Mitigating Accidental Backgrounds in LUX-Zeplin (LZ) using Machine Learning

Ishan Khurana (UCL)

IoP HEPP & APP Annual Conference 2022
20/03/2022

Overview

- The LUX Zeplin (LZ) collaboration, detector and experimental approach
- Accidental events in Dual Phase LXe detectors.
- Mitigation using XGBoost and a linear classifiers.
- First look at calibration data.
- Further work using convolutional neural networks.

Photo: LZers walking through the drift in the homestake mine at the 4850ft level.

LZ (LUX-ZEPLIN) Collaboration

@lzdarkmatter

https://lz.lbl.gov/

35 Institutions: 250 scientists, engineers, and technical staff

- Black Hills State University
- Brandeis University
- Brookhaven National Laboratory
- Brown University
- Center for Underground Physics
- Edinburgh University
- Fermi National Accelerator Lab.
- Imperial College London
- Lawrence Berkeley National Lab.
- Lawrence Livermore National Lab.
- LIP Coimbra
- Northwestern University
- Pennsylvania State University
- Royal Holloway University of London
- SLAC National Accelerator Lab.
- South Dakota School of Mines & Tech
- South Dakota Science & Technology Authority
- STFC Rutherford Appleton Lab.
- Texas A&M University
- University of Albany, SUNY
- University of Alabama
- University of Bristol
- University College London
- University of California Berkeley
- University of California Davis
- University of California Los Angeles
- University of California Santa Barbara
- University of Liverpool
- University of Maryland
- University of Massachusetts, Amherst
- University of Michigan
- University of Oxford
- University of Rochester
- University of Sheffield
- University of Wisconsin, Madison

US UK Portugal Korea

Thanks to our sponsors and participating institutions!

U.S. Department of Energy
Office of Science

UKRI Science and Technology Facilities Council

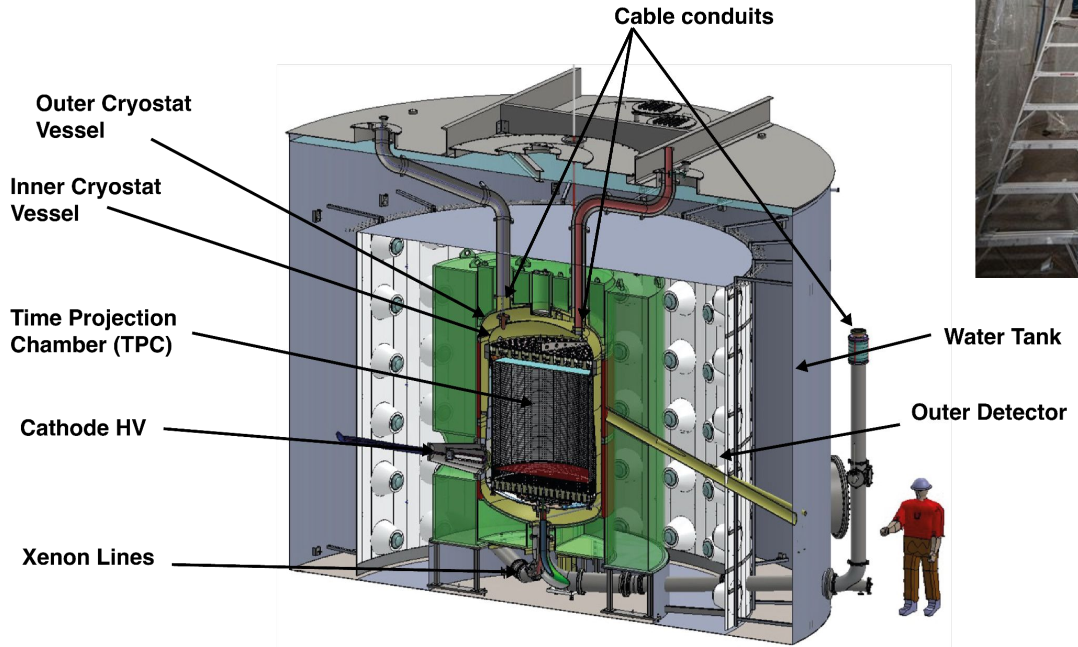
Sanford Underground Research Facility
South Dakota Science and Technology Authority

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

ibs Institute for Basic Science

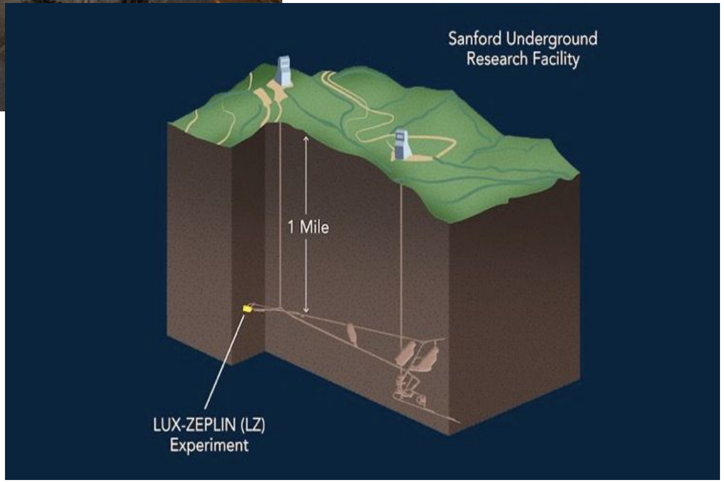
LZ Detector Overview

- Largest Dual Phase LXe TPC
- 7 active tonnes of Xe
- 5.6 tonne fiducial volume
- 494 TPC PMTs
- **Taking data now!**



ICV Installation

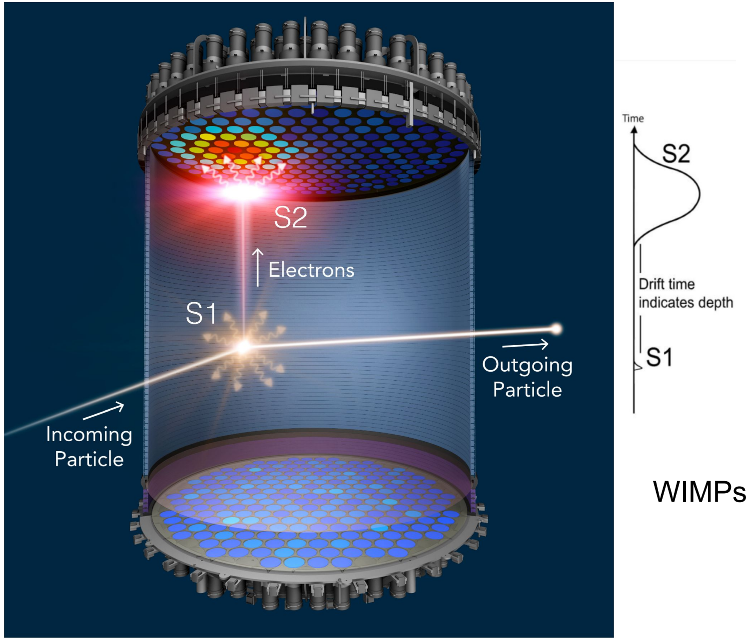
Assembled TPC



Detection Principle

Particle interactions in the LXe create:

- Prompt scintillation (S1)
- Electrons from ionisation
  - drifted upward to GXe
  - delayed proportional scintillation (S2)

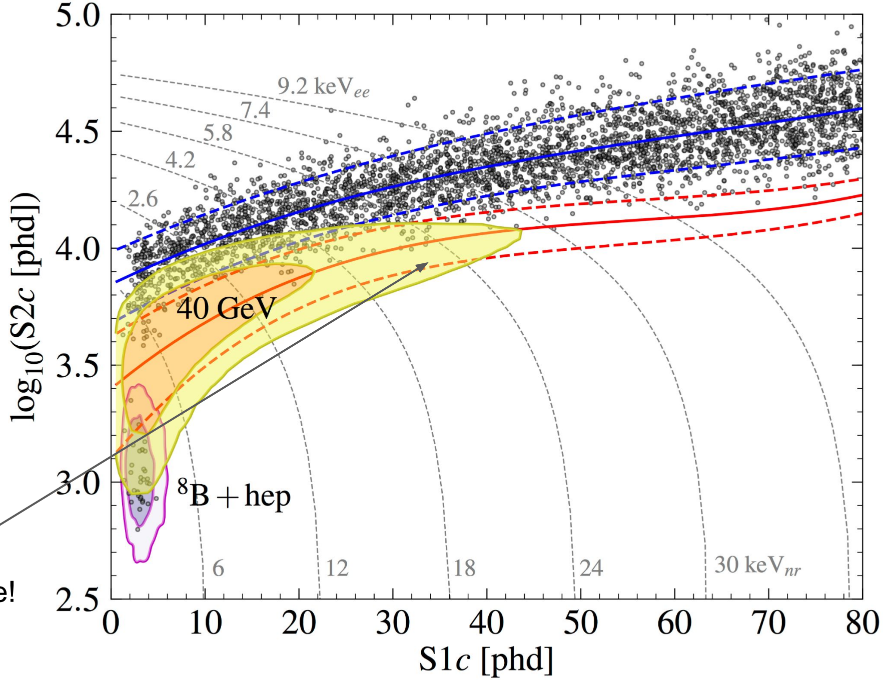


Two types of particle Interactions:

- Nuclear Recoils
- Electron Recoils

ER and NR events are discriminated from their different S2/S1 ratios

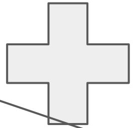
WIMPs found here!



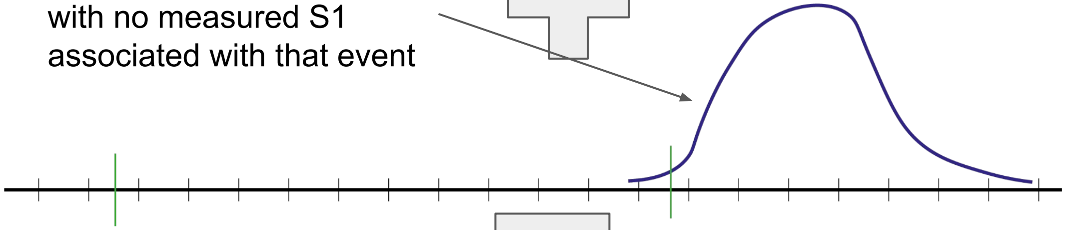
LZ simulated data set for a background-only 1000 live day run and a 5.6 tonne fiducial mass. Source: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.101.052002.

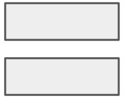
Accidental Background Events

S1 Only Event i.e. an S1 pulse with no measured S2 associated with that event.



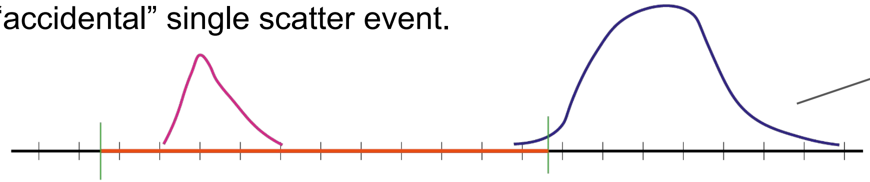
S2 Only Event i.e. an S2 pulse with no measured S1 associated with that event



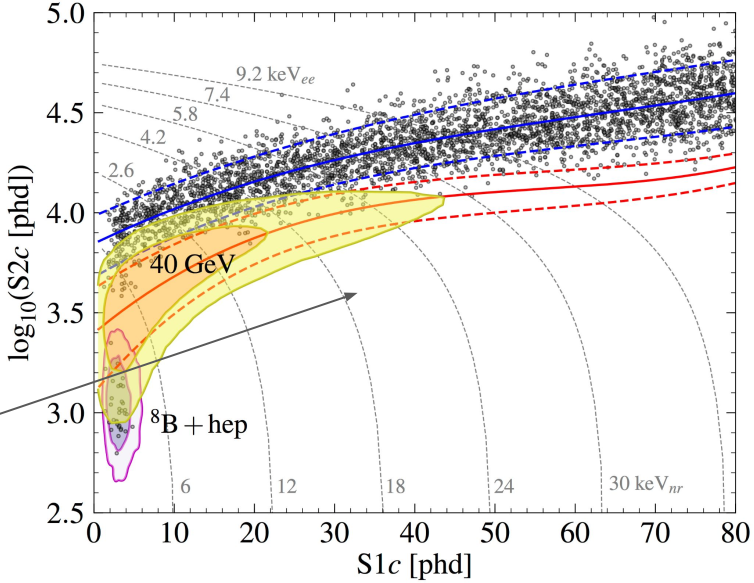


Example Accidental Event

The random coincidence of an S1 only pulse and an S2 only pulse to form an “accidental” single scatter event.



With 1000 day livetime, accidental coincidence events more likely to occur.

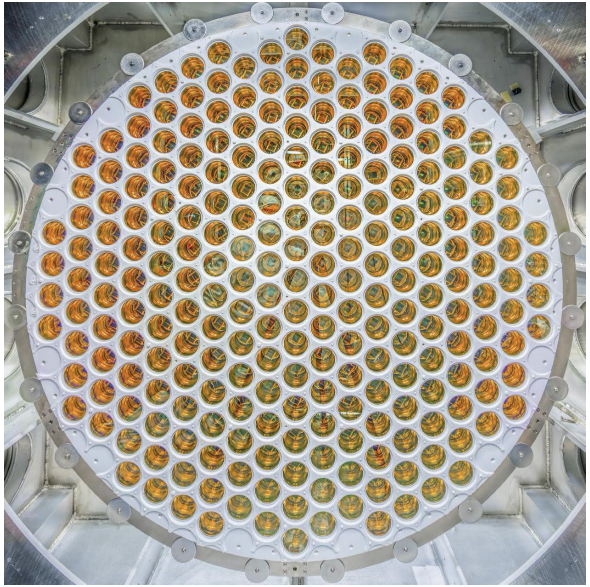


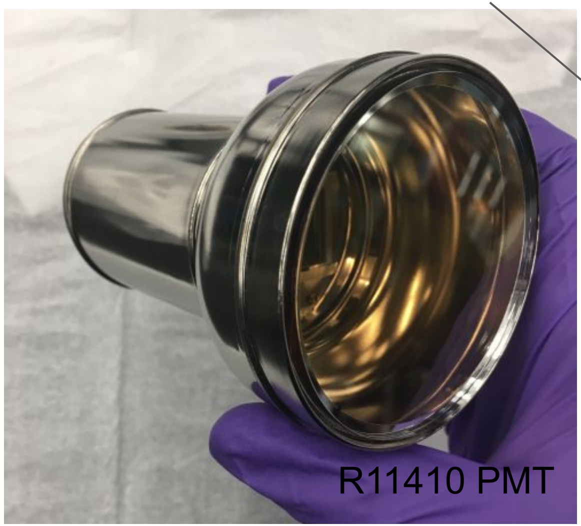
LZ simulated data set for a background-only 1000 live day run and a 5.6 tonne fiducial mass. Source: https://journals.aps.org/prd/abstract/10.1103/PhysRevD.101.052002.

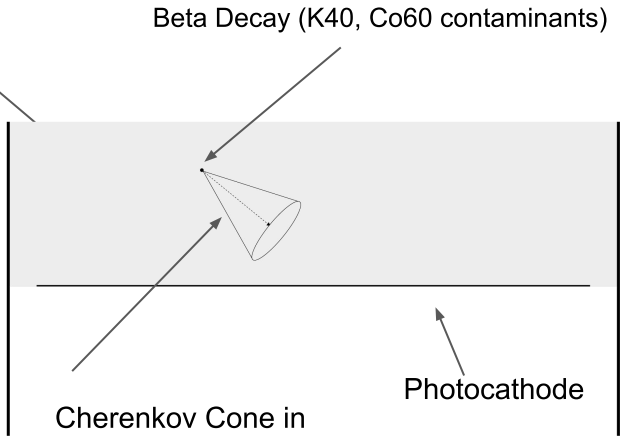
Cherenkov in TPC PMT Glass

LZ TPC is instrumented with 2 arrays of Hamamatsu
R11410-20 PMTs (494 total)

- Beta emission (^{40}K, ^{60}Co contaminants) in the PMT glass can lead to cherenkov.
- Leads to instant signal in the PMT
- Can fake an S1 if some photons escape the PMT of origin







Matthew Tan Bachelor's Thesis submitted to Brown University. Available here.

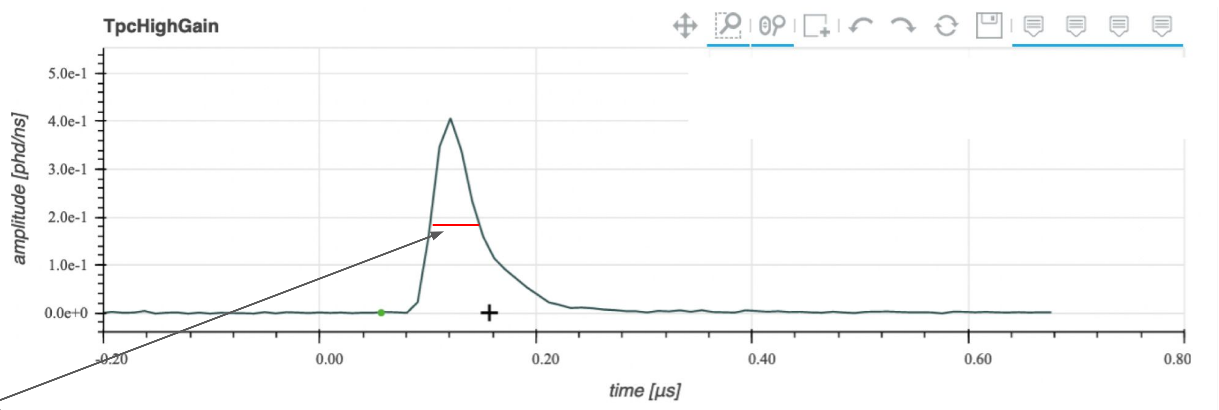
Classification Parameters

LZ pulse are characterised using Reduced Quantities (RQs).

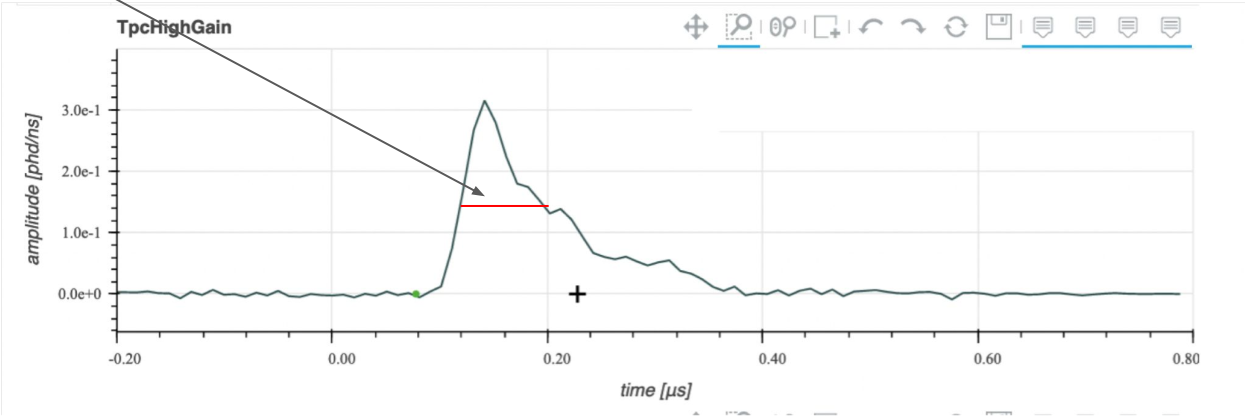
Four RQs are used to separate cherenkov ‘S1’ from S1s associated with TPC interactions.

- **Full Width Half Maximum (ns)**

Simulated Cherenkov ‘S1’



Simulated Nuclear Recoil S1



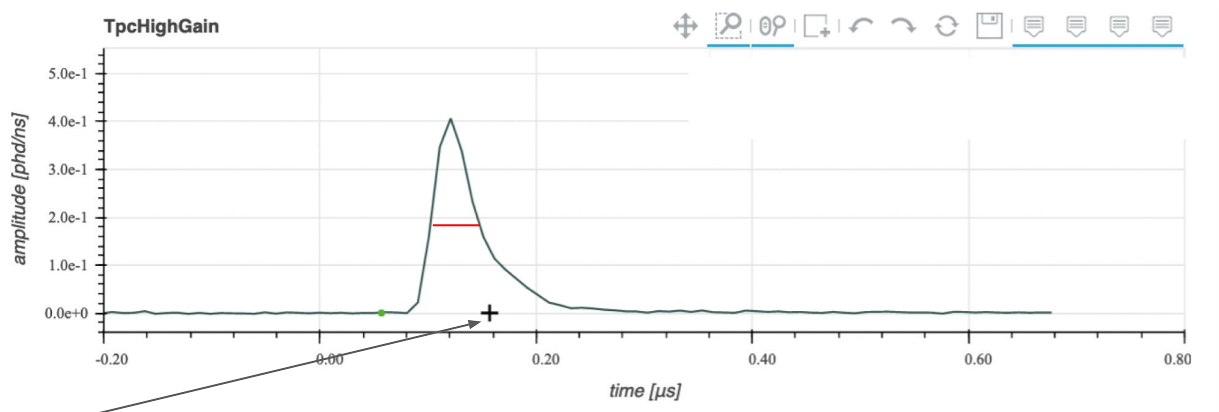
Classification Parameters

LZ pulse are characterised using Reduced Quantities (RQs).

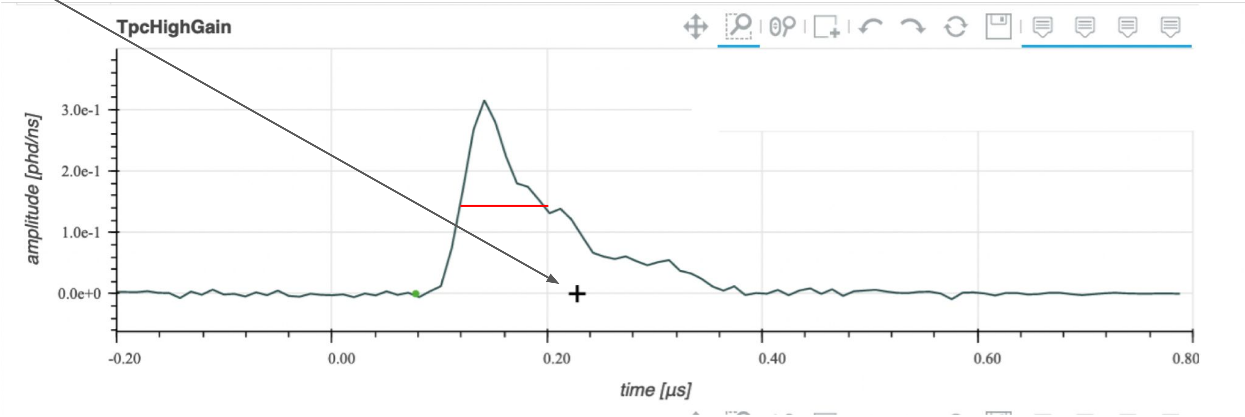
Four RQs are used to separate cherenkov ‘S1’ from S1s associated with TPC interactions.

- Full Width Half Maximum (ns)
- **75% Area Fraction Time (ns).**

Cherenkov ‘S1’



Nuclear Recoil S1



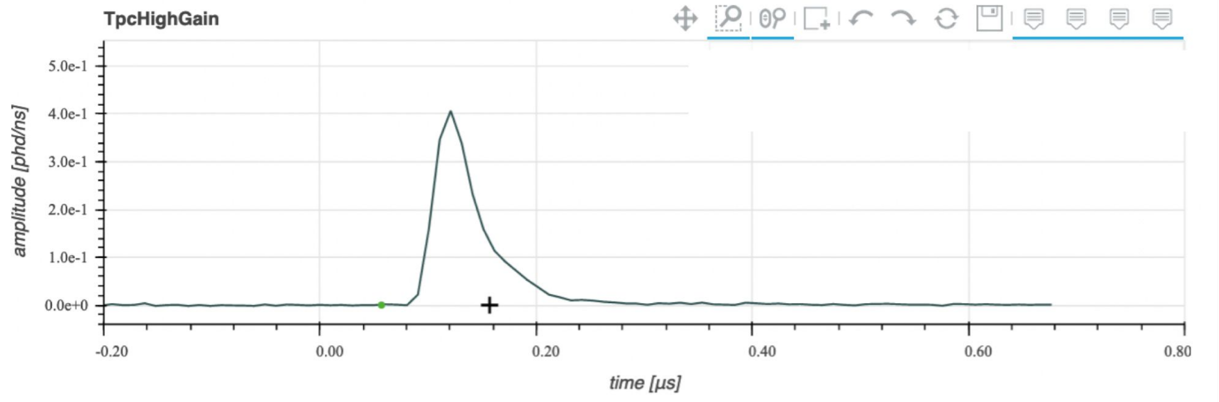
Classification Parameters

LZ pulse are characterised using Reduced Quantities (RQs).

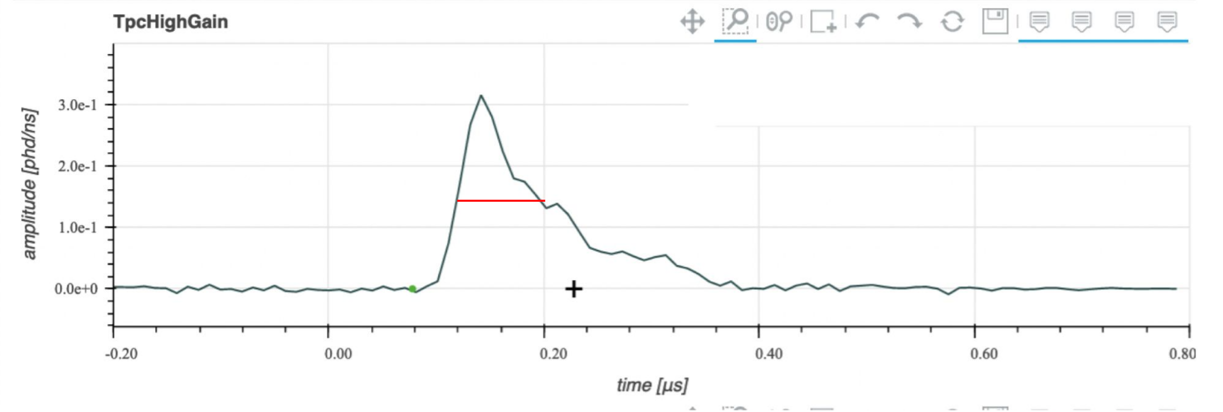
Four RQs are used to separate cherenkov ‘S1’ from S1s associated with TPC interactions.

- Full Width Half Maximum (ns)
- 75% Area Fraction Time (ns).
- **Prompt Fraction 50ns.**
    - **Fraction of the total photons detected in 50ns.**

Cherenkov ‘S1’



Nuclear Recoil S1

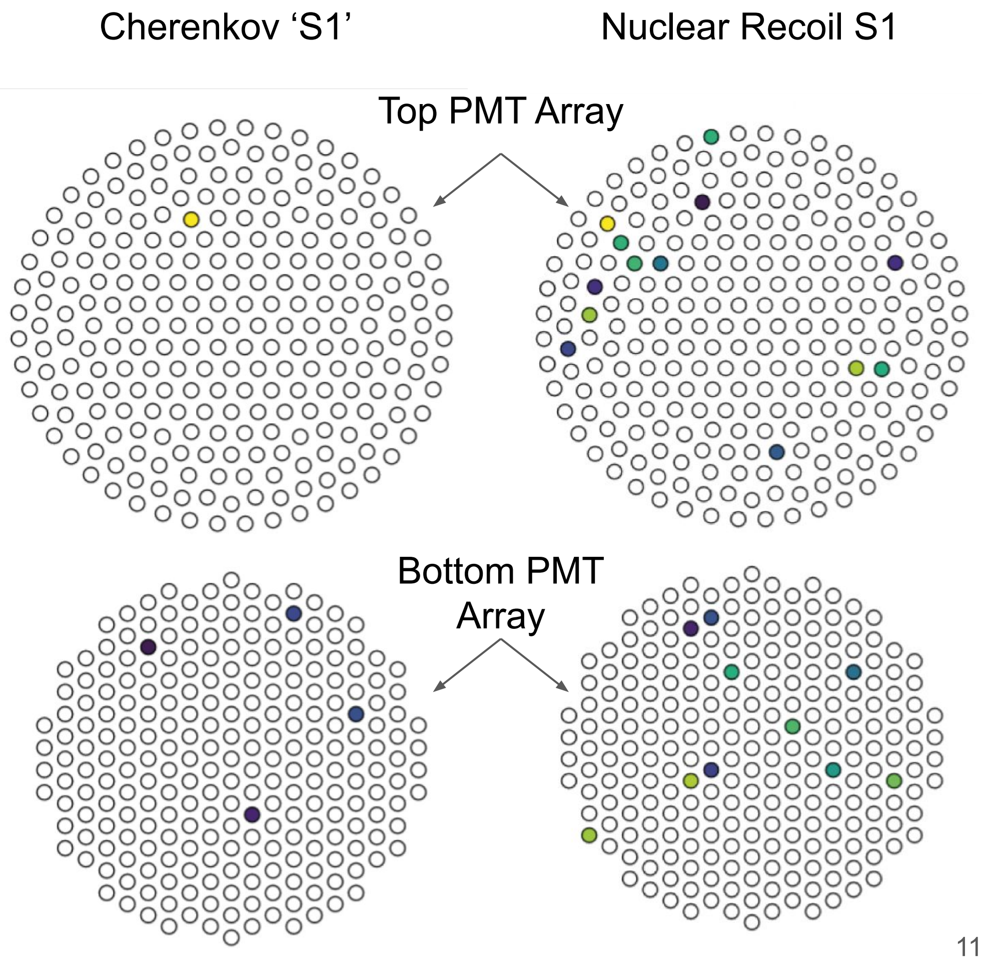


Classification Parameters

LZ pulse are characterised using Reduced Quantities (RQs).

Four RQs are used to separate cherenkov ‘S1’ from S1s associated with TPC interactions.

- Full Width Half Maximum (ns)
- 75% Area Fraction Time (ns).
- Prompt Fraction 50ns.
  - Fraction of the total photons detected in 50ns.
- **Max Channel Area Fraction**
  - **Fraction of the total photons in the pulse that are detected in the PMT with the largest signal.**



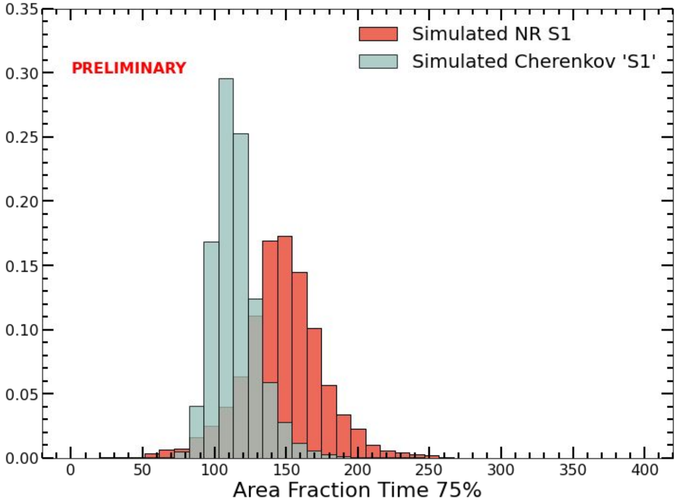
Classification Parameters

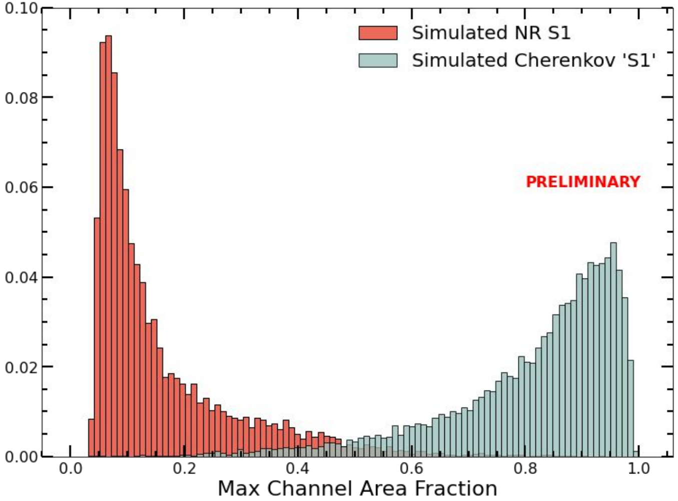
LZ pulse are characterised using Reduced Quantities (RQs).

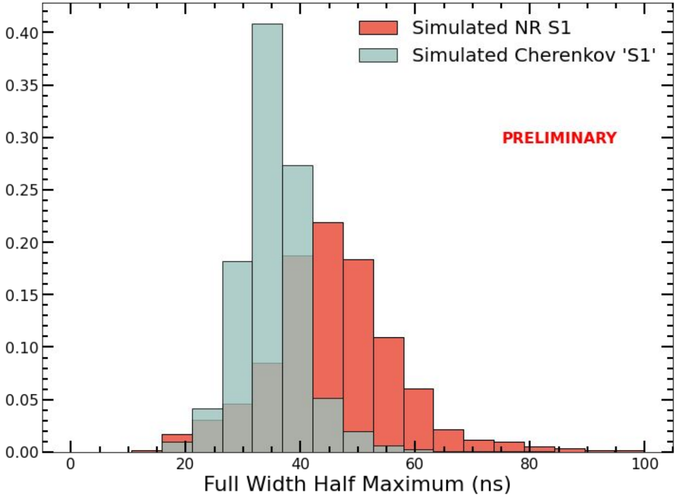
Four RQs are used to separate cherenkov ‘S1’ from S1s associated with Nuclear Recoils.

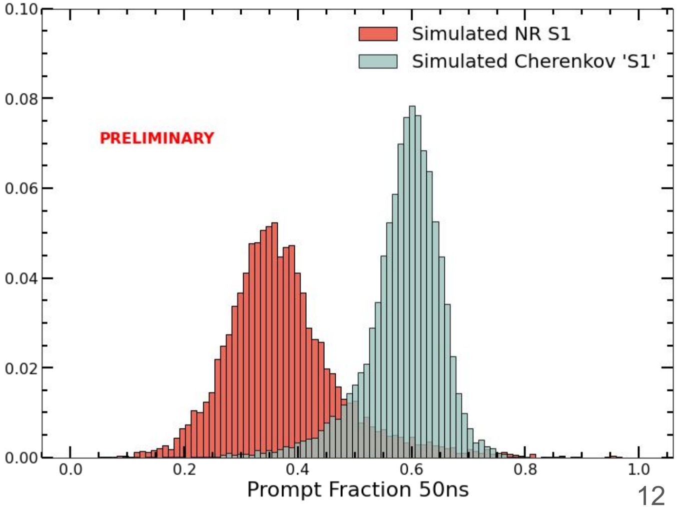
- 75% Area Fraction Time (ns).
- Max Channel Area Fraction
- Full Width Half Maximum
- Prompt Fraction 50ns.

Distributions produced from simulated NR S1s and simulated cherenkov ‘S1’

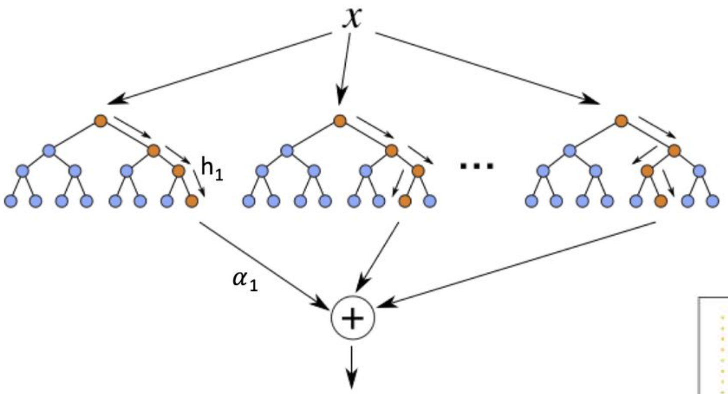








Boosted Decision Tree

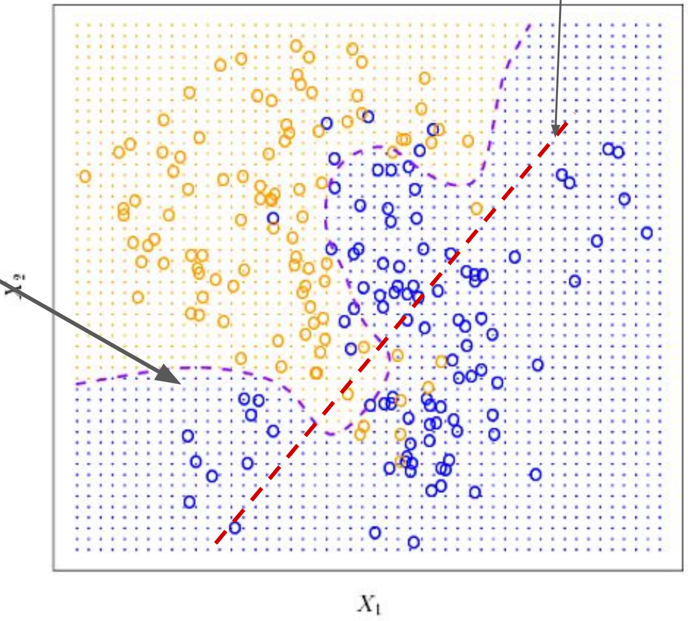


- Using XGBoost
- Outputs a classification score for each sample.
- Allows for a non-linear decision boundary.

Hyperparameters
Number of trees: 300
Max depth: 3
Learning rate: 0.1

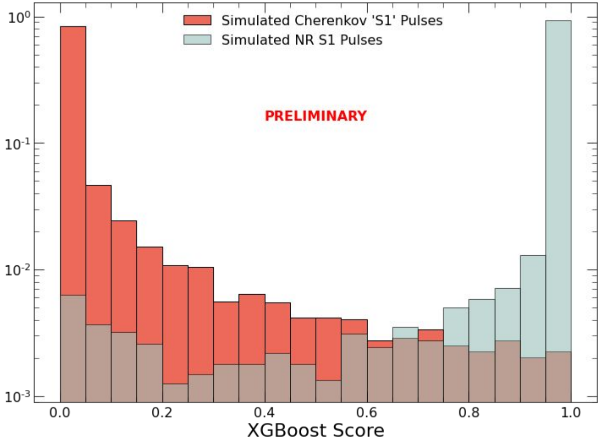
Linear Classifier

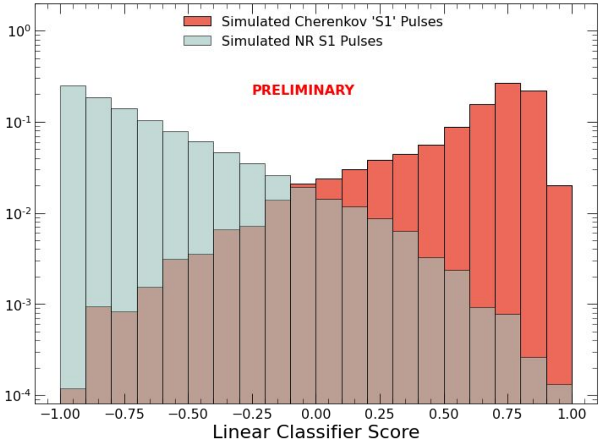
- Linear decision boundary (Line in 2D, hyperplane in 4D).
- Outputs signed euclidean distance from decision boundary.

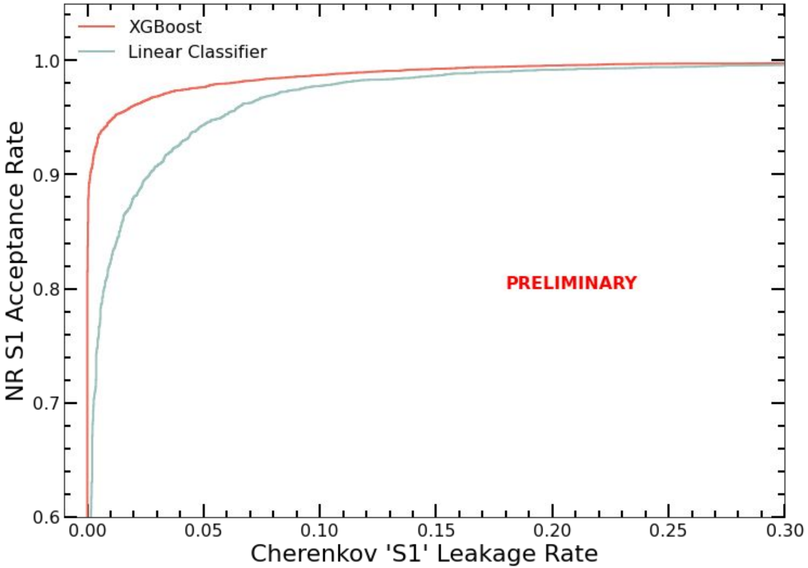


Stochastic Gradient Descent (SGD) to minimise a quadratic loss function (least squares

BDT Outperforms Linear Cut Approach







Boosted Decision Tree:

- For 95% NR S1 Acceptance, the cherenkov ‘S1’ leakage is 0.7%.

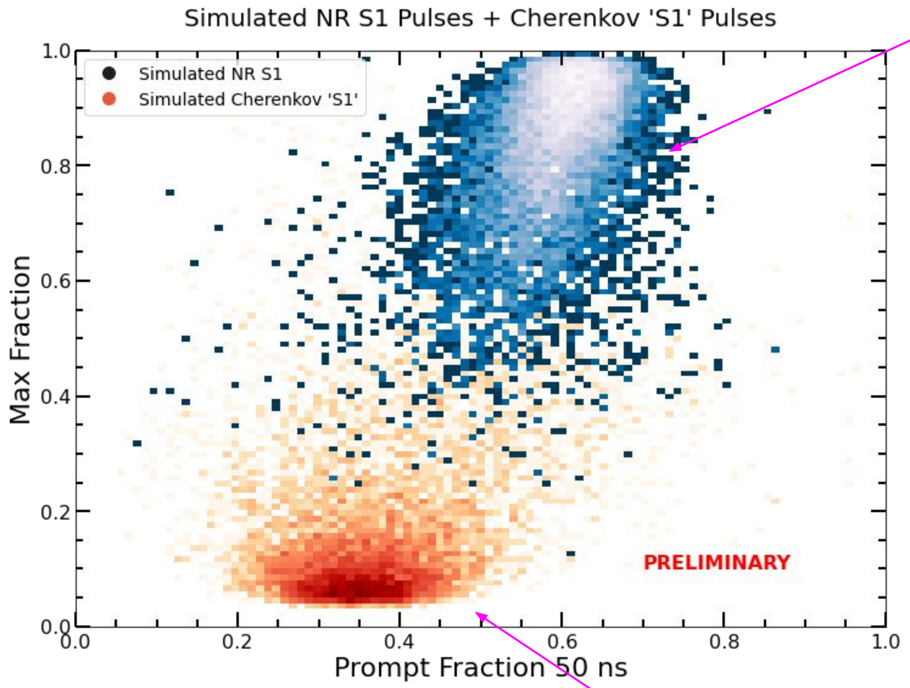
Linear Classifier:

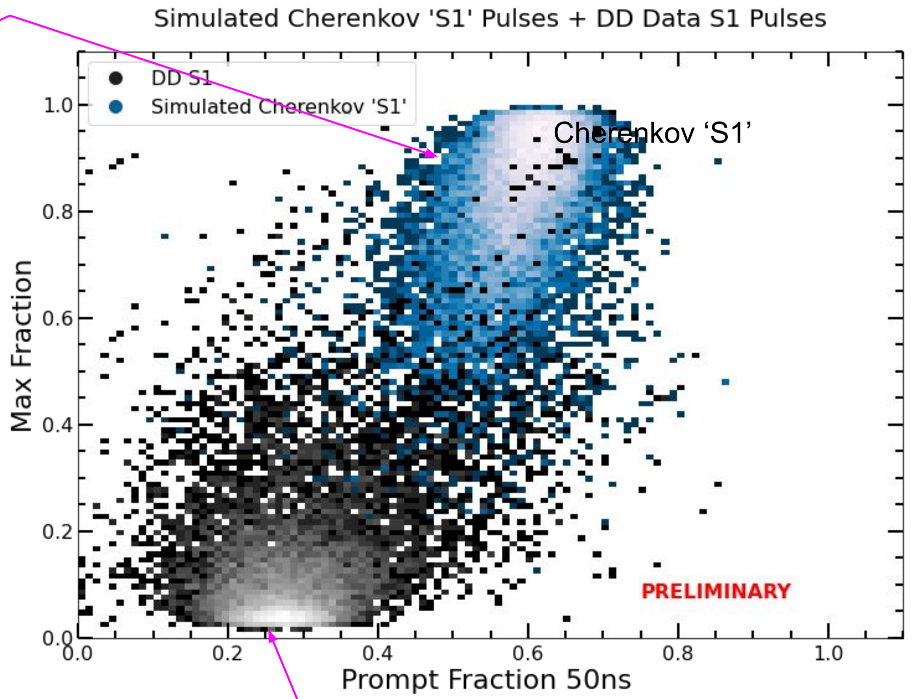
- For 95% NR S1 Acceptance, the cherenkov ‘S1’ leakage is 5%.

See: [Low-energy (0.7-74 keV) nuclear recoil calibration of the LUX dark matter experiment using D-D neutron scattering kinematics](#).

First look at LZ D-D Neutron Generator Calibration Data

Simulated Cherenkov ‘S1’



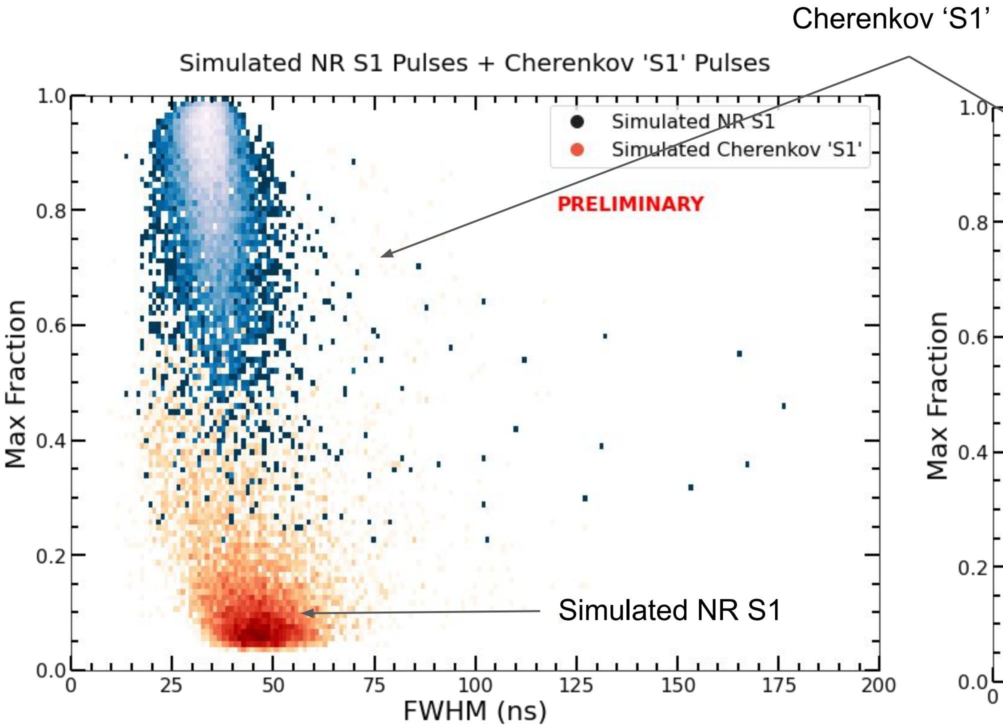


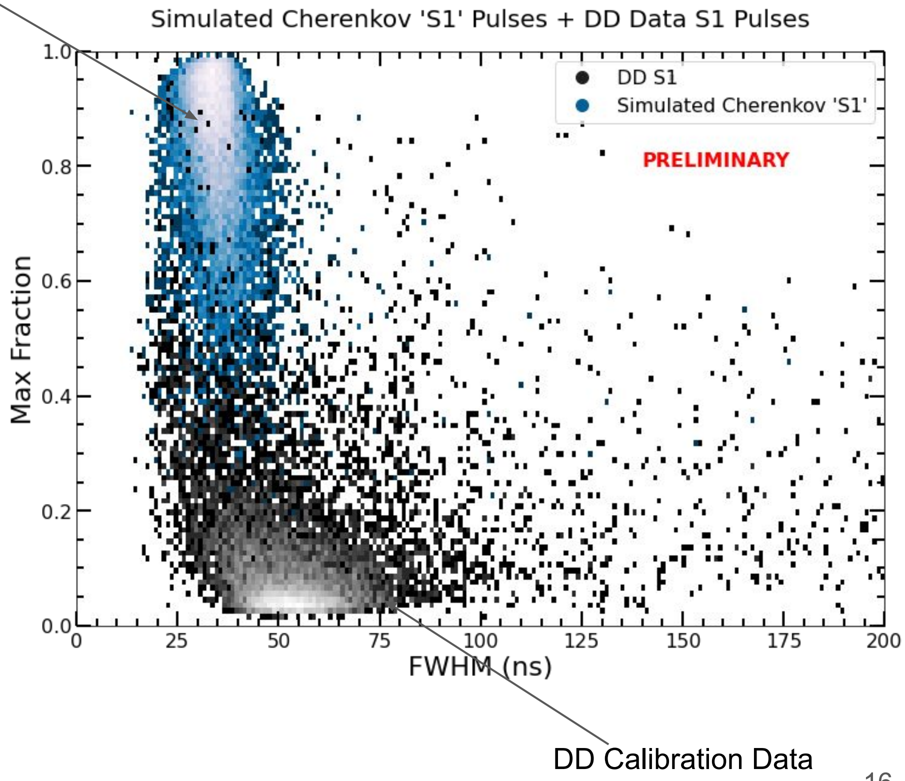
Simulated NR S1

DD Calibration Data

First look at D-D Neutron Generator Calibration Data

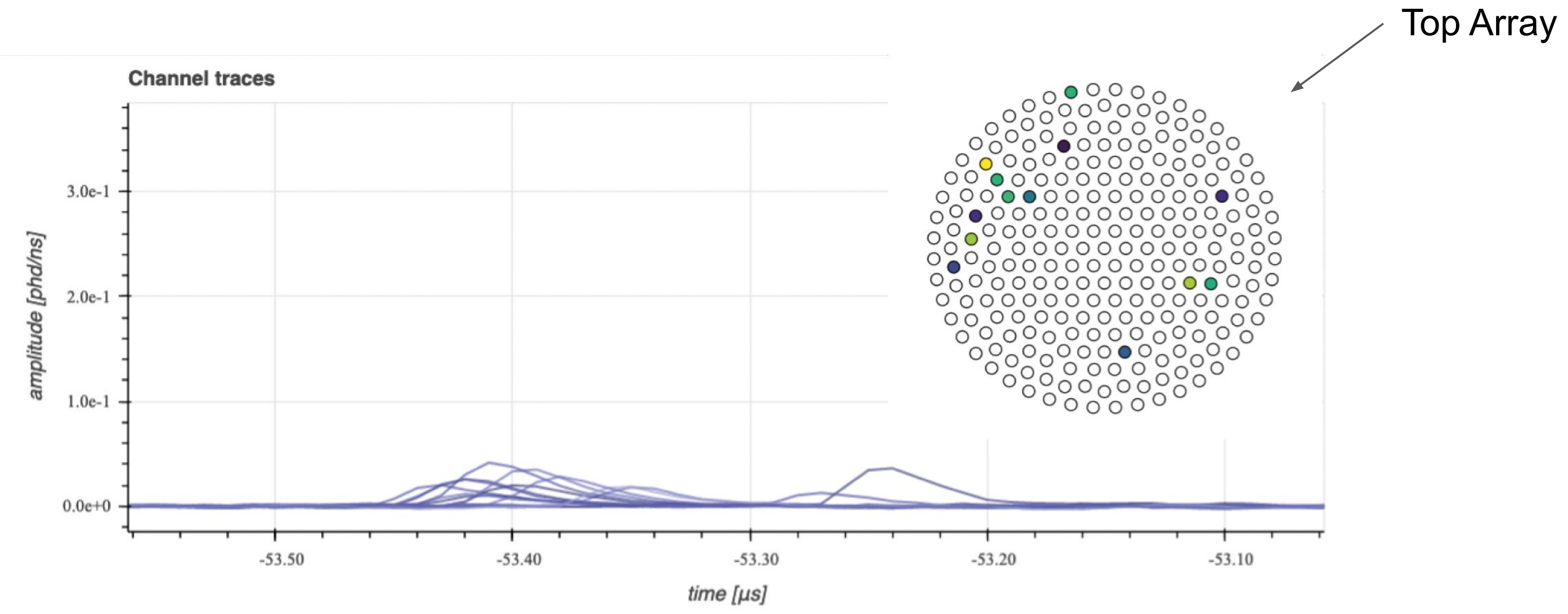
Suggests sims can be used, with tuning, to characterise and classify S1s.





Machine Learning on the Full Waveforms

- 494 TPC PMTs -> 494 Waveforms containing spatial (PMT positions) and temporal information. Suited to other algorithms like RNNs and CNNs.
- Requires more consideration to avoid bias e.g. Is the algorithm implicitly learning/cutting on a parameter used in the PLR (position or reconstructed energy).
- Simplifying problem for CNN by creating physics-informed reduced representation.



Analysis using Reduced Representation of Waveforms

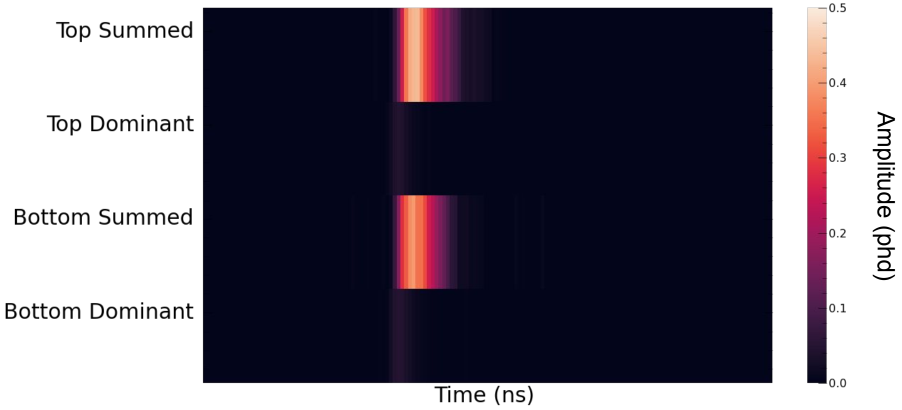
Parsing 4 waveforms with key information into a **Convolutional Neural Network** for binary classification:

- Top/bottom array all channels summed (pulse shape)
- Top/bottom array channel with with most light (light concentration and pulse shape).

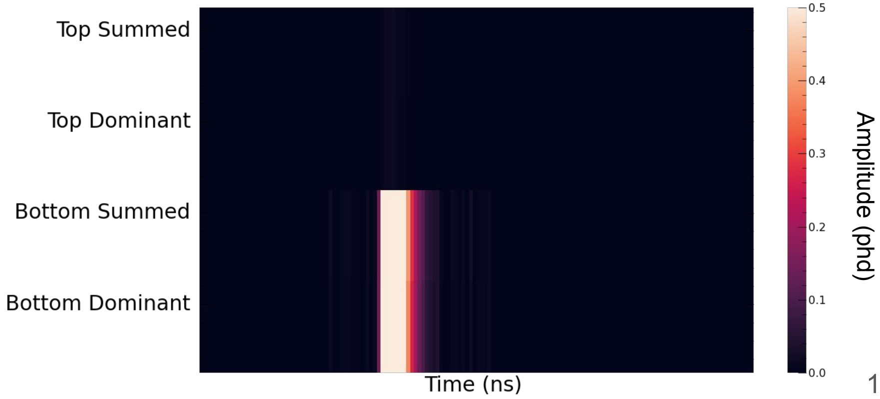
Cherenkov light is arrives immediately and is mostly in the PMT of origin.

Currently a work in progress!

S1 Pulse



Cherenkov “S1” Pulse



Summary

- With increasing livetimes, identifying accidental events will be crucial.
- XGBoost reduces cherenkov ‘s1’ leakage rate by a factor of 7 c.f a linear classifier.
- More gains to be made using full waveform.
- LZ taking physics data now! Preparing to apply ML methods.

Thanks to our sponsors and participating institutions!

U.S. Department of Energy Office of Science

UKRI Science and Technology Facilities Council

Sanford Underground Research Facility
South Dakota Science and Technology Authority

FCT Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA EDUCAÇÃO E CIÊNCIA

ibs Institute for Basic Science