Towards UHE Cosmic Ray Candidate Identification in RNO-G Deep Antennas Using Machine Learning

Bryan Hendricks

1/29/25

PennS

RNO-G

Radio Neutrino Observatory - Greenland

UHECRs and neutrinos

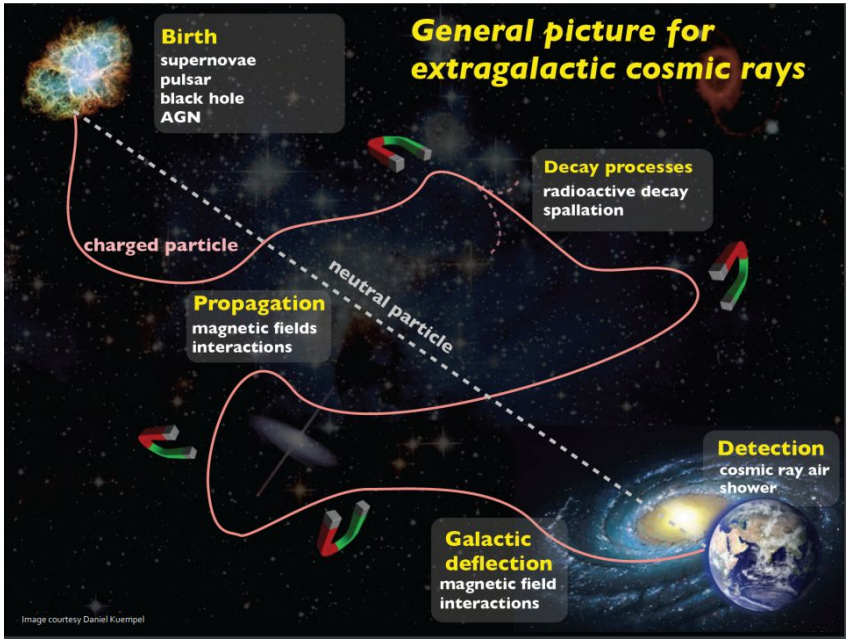
Cosmic rays with energies exceeding 1 EeV

Most energetic individual particles ever detected, yet sources still unclear

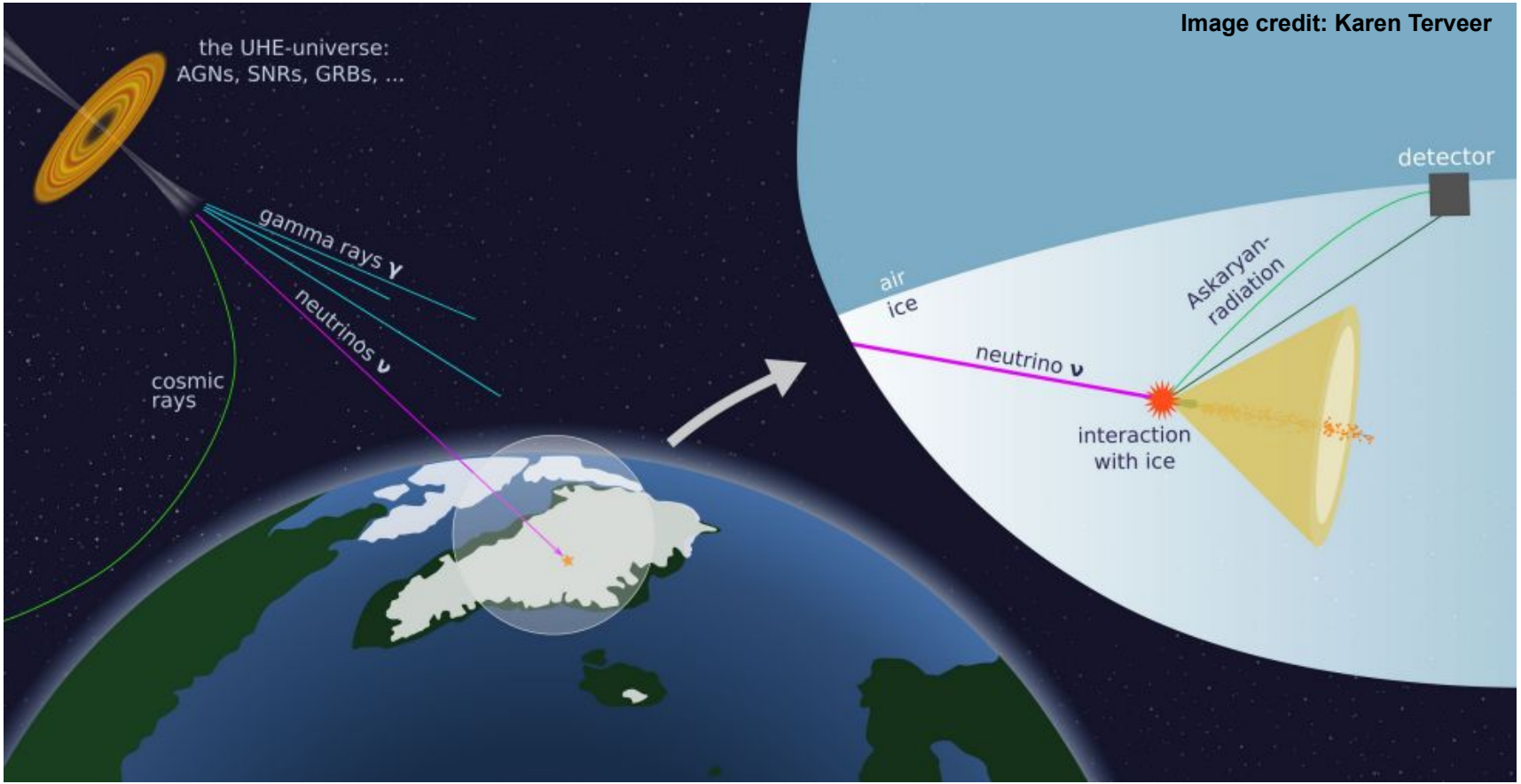
Source ID difficult due to curved path

UHE neutrinos expected to be produced as a result of same environments that produce UHECRs

Can travel undeflected and point back to source, potentially identifying source of UHECRs

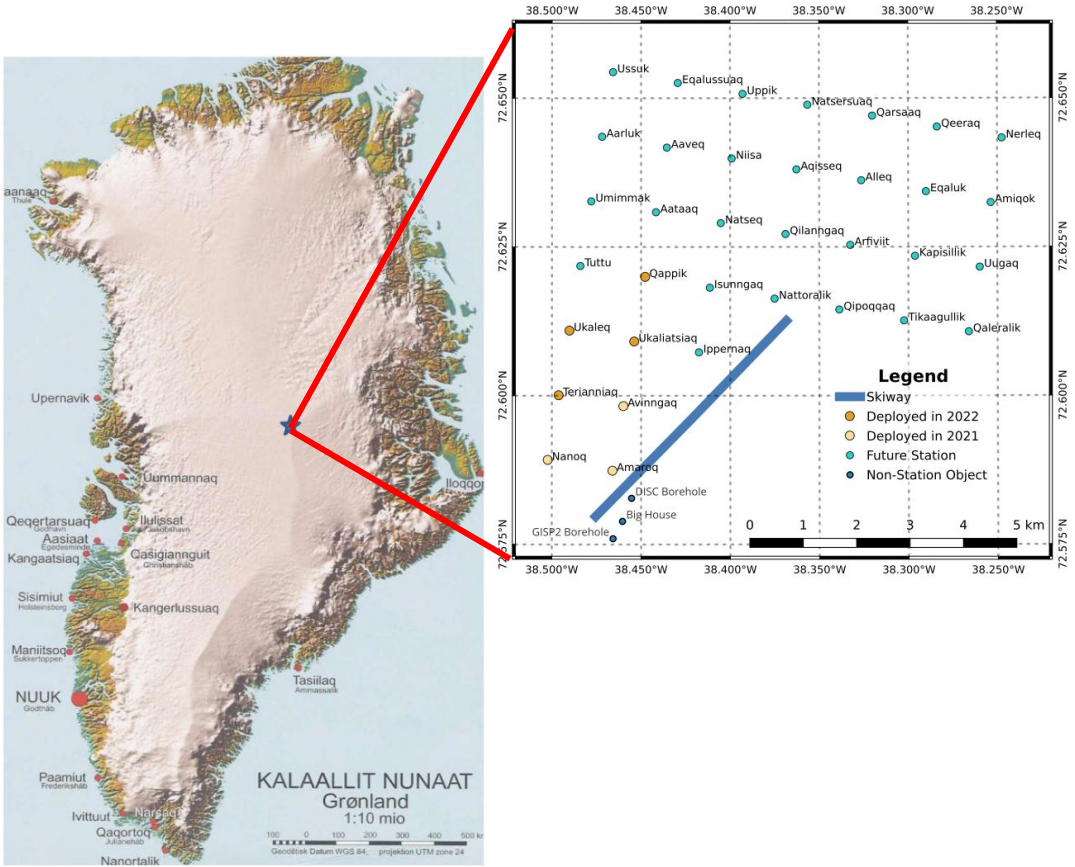


Radio Detection Method for UHE Neutrinos

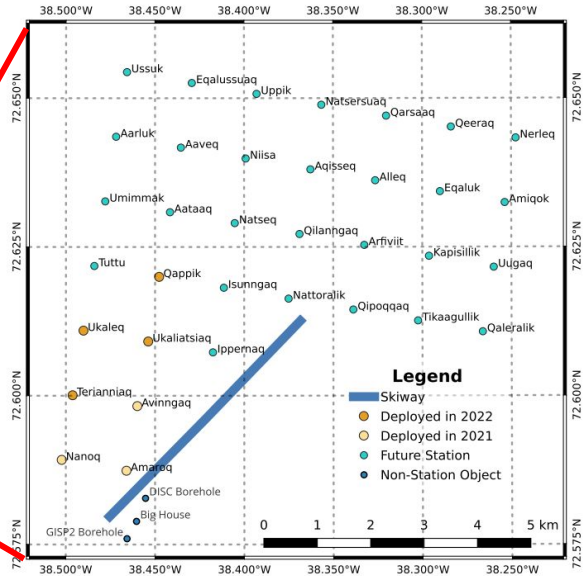


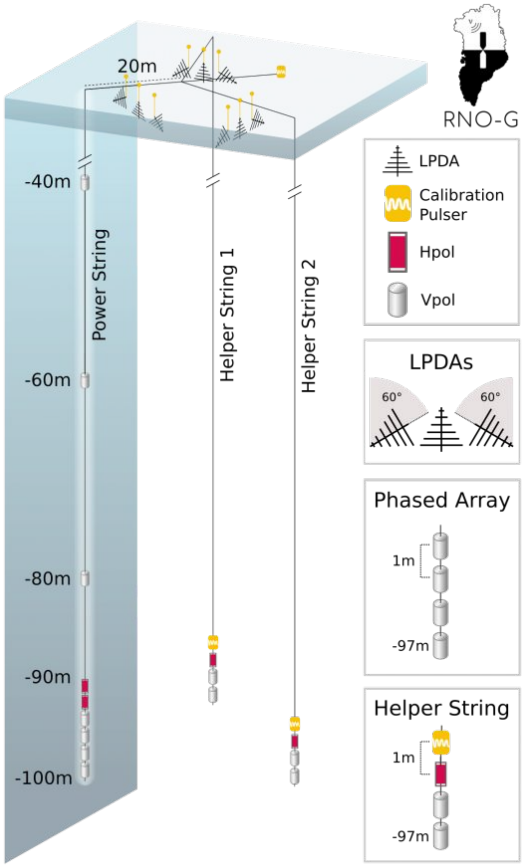
Radio Neutrino Observatory in Greenland (RNO-G)

Radio Neutrino Observatory in Greenland (RNO-G)



Radio Neutrino Observatory in Greenland (RNO-G)



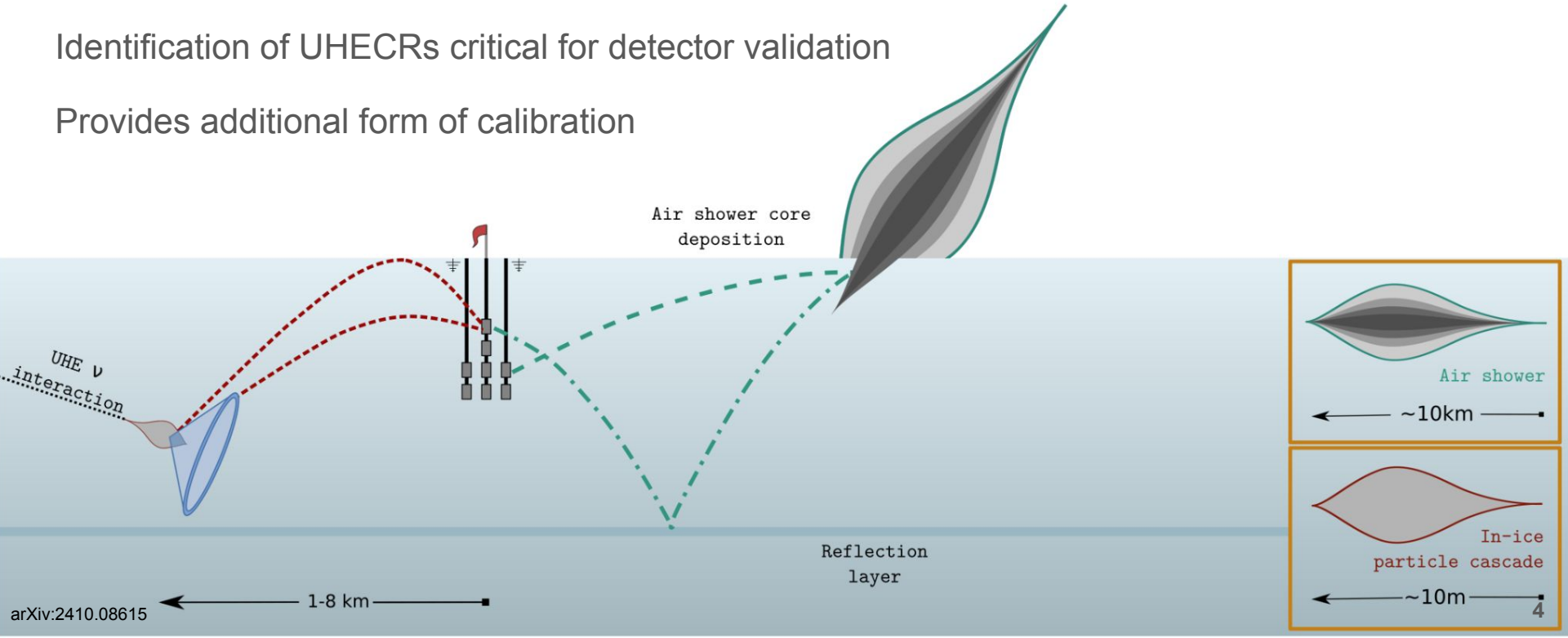


UHECRs as neutrino background

While we want to uncover the source of UHECRs, they can act as a background to our neutrino-based search, particularly in the deep channels

Identification of UHECRs critical for detector validation

Provides additional form of calibration



arXiv:2410.08615

Classifiers being investigated

Two primary classifiers are being pursued at the moment to begin identifying cosmic ray signals

Linear Discriminant

Finds linear combination of features to maximize separation between classes

Pros

- Simple, efficient, easy to interpret

Cons

- Reliance on distribution assumptions
- Difficulty in handling non-linearity

Will be the focus of this talk

Goal: identify CR-like events in data after removing known non-thermal backgrounds

Convolutional Neural Network

Pros

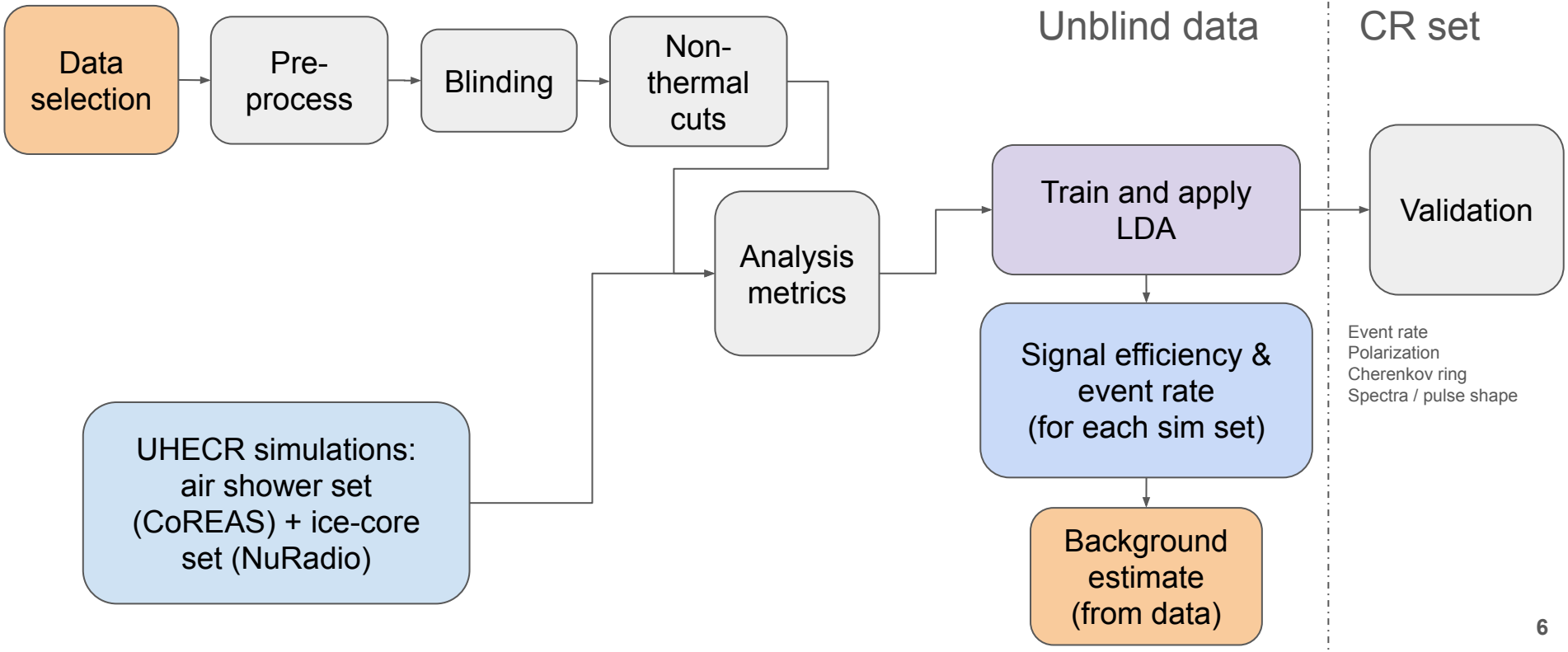
- No explicit assumptions on distributions
- Can handle more complex data patterns

Cons

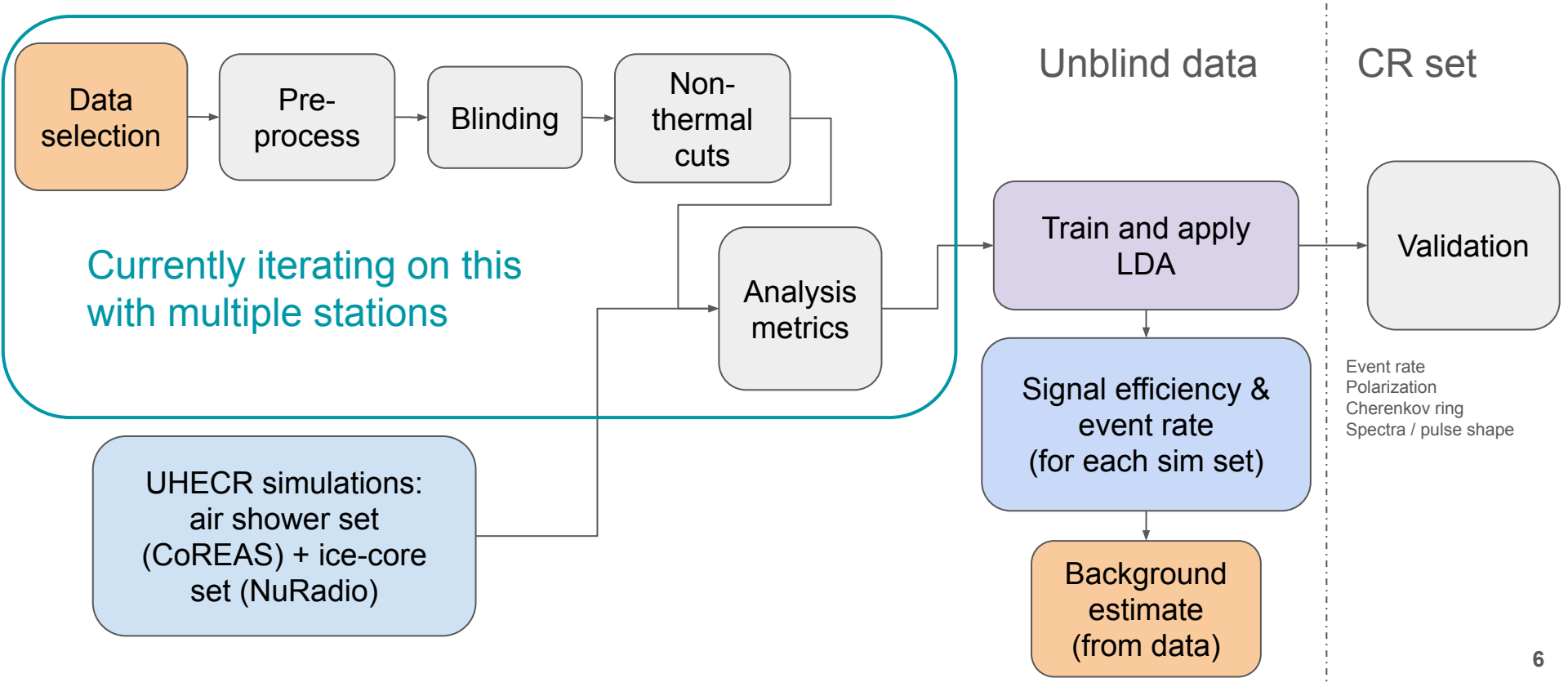
- Interpretability
- Relies on large datasets
- Less efficient

Exploratory analysis being developed in parallel for the neutrino search as well

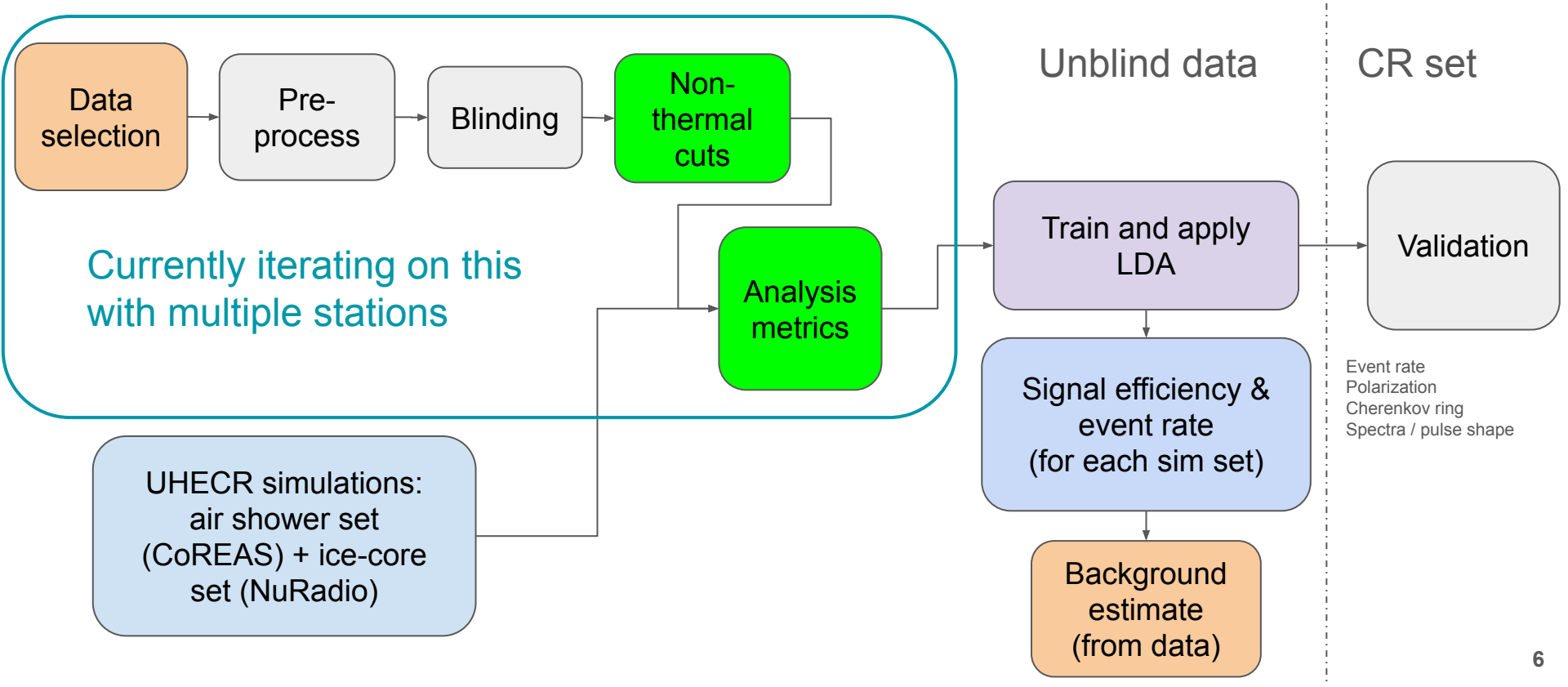
Analysis process with linear discriminant



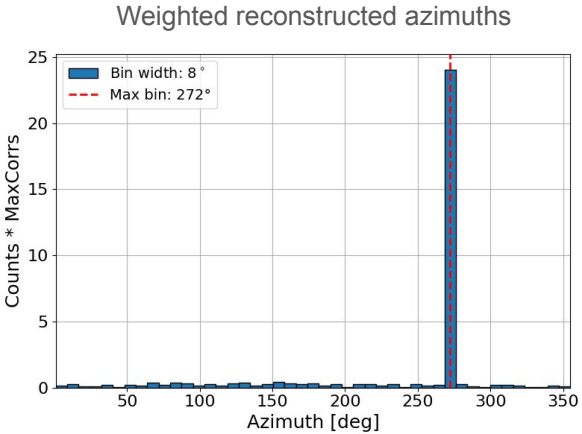
Analysis process with linear discriminant

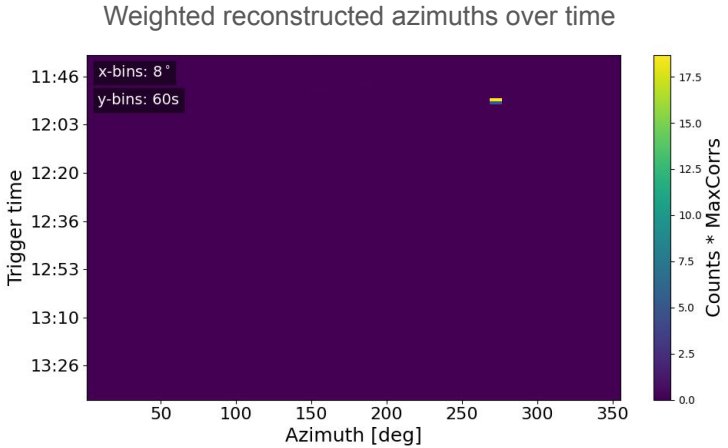


Analysis process with linear discriminant

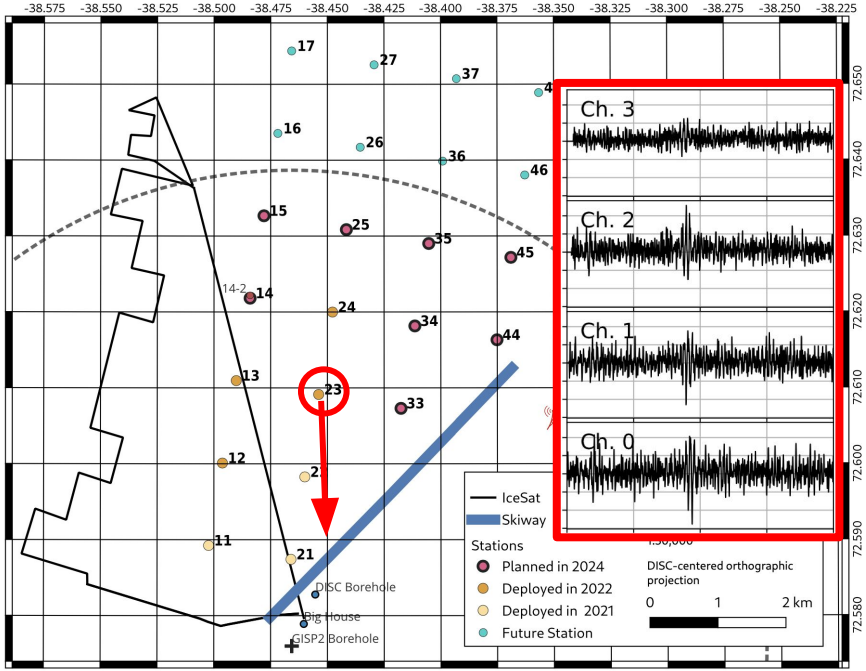


Non-thermal cut example: direction-based cut





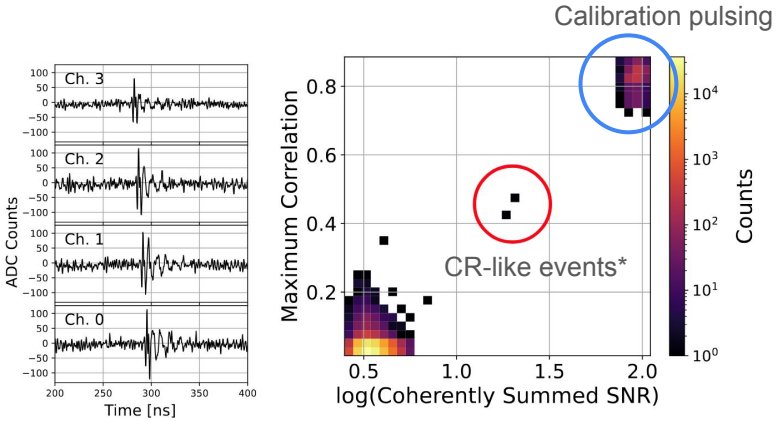
Bright repeated event pointing to Summit Camp indicating need for direction cut

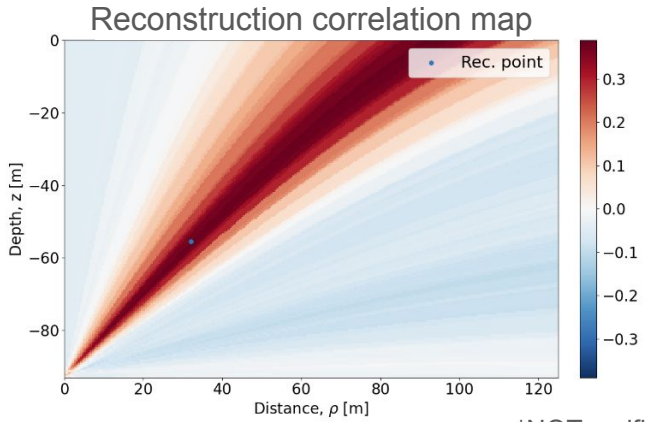


LDA features

Examples of analysis variable types which can separate impulsive, CR-like waveforms from thermal noise:

- Maximum amplitude
- Impulsivity
- SNR
- Surface correlation
- Direction





*NOT verified as CRs

Exploratory data analysis

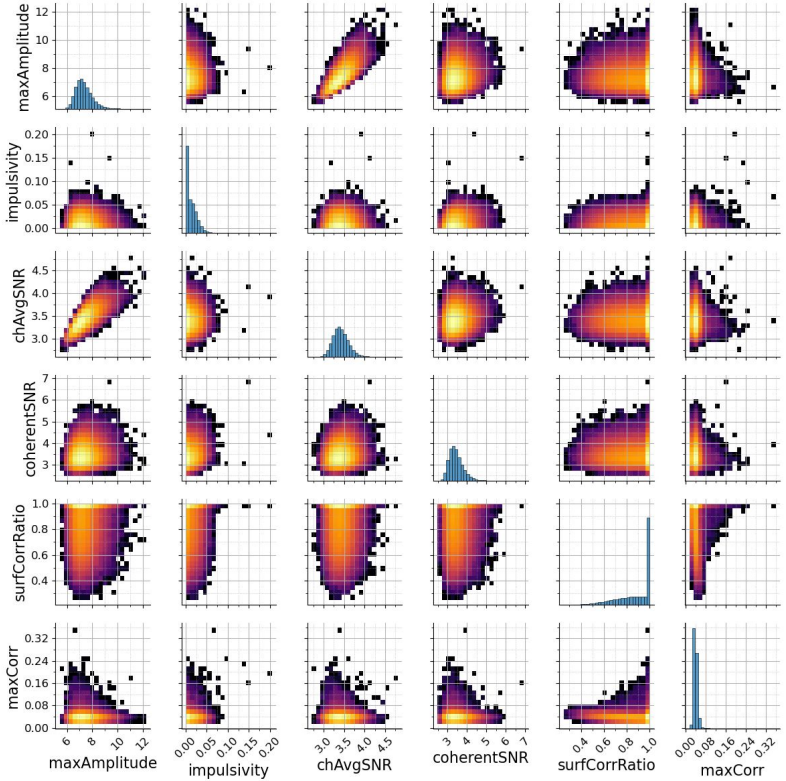
Distributions following cuts in the data

Results characteristic of thermal noise, with some outliers remaining

Exploring how to eliminate remaining outliers in systematic way as they are likely anthropogenic given the dataset size

Once able to demonstrate complete non-thermal background removal, and sims ready, will train on thermal data and simulated cosmic rays

Distributions of analysis variables on reduced dataset (~140k events)



Simulations and future work

FAERIE simulations being developed to provide full cosmic ray sims which will supply UHECR events for training

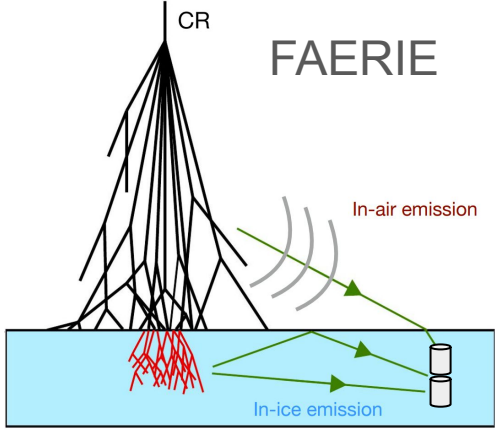
In the meantime, a simplified set of separated simulations will be used

- One set focuses on air shower emission using CoREAS and is propagated into the ice
- Another on ice-core emission using NuRadio

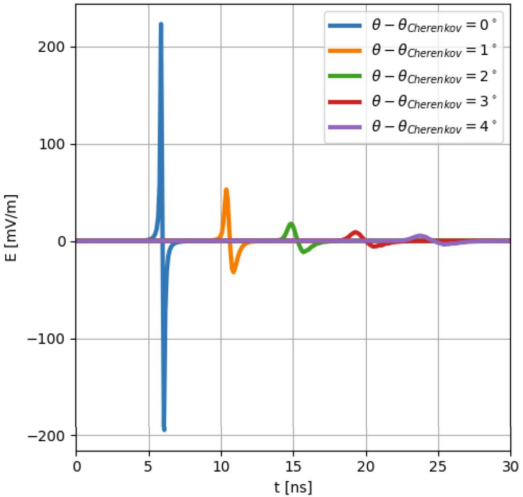
After training and applying classifier, candidates will be validated using

- Event rate
- Polarization
- Cherenkov ring
- Spectra / pulse shape

When full simulations available, goal will be to train classifier to distinguish between air shower emission, in-ice core emission, and thermal noise



arXiv:2409.02185v1



arXiv:2010.12279

Summary

UHECR sources can potentially be identified using neutrinos, but UHECRs themselves can be a difficult background for neutrino searches due to similar emission properties

First step in differentiating between neutrinos and CRs is to identify CRs in reduced feature space

LDA being set up to identify impulsive events against thermal noise in this reduced space

Clear separation shown between known impulsive events and thermal noise in current selected features for LDA

Simulations in development to provide UHECR training events and simplified set will be ready very soon