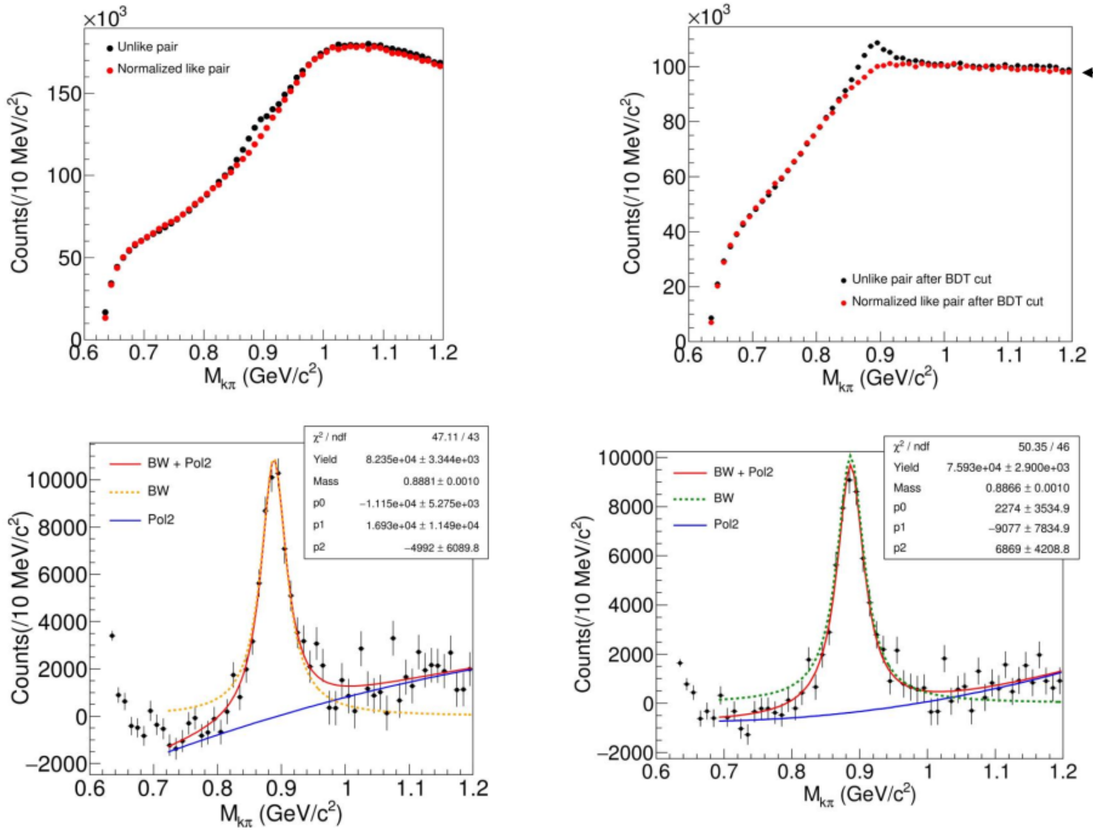
Machine Learning Applications in Dark Matter Search

-Viraj Thakkar,
NISER
24 April 2019

Previous work in High Energy Physics

- **Improving the significance of resonance signal (K^{*0}) using Machine Learning Techniques in high energy heavy-ion collisions.**

Invariant mass distribution: ALICE data



Signal is more visible!

Traditional approach

ML approach with BDT

Significance by traditional approach =36.01

Significance by ML approach =39.87

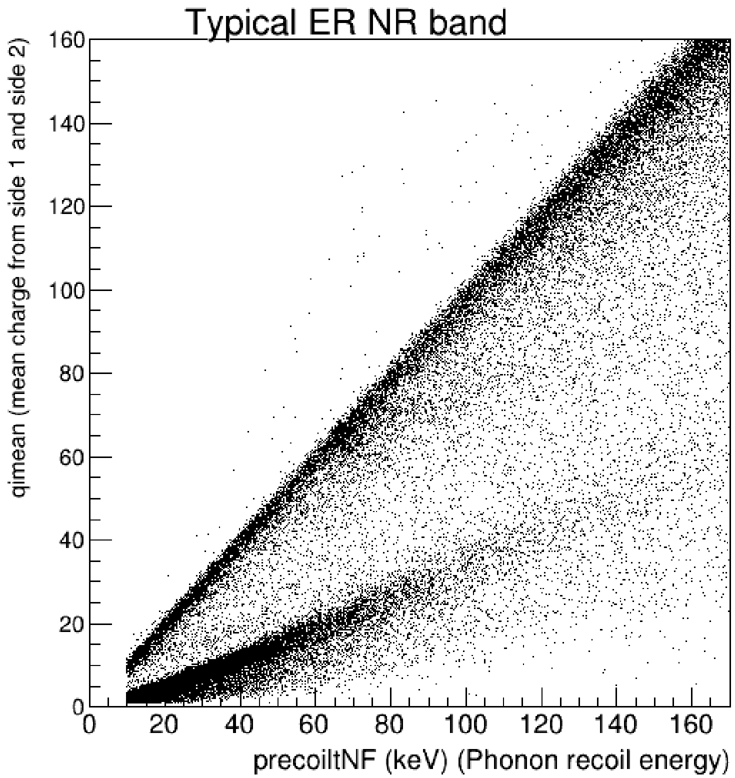
Ref. chap 3 and Appendix of thesis

Introduction to Dark Matter

- Compelling evidence from a number of astrophysical observations.
- Hypothetical form of matter thought to account for ~27% of the energy density of the universe.
- Does not interact with electromagnetic radiation, hence called “dark”.
- Non-baryonic in nature and possibly composed of yet undiscovered particles e.g Weakly interacting massive particles (WIMPs).
- They interact with gravity and other weak-like forces which are not a part of the Standard Model.
- We will be studying the dark matter detection experiment of SuperCDMS.

Dark Matter Search by SuperCDMS

- **Experiment**: Super Cryogenic Dark Matter Search (**SuperCDMS**).
- **Search**: WIMPs as a dark matter candidate.
- **Measures**: Phonon and charge signals.
- **Signal**: Nuclear Recoils (NR).
- **Background**: Electron Recoils (ER).
- **My work:** Separate ER and NR using Machine Learning Techniques.



Flow Chart of Analysis

Get dataset and apply preselection cuts

Traditional method

Fiducial Volume(FV) optimization
- Define phonon radial cut
- Define charge symmetric cut
- Define charge radial cut

ML method

Reconstruct relevant variables

Train and Test BDT

Apply cut on BDT score

Classification of ER and NR events

Data Set: Model

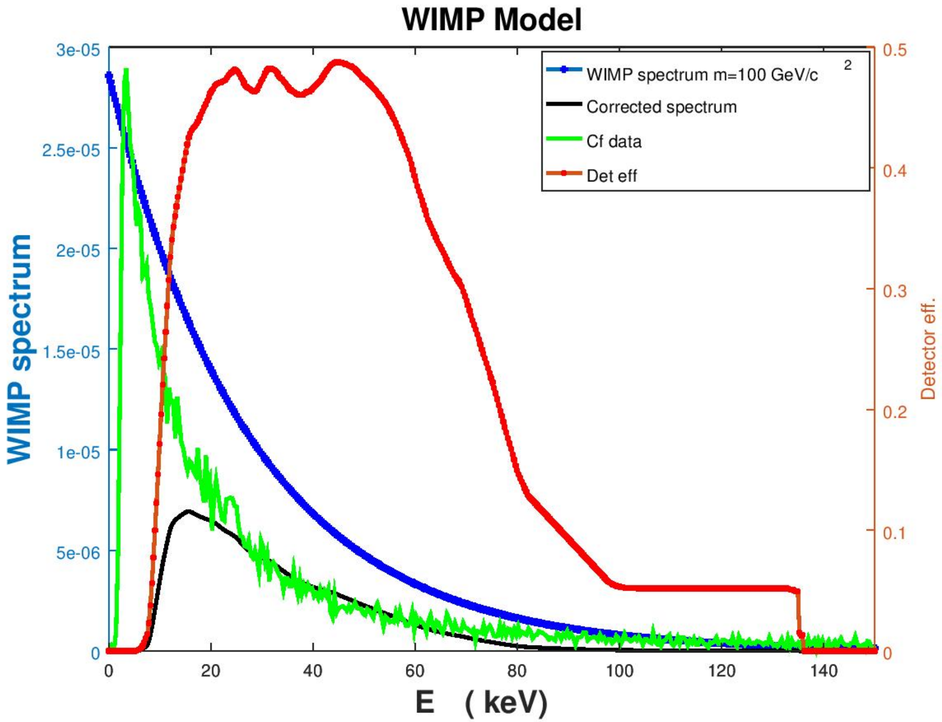
- **Events were taken from $^{252}Cf_{98}$ and $^{133}Ba_{56}$ data.**
- **As dark matter particles have not been discovered yet, we need a system to mimic the dark matter signal, and a system to mimic the background events.**
- **Dark matter signal: Neutrons from Cf source (Nuclear recoils)**
- **Background: Gamma from Ba source. (Electron recoils)**

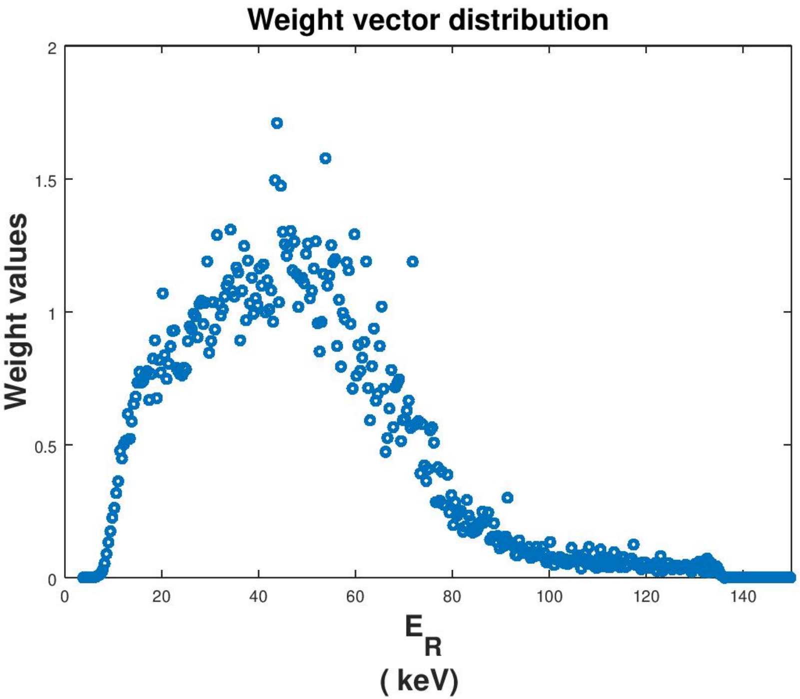
Data and Preselection cuts

- Data was collected during the period Aug 2013-Aug 2014.

Not Empty	Removes empty events in the detector with no trigger.
Good Flashtime	Removes events when the detector was not flashed for period greater than 3300 seconds.
Base Temperature	Removes events when the base temperature was not in the good range.
Voltage Bias	Ensures voltage was maintained within 4V range.
Good Start time	Removes events with Non Stationary Optimal Filter delay.
Analysis threshold	Phonon recoil energy greater than 10 keV.

WIMP Signal Model

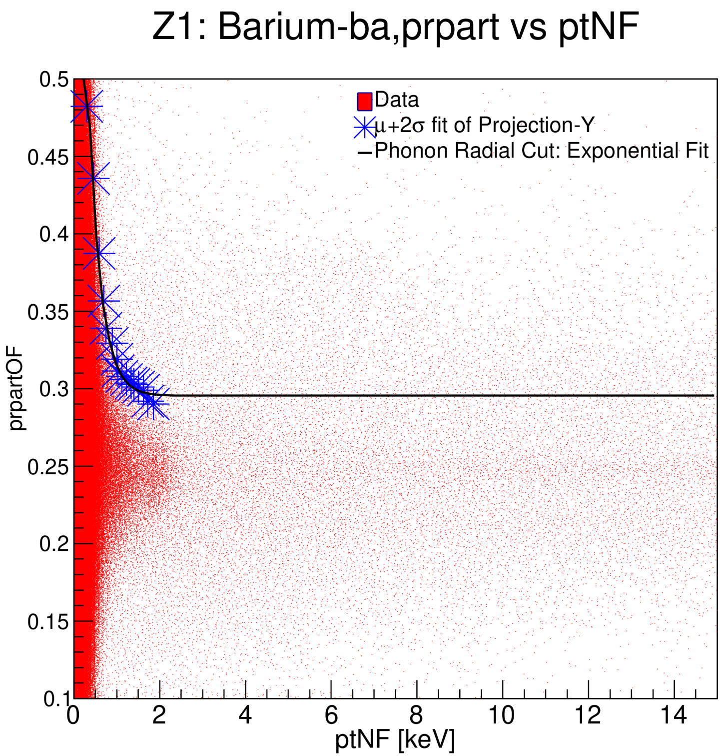




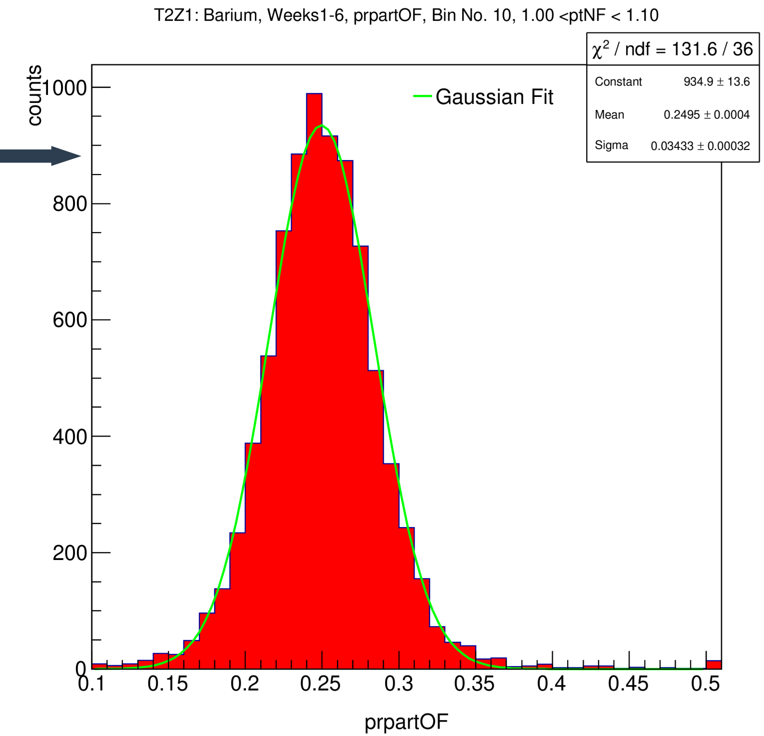
- **Blue** is WIMP spectra from theory.
- **Green** is Cf data distribution.
- **Red** is detector efficiency.
- **Black** is the corrected spectrum.

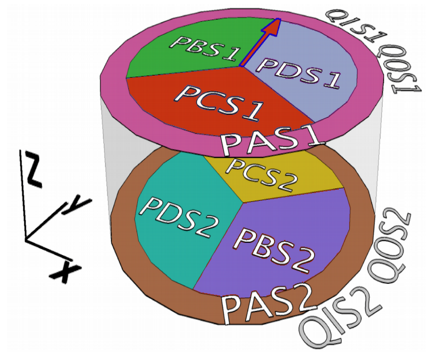
- Weight vector is calculated by taking the ratio of corrected sprectrum (**black**) and Cf data (**green**)
- Weigh each event of the Cf data by the corresponding weight in the weight vector.

FV: Phonon Radial Cut



Projection along Y for a ptNF bin, fitted with a gaussian function

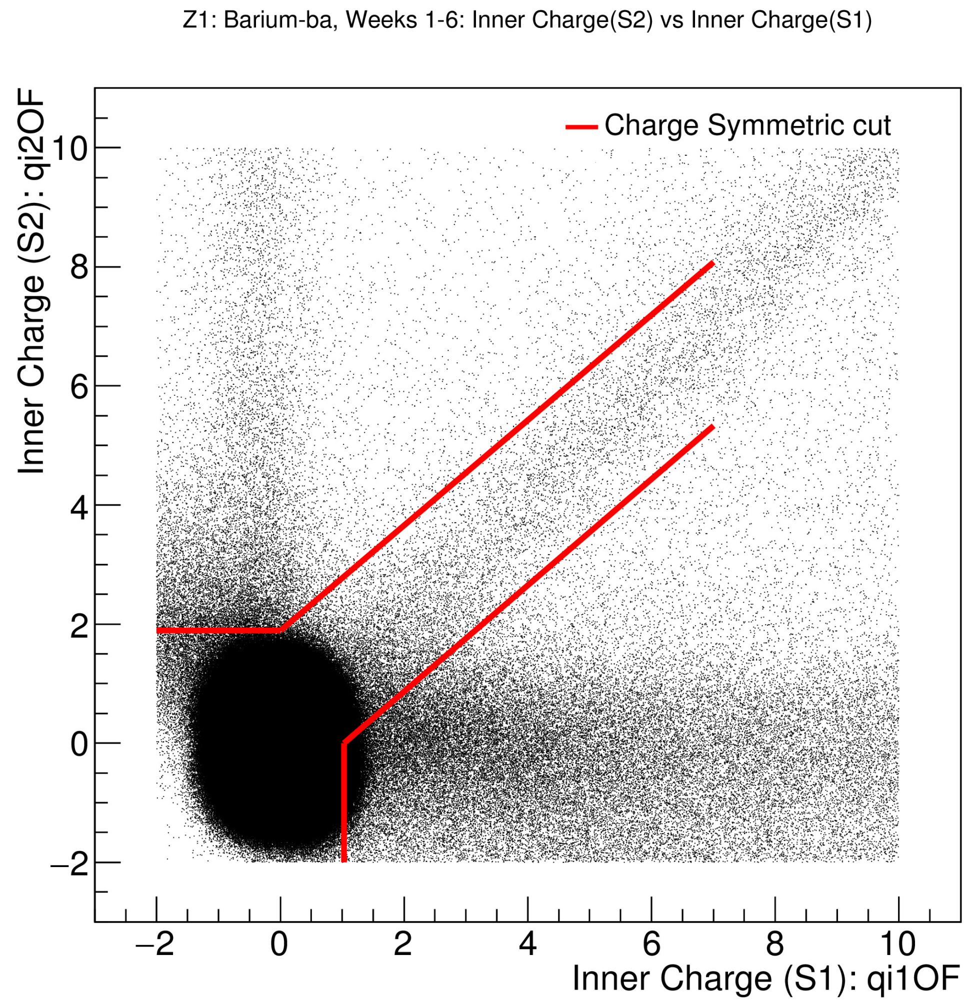




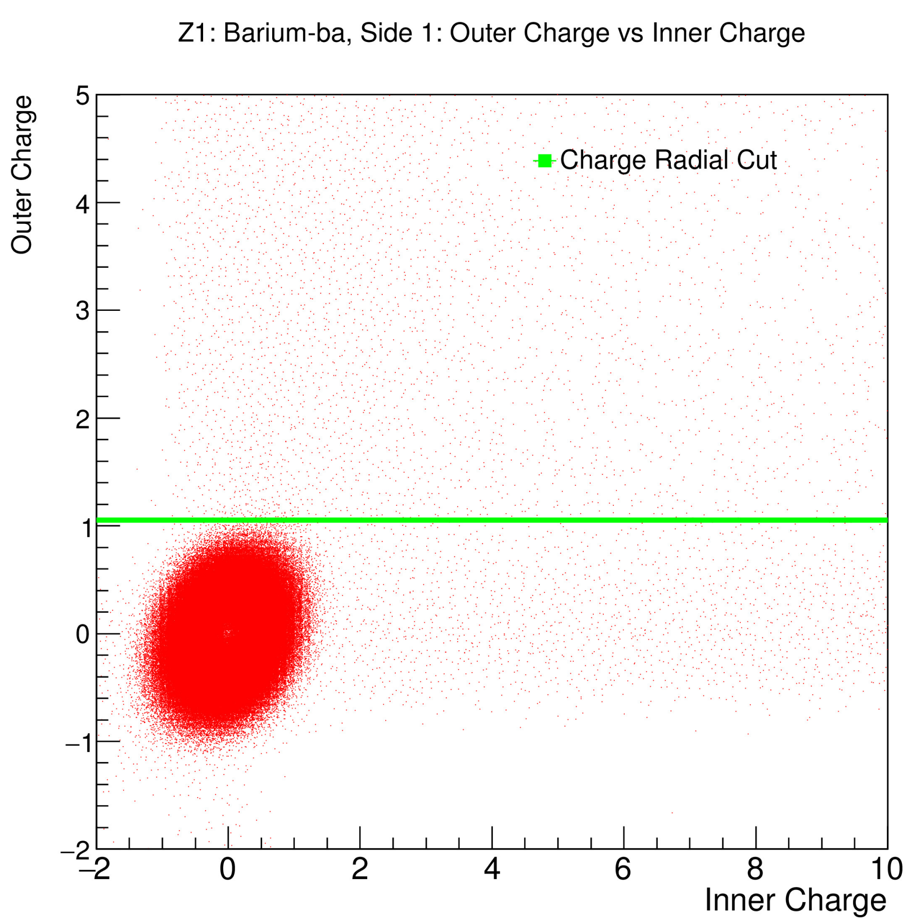
- **prpartOF is the ratio of outer phonon sum and total phonon energy. ptNF is total phonon energy**
- **Phonon radial cut removes high-radius phonon events i.e events above the Exponential fit.**

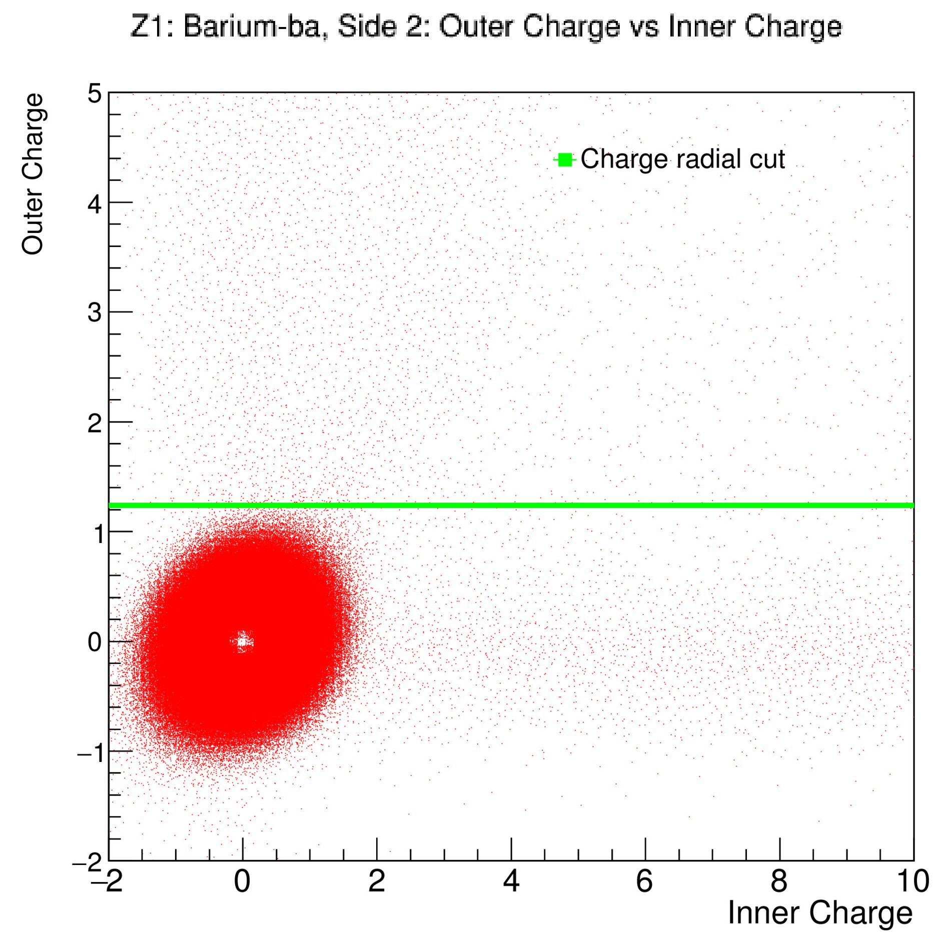
FV: Charge Symmetric Cut

- **Inner charges collected from side 1 and side 2.**
- **Geometry of the detector→ Asymmetric charge collection near surface, symmetric charge collection in the bulk region.**



FV: Charge Radial Cut





Removes events closer to the cylindrical sidewall of the detector i.e larger value of outer charges.

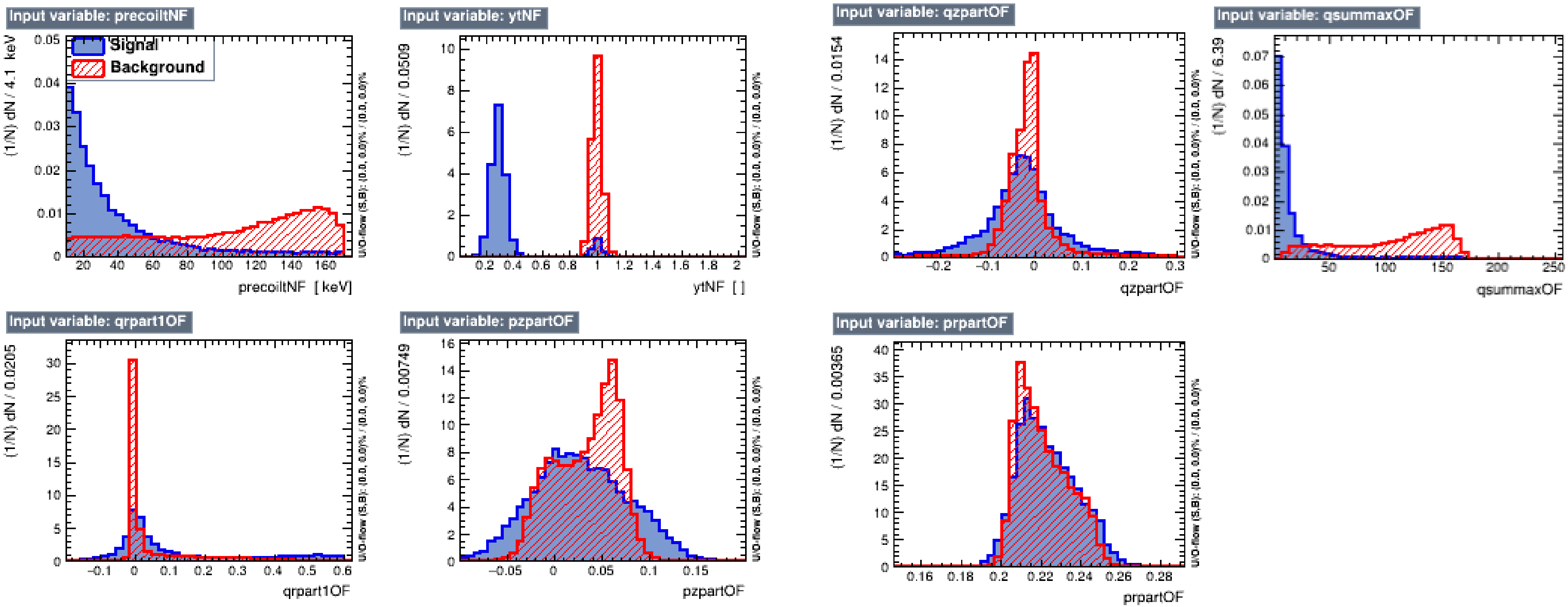
ML classification of NR and ER events

- **We have used Cf data as a source of NR events and Ba data as a source of ER events.**

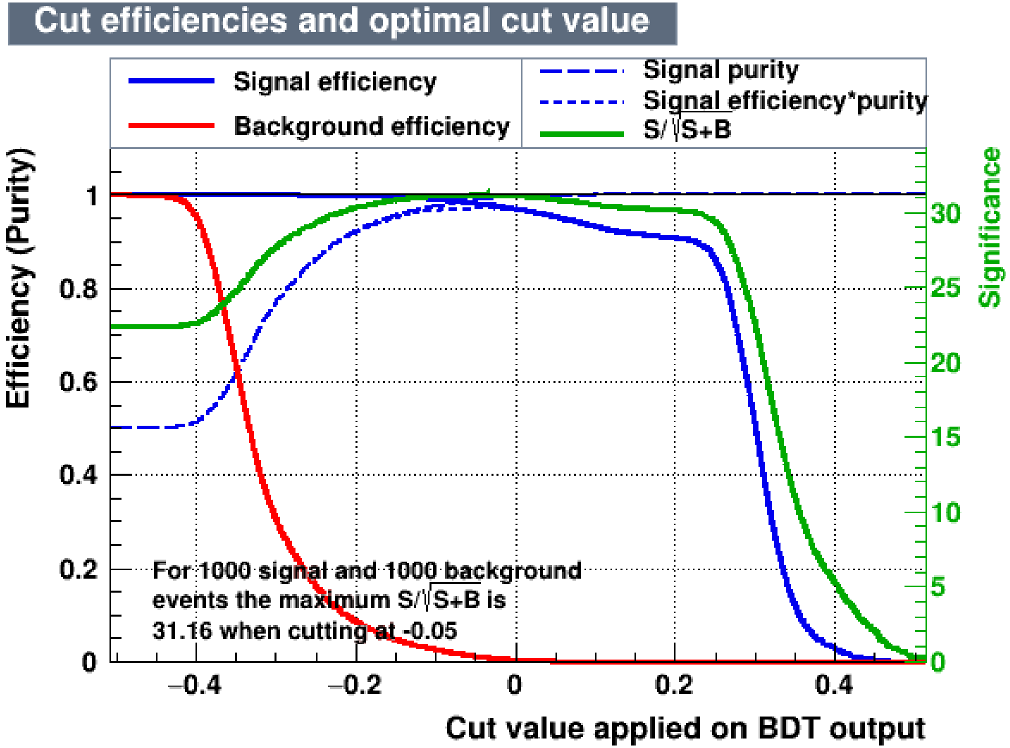
Features used for Training our BDT

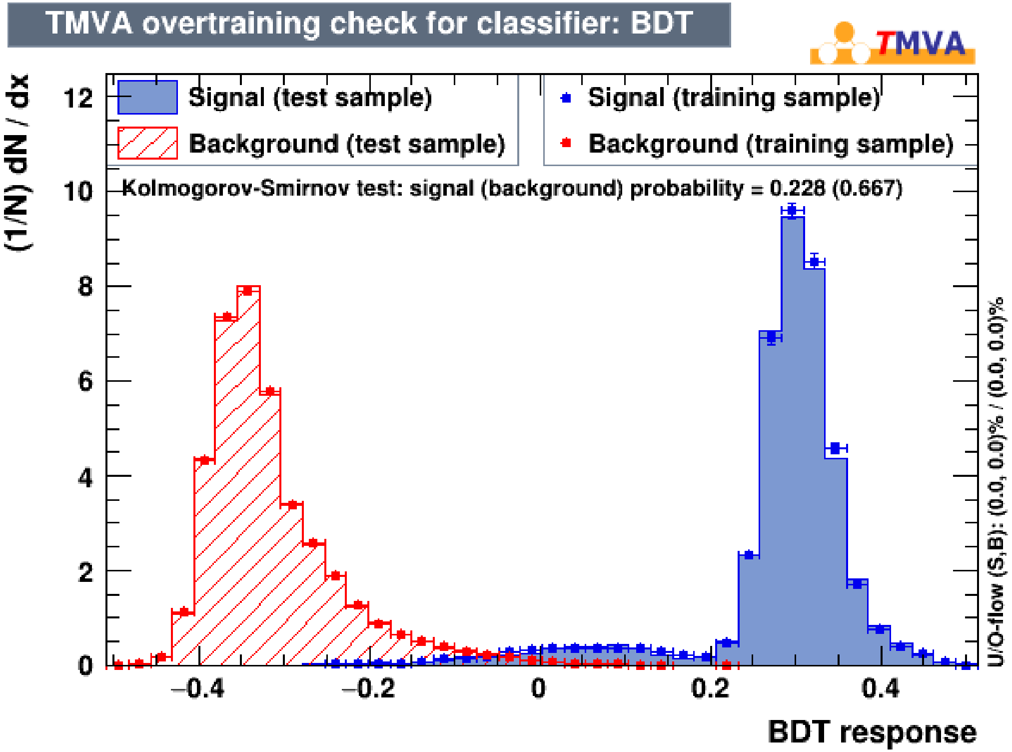
precoiltNF:	the non-luke phonon energy
qsummaxOF:	sum for the side with the maximum qsum (inner+outer charge)
ytNF:	Yield i.e qsummaxOF/precoiltNF
qzpartOF:	Z-direction fiducializing parameter for charge
qrpartOF:	Radial fiducializing parameter for charge
pzpartOF:	Z-direction fiducializing parameter for phonon
prpartOF:	Radial fiducializing parameter for phonon

Input features used for training of BDT



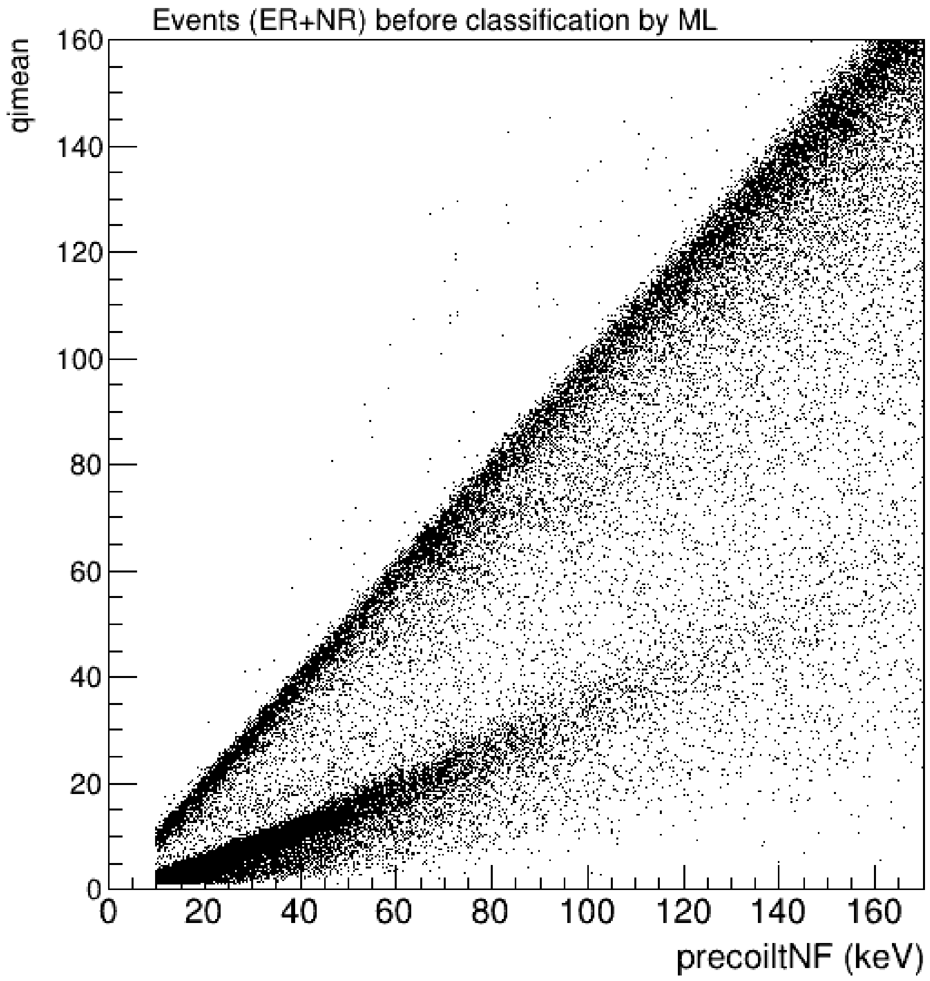
BDT distributions

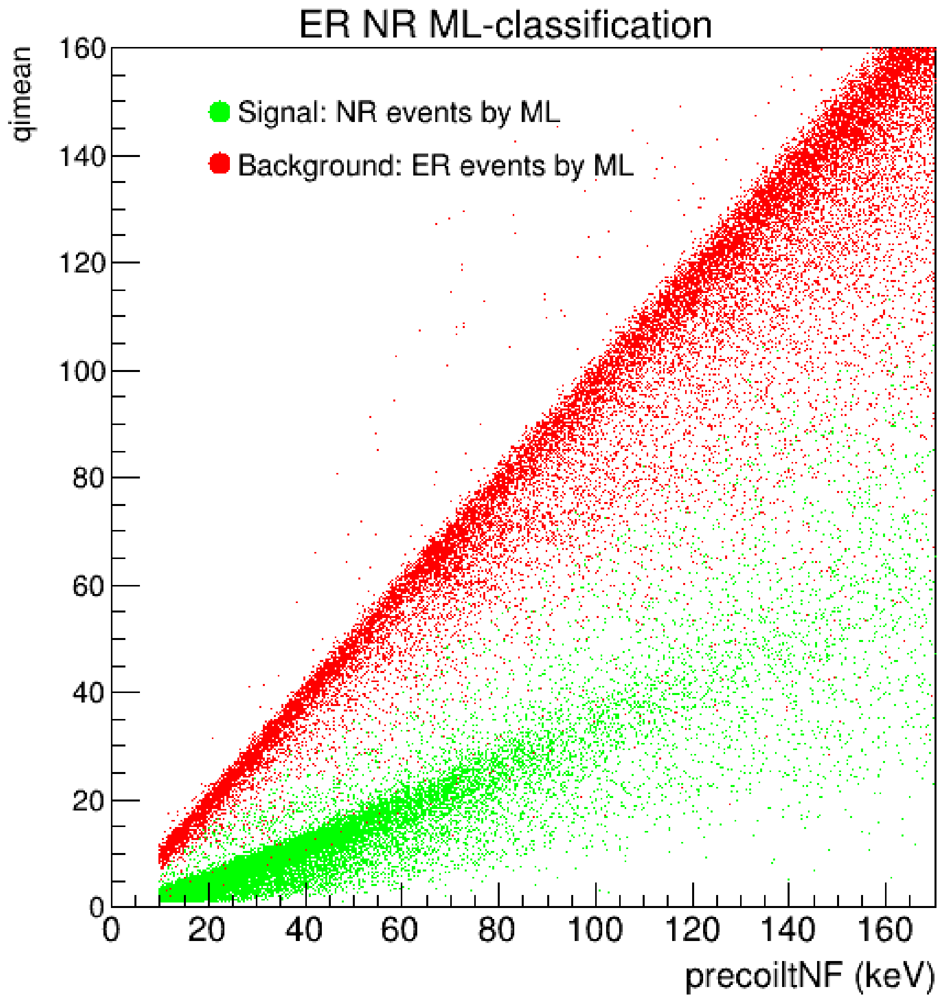




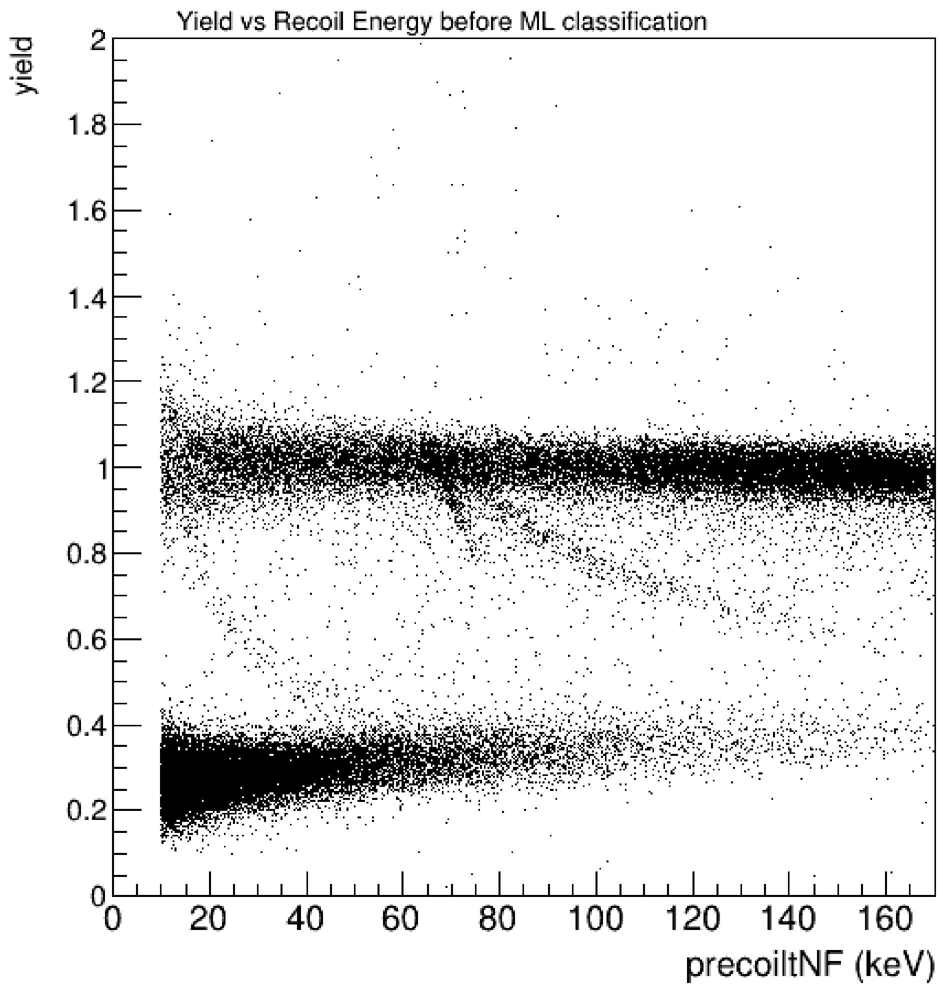
- We take the BDT cut value to be -0.0528
- Events with BDT response> BDT cut value are classified as signal (NR).

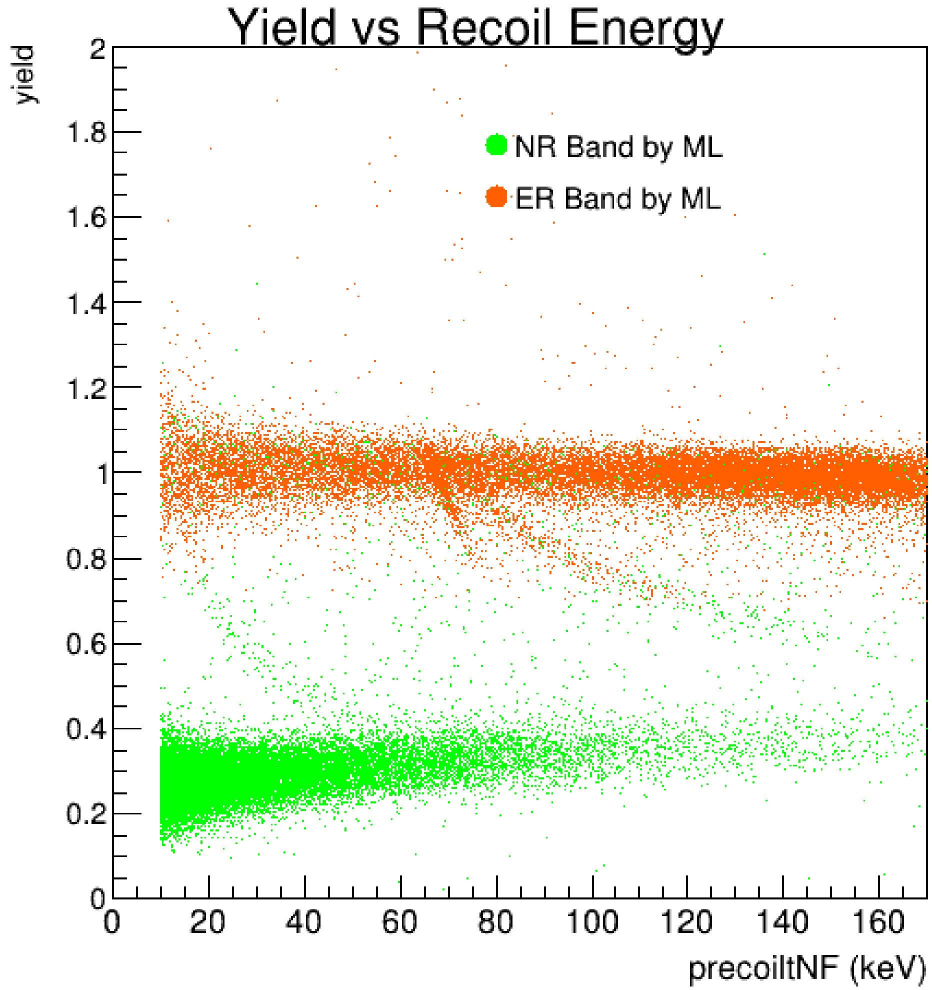
ML Application Phase results: ER NR bands



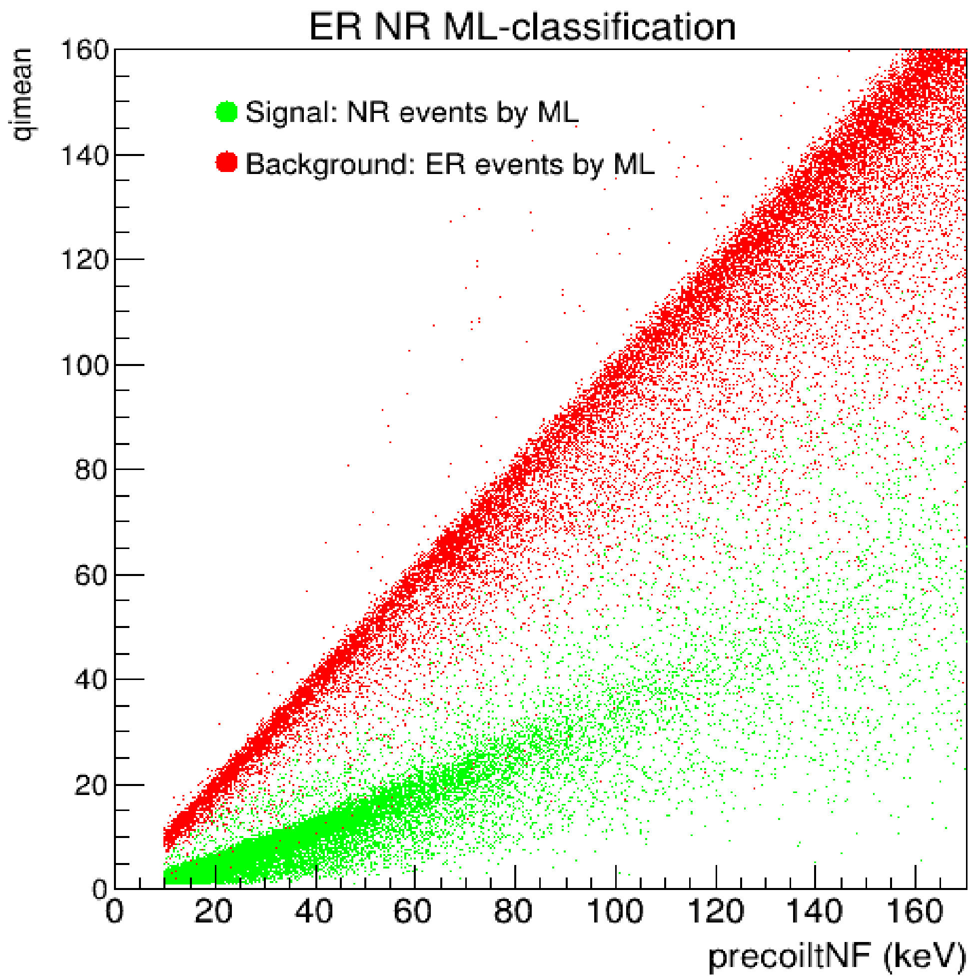


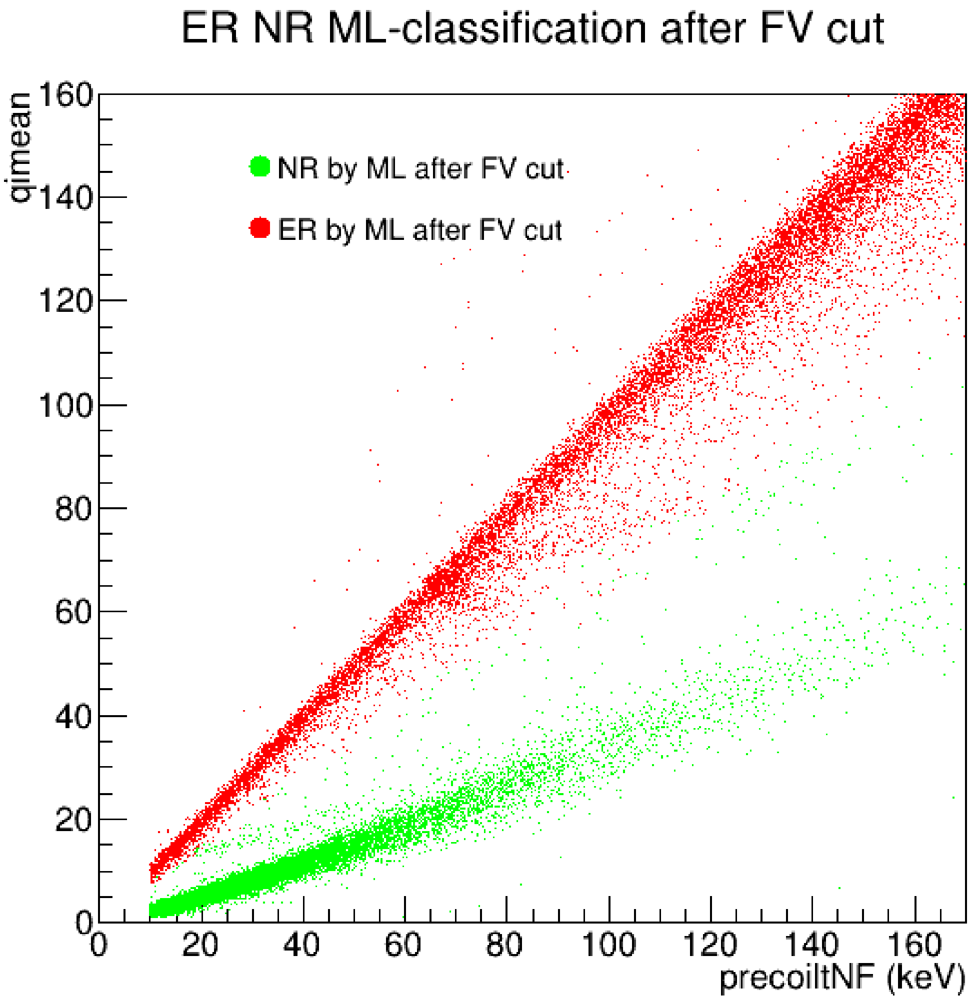
ML Application Phase results: Yield bands



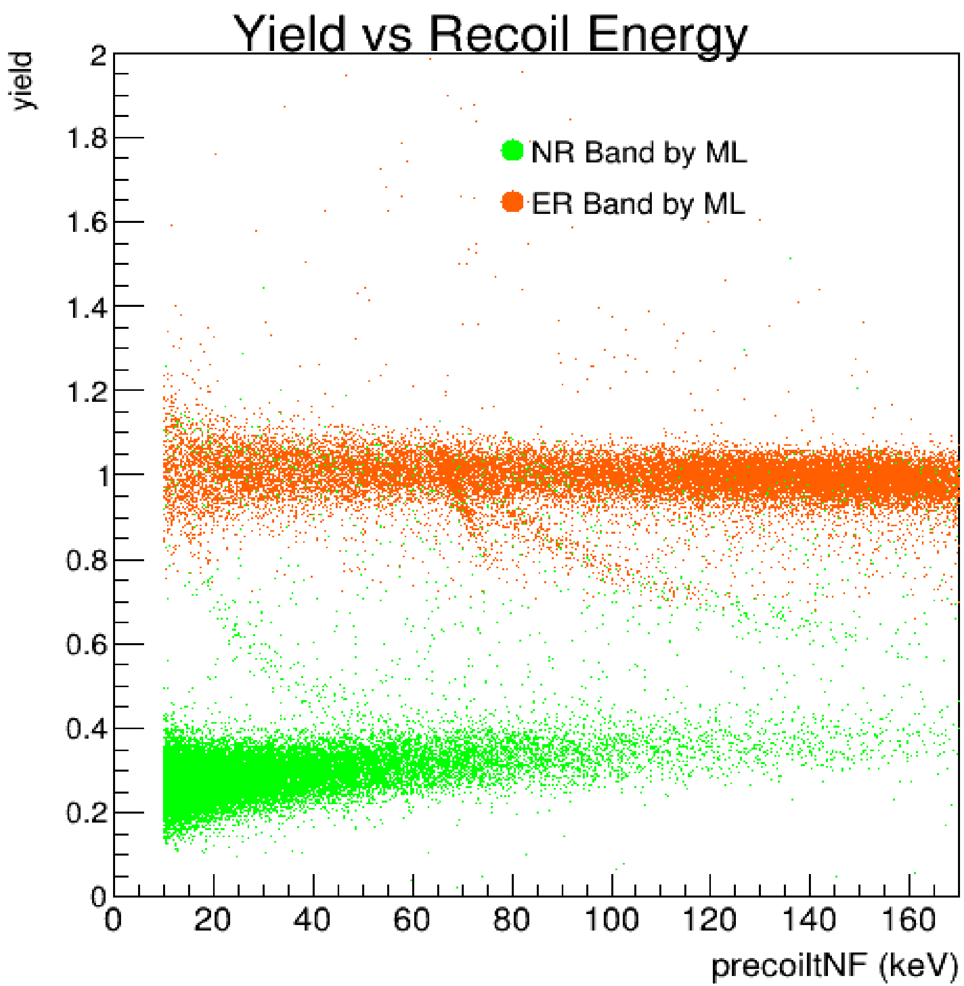


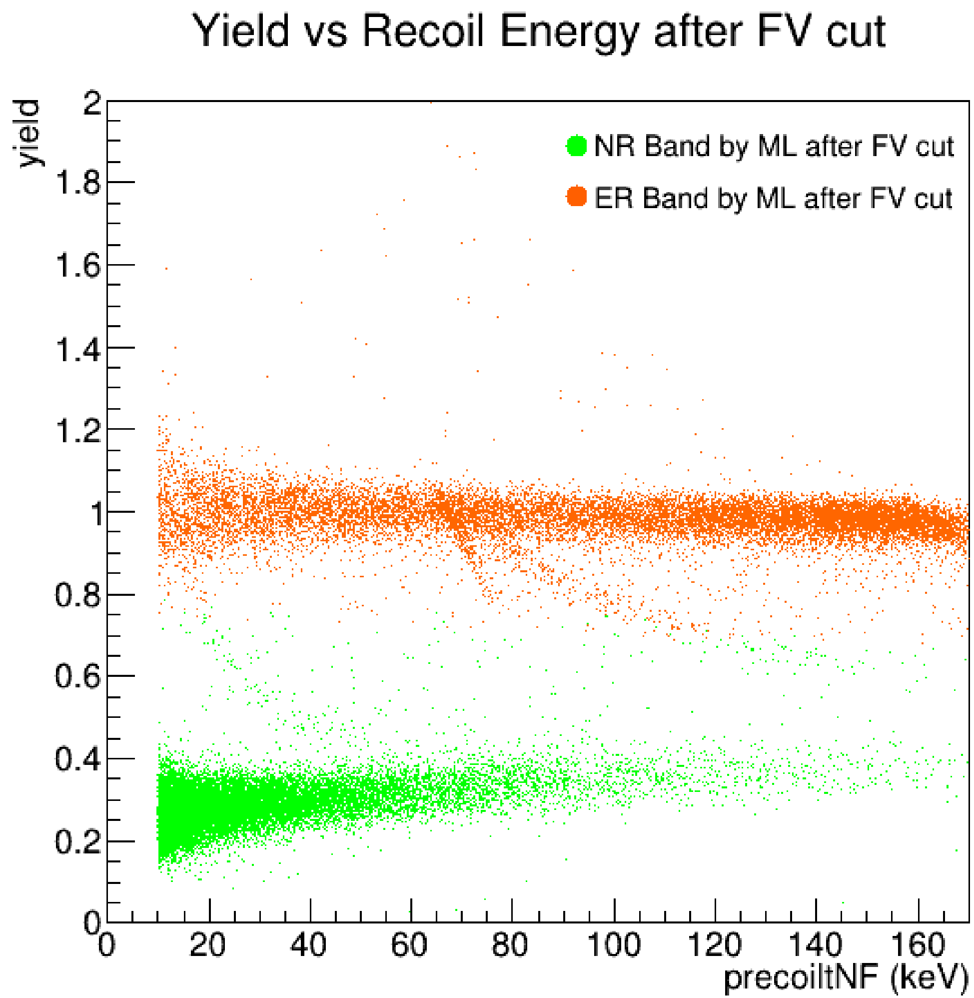
FV cuts and BDT cut





FV cuts and BDT cut: yield bands





Conclusion

- BDT cut provides an event by event separation of Electron Recoils and Nuclear Recoils.
- We achieve an enhanced NR band after FV cuts and BDT classification.
- We have demonstrated a way to get a pure NR band (and remove ER background contributions) in the detector using ML techniques, this can be adapted for DM search.

Outlook

- Further classification of the NR events which may have possibly come from WIMPs and neutrons.

Backup

SuperCDMS experiment detector

- The Super Cryogenic Dark Matter Search (**SuperCDMS**) experiment is designed to directly detect particle dark matter in the form of WIMPs.
- The Ge-made interleaved Z-sensitive Ionization and Phonon (iZIP) detectors aims at measuring and distinguishing ionization and phonon signals for nuclear recoils(NR) produced by WIMPs and electron recoils(ER) produced by background sources.

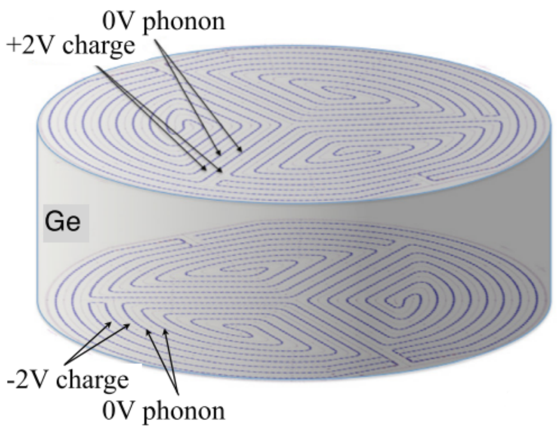


Figure reference: *Applied Physics Letters*, vol. 103, no. 16, p. 164105, 2013.

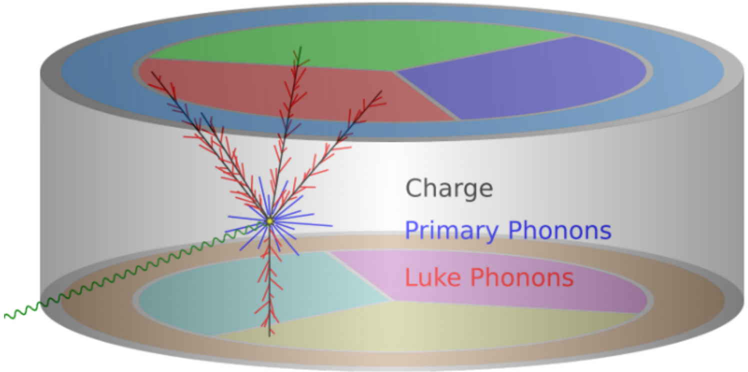


Figure reference: https://thesis.library.caltech.edu/11056/

Fiducial Volume Optimization

- **Fiducial Volume Optimization→ removes unwanted surface events and outer edge events. (“Fiducialization”)**
- **These unwanted events lead to misidentification of ERs with suppressed ionization collection as NRs.**
- **Fiducial Volume (FV) of the detector → Hypervolume which is set by optimizing cuts in various parameter space.**
- **Events inside the FV are called “bulk” events and events outside the FV are called “surface” events.**

Features chosen (Reasons)

precoiltNF:	Difference in distribution at low phonon energy
qsummaxOF:	sum for the side with the maximum qsum (inner+outer charge)
ytNF:	Yield i.e qsummaxOF/precoiltNF, a classic CDMS discriminator.
qzpartOF:	Z-direction fiducializing parameter for charge, removes near surface events
qrpartOF:	Radial fiducializing parameter for charge
pzpartOF:	Z-direction fiducializing parameter for phonon
prpartOF:	Radial fiducializing parameter for phonon

Theory spectra from Lewin and Smith

$$\frac{\mathbf{dR}}{\mathbf{dE_R}} = \frac{\mathbf{R_0}}{\mathbf{E_0 r}} \mathbf{e}^{-\mathbf{E_R}/\mathbf{E_0 r}}$$

$\rightarrow$ $\mathbf{E_R}$ is the Recoil Energy

$\rightarrow$ $\mathbf{E_0}$ is the most probably KE of dark matter of mass $\mathbf{M_D}$

$\rightarrow$ $\mathbf{R}$ is the event rate per unit mass

$\rightarrow$ $\mathbf{R_0}$ is the total event rate

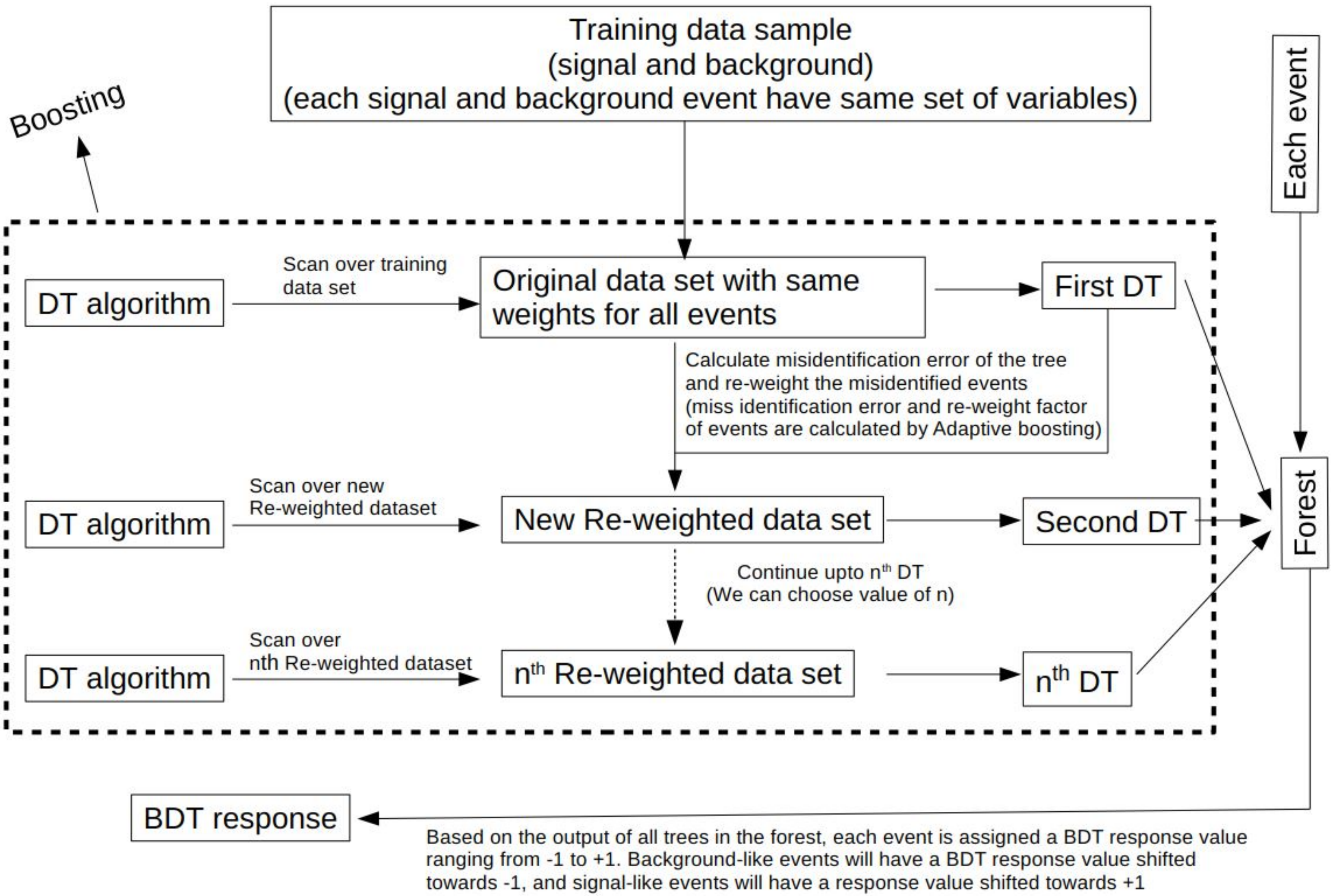
$\rightarrow$ $\mathbf{r = 4M_D M_T/(M_D + M_T)^2}$

- **Event rate is expressed in events/keV*kg*day or "dru".**
- **Reference: J.D. L.ewin, PE Smith/Astroparticle Physics 6 (1996)**

Final Numbers after FV and BDT cut

- **Passage fraction (ratio of number of events passing the FV cut and total number of events)=67.7 %**
- **The fraction of events classified as NR by ML improves from 54.49 % to 56.28 % after the application of FV and BDT cut.**

BDT algorithm flow chart



AdaBoost: learning ensemble

- Start same weight for all points: $\alpha_i = 1/N$
- For t = 1,…,T
  - Learn $f_t(\mathbf{x})$ with data weights α_i
  - $\boxed{\text{Compute coefficient } \hat{w}_t}$
  - $\boxed{\text{Recompute weights } \alpha_i}$
  - $\boxed{\text{Normalize weights } \alpha_i}$
- Final model predicts by:

$$\hat{y} = sign\left(\sum_{t=1}^{T} \hat{\mathbf{w}}_t f_t(\mathbf{x})\right)$$

$$\boxed{\hat{\mathbf{w}}_t = \frac{1}{2}\ln\left(\frac{1 - weighted\_error(f_t)}{weighted\_error(f_t)}\right)}$$

$$\alpha_i \leftarrow \begin{cases} \alpha_i\, e^{-\hat{w}_t}, & \text{if } f_t(\mathbf{x}_i) = y_i \\ \alpha_i\, e^{\hat{w}_t}, & \text{if } f_t(\mathbf{x}_i) \neq y_i \end{cases}$$

$$\boxed{\alpha_i \leftarrow \frac{\alpha_i}{\sum_{j=1}^{N} \alpha_j}}$$

BDT parameters

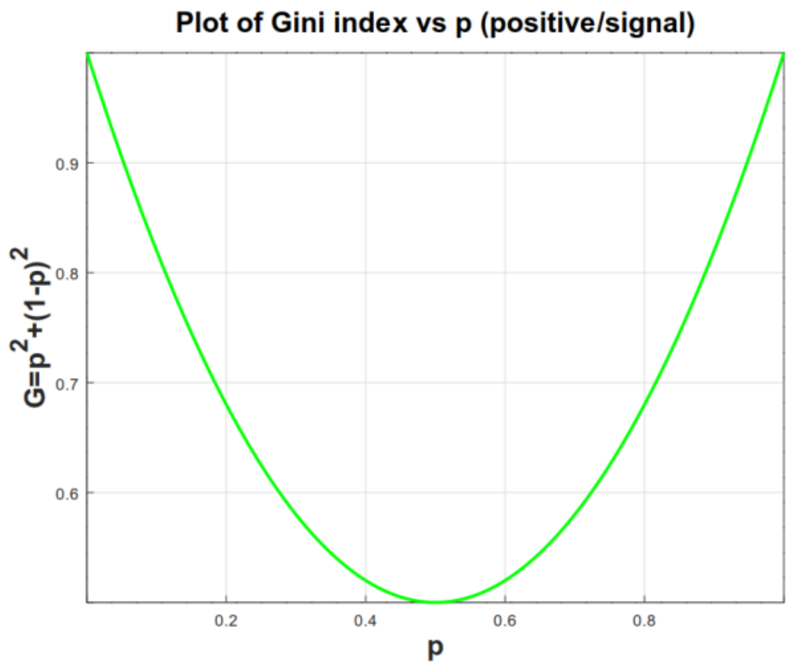
- **Number of Trees= 850**
- **Min node size i.e Minimum percentage of training events required in a leaf node= 2.5%**
- **Max depth i.e max depth of DT= 3**
- **Learning rate for AdaBoost algorithm adaboostbeta= 0.5**
- **Separation Type= Gini Index**
- **Ncuts Number of grid points in variable range used in finding optimal cut in node splitting=20**

Gini Index

How do I select which feature I take first?

Gini Index : Define as $p^2 + q^2$ (at a given node)

- p: fraction of positive (signal) events
- q=1-p : fraction of negative (background) events



Weighted Gini Split:

$$\mathbf{I_G = f_{left} * G_{left} + f_{right} * G_{right}}$$

where,

$$f_{left} = \text{fraction of events which go in the left split}$$

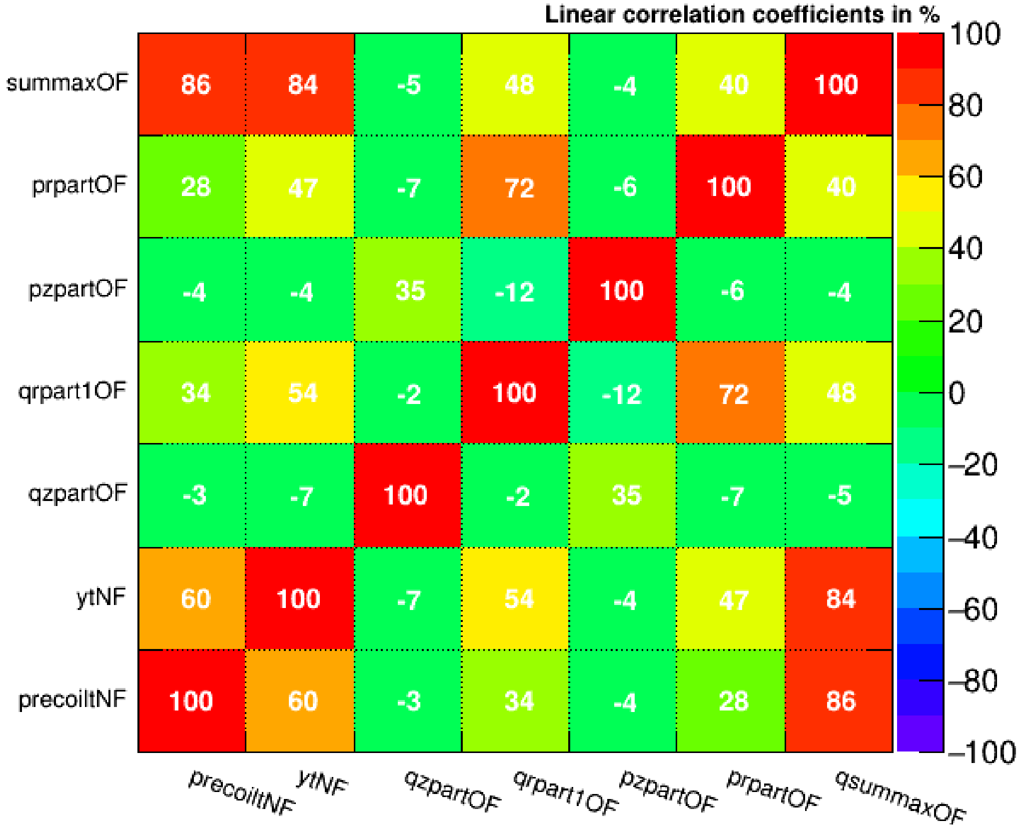
$$f_{right} = \text{fraction of events which go in the right split}$$

$$G_{left} = \text{Gini index of left node}$$

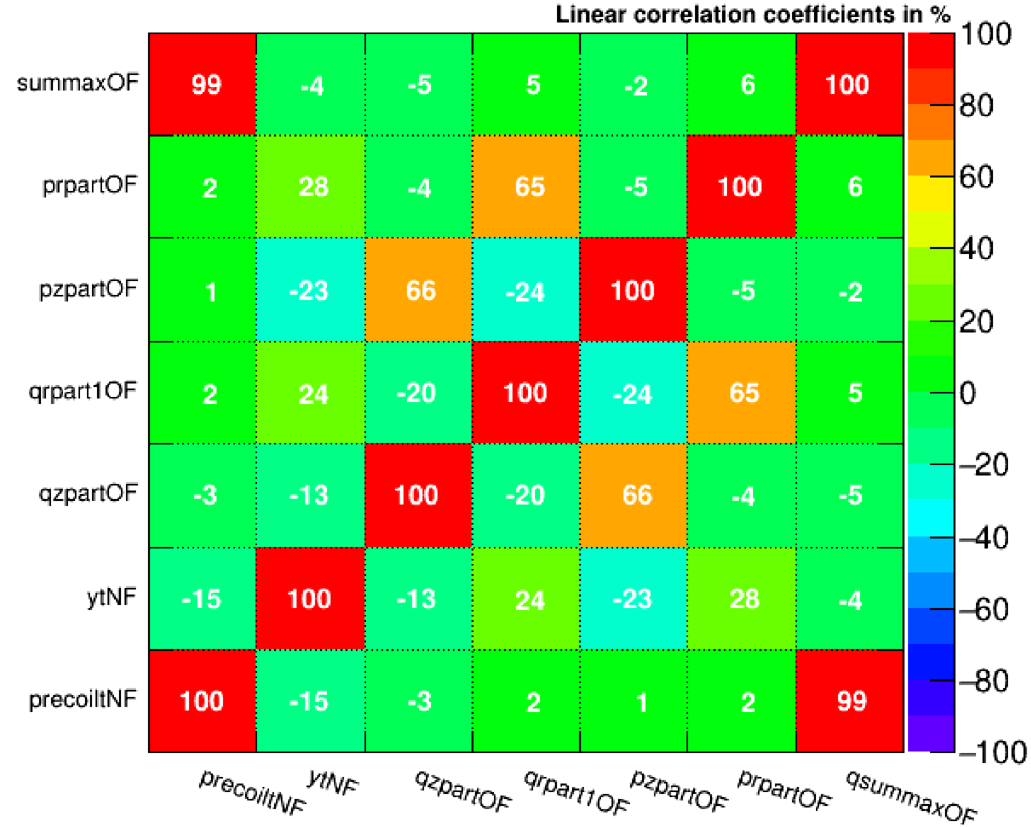
$$G_{right} = \text{Gini index of right node}$$

Correlations

Correlation Matrix (signal)



Correlation Matrix (background)



Additional content

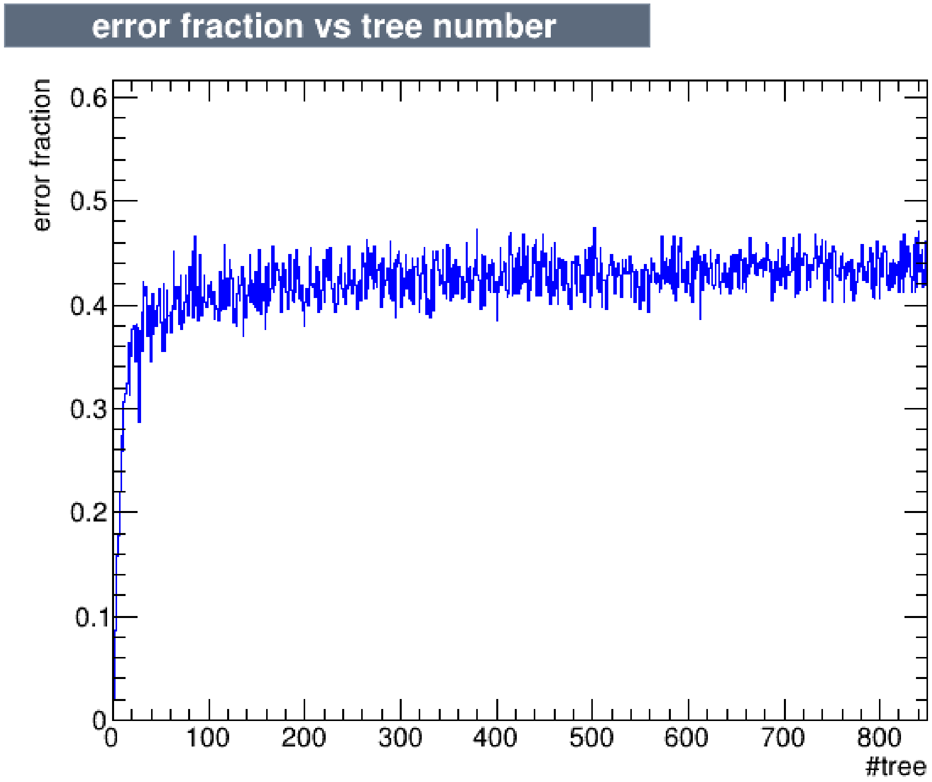
The extracted signal of K^{*0} is fitted with Breit-Wigner distribution plus a residual background function. Mass and width for K^{*0} can be obtained from fit parameters. Breit-Wigner function is defined below:

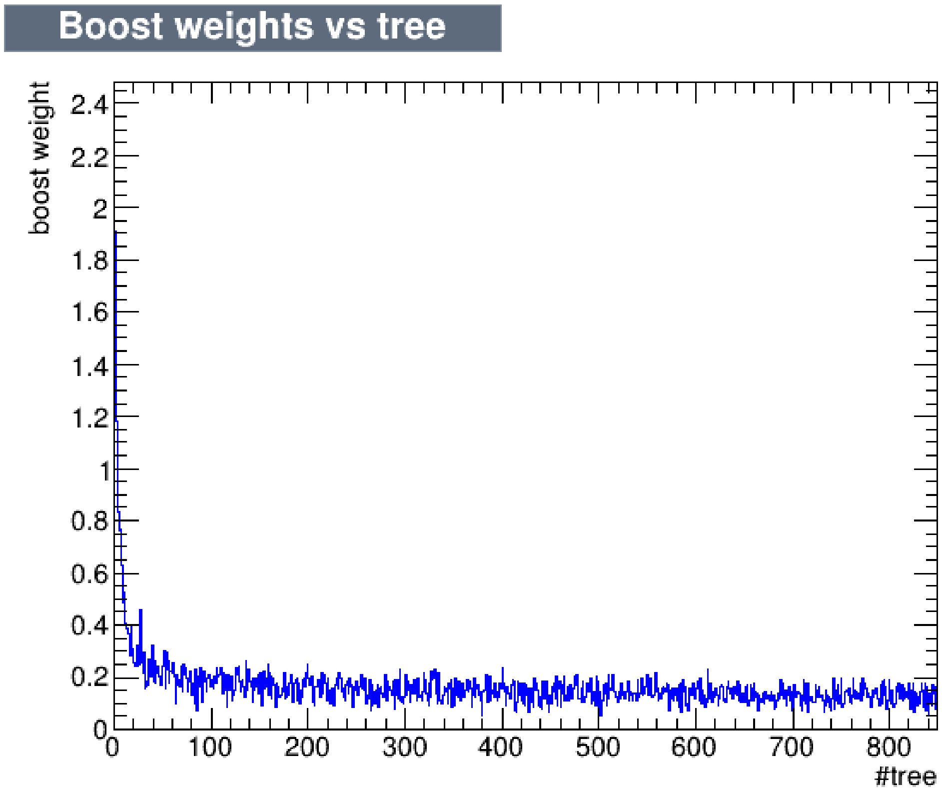
$$\frac{Y}{2\pi} \frac{\Gamma_0}{(M_{\pi K} - M_0)^2 + \frac{\Gamma_0^2}{4}} + Res.Bkg. \tag{1}$$

where M_0 and Γ_0 are the mass and width of the K^{*0}. $M_{\pi K}$ is πK invariant mass. The parameter Y gives the Breit-Wigner area. The last term is residual background function, which is taken as the polynomial of second order in invariant mass ($A M_{\pi K}^2 + B M_{\pi K} + C$). The parameter Y for each p_T bin gives the raw yield counts of K^{*0}. Extracted signal of K^{*0} is fitted with Breit-Wigner distribution and is shown in the right panel of Fig. 14 and Fig. 15 for two different cases (mixed background and like sign event background), where blue line indicates fit for signal with residual background and red line indicates fit for only residual background.

```
                         : Ranking input variables (method specific)...
BDT                      : Ranking result (top variable is best ranked)
                         : -------------------------------------------
                         : Rank : Variable    : Variable Importance
                         : -------------------------------------------
                         :    1 : ytNF        : 1.721e-01
                         :    2 : precoiltNF  : 1.578e-01
                         :    3 : qsummaxOF   : 1.445e-01
                         :    4 : pzpartOF    : 1.425e-01
                         :    5 : qzpartOF    : 1.389e-01
                         :    6 : qrpart1OF   : 1.334e-01
                         :    7 : prpartOF    : 1.108e-01
```

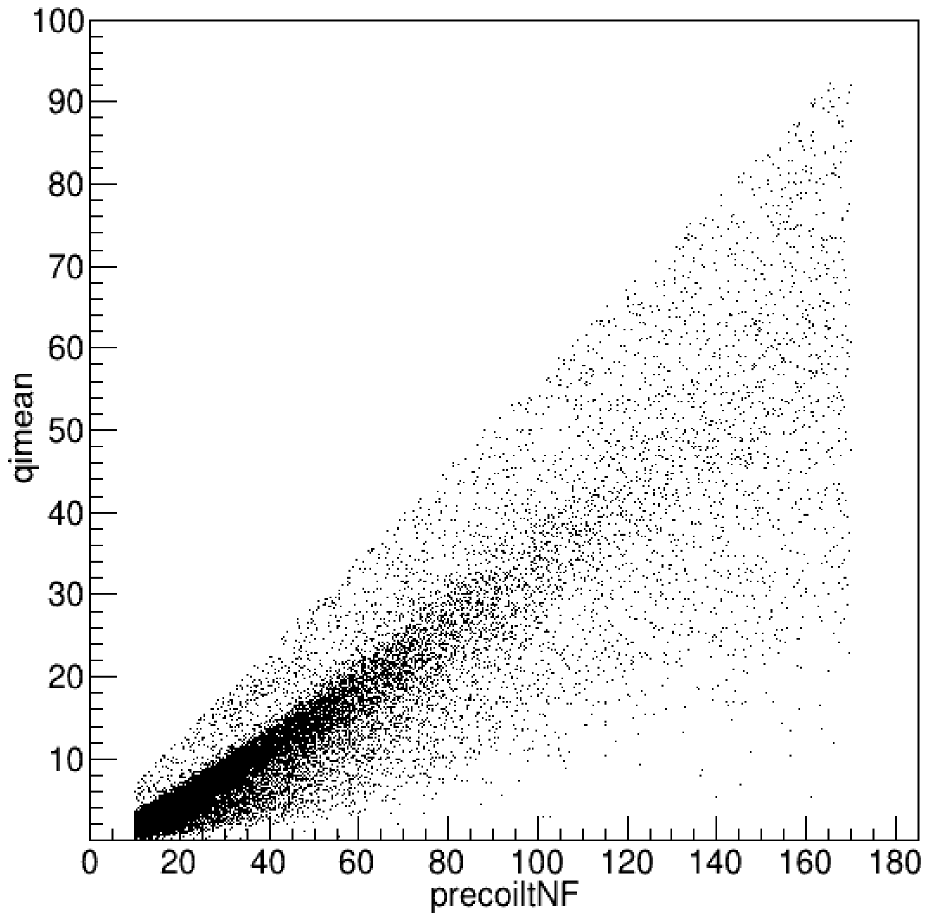
Boosted Decision Trees Control plots





Training events selection

NR events from Cf



ER events used from Ba