PMTs: Studies of the main variable distributions

CYGNO Analysis Meeting

David Marques

& PMT analysis working group:
F. Borra, M. Folcarelli, S. Piacentini,
A. Messina, E. Baracchini, E. Kemp

8/06/2023 - Coimbra

PMT group works rundown:

1. **On-going:**
   a. PMT reconstruction **TTree**
   b. **(L,x,y) fit**
      i. Bayesian fit approach
      ii. Machine learning
   c. **Z diffusion** from Fe spots ⇒ See Matteo's discussion
   d. **Association** of PMT & camera events
      i. One-to-one (camera & PMT) **golden dataset**.
      ii. **Closest neighbour approach** under study

PMT group works rundown:

1. **On-going:**
   - **a.** PMT reconstruction **TTree**
   - **b.** **(L,x,y) fit**
     - i. Bayesian fit approach
     - ii. Machine learning
   - **c.** **Z diffusion** from Fe spots ⇒ See Matteo's discussion
   - **d.** **Association** of PMT & camera events
     - i. One-to-one (camera & PMT) **golden dataset**.
     - ii. **Closest neighbour approach** under study
   - **e.** Study of the **α parameter in $L \propto 1/R^{\alpha}$**.
   - **f.** Optimize moving average to **filter high frequency noise** ⇒ important to understand constant **sources of noise**.
   - **g.** **Optimize *find_peaks()*** in order to identify better the peaks ⇒ Relevant especially for long tracks.
   - **h.** Compute **track angle / ΔZ** ⇒ First steps towards **3D reconstruction**.
     - i. Using **single peaks** approach
     - ii. Using **time over threshold** ⇒ A good/reasonable approach for straight tracks.

PMT group works rundown:

1. **On-going:**
   a. PMT reconstruction **TTree**
   b. **(L,x,y) fit**
      i. Bayesian fit approach
      ii. Machine learning
   c. **Z diffusion** from Fe spots ⇒ See Matteo's discussion
   d. **Association** of PMT & camera events
      i. One-to-one (camera & PMT) **golden dataset**.
      ii. **Closest neighbour approach** under study
   e. Study of the **α parameter in L ∝ 1/R$^{\alpha}$**.
   f. Optimize moving average to **filter high frequency noise** ⇒ important to understand constant **sources of noise**.
   g. **Optimize *find_peaks()*** in order to identify better the peaks ⇒ Relevant especially for long tracks.
   h. Compute **track angle / ΔZ** ⇒ First steps towards **3D reconstruction**.
      i. Using **single peaks** approach
      ii. Using **time over threshold** ⇒ A good/reasonable approach for straight tracks.

- **Future steps:**
  - Fully or partial **integration of x/y/L fit** in PMT reconstruction.
  - **Am-Be run** will provide us with a great new dataset!

PMT- RecoTTree

1. **Goal:**
    a. Create a **TTree** with the **PMT reconstructed variables**.
    b. Run **PMT** reconstruction together with **Camera** reconstruction on the cloud.
    c. Create a **framework** that allows the ***analysis of PMT events stand-alone and together with camera***.
        i. (L,x,y), Z coordinate, Z diffusion, 3D reco, NR vs ER, PID, etc.

PMT- RecoTTree

1. Goal:

a. Create a **TTree** with the **PMT reconstructed variables**.

b. Run **PMT** reconstruction together with **Camera** reconstruction on the cloud.

c. Create a **framework** that allows the ***analysis of PMT events stand-alone and together with camera***.

   i. (L,x,y), Z coordinate, Z diffusion, 3D reco, NR vs ER, PID, etc.

2. Steps already taken:

a. **Retrieve data** from midas data banks through ***cygno_libs*** -> Both ***fast*** and ***slow*** digitizers.

   i. All four PMTs and all triggers associated with a given picture.

   ii. Waveforms **automatically corrected**. (maximum digitizer offset of **-0.3V**)

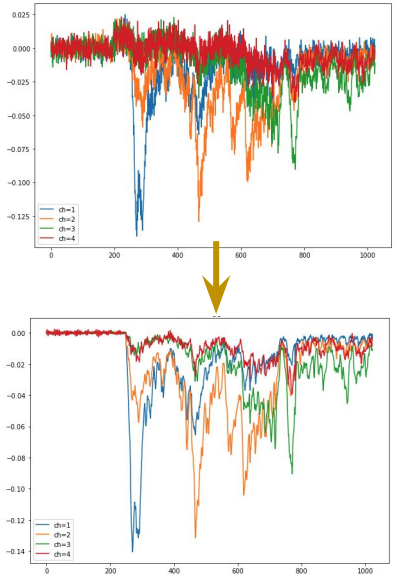
b. PMT-Reco working for both Run1 and Run2 (which had different midas files)

c. Tree created with basic variables.

d. Config files updated.

e. Bug-testing ongoing:

   i. Missing tests with pedestals and simulation reconstruction.



PMT- RecoTTree

➔ **Mostly completed!**

- Further tests will be conducted. ⇒ Instructions to install the libs and run the code will follow.
- Additions still being made and new ones can be requested.

PMT- RecoTTree

➔ **Mostly completed!**

- Further tests will be conducted. ⇒ Instructions to install the libs and run the code will follow.
- Additions still being made and new ones can be requested.

➔ **Rundown of the variables:**

- pmt_wf_[run/event/trigger/channel/sampling] ⇒ ***Basic*** info of each waveform
- pmt_wf_insideGE ⇒ Checks which triggers are inside the camera's ***Global Exposure***
- pmt_[baseline/RMS] ⇒ Baseline and RMS. Useful to check ***quality of data***
- pmt_tot_[integral/charge] ⇒ Sum and conversion of ***all samples*** of a waveform
- pmt_max_ampl ⇒ Waveform's max amplitude. Used to ***signal selection***
- pmt_nPeaks ⇒ Waveform's number of peaks. Useful to ***easily select*** ^{55}***Fe*** vs. others
- pmt_peak_Number ⇒ Useful for ***peak coincidence*** among PMTs
- *pmt_peak_[Position/Height/HalfWidth/FullWidth]* ⇒ ***Main analysis point. Basic implementation at the moment.***

PMT- RecoTTree

➔ **Mostly completed!**

- Further tests will be conducted. ⇒ Instructions to install the libs and run the code will follow.
- Additions still being made and new ones can be requested.

➔ **Rundown of the variables:**

- pmt_wf_[run/event/trigger/channel/sampling] ⇒ ***Basic*** info of each waveform
- pmt_wf_insideGE ⇒ Checks which triggers are inside the camera's ***Global Exposure***
- pmt_[baseline/RMS] ⇒ Baseline and RMS. Useful to check ***quality of data***
- pmt_tot_[integral/charge] ⇒ Sum and conversion of ***all samples*** of a waveform
- pmt_max_ampl ⇒ Waveform's max amplitude. Used to ***signal selection***
- pmt_nPeaks ⇒ Waveform's number of peaks. Useful to ***easily select*** ^{55}***Fe*** vs. others
- pmt_peak_Number ⇒ Useful for ***peak coincidence*** among PMTs
- *pmt_peak_[Position/Height/HalfWidth/FullWidth]* ⇒ ***Main analysis point. Basic implementation at the moment.***

The ***find_peaks()*** is the **most important function** which still needs optimization:

There's space here to some work!

PMT- Studies of the main variable distributions

⇒ ^{55}Fe Time length

PMT- Studies of the main variable distributions

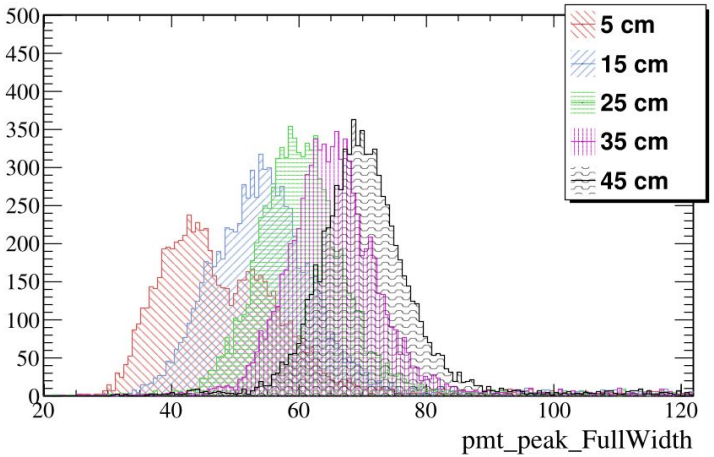
⇒ ^{55}Fe Time length

1. From the **tree**, now one can also easily look into the **^{55}Fe signal time distributions** with distance.
   a. The *find_peaks()* function works rather well for the simple ^{55}Fe peak.
   b. There is still some **discussion on how to get the time_width** given that the 4 PMTs see slightly different signals.

PMT- Studies of the main variable distributions

⇒ ^{55}Fe Time length

1. From the **tree**, now one can also easily look into the **^{55}Fe signal time distributions** with distance.
   a. The *find_peaks()* function works rather well for the simple ^{55}Fe peak.
   b. There is still some **discussion on how to get the time_width** given that the 4 PMTs see slightly different signals.



This was done with 25 calibration runs (5 positions x 5 runs).
For higher statistics, it's just a matter of running this on the cluster with many runs…

PMT- Studies of the main variable distributions

⇒ **Time/Amplitude over threshold (TOT)**

PMT- Studies of the main variable distributions

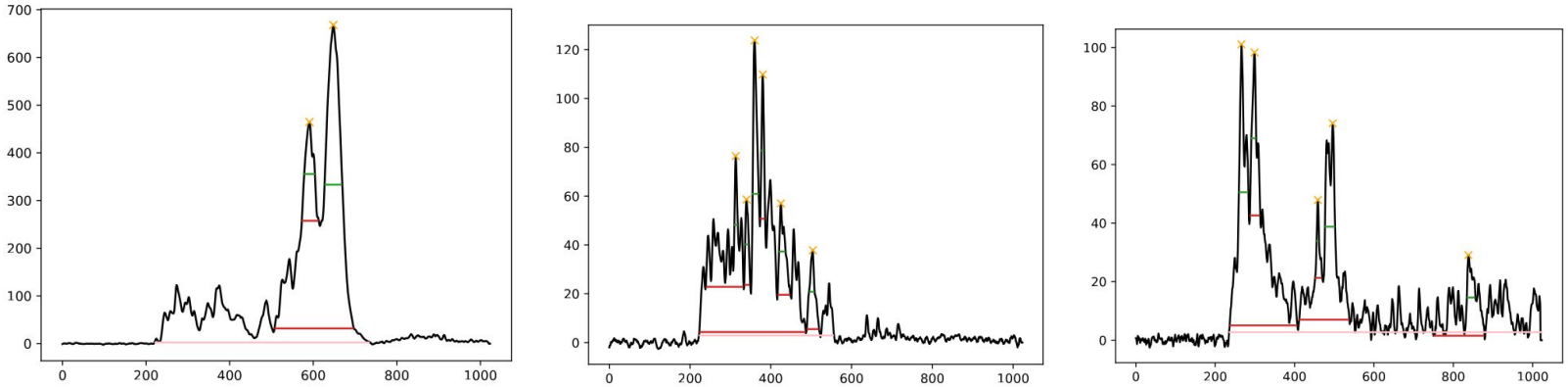
⇒ Time/Amplitude over threshold (TOT)

- Measurement of the **time length** of the signal which is **above a given threshold**.
  - Not achievable with find_peaks() function.
  - (For 55Fe both methods work since it's a single normal peak.)

PMT- Studies of the main variable distributions

⇒ Time/Amplitude over threshold (TOT)

- Measurement of the **time length** of the signal which is **above a given threshold**.
  - Not achievable with find_peaks() function.
  - (For 55Fe both methods work since it's a single normal peak.)

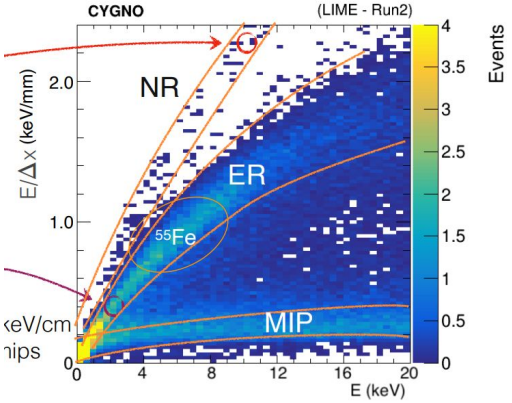


Legend: **Pink** = TOT; **Red** and **Green**: Full and Half widths of peaks retrieved with *find_peaks()*

⇒ **Time/Amplitude over threshold (TOT)**

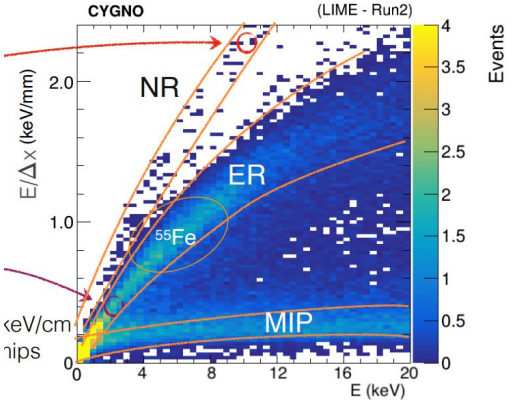
⇒ Time/Amplitude over threshold (TOT)

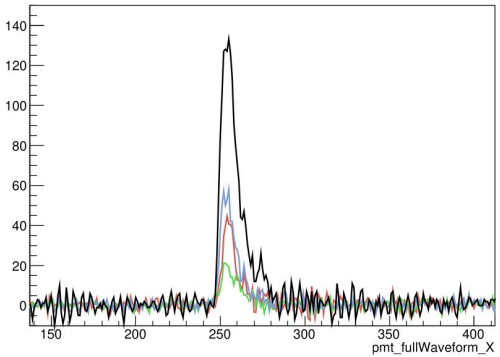
1. I tried to replicate Emanuele's **dE/dx vs. E** plot, but now with the **area as E** and **Δt as Δxy**.

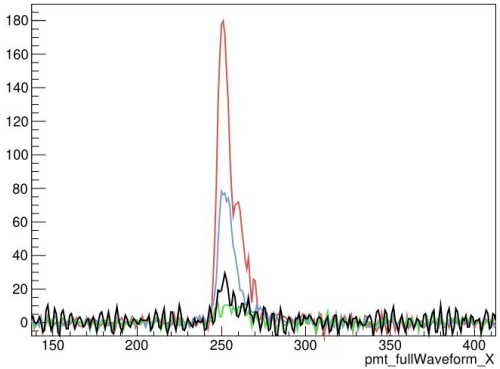


⇒ Time/Amplitude over threshold (TOT)

1. I tried to replicate Emanuele's **dE/dx vs. E** plot, but now with the **area as E** and **Δt as Δxy**.
2. Immediately found an issue:
   a. Each PMT sees a **different intensity** but approximately the same time extension ⇒ Results in a different dE/dx.

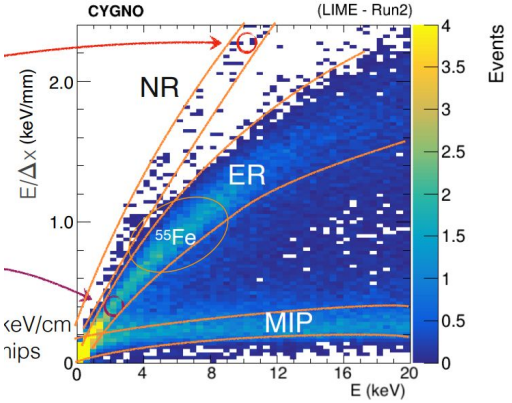


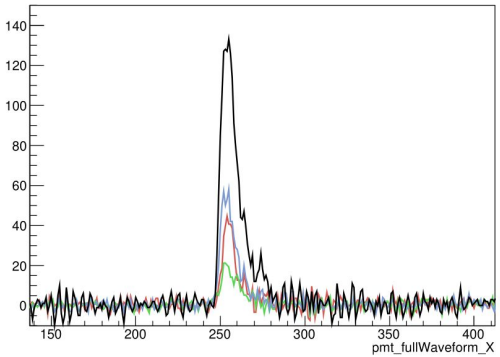


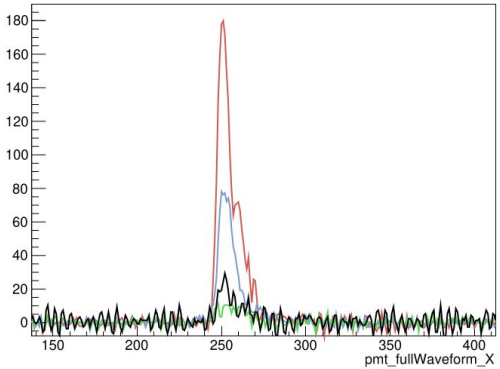


⇒ Time/Amplitude over threshold (TOT)

1. I tried to replicate Emanuele's **dE/dx vs. E** plot, but now with the **area as E** and **Δt as Δxy**.
2. Immediately found an issue:
   a. Each PMT sees a **different intensity** but approximately the same time extension ⇒ Results in a different dE/dx.
   b. This is solved using Francesco x/y/L fit that returns the **fitted luminosity** of a given trigger mixing the 4 signals.

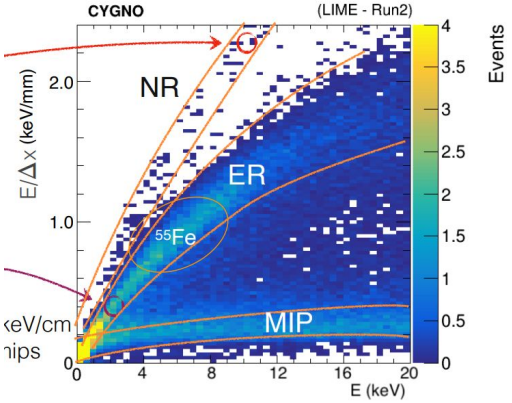


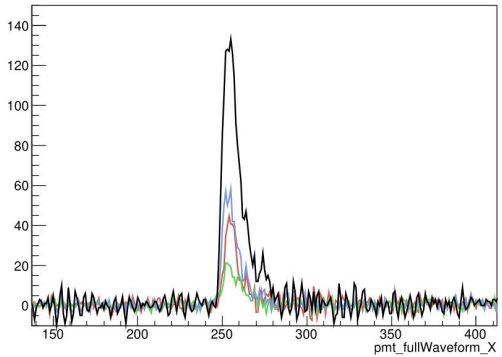


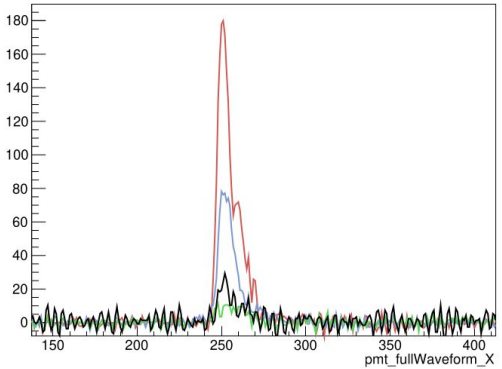


⇒ Time/Amplitude over threshold (TOT)

1. I tried to replicate Emanuele's **dE/dx vs. E** plot, but now with the **area as E** and **Δt as Δxy**.
2. Immediately found an issue:
   a. Each PMT sees a **different intensity** but approximately the same time extension ⇒ Results in a different dE/dx.
   b. This is solved using Francesco x/y/L fit that returns the **fitted luminosity** of a given trigger mixing the 4 signals.





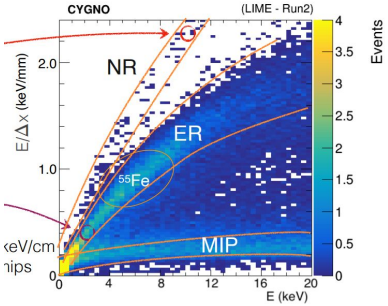


For the ***length*** I used a basic ***4 waveform average*** since the width is mostly constant.

↓

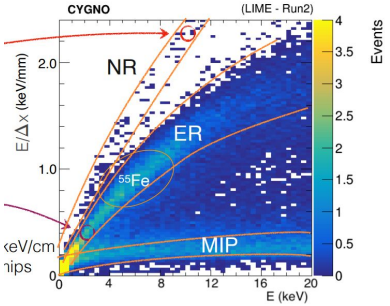
A more precise determination of the signal's width is undergoing.

⇒ dE/dx on ^{55}Fe signals



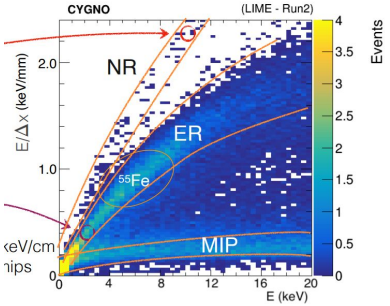
⇒ dE/dx on ^{55}Fe signals

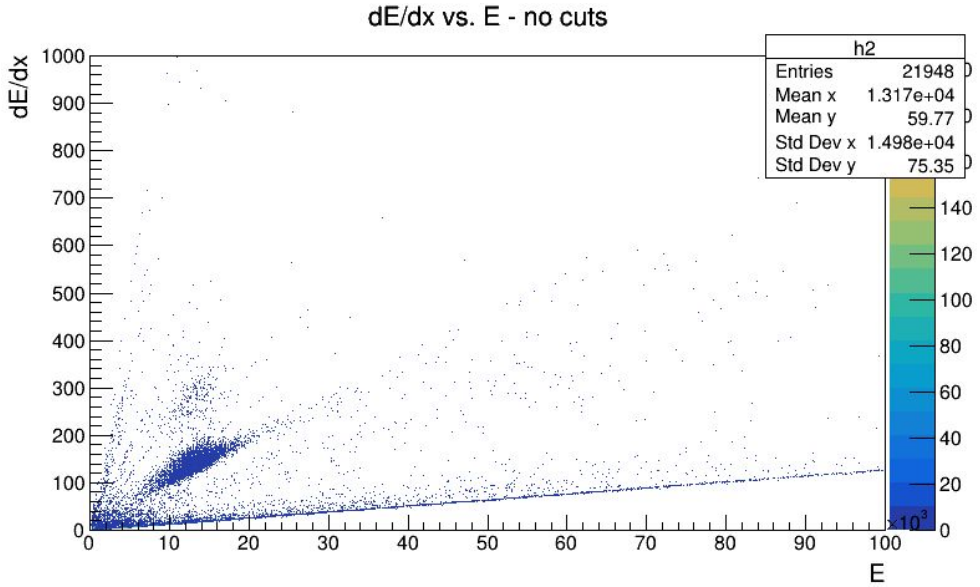
1. This is calculated using the the **fitted L** and **time over threshold (TOT)**
2. Different **cuts** were studies: **RMS** and **mj2p**



⇒ dE/dx on ^{55}Fe signals

1. This is calculated using the the **fitted L** and **time over threshold (TOT)**
2. Different **cuts** were studies: **RMS** and **mj2p**





⇒ **Diagonal line** ⇒ Digitizer sampling

↓

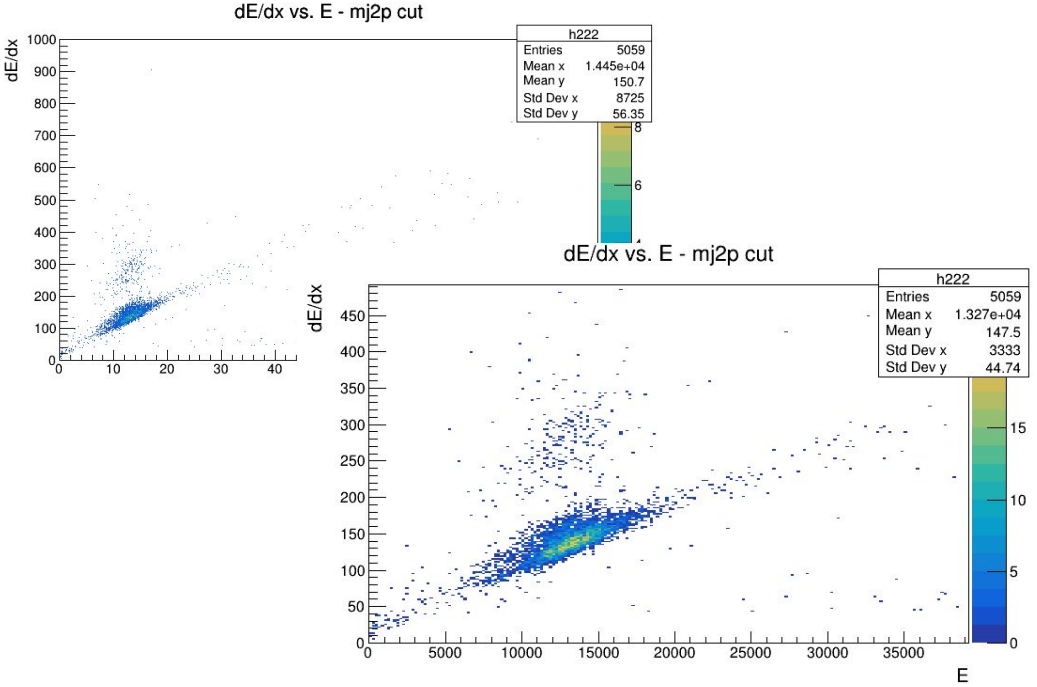
⇒ Most long track saturate the "time"

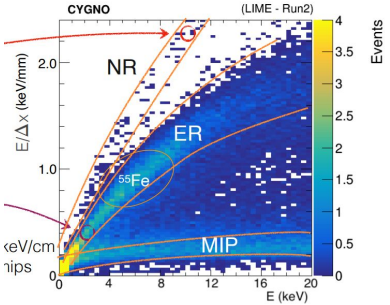
↓

⇒ Important to look at **slow digitizer** (more and longer samples)

⇒ dE/dx on ^{55}Fe signals

1. This is calculated using the the **fitted L** and **time over threshold (TOT)**
2. Different **cuts** were studies: **RMS** and **mj2p**





⇒ The **mj2p = 1 cut** requires at least two PMTs with a single peak in the ~same position

↓

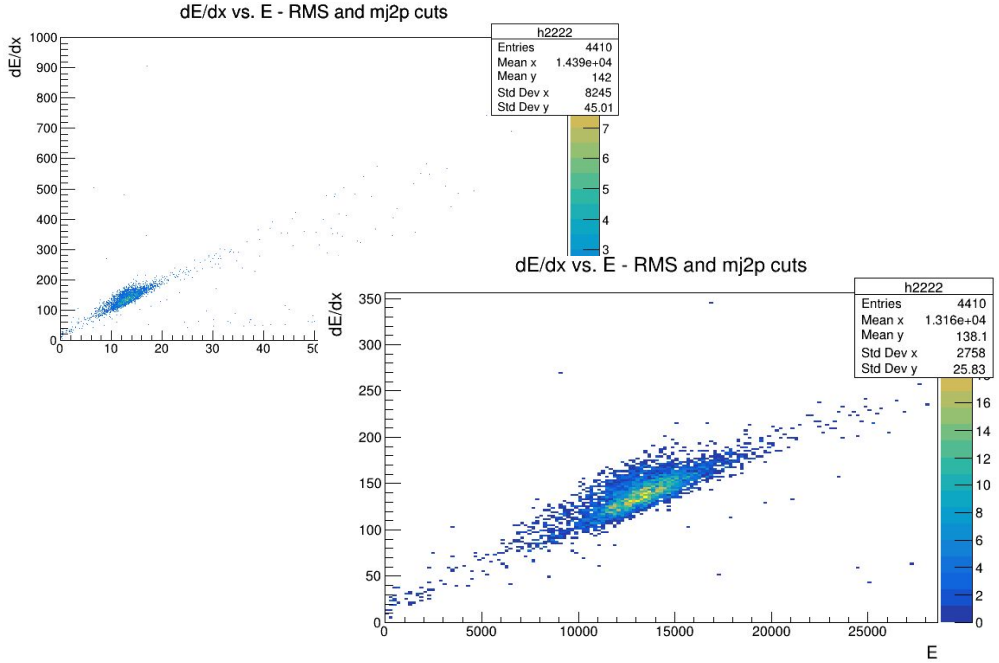
⇒ Most long track cut out

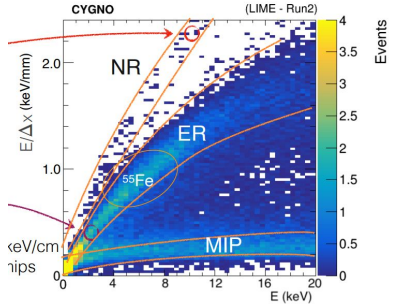
↓

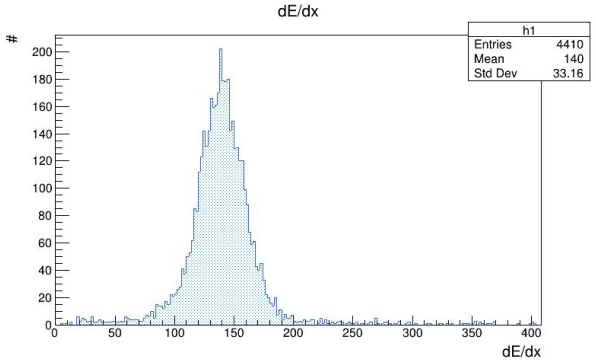
⇒ An ***interesting blob*** appears above the ^{55}Fe zone…

⇒ dE/dx on ^{55}Fe signals

1. This is calculated using the the **fitted L** and **time over threshold (TOT)**
2. Different **cuts** were studies: **RMS** and **mj2p**





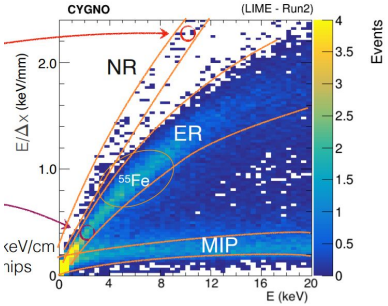


⇒ The **RMS cut** ***deletes completely the blob*** above 55Fe, leaving a rather **clean 55Fe dE/dx position**

↓

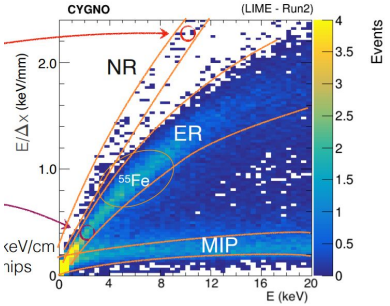
⇒ More on the RMS later

⇒ dE/dx on BG signals



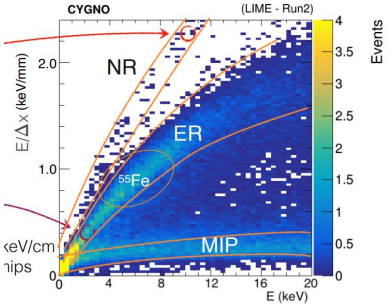
⇒ dE/dx on BG signals

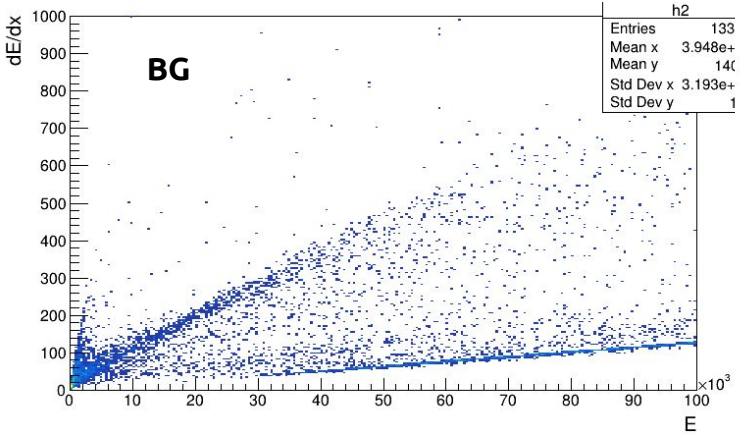
1. This is calculated using the the **sum of sliced fitted L** (see Francesco's presentation) and **time over threshold (TOT)**
2. Only cuts applied were the **RMS**

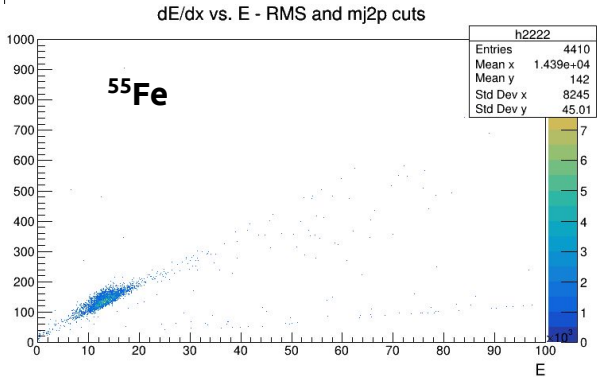


⇒ dE/dx on BG signals

1. This is calculated using the the **sum of sliced fitted L** (see Francesco's presentation) and **time over threshold (TOT)**
2. Only cuts applied were the **RMS**

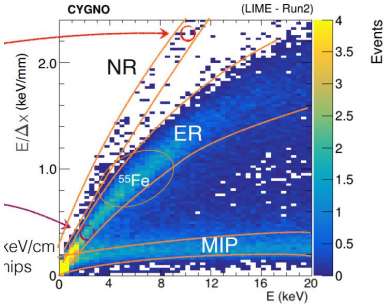


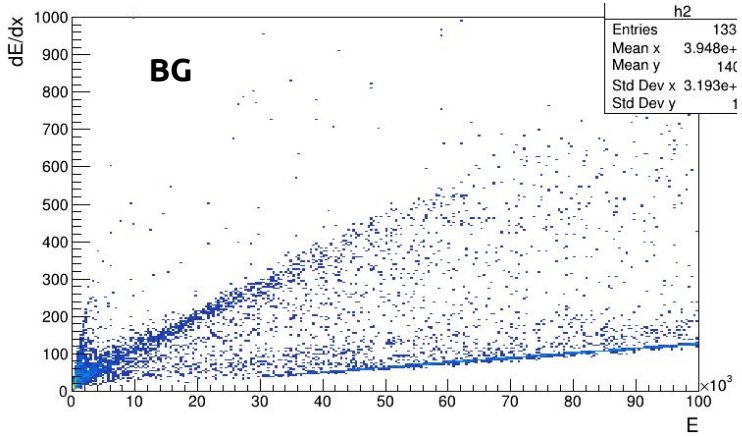


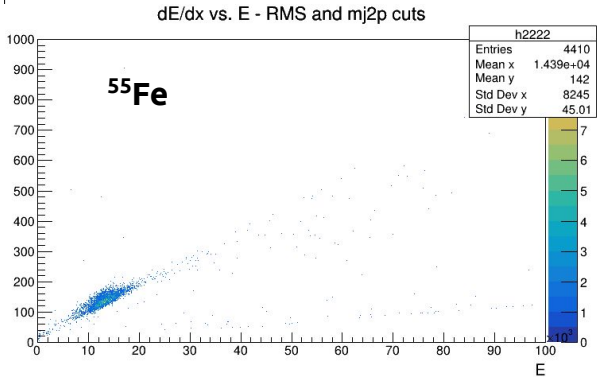


⇒ dE/dx on BG signals

1. This is calculated using the the **sum of sliced fitted L** (see Francesco's presentation) and **time over threshold (TOT)**
2. Only cuts applied were the **RMS**



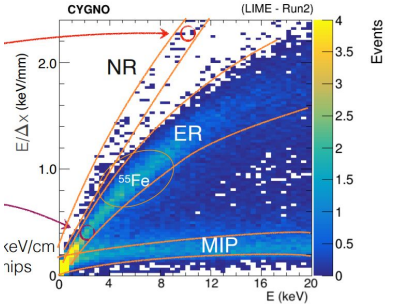


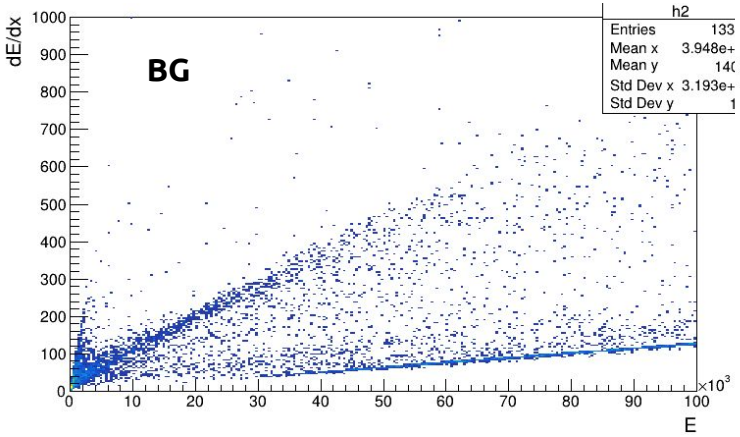


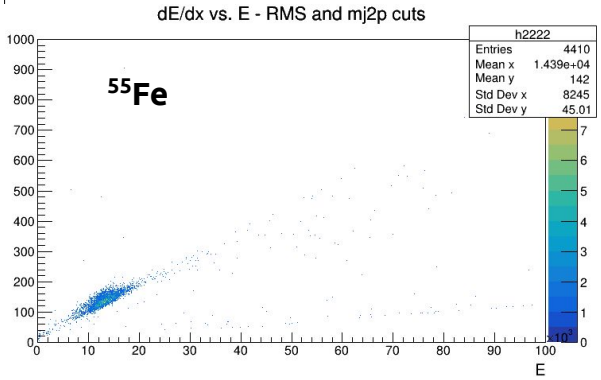
1. One visible **band** for supposedly **ER** ⇒ Where ^{55}Fe sits.
2. Eventual **band at low dE/dx** mixed with a diagonal cut ⇒ Most long track don't "fit" in the fast digitizer.
   a. Reconstruct and fit **slow waveforms** is important ⇒ Next step!

⇒ dE/dx on BG signals

1. This is calculated using the the **sum of sliced fitted L** (see Francesco's presentation) and **time over threshold (TOT)**
2. Only cuts applied were the **RMS**







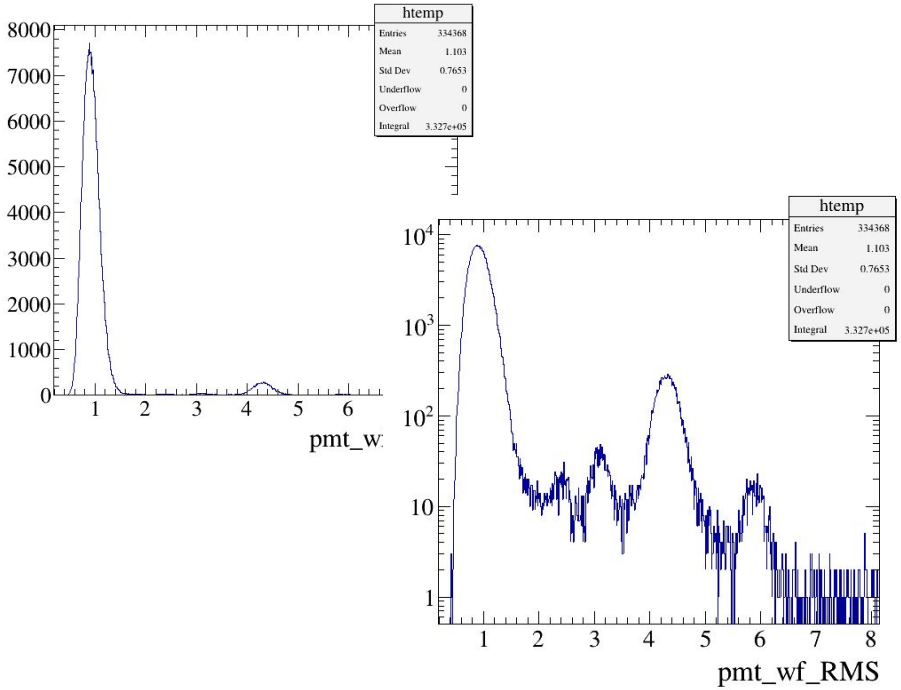
⇒ Eventual **high dE/dx and low energy band** ⇒ Must be cross-checked directly with pictures!

1. One visible **band** for supposedly **ER** ⇒ Where ^{55}Fe sits.
2. Eventual **band at low dE/dx** mixed with a diagonal cut ⇒ Most long track don't "fit" in the fast digitizer.
   a. Reconstruct and fit **slow waveforms** is important ⇒ Next step!

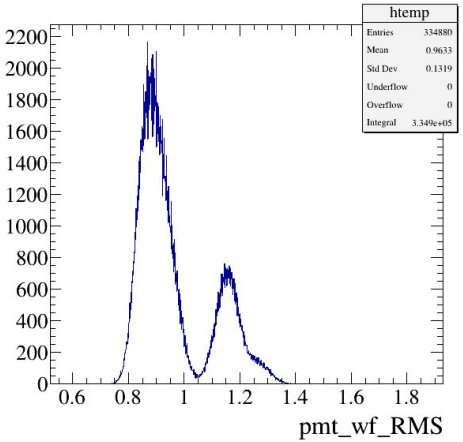
⇒ Some odd features:

⇒ Some odd features:

1. **RMS as a double peak** in the fast digitizer:
   a. Here I show all the RMS of all runs, not averaged.

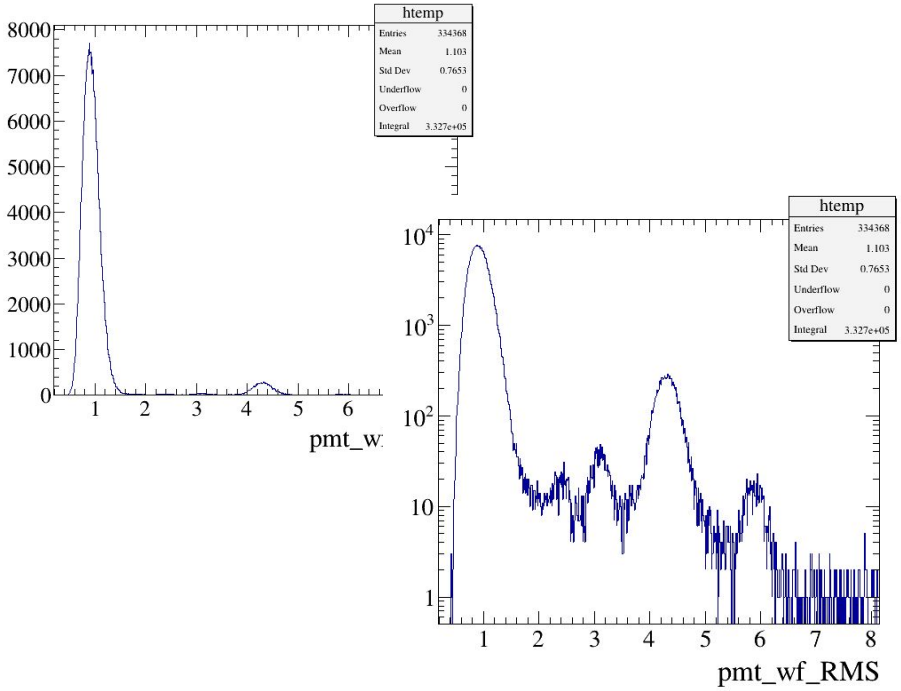


2. And the RMS of the slow digitizer:

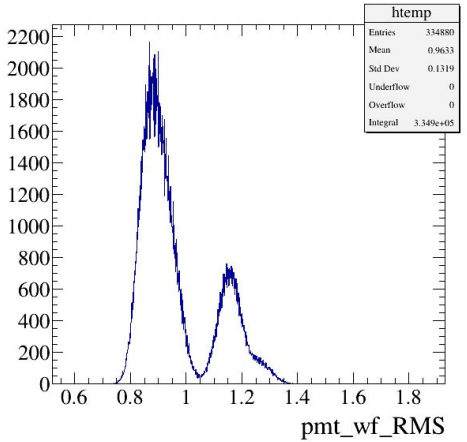


⇒ Some odd features:

1. **RMS as a double peak** in the fast digitizer:
   a. Here I show all the RMS of all runs, not averaged.



2. And the RMS of the slow digitizer:



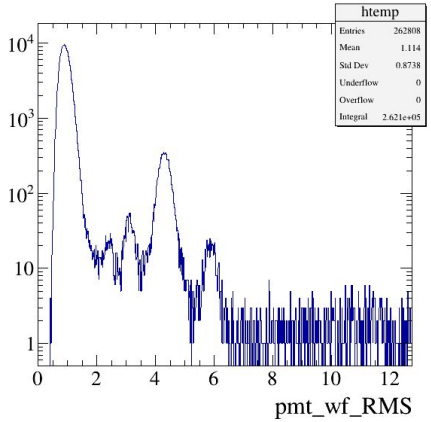
⇒ The **waveforms' RMSs** show some **weird features** that could spoil further studies.

↓

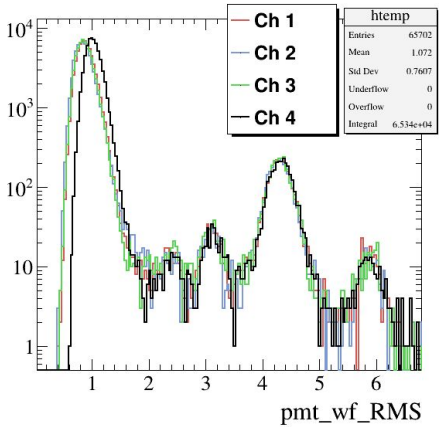
One must understand the **reason** behind this: is it an **environment** dependency (P,T,n), or does it result from **particular particle interaction**?

⇒ RMS – isolating the problem:

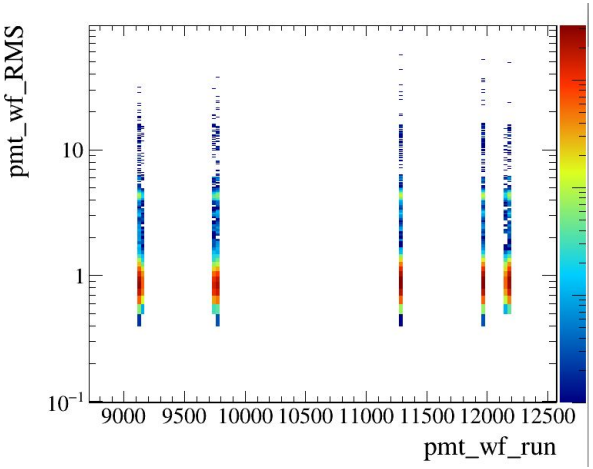
1. I start with the RMS of 25 calibration runs:



2. Divided by channel:



3. Divided by run:



⇒ The waveforms' **RMS** has a **second main peak** at **~4 samples intensity**

⇒ I will check with **other factors** like temperature, humidity, light yield, etc.

⇒ We also discussed the possibility of **higher RMS coming from hotpots** or **high energy particles** that seem to destabilize the PMTs.

PMT- Variables' analysis sum-up

➔ **Reco TTRee:**

- Some tests still required ⇒ Is being constantly updated ⇒ Once in the cluster, high statistics analysis can be done.

PMT- Variables' analysis sum-up

➔ **Reco TTRee:**

- Some tests still required ⇒ Is being constantly updated ⇒ Once in the cluster, high statistics analysis can be done.

➔ **Variables studies:**

- **Time over threshold (TOT) implemented**, whereas the fitted L is retrieved for Francesco's fit.
  - A better **integration** of ***reco and fit*** is necessary and is being discussed.
  - Test **TOT at different levels/amplitudes** to see if a new particle ID (PSD) is achievable.
  - Use **TOT as an approach for 3D** reconstruction of straight(ish) tracks.

PMT- Variables' analysis sum-up

➔ **Reco TTRee:**

- Some tests still required ⇒ Is being constantly updated ⇒ Once in the cluster, high statistics analysis can be done.

➔ **Variables studies:**

- **Time over threshold (TOT) implemented**, whereas the fitted L is retrieved for Francesco's fit.
  - A better **integration** of ***reco and fit*** is necessary and is being discussed.
  - Test **TOT at different levels/amplitudes** to see if a new particle ID (PSD) is achievable.
  - Use **TOT as an approach for 3D** reconstruction of straight(ish) tracks.
- Understand **RMS** weird behaviour
  - Alphas?
  - Humidity/LY?
- Compare dE/dx with other variables (like number of peaks)
  - Discuss more in detail how to assess TOT once fitted L is calculated.
- **Normalize plots to ^{55}Fe** (eventually LED (?)) to get a better grasp of the units being discussed.
- Understand how to differentiate and integrate the **analysis of the slow digitizer** ⇒ Crucial for long waveforms.

PMT- Variables' analysis sum-up

- **Cross-check with images:**
  - At this point is important to **look directly at some pictures** to understand if there are **specific and clear features** that we aren't seeing in the PMT.
  - It's easier to cross-check with the pictures which type of particle we are looking at.

PMT- Variables' analysis sum-up

➔ **Cross-check with images:**

- At this point is important to **look directly at some pictures** to understand if there are **specific and clear features** that we aren't seeing in the PMT.
- It's easier to cross-check with the pictures which type of particle we are looking at.

➔ **Eventual plans for the future (to keep in mind):**

- **PMT R&D:**
  - LED calibration at different positions.
  - Multi-particle/energy calibration
- **Analysis of PMT simulation:**
  - Would be relevant at this point to also start discussing with the PMT simulation group.
    - They can create clear data-set with different particles / energies / directions. etc..


Thank you for your attention!
The CYGNO Project counts with the collaboration of several international researchers coming from:
THANKS!