Quantum Information Science In High Energy Physics

9th Large Hadron Collider Physics Conference
7-12 June, 2021
Outline

- Motivations for Quantum Technologies
- Sensors
- Communications
- Quantum Computation
- Quantum / Machine Learning
- Summary and outlook

Impossible to do full justice to the topic in the allocated time.

Thanks to Heather Gray, Cristian Pena, Gabriel Perdue, Koji Terashi, Sofia Vallecorsa, Si Xie for material and input.
Why go Quantum

Scientific reasons:
- Classical computers are deterministic, quantum systems are probabilistic in nature.
- Quantum systems have rich evolution dynamics.
- Operating sensors at the highest level of precision possible.
- Quantum systems are fragile, cannot be tempered with, without information loss.

Sociological reasons:
- Quantum technologies are expanding fast.
- Outcomes with tremendous impact in industry.
- Funding for quantum technologies R&D.
- Forging the quantum minds.
Range of Applications

Quantum Information Science
- Quantum sensing
- Quantum networking
- Quantum cryptography
- Quantum simulation

Diagram H.M. Gray
Quantum Sensors

- Quantum systems are highly sensitive systems.
- Superconducting Radio-Frequency (SRF) cavities, superconducting nanowire single photon detectors (SNSPD), superconducting qubits,… operating in quantum regime can provide sensitive complementary to other experiments.
- Synergy with high precision readout R&D in quantum telecommunication.
Quantum Communications

- Secure communications, connect quantum devices, …
- Based on quantum teleportation protocol.
- No “faster than the speed of light” communication.
- Destructive eavesdropping mechanism.

DOI: 10.1103/PRXQuantum.1.020317
Quantum Computing Devices

Multiple platforms to implement quantum circuits available.
Multiple types of quantum substrates: superconducting qubits, ion-trap, photonics, ...
System size still somehow limited, evolving fast.
Quantum Volume is the metric for comparison (4M for IonQ)

Quantum annealing device by D-Wave, digital annealer by Fujitsu.
Thousands+ qubits available.


QIS in HEP, LHCP 2021, J-R Vlimant
qubit and qubit

Quantum Circuits
Series of quantum gates operating on a set of quantum states.

Quantum Annealing
Evolution of a quantum system to a low T Gibbs state

Limited to a small phase of all possible algorithms.
Not a generic computer.
Only tackle problems that can be cast in the system specific formulation.
Adiabatic Quantum Annealing

- System setup with trivial Hamiltonian $H(0)$ and ground state
- Evolve adiabatically the Hamiltonian towards the desired Hamiltonian $H_p$
- **Adiabatic theorem**: with a slow evolution of the system, the state stays in the ground state.

![Diagram showing adiabatic quantum annealing](image)

$H(t) = A(t)H(0) + B(t)H_p$

arxiv:0001106, arxiv:0104129
Quantum Circuit

Qubits two states quantum system (qutrits, etc possible). Quantum gates are the elementary operations on set of qubits. Quantum circuits are composed of gates, forming a unitary transformation.

Many quantum algorithms proposed with a provable advantage over classical counterpart:

- Prime factorization (Shore’s algorithm): [arxiv:9508027](https://arxiv.org/abs/9508027)
- Database search (Grover’s algorithm): [doi:10.1145/237814.237866](https://doi.org/10.1145/237814.237866)
- Linear system solver (HHL): [arxiv:0811.3171](https://arxiv.org/abs/0811.3171)
- …


Software and Toolkit available: [qiskit](https://qiskit.org), [cirq](https://cirq.readthedocs.io), [quil](https://projectq.readthedocs.io), [xacc](https://x-openqasm.github.io/xacc/).
QC in HEP

• (re)cast combinatorial problems in a quantum formulation
  ➡ Vertexing : arxiv:1903.08879, ...
  ➡ Jet clustering : arxiv:1908.08949, arxiv:2012.14514, arxiv:2101.05618, ...

• Using the probabilistic character of quantum device
  ➡ Unfolding : arxiv:1908.08519, ...

The multi-jet QUBO objective function is

\[
O_{\text{QUBO}}(\{x_{ij}\}) = \sum_{r=1}^{M} \sum_{i,j=1}^{N} \left( \frac{\vec{p}_i \cdot \vec{p}_j - E_i E_j \cos R}{1 - \cos R} \right) x_{ir} x_{jr} \\
+ \Lambda^2 \sum_{i=1}^{N} \left( 1 - \sum_{r=0}^{M} x_{ir} \right)^2.
\]

QIS in HEP, LHCP 2021, J-R Vlimant
Experimental Limitations

Several aspects of current quantum circuit devices are limiting

- **Decoherence**: noise and imperfect isolation of the quantum state perturbs the computation.
  - Mitigated in better devices, and error correction mechanism.
- **Connectivity**: most formulation would require full connectivity, while most device only implement limited qubit neighborhood
  - Mitigated in newer devices, and embedding or using qubit replica.
- **System size**: casting of computationally challenging problems often requires large number of qubits. Most device can still be simulated on classical computers.
  - Mitigated in newer devices.

\[ \text{doi:10.1103/PhysRevLett.121.220502} \]
Quantum Machine Learning

Deep learning is computing intensive, and de-facto enabled by use of GPU. People are looking for ways to leverage possible quantum advantage to accelerate machine learning techniques.

Main algorithms used in recent studies

- Variational Quantum Circuits (VQC)
- Quantum Support Vector Machine (QSVM)
- Quantum Restricted Boltzmann Machine (QRBM)
- Quantum Adiabatic Machine Learning (QAML)
- Quantum Generative Adversarial Network (QGAN)

Field in constant evolution. Embedding is crucial. Deep implications of kernel methods.

Software and toolkit available pennylane, tf-quantum

a. Training the embedding
b. Classification
QML in HEP

Applied where “classical machine learning” has already been applied

• Event reconstruction

• Classification:

➡ Quantum Support Vector Machine with a rich space of kernels.
➡ No hard evidence that QML > ML though (people keep looking).
➡ Not much littérature on “quantum generative models for HEP” yet.
➡ There might be advantages in doing QML on Quantum data.
Quantum computing, quantum algorithms, quantum machine learning are evolving fast. Several aspects are key to experimentation

• **Noise correction**: finding a way to mitigate the effect of decoherence and experimental noise in evaluating the quantum circuits.

• **Circuit optimization**: finding circuit layout that go beyond the arbitrary circuit Ansatz that most proposed algorithms make use of.

• **Circuit compression**: compiling a circuit of logical gates into an optimal hardware efficient physical operators. As short as possible due to decoherence.

Applicable methods do not necessarily involved machine learning.
Summary

• Quantum sensors/devices offer novel detection capabilities.
• Quantum communication is becoming reality.
• Quantum computers hold a great deal of potential.
• Devices are getting to useful size.
• Too early for practical usage of quantum computing in HEP; and yet a promising place for studying innovative ways of doing scientific computation.
Initiatives:
- Fermilab Superconducting Quantum Materials and Systems Center: [https://sqms.fnal.gov](https://sqms.fnal.gov)
- Fermilab Quantum Institute: [https://quantum.fnal.gov](https://quantum.fnal.gov)
- NSF QLCI: [https://hqan.illinois.edu/about](https://hqan.illinois.edu/about), [https://www.colorado.edu/research/qsense/](https://www.colorado.edu/research/qsense/)
- Alliance for Quantum Technologies: [https://inqnet.caltech.edu](https://inqnet.caltech.edu)
- CERN Quantum Technology Initiative: [https://quantum.cern](https://quantum.cern)
- Quantum Flagship: [https://qt.eu](https://qt.eu)

Lectures:
- Elias Fernandex-Combarro alvarez: [https://indico.cern.ch/event/970903/](https://indico.cern.ch/event/970903/)
- Maria Schuld: [https://indico.cern.ch/event/975609/](https://indico.cern.ch/event/975609/)
- Heather Gray: [https://indico.cern.ch/event/870515/](https://indico.cern.ch/event/870515/)

Reviews:
- QML in HEP: [https://doi.org/10.1088/2632-2153/abc17d](https://doi.org/10.1088/2632-2153/abc17d)
Quantum Derivatives

Objective based on quantum measurement. Parameters of a quantum circuits as weights. Trainable circuits for quantum machine learning.

Quantum Machine Learning [1611.09347]
Quantum Machine Learning Models are Kernel Methods [2101.11020]

Stochastic Gradient Descent

• Application of one gradient descent is expensive. Can be prohibitive with large datasets
• Following the gradient update from each and every sample of a dataset leads to tensions
  • In binary classification, samples from opposite categories would have “opposite gradients”
• Gradients over multiple samples are independent, and can be computationally parallelyzed
→ Estimate the effective gradient over a batch of samples

$$\nabla_{\text{eff}} f(x) = \frac{1}{N} \sum_{i \in \text{batch}} \nabla_i f(x)$$
Gradient Descent Optimization

For a differentiable loss function $f$, the first Taylor expansion gives

$$f(x + \varepsilon) = f(x) + \varepsilon \nabla f(x)$$

The direction to locally maximally decrease the function value is anti-collinear to the gradient

$$\varepsilon = -\gamma \nabla f(x)$$

Amplitude of the step $\gamma$ to be taken with care to prevent overshooting.
Bayesian Optimization

- Applicable to optimize function **without close form** and that are **expensive to call** (numerical gradient impractical)
- Approximate the objective function with **Gaussian processes** (GP)
- Start at random points, then sample according to optimized acquisition function
  - Expected improvement
    \[-EI(x) = -E(f_{GP}(x) - f(x_{best}))\]
  - Lower confidence bound
    \[LCB(x) = \mu_{GP}(x) + \kappa \sigma_{GP}(x)\]
  - Probability of improvement
    \[-PI(x) = -P(f_{GP}(x) \geq f(x_{best}) + \kappa)\]
Evolutionary Algorithms

- Applicable to function in high dimensions, with a non regular landscape
- Start from random population
- Estimate fittest fraction of individuals
- Bread and mutate individuals

- Direction of optimization is given by the cross-over and mutation definition
- Multiple over algorithms: particle swarn, ...
Simulated Annealing

- Monte-Carlo based method to find ground state of energy functions
- Random walk across phase space
  - accepting descent
  - accepting ascent with probability $e^{-\Delta E/kT}$
- Decrease $T$ with time
Non Analytical SGD

- Some valuable loss function might not be analytical and their gradients cannot be derived
- Used finite element method to estimate the gradient numerically
  \[ \nabla f(x) = \frac{f(x + \varepsilon) - f(x)}{\varepsilon} \]
- Method can be extended to using more sampling and better precision
- Quite expensive computationally in number of function calls and impractical in large dimension
- Robust methods available in most program library
Second Order Methods

- Newton-Raphson method defines a recursive procedure to find the root of a function, using its gradient.
- Finding optimum is equivalent to finding roots of the gradient, hence applying NR method to the gradient using the Hessian

\[ f(x + \varepsilon) = f(x) + \varepsilon \nabla f(x) + \frac{1}{2} \varepsilon^T H(x) \varepsilon \]

\[ \varepsilon \sim - H(x)^{-1} \nabla f(x) \]

- Convergence guaranteed in certain conditions
- Alternative numerical methods tackle the escape of saddle points and computation issue with inverting the Hessian
- In deep learning “hessian-free” methods are prohibitive computationally wise
How could quantum computing help with ML?

- **Data**
  - speed up sampling from data distributions
  - use fewer data samples (e.g., Arunachalam 1701.06806)

- **Optimisation**
  - speed up optimisation (Wiebe et al. 1204.5242, Rebentrost et al. 1307.0471, Denil & Freitas ~ 2012 cs.ubc.ca/~nando/)
  - find better solutions

- **Model**
  - speed up existing models (Pararo et al. 1401.4997, Low et al. 1402.7359, Allcock et al. 1812.03089)
  - design better models (Amin et al. 1601.02036, Benedetti et al. 1906.07682)

https://indico.cern.ch/event/975609/