# Intro to Quantum Computing ML for Neutrino Astronomy 

August 24th, 2023<br>NPML 2023<br>Tufts University

Pavel Zhelnin<br>With Jeff L. Et al.<br>Harvard University<br>pzhelnin@g.harvard.edu

## Machine Learning

- In the 1990s the initial exploration of ML in particle physics began: SNO experiment
- At the beginning, these neural networks did not outperform other statistical techniques but they did demonstrate capabilities
- However as expertise grew ML techniques began to surpass traditional reco
- Now ML has played a role in nearly every particle physics discovery and measurement since


## Observation of high-energy neutrinos from the Galactic plane <br>  +380 authors Authors Info \& Affilitions



## Machine Learning

- In the 1990s the initial exploration of ML in particle physics began: SNO experiment
- At the beginning, these neural networks did not outperform other statistical techniques but they did demonstrate capabilities
- However as expertise grew ML techniques began to surpass traditional reco
- Now ML has played a role in nearly every particle physics discovery and measurement since


## Observation of Measurement of the properties of Higgs boson plane production at $\sqrt{s}=13 \mathrm{TeV}$ in the $H \rightarrow \gamma \gamma$ channel using $139 \mathbf{~ f b}^{-1}$ of $p p$ collision data with the ATLAS experiment



## Machine Learning

- In the 1990s the initial exploration of ML in particle physics began: SNO experiment
- At the beginning, these neural networks did not outperform other statistical techniques but they did demonstrate capabilities
- However as expertise grew ML techniques began to surpass traditional reco
- Now ML has played a role in nearly every particle physics discovery and measurement since
- What's the next iteration?


## Observation of Measurement of the nronerties of Hiacs hoson plane production a The Results of a Neural Network Statistical Event channel usin Class Analysis <br> the ATLAS en s.ice Publased 1996 P Prysice



## A Growing Data Challenge

- ML is essential in analyzing a commonality among experiments now: large data size
- In fact cuts are needed to manage modern experiments
- Even after cuts, datasets are huge
- CERN produces > 300 TB of data per day
- IceCube produces ~ 1 TB
- Templates based on our current understanding filters data
- Furthermore, next generation experiments will increase data output by an order of magnitude

- Could new physics be hiding in cut data?


## A Growing Challenge cont.

- We could be missing new physics due to un-modeled interactions (streetlight effect)
- Allowing additional data may be necessary for new physics
- For this, a paradigmatic shift is needed in data management to process trigger-level data



## Quanturm Computing

- A computer whose computations can only be described with the laws of quantum theory
- Exponentially large Hilbert space
- Entanglement
- Superposition
- Interference
- $2^{N}$ advantage over classical computers
- E.g. 8 classical bits $\rightarrow 3$ "qubits", 64 bits $\rightarrow 6$ "qubits", can store all of Google Drive cloud storage in $\sim 60$ qubits


## Basics of Quantum Computing

- Qubits = basic unit of information in a QC (akin to a bit)
- Often represented by a Bloch sphere



## Basics of Quantum Computing

- Qubits = basic unit of information in a QC (akin to a bit)
- Often represented by a Bloch sphere

- Quantum gates $=$ most basic operation that can performed on a qubit (or set of qubits)
- Two basic quantum gates: Hadamard/CNOT
- Hadamard creates superpositions
- CNOT entangles
- Combination makes a Bell State


| input | output | input | output |
| :---: | :---: | :---: | :---: |
| y | $x y+x$ | $x$ y | $x$ y+x |
| \|0) |0| | $\|0\rangle\|0\rangle$ | 00 | 00 |
| \|0) |1] | \|0) |1] | 01 | 01 |
| \|1) |0> | \|1) |1) | 10 | 1 |
| \|1) |1) | \|1) $\|0\rangle$ | 11 | 10 |

## Basics of Quantum Computing

- Qubits = basic unit of information in a QC (akin to a bit)
- Often represented by a Bloch sphere

- Quantum gates = most basic operation that can performed on a qubit (or set of qubits)
- Two basic quantum gates: Hadamard/CNOT
- Hadamard creates superpositions
- CNOT entangles
- Combination makes a Bell State


| input | output | input | output |
| :---: | :---: | :---: | :---: |
| y | $x \mathrm{y}+\mathrm{x}$ | $x$ y | $x$ y+x |
| \|0) |0才 | $\|0\rangle\|0\rangle$ | 00 | 00 |
| \|0) |1) | \|0) |1] | 01 | 01 |
| \|1) |0> | \|1) |1> | 10 | 11 |
| \|1) |1] | \|1) $\|0\rangle$ | 11 | 10 |

- Quantum circuit = a model for quantum computation in which a sequence of quantum gates are applied to a set of $n$ qubits



## A Data Processing Pipeline Using QC

- We don't want to just store data on a QC, we want to process it as well
- Its runtime is costly to do: classical $\rightarrow$ quantum or quantum $\rightarrow$ classical transfers of data
- We want a fully "quantum pipeline", no classical preprocessing
E.g. ...

Classical Data $\rightarrow$

$\rightarrow$ Classification

## Rest of the tallk: QML with a Variational Quantum Circuit

A VOC is a low depth, low width choice suitable for ML applications on current quantum computers.


## Rest of the tallk: QML



## Data Encoding

- "Data encoding is often the most crucial step in QML with classical data: it influences potential quantum advantage, learning performance and runtime."
- Most other QML encoding schemes involve some classical preprocessing then using either amplitude/basis encoding (arXiv:2012.11560, arXiv.1907.00397, arXiv.2010.07335)
- We want to avoid classical preprocessing while still working within the constraints of Near-Intermediate Scale Quantum (NISQ) computers


## Background: Data Encoding

- Amplitude encoding can store information with $2^{N}$ efficiency
- Susceptible to decoherence
$\mathbf{x}=\left[\begin{array}{c}\frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ -\frac{1}{2}\end{array}\right] \underbrace{\left|\psi_{\mathbf{x}}\right\rangle=\sum_{i=0}^{N-1} x_{i}|i\rangle .}_{|\mathbf{x}\rangle=\frac{1}{2}|00\rangle+\frac{1}{2}|01\rangle-\frac{1}{2}|10\rangle-\frac{1}{2}|11\rangle}$


## Background: Data Encoding

- Amplitude encoding can store information with $2^{N}$ efficiency
- Susceptible to decoherence
- Basis encoding is the simplest encoding
- No quantum advantage, a 1 to 1 mapping

$$
x=1011 \rightarrow|1011\rangle
$$

$$
|\mathbf{x}\rangle=\frac{1}{2}|00\rangle+\frac{1}{2}|01\rangle-\frac{1}{2}|10\rangle-\frac{1}{2}|11\rangle
$$

## Background: Data Encoding

- Amplitude encoding can store information with $2^{N}$ efficiency
- Susceptible to decoherence
- Basis encoding is the simplest encoding
- No quantum advantage, a 1 to 1 mapping

$$
x=1011 \rightarrow|1011\rangle
$$

- Angle encoding
- Rotations around principle axes of Bloch sphere
- Principle encoding scheme by others in QML HEP

$$
|\mathbf{x}\rangle=\bigotimes^{n} R\left(\mathbf{x}_{i}\right)\left|0^{n}\right\rangle
$$

Rotation by pi around y axis on Bloch sphere
$\left.\left.\mathbf{x}=\left[\begin{array}{l}\pi \\ \pi \\ \pi\end{array}\right] \rightarrow \right\rvert\, 工 \downarrow\right\}$

## Recap so far

- We want to use quantum computers because they can handle computational challenges of increase data loads of upcoming experiments
- This way we can investigate more data
- We don't want to reduce the complexity of our data (no PCA, no classical dimension reduction)
- Traditional quantum encoding schemes either
 don't use quantum advantage or are overly susceptible to decoherence
- So we want a near lossless quantum encoding scheme with quantum advantage


## Quantum Random Access Codes (QRAC)

- We want a resilient data encoding scheme that still occupies some advantage over classical systems
- Encodes digital information in correlations between qubits
- ~a $1.5^{N}$ advantage over classical systems
- It's resilient: for $\mathrm{N} \geq 18$ nearly lossless (0.999) recovery rate
- For $\mathrm{N} \geq 14$ ORAC has greater success than classical counterparts

Random access codes via quantum contextual redundancy


Giancarlo Gatti, ${ }^{1,2, *}$ Daniel Huerga, ${ }^{1, \dagger}$ Enrique Solano, ${ }^{1,3,4,5, \ddagger}$ and Mikel Sanz ${ }^{1,4,5, ~ § ~}$

## Digital Quantum Encoding

- For example, I want to encode the bit string '1011'
- Option 0: One-to-one mapping to z-spin

- Option 1: Two-to-one mapping to z-spins (have two options to encode)



## Correlation-based digital encoding

- The set $\{X, Y, Z\}^{\otimes N}$ are parity observables (POs) where N is number of qubits
- Measuring using POs always yield a $\pm 1$
- Instead of assigning a bit to each PO, assign a bit to each pair of POs, $0 \leftrightarrow=$ and $1 \leftrightarrow \neq$
- We create eigenstates of sets of commuting POs, these are our compressed states that when measured later recover our input data

| $b^{\text {target }}$ | 1 |  | 0 |  | 1 |  | 1 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| PO relationship | $\neq$ |  | = |  | $\neq$ |  | $\neq$ |  | ZZ |
| POs | $X X$ | XY | XZ | YX | $Y Y$ | $Y Z$ | ZX | ZY | NA |
| Option 1 | +1 | -1 | +1 | +1 | +1 | -1 | +1 | -1 | NA |
| Option 2 | -1 | +1 | -1 | -1 | -1 | +1 | -1 | +1 | NA |

## Example for encoding '1011' in two qubits

0 Let us use n-qubit systems
(and draw figures for $\mathrm{n}=4$ )
Alice has $\mathbf{m}$ bits of data

| 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 |
| 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 |
| 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |

that she wants Bob to randomly access $m=\left(3^{n}-1\right) / 2$

2 Alice maps her data to outcomes of n-body Pauli observables
I
$\left\{X_{1} X_{2} X_{3} X_{4}\right.$ yields +1 $X_{1} X_{2} X_{3} Y_{4}$ yields -1 yields -1
(more than +1) ${ }_{1}\left\{X_{1} X_{2} X_{3} Z_{4}\right.$ yields -1 $X_{1} X_{2} Y_{3} X_{4}$ yields +1
$\left\{X_{1} X_{2} Y_{3} Y_{4} \begin{array}{c}\text { yields-1 } \\ \text { (more than }\end{array}\right.$
$0 \begin{cases}X_{1} X_{2} Y_{3} Y_{4} & \begin{array}{l}\text { (more than }+1) \\ X_{1} X_{2} Y_{3} Z_{4} \\ \text { yeids } \\ \text { (more than }\end{array}\end{cases}$
! $3^{n}$ observables

## Alice prepares a group of n-qubit

 states which collectively have those outcome preferences$$
\begin{aligned}
\left|\psi_{1}\right\rangle & =\frac{1}{\sqrt{2}}(|0000\rangle+|1111\rangle) \\
\left|\psi_{2}\right\rangle & =\frac{1}{\sqrt{2}}(|0001\rangle+i|1110\rangle) \\
\left|\psi_{3}\right\rangle & =\frac{1}{\sqrt{2}}(|000+\rangle-|111-\rangle)
\end{aligned}
$$

$$
\vdots O\left(n(3 / 2)^{n}\right) \text { states }
$$

She sends a few copies of each state to Bob ( $k / n$ states in total)
E.g., the state $\left|\psi_{1}\right\rangle=\frac{1}{\sqrt{2}}(|0000\rangle+|1111\rangle)$ yields
+1 in $X_{1} X_{2} X_{3} X_{4},-1$ in $X_{1} X_{2} Y_{3} Y_{4}$,
-1 in $X_{1} Y_{2} X_{3} Y_{4}$, -1 in $X_{1} Y_{2} Y_{3} X_{4}$,
-1 in $Y_{1} X_{2} X_{3} Y_{4},-1$ in $Y_{1} X_{2} Y_{3} X_{4}$,
-1 in $Y_{1} Y_{2} X_{3} X_{4},+\mathbf{1}$ in $Y_{1} Y_{2} Y_{3} Y_{4}$
and +1 in $Z_{1} Z_{2} Z_{3} Z_{4}$, with probability 1


Bob measures the states with some of the observables


He finds their preferred outcomes

Bob reconstructs a fragment of Alice's original data


## Optimizations

- Which set of preferred parities to chose from out of $2^{N}$ choices where N are number of couples of POs
- First compute all compatible well-defined outcomes for all possible eigenstates
- Compare eigenvalues from this optimization to compatible preferred parities
- Which eigenstates best represents that chosen preferred parity order
- We want a low sampling requirement
- Least number of states


## Neutrino Astronomy: IceCube Events



| $X$ | 95.3 m | 01000010101111101001100110011010 |
| :--- | :--- | :--- |
| $\mathcal{Y}$ | 75.8 m | 11000010100101111001100110011010 |
| $Z$ | 484.6 m | 01000011111100100100110011001101 |
| $Q_{\text {tot }}$ | 2.84 PE | 01000000001101011100001010001111 |
| $\bar{t}$ | 26.2 ns | 01000001110100011001100110011010 |

0100001010111110100110011001101011000010100 1011110011001100110100100001111110010010011 0011001101010000000011010111000010100011110 1000001110100011001100110011010
Digitization Scheme: takes Optical Module (OM) position, light and time information and converts to binary. Each circle in image is a OM, size of circle indicates amount of light, color indicates time (red $\rightarrow$ purple)
This is our input to our QRAC

## Event example:



- 01000010101111101001100110011010110000101001011110011001100110100100001111110010010011001100110101 01110001000011010100001100001110111011101100001101111011110000110111001000100010001011011001100011 0101101000110101111011110100110000011010100011101101001001001000111111111000101100011000000110010 0011101000100110001011011000110100110001100110011010010000100101011110100010110010001111101011000 0111110101001111110100011001000110010111001111001001010110001011011100101010001000011000010100110 01111000011101000111101001010010001001000100111001100100010100011001010011111101000110010110101010 0011111000011101111011110001010011010001011111111101111011100111010011110001100011100101111101100 11001110001011001010101011001001111011110101100111010011001010011110101000001111011101111001100100 1010001010000110101001010000100010111001110010000110010011101011101011010000101001101111011001001 10001101000100101101100100101100110111101111100110011110100110111000111010100110101111000000010100
 1000110000100011101011000001101001111110011110111101111101001011010010000011111101000100001010111 01001111111101101011000110010011111110010011111111001100000001110011001001011101000100011101000110 10000000000100000010010001010100010100000000000111000000110101011110100011010001001111100101011010 1001011110011001100111110111100100000010001100100100110111110011110101001101010010110001110111011 0101010010101011011010010110111001100101011100100011000011101010111101110011000100101010110010011 1111010110111010101001101111100111010011111011000111100101011101111100011001010010010000101101001 0100100010100100001111001000101100110100010000010000001111000000011111111001011111010110011111100 1111101110101111001111001010100110100011111111010101101100000111010111101000001110101101010101000 11000100010011011001110111100010100000000101010101100100100101101110011100000110010101010010000010 11001111011101001000100010001111101100001010100001011010010111011111110100100000001101011101010010
1101001101100111000001000100110101101110010001100100010000111010111110000110011010101110100100000 10100111111110010010101101101000101101110000111101001101100000000101000111101101011010100101001111 10011011101110001000110010110001001010100001000111101100111110100110001011001001110011110100000101 01111111101001111100000010001000110001011010010010100000110100000010111111001010111101001111000 01111101010101001011001011001000010010010111011110000110000011010011001100110101111001111001111111 01000001010100111111011010001100100110101101010111111000001111111000101001111100010111100100110010 00001100101100011001000000111100001001000011110101010110001001000001010100101011111010011011011110 0000100011110001001000000011101000000110011101001110011110101100001111010010001111101101000111100 1100010100111101101001001111010010110111010111001001001110011101111010010011000100100100010000000 111010110111011000101010101101001110100111100101011101111011110100110110010111011001011101010010 011100110010011010111100101010111011011110101100001111100000000110110000100101000011001101100000 0100011010100101101010100011101001111101000010001001110000110101010110000001111101010011110110000 11000001000000001000011111010010010011000000111111011100100011100100100100011111110011100011100000 1000101101110000001100011001001000010011100000001101111110101101100000101011001000010110000111010 11010000100011111011001001000111011001001010011001111111101100001111011100000101111011111111011100 0000101010111011010000110110101010110010100011110010100101000100011110110110100101111000001010110 10111110011110001100111011011010001011110111100001010100110110000001100100111100101010011101001000 0100100101110000110100101111110110100101111101101111101101111001000011101001111000100010110110101 10010001100111011111101100010000100111101111110110100111101010110110100110011101111011111001011101 000000101111001110001100110011001110101101011110100110100001100010101011011010101101001010110100100 00100001000010101010010110100111111000100001000011100001100101010010010000010000101110000001111010 10001101010110010001000100101010100101011101011100110110011000000101110011000110000110111000101011
110101000011011000100111101100011101110100111011101001100110110010110101010110101010011100011101 01011000110110010110001111000111100001010001111010111100111000110001011011111101110011101011101011 0001110101000111101110111111100010110001100000100011001101011000111000010101001110100001111101101 0101011010001111001011101000011000111000000111100111111011010110011110101010000000100011111000100111 11110000010011000110010000110110101101100110110101110001101001011101010001011010011000010000011001


## QRAC: What has been done so far

- Have achieved compression with 8 qubits
- Almost demonstrated storage/retrieval of IceCube simulation data with 14 qubits


## Plan with IceCube data

- Take set of IceCube data: tracks/ cascades ( $\nu_{\mu} \mathrm{CC} / \nu_{e} \mathrm{CC}$ )
- Input into our QRAC
- Take states from QRAC, input them to train Variational Quantum Circuit
- First step, will be to investigate if VQC can classify tracks/cascades

Neutral-current / Ve


## Variational Quantum Circuit



## Variational Quantum Circuit



## Model Circuit

- Is the variational part, where machine learning happens, called the "ansatz"
- The ansatz is parameterized by a set of free parameters $\theta$ that will be updated during training
- The structure of the ansatz, entanglement, type of rotations, number of parameters, number of gates, are all tunable



## Variational Quantum Circuit



## Loss Function

- Default: is cross-entropy loss

$$
\text { CrossEntropyLoss }(\text { predict }, \text { target })=-\sum_{i=0}^{N_{\text {claws }}} \operatorname{target}_{i} * \log \left(\text { predict }_{i}\right) .
$$

- The difference between the ideal distribution (the true labels/target) and the measured distribution
- Calculated across a batch of samples, and the average is taken across the batch to obtain the final loss value for that iteration
- Others like MSE are also used
- Inherently a classical operation


## Optimization

- Optimization performed using classical algorithms: COBYLA, ADAM, SPSA, etc.
- Numerical Differentiation has been the norm for a while using parameter-shift differentiation
- It's unclear how intermediate derivatives could be stored/reused inside a quantum computation
- Investigating if we can use quantum optimization techniques like quantum annealing
- A lot is still unknown about what is optimal, more of an art to test
- "There is often little theoretical motivation for any given optimizer, although recent work has been to analyze the training dynamics of OVCs (VOCs)"


## Proof of Concept using Qiskit

- MNIST: image classification 0s and 1 s
- Input image via amplitude encoding into 8 qubit VOC
- ~about 100 tunable parameters
- Iterations are a hyper parameter
- COBYLA, cross-entropy loss function
- Training score: 97\%
- Testing score: 98\%



## What's next?

- Comparing VQC structure with amplitude encoding of IceCube simulation data with QRAC encoding
- Takes $\sim 16$ qubits to encode largest IceCube events via amplitudes
- Need larger quantum computers than publicly available (applied for ORNL grant)
- Now have access to $\sim 100$ qubit QCs
- Choose optimal qubits in target QCs, implement custom feature maps, testing ansatz types, optimizations, loss functions etc.
- Transition away from Qiskit to PennyLane (a QML framework)


## Conclusion/Takeaway

- ML in neutrino physics has been very successful the past two decades and will continue to contribute
- Yet with concerns about the streetlight effect and increasing data loads in next generation experiments/observatories, we might need new tools
- QC can take advantage of additional degrees of freedom to enhance data compression
- QCML is still a burgeoning field; much is still unknown about optimizations, actual quantum speedups, feasibility on current NISQ computers, etc.
- "The question on whether quantum computers can really play a role in identifying practical ML applications is still wide open, and it is unlikely to be decided by theoretical proofs..."


## Gracias! Thank you! Preguntas? Questions?

## New arm of hectopus:

 Quantum Computing!

## Measuring Context Eigenstates

- The values for one context are all related. Only need to calculate one "fingerprint" and figure out mapping
- Scales like $4 \wedge \mathrm{~N}$
- For 12 qubits this can be done in a few hours and stored in 2 MB



## Development Roadmap

| $2019 \odot$ | $2020 \Theta$ |
| :--- | :--- |
| Run quantum circuits <br> on the IBM Cloud | Demonstrate and <br> protototpe euantum <br> alogiths and <br> applications |

Executed by IBM ©
On target ソ)


## Overview

- Introduction to Quantum Computing
- Application of QC to HEP
- QCML in HEP
- QCML for neutrino astronomy
- Variational Quantum Circuits
- Conclusion

