolver (VQSE) Cerezo et al. [2022] algorithms require, respectively, only one or two copies of the density matrices ρ. This makes them much better suited for near-term quantum

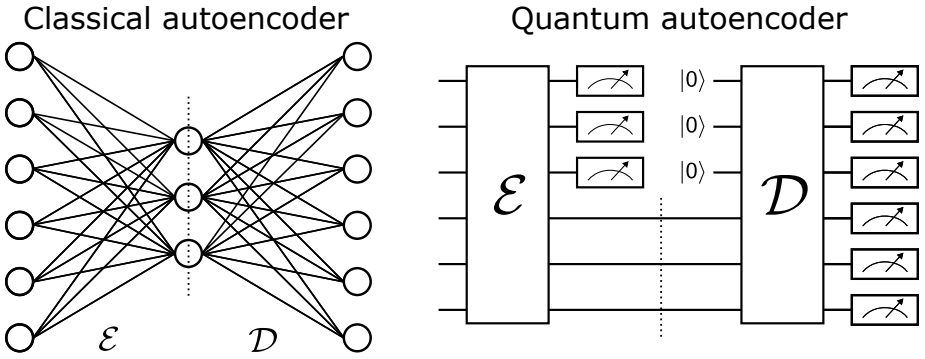


Figure 24: *Left:* Classical graphical representation of a 6-bit autoencoder with a 3-bit latent space. The map $\mathcal{E}$ encodes a 6-bit input into a 3-bit intermediate state, after which the decoder $\mathcal{D}$ attempts to reconstruct the input bits at the output. The bottleneck is represented by a dashed line. *Right:* Circuit implementation of a quantum algorithm inspired by the classical autoencoder on the left.

devices. In Figure 23 the application of qPCA for simulated ten-qubit system in Gordon et al. [2022] is presented.

27.2 Quantum autoencoders

The classical autoencoder Kramer [1991] is a type of neural network used to learn an efficient representation of unlabeled data. Its distinct feature is an encoder-decoder architecture. An encoder function $\mathcal{E}$ takes the $(l+k)$ bit input data x and transforms it into a lower-dimensional l bit representation $\mathcal{E} : x \to \tilde{x}$, effectively compressing the information contained in x. A decoder function $\mathcal{D}$ recreates the input data from the encoded representation $\tilde{x} \to \hat{x}$. The typical graphical representation (Figure 24a) of the autoencoder consists of a convergent neural network with a decreasing number of neurons with consecutive layers, followed by a set of layers of neural networks growing in width. The narrowest layer in the network, the bottleneck, includes the representation of data in latent space $\tilde{x}$. During training, an autoencoder tries to minimize the difference between input data x and output $\hat{x}$. The trained autoencoder, with the use of an encoding function, allows for efficient dimensionality reduction or feature extraction of the data. Unlike PCA, autoencoders can learn non-linear relationships and therefore perform better at compressing data.

Inspired by the erasure of k bits of information in the classical encoder function, the idea of quantum autoencoders was introduced in Romero et al. [2017]. Now, the input data is stored in $l+k$ qubits and processed by the variational layers in $\mathcal{E}$. k-qubits are then traced out (indicated by measurement in Figure 24b). The newly prepared k qubits with latent l qubits are then processed with the decoder function $\mathcal{D}$. The natural task for training the model would be to use the classical optimization of fidelity between input and output states $F(|\psi_i\rangle, \rho_i^{out}) = \langle\psi_i|\rho_i^{out}|\psi_i\rangle$ to train the model. The authors of Romero et al. [2017] argue that invoking a reference state and performing a swap test on it and k trash qubits leads (for high fidelities) to the equivalent optimization task, while training on only the trash state. In the concrete realization (see Figure 25) of the quantum autoencoder circuit, one initializes l qubits in the latent space, k qubits in the trash space and for reference state and one auxilliary qubit needed for the swap test.

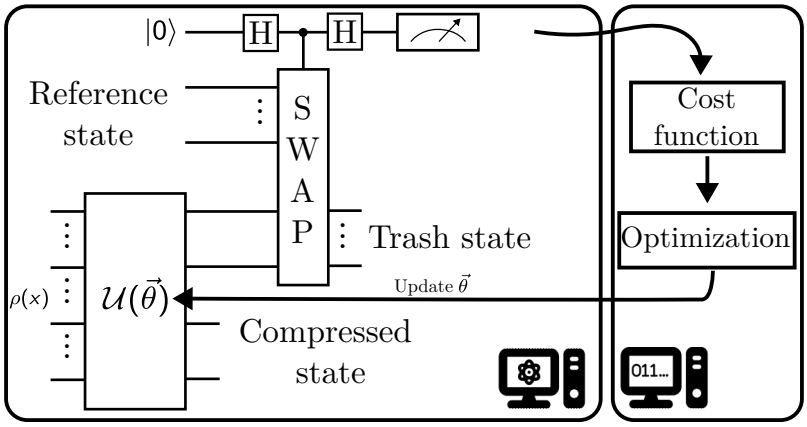


Figure 25: Schematic representation of the variational circuit for training a quantum autoencoder proposed by Romero et al. [2017]. The input state, carrying classical data x is transformed by a parameterized unitary $\mathcal{U}(\vec{\theta})$. For the computation of the cost function, SWAP test is employed. The quality of the compression of the initial state is optimized on a classical computer.

28 Sizing quantum machines for feature selection and feature extraction of hyperspectral imagery data

One of the most promising quantum hardware architectures is the circuit-based model, where quantum algorithms are implemented as sequences of quantum gates. This architecture allows for the efficient implementation of quantum feature selection and quantum feature extraction algorithms, such as quantum SVM-based RFE or quantum PCA. Various hardware implementations of gate-based quantum computers are currently under development, including superconducting quantum computers, trapped ions quantum computers, photonic quantum computers, spin qubit quantum computers, or topological quantum computers.

While the most advanced approach is currently the one based on superconducting qubits, it faces numerous difficulties, such as short coherence times and low interconnectivity of qubits. Consequently, only relatively shallow quantum circuits can be evaluated using the technology. The allowed scalability of the approach with the number of qubits partially compensates for this weakness. Nevertheless, because the quantum algorithms investigated in this Section are expected to utilize deep quantum circuits for real-world hyperspectral data, applying gate-based superconducting quantum computers may pose a difficulty.

Both the interconnectivity and decoherence properties of qubits can be substantially improved by applying other techniques, such as those based on trapped ions and photons. While promising, these technologies face challenges concerning their scalability. Therefore, except in the cases of multispectral rather than hyperspectral data or low-complexity feature maps, it is rather unlikely that gate-based quantum computers will be ready to implement successfully the discussed algorithms for real-world hyperspectral data earlier than in the 15 years time horizon.

Another important architecture is adiabatic quantum computing (quantum annealing) model, which is well-suited for solving optimization problems. As discussed in Sec. 26.3, adiabatic quantum computers can be utilized for quantum feature selection by formulating the problem as an optimization task and mapping it onto the Ising-type model.

The current and expected to develop in 3-5 years time horizon architectures of quantum annealers have a chance to meet the requirements of real-world feature selection problems. However, adiabatic quantum

computers, despite their potential, face several limitations. In particular, the requirement of adiabatic evolution imposes constraints on the execution time and the scalability of the algorithms. The adiabatic process needs to be slow enough to ensure the system remains in its ground state, which can result in longer computation times compared to classical methods or other quantum computing models. Furthermore, because the currently most advanced quantum annealers utilize the superconducting qubits technology, the approach faces challenges with maintaining coherence and providing a high level of interconnectivity. In consequence, even if potentially possible to apply to real-world hyperspectral data, superconducting quantum annealers remain difficult to assess with respect to their accuracy and time performance.

Regarding the feature extraction on universal quantum computers, for a hyperspectral dataset with approximately 200-300 spectral bands, the amplitude encoding of the ensemble average density matrix $\bar{\rho}$ requires the use of 8-qubit registers. In the original implementation of qPCA by Lloyd et al. Lloyd et al. [2014], one needs to prepare $n \in \mathcal{O}(1/\epsilon^3)$ copies of $\bar{\rho}$ to solve the eigenproblem with an accuracy of ϵ. On the other hand, the VQSD algorithm by Larose et al. LaRose et al. [2019] requires two copies of $\bar{\rho}$, resulting in a quantum architecture consisting of 16 qubits. In the case of VQSE proposed by Cerezo et al., Cerezo et al. [2022], only one copy of $\bar{\rho}$ is needed, hence requiring 8 qubits.

In the current implementation of autoencoders proposed by Romero et al. Romero et al. [2017], a single set of quantum registers is used for the latent space, along with two sets of quantum registers for the trash space and an additional qubit register for the swap test. This results in a total of $l + 2k + 1$ qubits. Considering the amplitude encoding of 8 input qubits, one would require quantum architectures consisting of 10 to 16 qubits.

28.1 Currently

Currently, the most sophisticated opportunities for implementing quantum algorithms in feature selection and extraction lie within the capabilities of classical emulators of quantum computers. Specifically, these emulators facilitate the investigation of hybrid feature selection methods for multispectral datasets comprising several dozens of bands. While the methodology applies amplitude encoding, it also possesses the potential to accommodate hyperspectral data with hundreds of bands. However, amplitude encoding offers limited noise control within genuine quantum implementations. Consequently, this approach may not seamlessly translate to future quantum computing technology.

The computational limitations of emulators pose substantial challenges for predicting the feasibility of hybrid quantum algorithms, especially those based on variational methods. Indeed, drawing definitive conclusions about the effectiveness of these hybrid techniques in managing real-world hyperspectral data is an intricate task, bordering on impossible.

In addition, the existing quantum computers' opportunities are considerably restricted due to the absence of practical quantum correction code implementations. Thus, only relatively uncomplicated tasks with limited practical value can be achieved. An example of this would be hybrid feature selection for low-dimensional (a few bands) data vectors.

At present, some opportunities are accessible through quantum annealers. This particularly refers to the feature selection method outlined in Section 26.3. However, there is no evidence yet that currently available quantum annealers offer a computational advantage.

28.2 3-5 years

Given the relatively short forecast horizon of 3-5 years, significant advancements in quantum computing capabilities are not expected. Nonetheless, this period may witness the deployment of the initial practical implementations equipped with quantum error correction codes. This development would pave the way for the usage of noisy superconducting quantum processors to carry out more complex quantum algorithms.

In light of our current knowledge, it is anticipated that quantum or hybrid feature selection and extraction for multispectral data encompassing approximately ten bands could be realized on actual quantum computers. However, this advancement is not expected to surpass or even match the capabilities of classical computers.

28.3 15 years

The next 15 years hold immense potential for significant advancements compared to the current state of the art. Notably, the successful implementation of quantum correction codes on superconducting quantum computers comprising thousands of qubits may pave the way for hybrid and quantum feature selection and extraction, enabling the utilization of approximately 100 bands. However, it is crucial to emphasize that this does not guarantee a definitive advantage in terms of time performance or accuracy. There is still limited knowledge regarding the algorithms employed and the future capabilities of the hardware, making it challenging to draw conclusive statements in this regard. Nonetheless, certain advantages can be anticipated, e.g., in terms of the energy consumption of quantum computers when compared to classical computing units. The implementation of Recursive Feature Elimination (RFE) based on Quantum Support Vector Machine (QSVM), which is expected to offer exponential speed-up, will possibly exceed the given 15-year time frame.

29 SWOT Analysis

29.1 Strengths

- High level of motivation resulting from the computational intricacy of the underlying problem.
- The availability of real-world datasets for training and testing.
- Documented evidence of exponential speed-up for Recursive Feature Elimination (RFE) based on Quantum Support Vector Machine (QSVM).

29.2 Weaknesses

- Implementation challenges of the RFE using QSVM due to the required number of logical qubits being of the order of the size of the training dataset (typically, tens of thousands). This poses significant difficulties in practical applications.
- Absence of definitive proof showcasing a clear edge for quantum kernel-based methods.

29.3 Opportunities

- Potential to utilize hybrid approaches that require a relatively small number of qubits (of the order on 10^2 logical qubits), thereby increasing feasibility.
- Possible advancements pertaining to accuracy levels of selection and feature extraction, alongside the prospect of reducing computational energy consumption.

29.4 Threats

- Typically, the deep quantum circuits under consideration require long coherence times for execution. This imposes implementation difficulties on superconducting quantum computers.
- The highly dimensional feature spaces offered in the quantum approach may not be needed in many relevant scenarios.

Table 3: Summary of the identified feature selection methods for hyperspectral imagery data.

Method	QSVM RFE Rebentrost et al. [2014]	VQC RFE Havlíček et al. [2019]	Quantum optimization Otgonbaatar and Datcu [2021a]
Resources	high $\gtrsim 10^5$ logical qubits	moderate $\sim 10^2$ logical qubits	low/moderate $\sim 10^2$ logical qubits
Time horizon	>15 years	3-5 years	now/3-5 years
Architecture	gate-based quantum	gate-based hybrid	annealing/gate-based hybrid
Speedup	exponential	polynomial	polynomial/exponential

Table 4: Summary of the identified feature extraction methods for hyperspectral imagery data.

Method	qPCA Lloyd et al. [2014]	variational qPCA Gordon et al. [2022]	qAutoencoders Romero et al. [2017]
Resources	high $\gtrsim 10^3$ logical qubits	moderate $\sim 10^2$ logical qubits	low/moderate $\sim 10^1 - 10^2$ logical qubits
Time horizon	15 years	3-5 years	3-5 years
Architecture	gate-based quantum	gate-based hybrid	gate-based hybrid
Speedup	polynomial	polynomial	polynomial

30 Conclusions

In this Section, a set of quantum and hybrid (classical-quantum) algorithms for feature selection and feature extraction was presented. The selected set includes long-term and near-term algorithms. The greatest opportunity for quantum advantage lies in the long-term routines for fault-tolerant quantum computers. These include the following: Recursive Feature Elimination based on the full Quantum Support Vector Machines, feature selection based on quantum kernel methods, and original Quantum Principal Component Analysis. To harness the advantage of these algorithms, long-term technological development of quantum computers is needed. A key issue is to be able to access a large ($\sim 10^2$) number of logical qubits, increase their coherence time, and improve the precision of the unitary gates used in quantum circuits. Therefore, we estimate at least a 15-year time horizon for such methods to become available. On the shorter time scale (5-year horizon), promising candidates for quantum advantage include hybrid algorithms based on variational architectures. These include feature selection methods based on Variational Quantum Classifiers, quantum Principal Component Analysis using Variational Quantum State Diagonalization or Variational Quantum State Eigensolver, and quantum autoencoders. The identified quantum and hybrid approaches to feature selection have been summarised in Tab. 3, while the identified approaches to feature extraction have been collected in Tab. 4

The crucial issue in the context of hyperspectral data analysis is the state preparation for the above algorithms. On the one hand, it is possible to employ an amplitude encoding scheme, which allows one to operate some of the selected algorithms on the order of 10^1 qubits. However, this requires operating with high precision on the quantum states. On the other hand, other encoding schemes require access to at least hundreds of good-quality qubits. No quantum advantage in practical tasks has been confirmed so far, but we can expect it with the technical development of quantum computers. Answering whether such an advantage will realistically provide a net gain for Earth observation is an issue that requires further in-depth studies.

Summary

Quantum machines promise to solve a certain class of computational problems faster than conventional machines. In particular, the hardness of the computational problems can be measured from the perspective of a computational complexity theory. Hence, this study identifies five intractable practical problems which can not be efficiently solved on classical computers, but quantum machines promise to find their solutions faster than their classical counterparts. In addition, we assess and examine distinct quantum machines including a quantum annealer, a quantum simulator, and universal quantum computers for their practicality. For our identified EO UCs, we propose climate adaptation digital twin HPC+QC workflow (UCI), uncertainty quantification for remotely-sensed datasets (UCII) variational quantum algorithms (UCIII), a quantum feature selection and reduction for large-scale datasets (UCIV). In particular, we analyzed and evaluated the hardness of our EO UCs based on a theoretical computational complexity theory. In addition, we selected and proposed some quantum machines for each EO UC by closely sticking to their quantum development roadmap in the reference time-frame of 3-5 to 15 years, since distinct quantum machines are designed to solve a specific class and form of hard computational problems.

Difficulty to assess the sizing of quantum computers for applications

Recently, IBM researchers reported in Nature vol. 618 in June 2023 the "Evidence for the utility of quantum computing before fault tolerance" Kim et al. [2023]. In this work, the authors use a 127 qubit superconducting quantum processor with 2880 CNOT gates in order to simulate the quantum dynamics of a particular Hamiltonian. The problem they choose is not useful per se but interesting from the standpoint of study of complex quantum systems. The exact topology of the problem they choose is exactly mapped to the topology of the quantum processor itself, which makes the simulation feasible.

The authors compare their results obtained from the quantum processor with a variety of the state-of-the-art methods that are currently used to perform simulation of the dynamics of similar Hamiltonians using a classical computer. The IBM team was able to show that using their device and error mitigation techniques, they are able to obtain results that they could not replicate in simulation. It is important to mention that the authors provide wall-clock run time for computation. It was approximately 4 h for the first experiment and 9.5 h for the second, but the computation time was dominated by classical processing delays. They estimate that sampling time from the quantum computer was only roughly 5 min. The paper presented by the IBM team is interesting and provides insights into the capabilities of currently-existing quantum computing technology. Interestingly, about two weeks after the publication of the aforementioned paper in Nature two preprints were published on ArXiv Begušić and Chan [2023]; Tindall et al. [2023] that recreate the calculations performed by the IBM team using regular commodity computers in roughly 5–7 minutes. This story illustrates how difficult it is to discuss quantum advantage even for tasks that are purely academic and have no practical applications. Therefore, currently any sizing of quantum hardware for particular tasks related to practical applications such as e.g. Earth observation is doomed to fail. As an alternative, we provide recommendations of studies that we believe should be performed to explore the applicability of future quantum computers for Earth observations. A research program lasting 3-5 years should allow the scientific community to identify pathways towards quantum computing advantage for Earth observations. Only after these studies, and similar, are finalized, a proper assessment of the future research directions can be made.

Recommendations

The authors recommend ESA to implement a series of actions that will support the development of the European community and network of institutions and companies studying and implementing quantum advantage for applications in Earth observation. We recommend investigating deeper ideas presented below.

Idea 1: Hyper-spectral Image segmentation using deep RBMs/QBMs.

Restricted Boltzmann Machines (RBMs) and Deep Boltzmann Machines (DBMs) can be used for the analysis and processing of Remote Sensing (RS) satellite imagery for Earth Observation (EO) applications. Quantum annealers are specific purpose quantum computers aiming at solving combinatorial optimization problems. In the context of Earth observation it is important to determine whether the use of quantum annealers as co-processors to assist in Machine Learning (ML) training can enhance the analysis and interpretation of complex EO data sets and expedite RS data processing. Research in this area should provide responses to the following questions. Can DBMs effectively facilitate semantic segmentation of multispectral images for EO applications? What is the feasibility of training RBMs and DBMs on satellite imagery using quantum computing in general and quantum annealers in particular? How does the performance of DBMs compare to that of deep neural networks trained on classical GPUs when applied to spectral image segmentation tasks? Is using quantum annealers as co-processors in RS data processing pipelines for ML training a viable solution for EO applications?

The application of quantum annealing and other modern hardware-assisted techniques for training RBMs is an active research area. It was demonstrated that Distributed Boltzmann Networks (DBNs) trained on sparse Ising machines achieve high values of accuracy quickly by employing the Ising machines as samplers from the Boltzmann distribution. It is postulated that DBNs trained using hardware-assisted methods could become a viable alternative for DNNs. Quantum annealers have the potential to significantly speed up the training process compared to the use of standalone HPC resources and improve the quality of learned representations in generative and classification models. While early studies have demonstrated promising results, the practical implementation and evaluation of quantum annealing for RBNs in the context of EO image segmentation was not reported in the literature yet.

Idea 2: Application of probabilistic graphical models for segmentation and post-processing

The goal of the research is to provide limited quantum advantage by designing, implementing and evaluating algorithms for performing training and inference in the Probabilistic Graphical Models (PGMs) aimed to be executed at a digital quantum computer and quantum annealers. The advantage provided will be related to the area of Earth observation. PGMs have been extensively used in EO applications (before the advent of deep learning) to model spatio-temporal RS data. PGMs augment solution finding process using diagrammatic representations of probability and their framework offer suitable properties for spatial-temporal crop classification with phenology.

The goal of this research would be to provide a technical solution for EO applications which require processing of huge amounts of data. A quantum-based image processing solution which uses dynamic PGMs expressed in the mathematical language of factor graphs could be proposed. Quantum annealers are specific purpose quantum computers aiming at approximately solving Quadratic Unconstrained Binary Optimization (QUBO) problems. While they are able to find low energy solutions for specific QUBO problems quickly their applicability is limited by the topology of qubit interactions. Many hard QUBO problems require very dense graphs of interactions between binary variables which require use of complex embeddings of QUBO problems onto the annealer topology. The requirement to use embeddings effectively reduces the usability of quantum annealers, since many physical qubits have to be used to represent a single logical qubit. Therefore, it is unlikely that quantum annealers will find practical applications to solve dense QUBO problems.

A major effort in applying quantum annealers to particular problems is to find highly efficient ways to encode a problem onto the device. Such encodings should be informed by both the structure of the problem and the structure of the hardware graph. While the hardware graphs of future generations of superconducting circuit based quantum annealers are likely to become more connected in the future, it is likely the connections will be sparse and have a quasi-planar topology, meaning that all coupled qubits are physically nearby on a 2D plane.

Therefore, it is important to study how use-cases can drive the development of quantum annealers and how to adapt the connectivity of the quantum annealing hardware to provide advantage for a particular task.

Idea 3: Change detection

We propose calls for projects whose focal point should be the construction of hybrid quantum-classical machine learning models that can effectively process and analyze satellite imagery. Proposed approaches should explicitly detail how conventional machine learning techniques, such as autoencoders, will be utilized for tackling the high-dimensionality commonly associated with satellite imagery. Further, proposals must delineate how quantum computing structures, including but not limited to D-Wave quantum annealers and Deep Belief Networks (DBNs), will be harnessed for efficient model training.

We express a specific interest in projects that probe the utilization of VQAs for the training of Quantum Neural Networks (QNNs) and their subsequent implementation in the detection of changes in Earth observations. Proposals that convincingly demonstrate how these quantum computing techniques can provide an advantage over traditional methods in terms of efficiency, precision, and cost-effectiveness will be viewed favourably.

The ultimate objective of the projects should be the aim to unlock the potential of quantum computing to create robust and efficient approaches to change detection in Earth observations. Applications that show promise in delivering real-world impact, such as expedited response times to natural disasters, enhanced accuracy in climate change monitoring, and superior land-use classification techniques, are strongly encouraged.

These calls should be opened to applications from consortia that bring together academic institutions and startup companies eager to pioneer in the quantum computing field. Applicants should be able to demonstrate a strong balance between academic rigor, innovative ideation, and the agility and speed of a startup in prototyping and scaling solutions. This initiative presents a prime opportunity to push the frontiers of quantum computing applications and contribute to the evolution of more sustainable and resilient Earth observation systems.

Idea 4: Application of physics inspired simulated bifurcation, and other hybrid algorithms combined with quantum annealers for EO

In the context of Earth Observation (EO), combining simulated bifurcation and, for instance, quantum annealing algorithms can lead to improved accuracy and efficiency in solving combinatorial problems. Simulated bifurcation is a classical optimization algorithm that uses a perturbation approach to find the optimal solution. On the other hand, quantum annealing algorithms use quantum mechanics to search for the optimal solution of an optimization problem. The idea behind combining these two algorithms is to use simulated bifurcation to pre-process the data and provide initializations for the quantum annealer. This can help to reduce the search space and make the overall computation more efficient. One can also approach the problem by searching iteratively in an alternating manner, using both quantum and classical solvers.

Yet another way to combine these algorithms is to use the simulated bifurcation algorithm to identify regions of the search space that have a high probability of containing the optimal solution. These regions can then be used to define the initial state of the quantum annealer, which will then search for the optimal solution using its quantum properties.

For example, in EO, this approach can be used for land cover classification, where the algorithm can combine high-resolution LiDAR data with low-resolution satellite imagery to accurately classify land cover types. The simulated bifurcation algorithm can be used to identify regions of interest in the LiDAR data that are likely to correspond to a specific land cover type, which can then be used to initialize the quantum annealer. The quantum annealer can then search for the optimal solution, which corresponds to the optimal classification of land cover types in the satellite imagery.

Overall, the combination of simulated bifurcation and quantum annealing algorithms in EO can help to improve the accuracy and efficiency of data analysis, particularly in problems that involve complex optimization challenges. By combining these algorithms, researchers can leverage the benefits of both classical and quantum computing to arrive at better solutions for challenging EO problems.

The hardware necessary for combining simulated bifurcation and quantum annealing algorithms in the context of Earth Observation would include both classical and quantum computing hardware. Classical computing hardware includes a high-performance server or cluster with significant computing power and memory. Simulated bifurcation algorithms are computationally intensive and require a lot of computational resources, especially for large datasets. Quantum computing hardware includes a quantum annealer, such as D-Wave Systems' quantum annealer, which can perform optimization problems using quantum mechanics.

References

A. Abbas, D. Sutter, C. Zoufal, A. Lucchi, A. Figalli, and S. Woerner. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409, Jun 2021. ISSN 2662-8457. doi: 10.1038/s43588-021-00084-1. URL https://doi.org/10.1038/s43588-021-00084-1.

R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush, D. Bacon, J. C. Bardin, J. Basso, A. Bengtsson, S. Boixo, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, D. M. Debroy, A. Del Toro Barba, S. Demura, A. Dunsworth, D. Eppens, C. Erickson, L. Faoro, E. Farhi, R. Fatemi, L. Flores Burgos, E. Forati, A. G. Fowler, B. Foxen, W. Giang, C. Gidney, D. Gilboa, M. Giustina, A. Grajales Dau, J. A. Gross, S. Habegger, M. C. Hamilton, M. P. Harrigan, S. D. Harrington, O. Higgott, J. Hilton, M. Hoffmann, S. Hong, T. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, P. Juhas, D. Kafri, K. Kechedzhi, J. Kelly, T. Khattar, M. Khezri, M. Kieferová, S. Kim, A. Kitaev, P. V. Klimov, A. R. Klots, A. N. Korotkov, F. Kostritsa, J. M. Kreikebaum, D. Landhuis, P. Laptev, K.-M. Lau, L. Laws, J. Lee, K. Lee, B. J. Lester, A. Lill, W. Liu, A. Locharla, E. Lucero, F. D. Malone, J. Marshall, O. Martin, J. R. McClean, T. McCourt, M. McEwen, A. Megrant, B. Meurer Costa, X. Mi, K. C. Miao, M. Mohseni, S. Montazeri, A. Morvan, E. Mount, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, H. Neven, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, M. Y. Niu, T. E. O'Brien, A. Opremcak, J. Platt, A. Petukhov, R. Potter, L. P. Pryadko, C. Quintana, P. Roushan, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, M. J. Shearn, A. Shorter, V. Shvarts, J. Skruzny, V. Smelyanskiy, W. C. Smith, G. Sterling, D. Strain, M. Szalay, A. Torres, G. Vidal, B. Villalonga, C. Vollgraff Heidweiller, T. White, C. Xing, Z. J. Yao, P. Yeh, J. Yoo, G. Young, A. Zalcman, Y. Zhang, N. Zhu, and G. Q. AI. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614(7949):676–681, Feb 2023. ISSN 1476-4687. doi: 10.1038/s41586-022-05434-1. URL https://doi.org/10.1038/s41586-022-05434-1.

A. Alcolea and J. Resano. Bayesian neural networks to analyze hyperspectral datasets using uncertainty metrics. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–10, 2022. doi: 10.1109/TGRS.2022.3205119.

J. Allcock, C.-Y. Hsieh, I. Kerenidis, and S. Zhang. Quantum algorithms for feedforward neural networks. *ACM Transactions on Quantum Computing*, 1(1), oct 2020. ISSN 2643-6809. doi: 10.1145/3411466. URL https://doi.org/10.1145/3411466.

D. An, N. Linden, J.-P. Liu, A. Montanaro, C. Shao, and J. Wang. Quantum-accelerated multilevel Monte Carlo methods for stochastic differential equations in mathematical finance. *Quantum*, 5:481, June 2021. ISSN 2521-327X. doi: 10.22331/q-2021-06-24-481. URL https://doi.org/10.22331/q-2021-06-24-481.

S. Arora and B. Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

D. Arthur and P. Date. A hybrid quantum-classical neural network architecture for binary classification. *arXiv*, 2022. URL https://arxiv.org/abs/2201.01820.

S. Bartolucci, P. Birchall, H. Bombín, H. Cable, C. Dawson, M. Gimeno-Segovia, E. Johnston, K. Kieling, N. Nickerson, M. Pant, F. Pastawski, T. Rudolph, and C. Sparrow. Fusion-based quantum computation. *Nature Communications*, 14(1):912, Feb 2023. ISSN 2041-1723. doi: 10.1038/s41467-023-36493-1. URL https://doi.org/10.1038/s41467-023-36493-1.

T. Begušić and G. K.-L. Chan. Fast classical simulation of evidence for the utility of quantum computing before fault tolerance, June 2023. URL http://arxiv.org/abs/2306.16372. arXiv:2306.16372 [quant-ph].

R. Bellman. Dynamic programming, princeton univ. *Press Princeton, New Jersey*, 1957.

M. Benedetti, B. Coyle, M. Fiorentini, M. Lubasch, and M. Rosenkranz. Variational inference with a quantum computer. *Phys. Rev. Appl.*, 16:044057, Oct 2021. doi: 10.1103/PhysRevApplied.16.044057. URL https://link.aps.org/doi/10.1103/PhysRevApplied.16.044057.

V. Bergholm, J. Izaac, M. Schuld, C. Gogolin, S. Ahmed, V. Ajith, M. S. Alam, G. Alonso-Linaje, B. Akash-Narayanan, A. Asadi, J. M. Arrazola, U. Azad, S. Banning, C. Blank, T. R. Bromley, B. A. Cordier, J. Ceroni, A. Delgado, O. D. Matteo, A. Dusko, T. Garg, D. Guala, A. Hayes, R. Hill, A. Ijaz, T. Isacsson, D. Ittah, S. Jahangiri, P. Jain, E. Jiang, A. Khandelwal, K. Kottmann, R. A. Lang, C. Lee, T. Loke, A. Lowe, K. McKiernan, J. J. Meyer, J. A. Montañez-Barrera, R. Moyard, Z. Niu, L. J. O'Riordan, S. Oud, A. Panigrahi, C.-Y. Park, D. Polatajko, N. Quesada, C. Roberts, N. Sá, I. Schoch, B. Shi, S. Shu, S. Sim, A. Singh, I. Strandberg, J. Soni, A. Száva, S. Thabet, R. A. Vargas-Hernández, T. Vincent, N. Vitucci, M. Weber, D. Wierichs, R. Wiersema, M. Willmann, V. Wong, S. Zhang, and N. Killoran. Pennylane: Automatic differentiation of hybrid quantum-classical computations, 2022.

N. Berner, V. Fortuin, and J. Landman. Quantum bayesian neural networks, 2021. URL https://arxiv.org/abs/2107.09599.

M. E. Beverland, P. Murali, M. Troyer, K. M. Svore, T. Hoefler, V. Kliuchnikov, G. H. Low, M. Soeken, A. Sundaram, and A. Vaschillo. Assessing requirements to scale to practical quantum advantage, 2022.

K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, 94(1):015004, Feb 2022. ISSN 0034-6861, 1539-0756. doi: 10.1103/RevModPhys.94.015004.

J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

C. M. Bishop and N. M. Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

D. M. Blei, A. Kucukelbir, and J. D. McAuliffe. Variational inference: A review for statisticians. *Journal of the American Statistical Association*, 112(518):859–877, apr 2017. doi: 10.1080/01621459.2017.1285773. URL https://doi.org/10.1080%2F01621459.2017.1285773.

R. Booysen, S. Lorenz, S. T. Thiele, W. C. Fuchsloch, T. Marais, P. A. Nex, and R. Gloaguen. Accurate hyperspectral imaging of mineralised outcrops: An example from lithium-bearing pegmatites at uis, namibia. *Remote Sensing of Environment*, 269:112790, 2022.

M. J. Bremner, R. Jozsa, and D. J. Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2126):459–472, aug 2010. doi: 10.1098/rspa.2010.0301. URL https://doi.org/10.1098%2Frspa.2010.0301.

S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, editors. *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC, may 2011. doi: 10.1201/b10905. URL https://doi.org/10.1201%2Fb10905.

Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien. Quantum error mitigation. *arXiv*, 2023. URL https://arxiv.org/abs/2210.00921.

Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, aug 2019. doi: 10.1021/acs.chemrev.8b00803. URL https://doi.org/10.1021%2Facs.chemrev.8b00803.

G. Cavallaro, D. B. Heras, Z. Wu, M. Maskey, S. Lopez, P. Gawron, M. Coca, and M. Datcu. High-performance and disruptive computing in remote sensing: Hdcrs—a new working group of the grss earth science informatics technical committee [technical committees]. *IEEE Geoscience and Remote Sensing Magazine*, 10(2):329–345, 2022.

M. Cerezo, K. Sharma, A. Arrasmith, and P. J. Coles. Variational quantum state eigensolver. *npj Quantum Information*, 8(1):113, 2022.

C.-I. Chang. *Hyperspectral imaging: techniques for spectral detection and classification*, volume 1. Springer Science & Business Media, 2003.

M. Chantry, H. Christensen, P. Dueben, and T. Palmer. Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200083, 2021. doi: 10.1098/rsta.2020.0083. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2020.0083.

B. Cheng, X.-H. Deng, X. Gu, Y. He, G. Hu, P. Huang, J. Li, B.-C. Lin, D. Lu, Y. Lu, C. Qiu, H. Wang, T. Xin, S. Yu, M.-H. Yung, J. Zeng, S. Zhang, Y. Zhong, X. Peng, F. Nori, and D. Yu. Noisy intermediate-scale quantum computers. *Frontiers of Physics*, 18(2), Mar. 2023a. doi: 10.1007/s11467-022-1249-z. URL https://doi.org/10.1007/s11467-022-1249-z.

B. Cheng, X.-H. Deng, X. Gu, Y. He, G. Hu, P. Huang, J. Li, B.-C. Lin, D. Lu, Y. Lu, C. Qiu, H. Wang, T. Xin, S. Yu, M.-H. Yung, J. Zeng, S. Zhang, Y. Zhong, X. Peng, F. Nori, and D. Yu. Noisy intermediate-scale quantum computers. *Frontiers of Physics*, 18(2):21308, Apr 2023b. ISSN 2095-0462, 2095-0470. doi: 10.1007/s11467-022-1249-z.

G. Cheng, J. Han, and X. Lu. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10):1865–1883, 2017. doi: 10.1109/JPROC.2017.2675998.

G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia. Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:3735–3756, 2020. doi: 10.1109/JSTARS.2020.3005403.

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

R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, 100(1):32–74, Dec. 1928. doi: 10.1007/bf01448839. URL https://doi.org/10.1007/bf01448839.

B. Coyle, D. Mills, V. Danos, and E. Kashefi. The born supremacy: quantum advantage and training of an ising born machine. *npj Quantum Information*, 6(1):60, Dec 2020. ISSN 2056-6387. doi: 10.1038/s41534-020-00288-9.

D-Wave Systems Inc. Clarity: A roadmap for building, testing, and understanding quantum computers. https://www.dwavesys.com/media/xvjpraig/clarity-roadmap_digital_v2.pdf, 2021.

Y. Deng. Generalized evidence theory. *Applied Intelligence*, 43(3):530–543, Oct 2015. ISSN 1573-7497. doi: 10.1007/s10489-015-0661-2. URL https://doi.org/10.1007/s10489-015-0661-2.

S. Ding, H. Zhu, W. Jia, and C. Su. A survey on feature extraction for pattern recognition. *Artificial Intelligence Review*, 37:169–180, 2012.

D. P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000.

V. Dixit, R. Selvarajan, M. A. Alam, T. S. Humble, and S. Kais. Training and classification using a restricted boltzmann machine on the d-wave 2000q. *arXiv preprint arXiv:2005.03247*, 2020.

V. Dixit, R. Selvarajan, M. A. Alam, T. S. Humble, and S. Kais. Training restricted boltzmann machines with a d-wave quantum annealer. *Frontiers in Physics*, 9:589626, 2021.

DLR. German hyperspectral satellite: EnMAP, 2022. URL `https://www.enmap.org/mission/`.

ESA. European Space Agency, 2022. URL `https://eo4society.esa.int/2021/12/22/openeo-platform-a-truly-federated-and-open-environment-for-the-eo-community/`.

N. Ezzell, Z. Holmes, and P. J. Coles. The quantum low-rank approximation problem. *arXiv preprint arXiv:2203.00811*, 2022.

J. Fajtl, V. Argyriou, D. Monekosso, and P. Remagnino. Latent bernoulli autoencoder. In *International Conference on Machine Learning*, pages 2964–2974. PMLR, 2020.

E. Farhi and A. W. Harrow. Quantum supremacy through the quantum approximate optimization algorithm, 2016. URL `https://arxiv.org/abs/1602.07674`.

E. Farhi and H. Neven. Classification with quantum neural networks on near term processors, 2018.

E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum computation by adiabatic evolution. *arXiv preprint quant-ph/0001106*, 2000.

R. Fok, A. An, and X. Wang. Spontaneous symmetry breaking in neural networks, 2017.

A. G. Fowler and C. Gidney. Low overhead quantum computation using lattice surgery, 2019.

A. Fridlind, M. van Lier-Walqui, G. Elsaesser, M. Kelley, A. Ackerman, G. Cesna, and G. Schmidt. A grand challenge "uncertainty project" to accelerate advances in earth system predictability: AI-enabled concepts and applications. Technical report, United States, 2021. URL `https://doi.org/10.2172/1769643`. Contract Number: 35604.1.

Y. Gal, P. Koumoutsakos, F. Lanusse, G. Louppe, and C. Papadimitriou. Bayesian uncertainty quantification for machine-learned models in physics. *Nature Reviews Physics*, 4(9):573–577, Sep 2022. ISSN 2522-5820. doi: 10.1038/s42254-022-00498-4. URL `https://doi.org/10.1038/s42254-022-00498-4`.

J. Gawlikowski, C. R. N. Tassi, M. Ali, J. Lee, M. Humt, J. Feng, A. Kruspe, R. Triebel, P. Jung, R. Roscher, M. Shahzad, W. Yang, R. Bamler, and X. X. Zhu. A survey of uncertainty in deep neural networks, 2022.

P. Gawron and S. Lewiński. Multi-spectral image classification with quantum neural network. In *IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium*, pages 3513–3516. IEEE, 2020.

C. Gidney and A. G. Fowler. Efficient magic state factories with a catalyzed $|CCZ\rangle$ to $2|T\rangle$ transformation. *Quantum*, 3:135, Apr. 2019. ISSN 2521-327X. doi: 10.22331/q-2019-04-30-135. URL `https://doi.org/10.22331/q-2019-04-30-135`.

A. Gilyén, Y. Su, G. H. Low, and N. Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. ACM, jun 2019. doi: 10.1145/3313276.3316366. URL `https://arxiv.org/abs/1806.01838`.

A. Glos, A. Nykänen, E.-M. Borrelli, S. Maniscalco, M. A. C. Rossi, Z. Zimborás, and G. García-Pérez. Adaptive POVM implementations and measurement error mitigation strategies for near-term quantum devices. *arXiv*, 2022. URL https://arxiv.org/abs/2208.07817.

I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

Google Quantum AI. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614: 676–681, 2023. URL https://doi.org/10.1038/s41586-022-05434-1.

M. H. Gordon, M. Cerezo, L. Cincio, and P. J. Coles. Covariance matrix preparation for quantum principal component analysis. *PRX Quantum*, 3(3):030334, 2022.

H. Goto. Combinatorial optimization by simulating adiabatic bifurcations in nonlinear hamiltonian systems. *Science Advances*, 7(14), 2021. doi: 10.1126/sciadv.abe7953.

H. Goto, K. Endo, M. Suzuki, Y. Sakai, T. Kanao, Y. Hamakawa, R. Hidaka, M. Yamasaki, and K. Tatsumura. High-performance combinatorial optimization based on classical mechanics. *Science Advances*, 7(6): eabe7953, 2021.

H. Grahn and P. Geladi. *Techniques and applications of hyperspectral image analysis*. John Wiley & Sons, 2007.

L. Grover and T. Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. *arXiv preprint quant-ph/0208112*, 2002.

L. K. Grover. Synthesis of quantum superpositions by quantum computation. *Physical review letters*, 85(6): 1334, 2000.

A. Grundner, T. Beucler, P. Gentine, F. Iglesias-Suarez, M. A. Giorgetta, and V. Eyring. Deep learning based cloud cover parameterization for ICON. *Journal of Advances in Modeling Earth Systems*, 14(12), dec 2022. doi: 10.1029/2021ms002959. URL https://doi.org/10.1029/2021ms002959.

M. K. Gupta, M. Beseda, and P. Gawron. How quantum computing-friendly multispectral data can be? In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*, pages 4153–4156. IEEE, 2022.

M. K. Gupta, M. Romaszewski, and P. Gawron. Potential of quantum machine learning for processing multispectral earth observation data. 2023.

I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1):389–422, 2002. doi: 10.1023/A:1012487302797. URL https://doi.org/10.1023/A:1012487302797.

C. Gyurik and V. Dunjko. On establishing learning separations between classical and quantum machine learning with classical data, August 2022. URL http://arxiv.org/abs/2208.06339. arXiv:2208.06339 [quant-ph].

A. W. Harrow, A. Hassidim, and S. Lloyd. Quantum algorithm for linear systems of equations. *Physical Review Letters*, 103(15), oct 2009. doi: 10.1103/physrevlett.103.150502. URL https://doi.org/10.1103/physrevlett.103.150502.

V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019. doi: 10.1038/s41586-019-0980-2. URL https://doi.org/10.1038/s41586-019-0980-2.

V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212, 2019. URL https://doi.org/10.1038/s41586-019-0980-2.

L. Henriet, L. Beguin, A. Signoles, T. Lahaye, A. Browaeys, G.-O. Reymond, and C. Jurczak. Quantum computing with neutral atoms. *Quantum*, 4:327, sep 2020. doi: 10.22331/q-2020-09-21-327. URL https://doi.org/10.22331%2Fq-2020-09-21-327.

N. Herrmann, D. Arya, M. W. Doherty, A. Mingare, J. C. Pillay, F. Preis, and S. Prestel. Quantum utility – definition and assessment of a practical quantum advantage. (arXiv:2303.02138), Mar 2023. URL http://arxiv.org/abs/2303.02138. arXiv:2303.02138 [quant-ph].

T. Hoefler, T. Häner, and M. Troyer. Disentangling hype from practicality: On realistically achieving quantum advantage. *Commun. ACM*, 66:82–87, 2023. URL https://doi.org/10.1145/3571725.

M. D. Hoffman and A. Gelman. The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo, 2011. URL https://arxiv.org/abs/1111.4246.

C. Hohenegger, P. Korn, L. Linardakis, R. Redler, R. Schnur, P. Adamidis, J. Bao, S. Bastin, M. Behravesh, M. Bergemann, J. Biercamp, H. Bockelmann, R. Brokopf, N. Brüggemann, L. Casaroli, F. Chegini, G. Datseris, M. Esch, G. George, M. Giorgetta, O. Gutjahr, H. Haak, M. Hanke, T. Ilyina, T. Jahns, J. Jungclaus, M. Kern, D. Klocke, L. Kluft, T. Kölling, L. Kornblueh, S. Kosukhin, C. Kroll, J. Lee, T. Mauritsen, C. Mehlmann, T. Mieslinger, A. K. Naumann, L. Paccini, A. Peinado, D. S. Praturi, D. Putrasahan, S. Rast, T. Riddick, N. Roeber, H. Schmidt, U. Schulzweida, F. Schütte, H. Segura, R. Shevchenko, V. Singh, M. Specht, C. C. Stephan, J.-S. von Storch, R. Vogel, C. Wengel, M. Winkler, F. Ziemen, J. Marotzke, and B. Stevens. ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, 16(2): 779–811, 2023. doi: 10.5194/gmd-16-779-2023. URL https://gmd.copernicus.org/articles/16/779/2023/.

Y. Hua, J. Guo, and H. Zhao. Deep belief networks and deep learning. In *Proceedings of 2015 International Conference on Intelligent Computing and Internet of Things*, pages 1–4. IEEE, 2015.

H.-Y. Huang, M. Broughton, M. Mohseni, R. Babbush, S. Boixo, H. Neven, and J. R. McClean. Power of data in quantum machine learning. *Nature Communications*, 12(1):2631, May 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-22539-9. URL https://doi.org/10.1038/s41467-021-22539-9.

G. Hughes. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, 14(1):55–63, 1968.

E. Hüllermeier and W. Waegeman. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 110(3):457–506, Mar 2021. ISSN 1573-0565. doi: 10.1007/s10994-021-05946-3. URL https://doi.org/10.1007/s10994-021-05946-3.

T. S. Humble, A. McCaskey, D. I. Lyakh, M. Gowrishankar, A. Frisch, and T. Monz. Quantum computers for high-performance computing. *IEEE Micro*, 41:15–23, 2021. URL https://doi.org/10.1109/MM.2021.3099140.

M. P. Johansson, E. Krishnasamy, N. Meyer, and C. Piechurski. Quantum computing – a European perspective. *PRACE Technical Report*, 2021. URL https://doi.org/10.5281/zenodo.5547407.

M. Joshaghani, A. Davari, F. N. Hatamian, A. Maier, and C. Riess. Bayesian convolutional neural networks for limited data hyperspectral remote sensing image classification, 2022. URL https://arxiv.org/abs/2205.09250.

L. V. Jospin, H. Laga, F. Boussaid, W. Buntine, and M. Bennamoun. Hands-on bayesian neural networks—a tutorial for deep learning users. *IEEE Computational Intelligence Magazine*, 17(2):29–48, may 2022. doi: 10.1109/mci.2022.3155327. URL https://doi.org/10.1109%2Fmci.2022.3155327.

J. Jäger and R. V. Krems. Universal expressiveness of variational quantum classifiers and quantum kernels for support vector machines. *Nat. Commun.*, 14:576, 2023. URL https://doi.org/10.1038/s41467-023-36144-5.

K. P. Kalinin, S. Yun, A. V. Goltsev, and S. Mandrà. Simulated bifurcation for combinatorial optimization. *Physical Review Letters*, 126(10):100503, 2021. doi: 10.1103/PhysRevLett.126.100503.

P. Kansakar and F. Hossain. A review of applications of satellite earth observation data for global societal benefit and stewardship of planet earth. *Space Policy*, 36:46–54, 2016.

M. S. Karoui, F. Z. Benhalouche, Y. Deville, K. Djerriri, X. Briottet, T. Houet, A. Le Bris, and C. Weber. Partial linear nmf-based unmixing methods for detection and area estimation of photovoltaic panels in urban hyperspectral remote sensing data. *Remote Sensing*, 11(18):2164, 2019.

Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. van den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, and A. Kandala. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618:500–505, 2023. URL https://doi.org/10.1038/s41586-023-06096-3.

D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, and S. Hoyer. Machine learning–accelerated computational fluid dynamics. *Proceedings of the National Academy of Sciences*, 118(21):e2101784118, 2021. doi: 10.1073/pnas.2101784118. URL https://www.pnas.org/doi/abs/10.1073/pnas.2101784118.

C. Koller, G. Kauermann, and X. X. Zhu. Going beyond one-hot encoding in classification: Can human uncertainty improve model performance?, 2022.

M. A. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *AIChE journal*, 37(2):233–243, 1991.

J. Kubečka, Y. Knattrup, M. Engsvang, A. B. Jensen, D. Ayoubi, H. Wu, O. Christiansen, and J. Elm. Current and future machine learning approaches for modeling atmospheric cluster formation. *Nat. Comput. Sci*, 3:495–503, 2023. URL https://doi.org/10.1038/s43588-023-00435-0.

O. Kyriienko, A. E. Paine, and V. E. Elfving. Solving nonlinear differential equations with differentiable quantum circuits. *Phys. Rev. A*, 103:052416, May 2021. doi: 10.1103/PhysRevA.103.052416. URL https://link.aps.org/doi/10.1103/PhysRevA.103.052416.

L. Lapworth. Implicit hybrid quantum-classical CFD calculations using the HHL algorithm. 2022.

R. LaRose, A. Tikku, É. O'Neel-Judy, L. Cincio, and P. J. Coles. Variational quantum state diagonalization. *npj Quantum Information*, 5(1):57, 2019.

D. Layden, G. Mazzola, R. V. Mishmash, M. Motta, P. Wocjan, J.-S. Kim, and S. Sheldon. Quantum-enhanced markov chain monte carlo, 2022. URL https://arxiv.org/abs/2203.12497.

Y. Li, M. Tian, G. Liu, C. Peng, and L. Jiao. Quantum optimization and quantum learning: A survey. *IEEE Access*, 8:23568–23593, 2020. doi: 10.1109/ACCESS.2020.2970105.

D. Litinski. A game of surface codes: Large-scale quantum computing with lattice surgery. *Quantum*, 3:128, mar 2019. doi: 10.22331/q-2019-03-05-128. URL https://doi.org/10.22331%2Fq-2019-03-05-128.

Y. Liu, S. Arunachalam, and K. Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 17(9):1013–1017, 2021.

S. Lloyd, M. Mohseni, and P. Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9): 631–633, 2014.

W. Lv and X. Wang. Overview of hyperspectral image classification. *Journal of Sensors*, 2020, 2020.

W. Ma, C. Gong, Y. Hu, P. Meng, and F. Xu. The hughes phenomenon in hyperspectral classification based on the ground spectrum of grasslands in the region around qinghai lake. In *International Symposium on Photoelectronic Detection and Imaging 2013: Imaging Spectrometer Technologies and Applications*, volume 8910, pages 363–373. SPIE, 2013.

D. J. MacKay. Bayesian methods for adaptive models. dissertation (ph.d.). *California Institute of Technology*, 1, 1992. doi: doi:10.7907/H3A1-WM07. URL `https://resolver.caltech.edu/CaltechETD:etd-01042007-131447`.

C. Mallet and F. Bretar. Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64(1):1–16, 2009.

J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812, November 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-07090-4. URL `https://www.nature.com/articles/s41467-018-07090-4`. Number: 1 Publisher: Nature Publishing Group.

G. A. Meehl, C. A. Senior, V. Eyring, G. Flato, J.-F. Lamarque, R. J. Stouffer, K. E. Taylor, and M. Schlund. Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, 6(26):eaba1981, 2020. doi: 10.1126/sciadv.aba1981. URL `https://www.science.org/doi/abs/10.1126/sciadv.aba1981`.

A. Miroszewski, J. Mielczarek, G. Czelusta, F. Szczepanek, B. Grabowski, B. L. Saux, and J. Nalepa. Detecting Clouds in Multispectral Satellite Images Using Quantum-Kernel Support Vector Machines. 2 2023.

A. Montanaro. Quantum speedup of monte carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150301, sep 2015. doi: 10.1098/rspa.2015.0301. URL `https://doi.org/10.1098%2Frspa.2015.0301`.

M. Moroni, E. Lupo, E. Marra, and A. Cenedese. Hyperspectral image analysis in environmental monitoring: setup of a new tunable filter platform. *Procedia Environmental Sciences*, 19:885–894, 2013.

S. Mücke, R. Heese, S. Müller, M. Wolter, and N. Piatkowski. Feature selection on quantum computers. *Quantum Machine Intelligence*, 5(1):11, 2023. doi: 10.1007/s42484-023-00099-z. URL `https://doi.org/10.1007/s42484-023-00099-z`.

K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012. ISBN 0262018020.

J. Nalepa, B. Le Saux, N. Longépé, L. Tulczyjew, M. Myller, M. Kawulok, K. Smykala, and M. Gumiela. The hyperview challenge: Estimating soil parameters from hyperspectral images. In *2022 IEEE International Conference on Image Processing (ICIP)*, pages 4268–4272. IEEE, 2022.

R. M. Neal. Bayesian learning for neural networks. 1995.

A. Olivier, M. D. Shields, and L. Graham-Brady. Bayesian neural networks for uncertainty quantification in data-driven materials modeling. *Computer Methods in Applied Mechanics and Engineering*, 386: 114079, 2021. ISSN 0045-7825. doi: https://doi.org/10.1016/j.cma.2021.114079. URL https://www.sciencedirect.com/science/article/pii/S0045782521004102.

S. Otgonbaatar and M. Datcu. A quantum annealer for subset feature selection and the classification of hyperspectral images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14:7057–7065, 2021a. doi: 10.1109/JSTARS.2021.3095377.

S. Otgonbaatar and M. Datcu. Classification of remote sensing images with parameterized quantum gates. *IEEE Geoscience and Remote Sensing Letters*, pages 1–5, 2021b. doi: 10.1109/LGRS.2021.3108014.

S. Otgonbaatar and M. Datcu. Natural embedding of the stokes parameters of polarimetric synthetic aperture radar images in a gate-based quantum computer. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1–8, 2021c. doi: 10.1109/TGRS.2021.3110056.

M. Pandey, M. Fernandez, F. Gentile, O. Isayev, A. Tropsha, A. C. Stern, and A. Cherkasov. The transformational role of GPU computing and deep learning in drug discovery. *Nature Machine Intelligence*, 4(3):211–221, Mar 2022. ISSN 2522-5839. doi: 10.1038/s42256-022-00463-x. URL https://doi.org/10.1038/s42256-022-00463-x.

M. Paoletti, J. Haut, J. Plaza, and A. Plaza. Deep learning classifiers for hyperspectral imaging: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 158:279–317, 2019. ISSN 0924-2716. doi: https://doi.org/10.1016/j.isprsjprs.2019.09.006. URL https://www.sciencedirect.com/science/article/pii/S0924271619302187.

K. Pearson and O. Lines. Planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci*, 2:559–572, 1901.

A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, Jul 2014. ISSN 2041-1723. doi: 10.1038/ncomms5213. URL https://doi.org/10.1038/ncomms5213.

M. Plesch and Č. Brukner. Quantum-state preparation with universal gate decompositions. *Physical Review A*, 83(3):032302, 2011.

I. Pogorelov, T. Feldker, C. D. Marciniak, L. Postler, G. Jacob, O. Krieglsteiner, V. Podlesnic, M. Meth, V. Negnevitsky, M. Stadler, B. Höfer, C. Wächter, K. Lakhmanskiy, R. Blatt, P. Schindler, and T. Monz. Compact ion-trap quantum computing demonstrator. *PRX Quantum*, 2:020343, Jun 2021. doi: 10.1103/PRXQuantum.2.020343. URL https://link.aps.org/doi/10.1103/PRXQuantum.2.020343.

J. Preskill. Fault-tolerant quantum computation, 1997. URL https://arxiv.org/abs/quant-ph/9712048.

J. Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018a. ISSN 2521-327X. doi: 10.22331/q-2018-08-06-79. URL https://doi.org/10.22331/q-2018-08-06-79.

J. Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018b.

Qinsider. An action plan for a universal quantum computer, 2023. URL https://thequantuminsider.com/2023/05/03/germany-announces-3-billion-euro-action-plan-for-a-universal-quantum-computer/.

QRM. Quantum roadmap, 2023. URL https://databaseline.tech/quantum.html.

Y. Quek, D. S. França, S. Khatri, J. J. Meyer, and J. Eisert. Exponentially tighter bounds on limitations of quantum error mitigation. (arXiv:2210.11505), Oct 2022. URL http://arxiv.org/abs/2210.11505. arXiv:2210.11505 [math-ph, physics:quant-ph].

S. Rainjonneau, I. Tokarev, S. Iudin, S. Rayaprolu, K. Pinto, D. Lemtiuzhnikova, M. Koblan, E. Barashov, M. Kordzanganeh, M. Pflitsch, and A. Melnikov. Quantum algorithms applied to satellite mission planning for earth observation. (arXiv:2302.07181), 2023.

M. Rast and T. H. Painter. Earth observation imaging spectroscopy for terrestrial systems: An overview of its history, techniques, and applications of its missions. *Surveys in Geophysics*, 40(3):303–331, 2019.

P. Rebentrost, M. Mohseni, and S. Lloyd. Quantum support vector machine for big data classification. *Phys. Rev. Lett.*, 113:130503, Sep 2014. doi: 10.1103/PhysRevLett.113.130503. URL https://link.aps.org/doi/10.1103/PhysRevLett.113.130503.

M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences*, 114(29):7555–7560, 2017. doi: 10.1073/pnas.1619152114. URL https://www.pnas.org/doi/abs/10.1073/pnas.1619152114.

D. Ries, J. Adams, and J. Zollweg. Target detection on hyperspectral images using mcmc and vi trained bayesian neural networks. In *2022 IEEE Aerospace Conference (AERO)*, pages 1–9, 2022. doi: 10.1109/AERO53065.2022.9843344.

RM. An action plan for a quantum technology, 2023. URL https://www.researchandmarkets.com/reports/5317365/quantum-technology-market-by-computing?utm_source=GNOM&utm_medium=PressRelease&utm_code=xt7dw4&utm_campaign=1668311+-+Global+Quantum+Technology+Market+Research+Report+2022-2027%3a+Quantum+Dots+Market+will+Reach+%2413.25+Billion+by+2027%2c+Growing+a+25.1%25+CAGR+and+Led+by+Displays&utm_exec=chdo54prd.

J. Romero, J. P. Olson, and A. Aspuru-Guzik. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2(4):045001, 2017.

C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. 2018. doi: 10.48550/ARXIV.1811.10154. URL https://arxiv.org/abs/1811.10154.

M. Schuld and N. Killoran. Quantum machine learning in feature hilbert spaces. *Phys. Rev. Lett.*, 122:040504, Feb 2019. doi: 10.1103/PhysRevLett.122.040504. URL https://link.aps.org/doi/10.1103/PhysRevLett.122.040504.

M. Schuld and N. Killoran. Is quantum advantage the right goal for quantum machine learning? *PRX Quantum*, 3(3), jul 2022. doi: 10.1103/prxquantum.3.030101. URL https://doi.org/10.1103/prxquantum.3.030101.

M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe. Circuit-centric quantum classifiers. *Physical Review A*, 101(3), Mar 2020. ISSN 2469-9934. doi: 10.1103/physreva.101.032308. URL http://dx.doi.org/10.1103/PhysRevA.101.032308.

M. Sensoy, L. Kaplan, and M. Kandemir. Evidential deep learning to quantify classification uncertainty, 2018. URL https://arxiv.org/abs/1806.01768.

P. Stano and D. Loss. Review of performance metrics of spin qubits in gated semiconducting nanostructures. *Nature Reviews Physics*, 4(10):672–688, aug 2022. doi: 10.1038/s42254-022-00484-w. URL https://doi.org/10.1038%2Fs42254-022-00484-w.

B. Stevens, M. Satoh, L. Auger, and et al. DYAMOND: the DYnamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, 6, 2019. doi: https://doi.org/10.1186/s40645-019-0304-z.

D. Stilck França and R. García-Patrón. Limitations of optimization algorithms on noisy quantum devices. *Nature Physics*, 17(11):1221–1227, November 2021. ISSN 1745-2481. doi: 10.1038/s41567-021-01356-3.

D.-W. Systems. D-wave quantum computing. `https://www.dwavesys.com/`, 2023.

E. Tang. Quantum principal component analysis only achieves an exponential speedup because of its state preparation assumptions. *Physical Review Letters*, 127(6):060503, 2021a.

E. Tang. Quantum Principal Component Analysis Only Achieves an Exponential Speedup Because of Its State Preparation Assumptions. *Physical Review Letters*, 127(6):060503, August 2021b. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.127.060503. URL `https://link.aps.org/doi/10.1103/PhysRevLett.127.060503`.

S. Thanasilp, S. Wang, M. Cerezo, and Z. Holmes. Exponential concentration and untrainability in quantum kernel methods, August 2022. URL `http://arxiv.org/abs/2208.11060`. arXiv:2208.11060 [quant-ph, stat].

S. Theodoridis and K. Koutroumbas. *Pattern recognition*. Elsevier, 2006.

J. Tindall, M. Fishman, M. Stoudenmire, and D. Sels. Efficient tensor network simulation of IBM's kicked Ising experiment, June 2023. URL `http://arxiv.org/abs/2306.14887`. arXiv:2306.14887 [quant-ph].

D. Tuia, K. Schindler, B. Demir, G. Camps-Valls, X. X. Zhu, M. Kochupillai, S. Džeroski, J. N. van Rijn, H. H. Hoos, F. D. Frate, M. Datcu, J.-A. Quiané-Ruiz, V. Markl, B. L. Saux, and R. Schneider. Artificial intelligence to advance earth observation: a perspective. *arXiv*, 2023. URL `https://arxiv.org/abs/2305.08413`.

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

D. Watson-Parris. Machine learning for weather and climate are worlds apart. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200098, 2021. doi: 10.1098/rsta.2020.0098. URL `https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2020.0098`.

J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar. Integrating scientific knowledge with machine learning for engineering and environmental systems, 2020. URL `https://arxiv.org/abs/2003.04919`.

R. Xia and S. Kais. Hybrid quantum-classical neural network for calculating ground state energies of molecules. *Entropy*, 22(8):828, jul 2020. doi: 10.3390/e22080828. URL `https://doi.org/10.3390/e22080828`.

C. Zhang, J. Butepage, H. Kjellstrom, and S. Mandt. Advances in variational inference, 2017. URL `https://arxiv.org/abs/1711.05597`.

Q. Zhao, L. Yu, Z. Du, D. Peng, P. Hao, Y. Zhang, and P. Gong. An overview of the applications of earth observation satellite data: impacts and future trends. *Remote Sensing*, 14(8):1863, 2022.

C. Zoufal. Generative quantum machine learning, 2021.