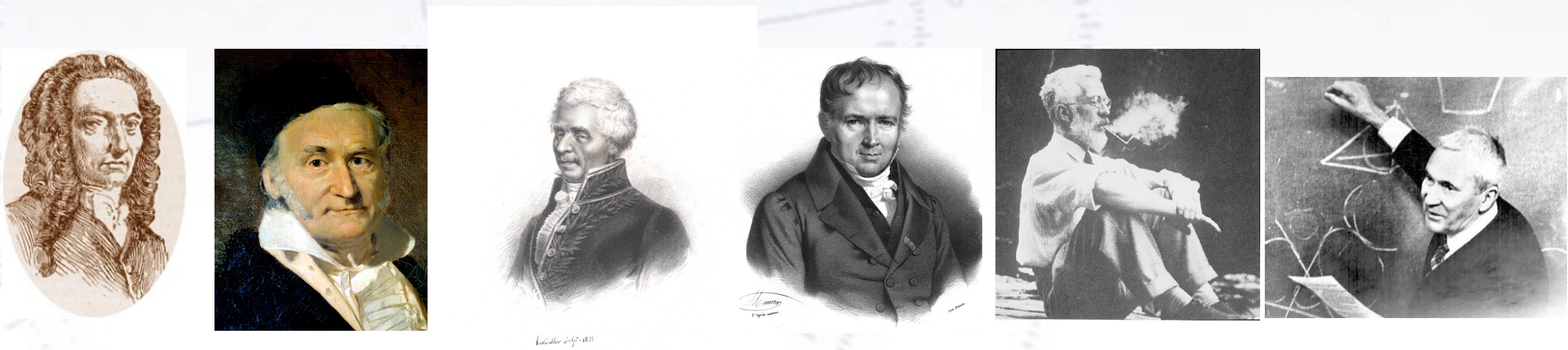
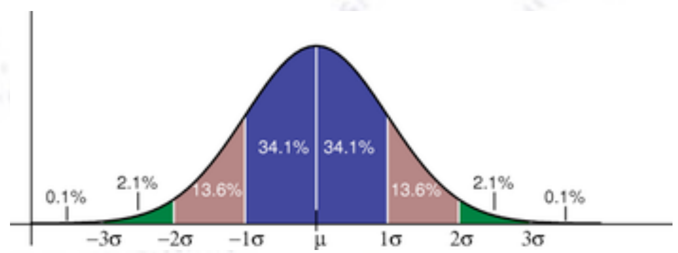
Applied Statistics

ATLAS test beam data analysis



Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

Ahhh… the smell of real data in the morning!

Ahhh… the smell of real data!

The data consists of a large number of events, where a charged particle (electron or pion) passed through a slice of the ATLAS detector. Each passage is recorded by different detectors - essentially **three independent detectors** - boiling down to **eleven numbers** (some more relevant than others).

The exercise is to separate electron and pion events based on these numbers, and in turn use this information to measure the interaction of pions and electrons separately.

Though the data is from particle physics, it could in principle have been from **ANY other source**, and the eleven numbers could for example have been indicators of cancer, key numbers for investors, or index numbers for identifying potential customers. It doesn't really matter - what matters is that…

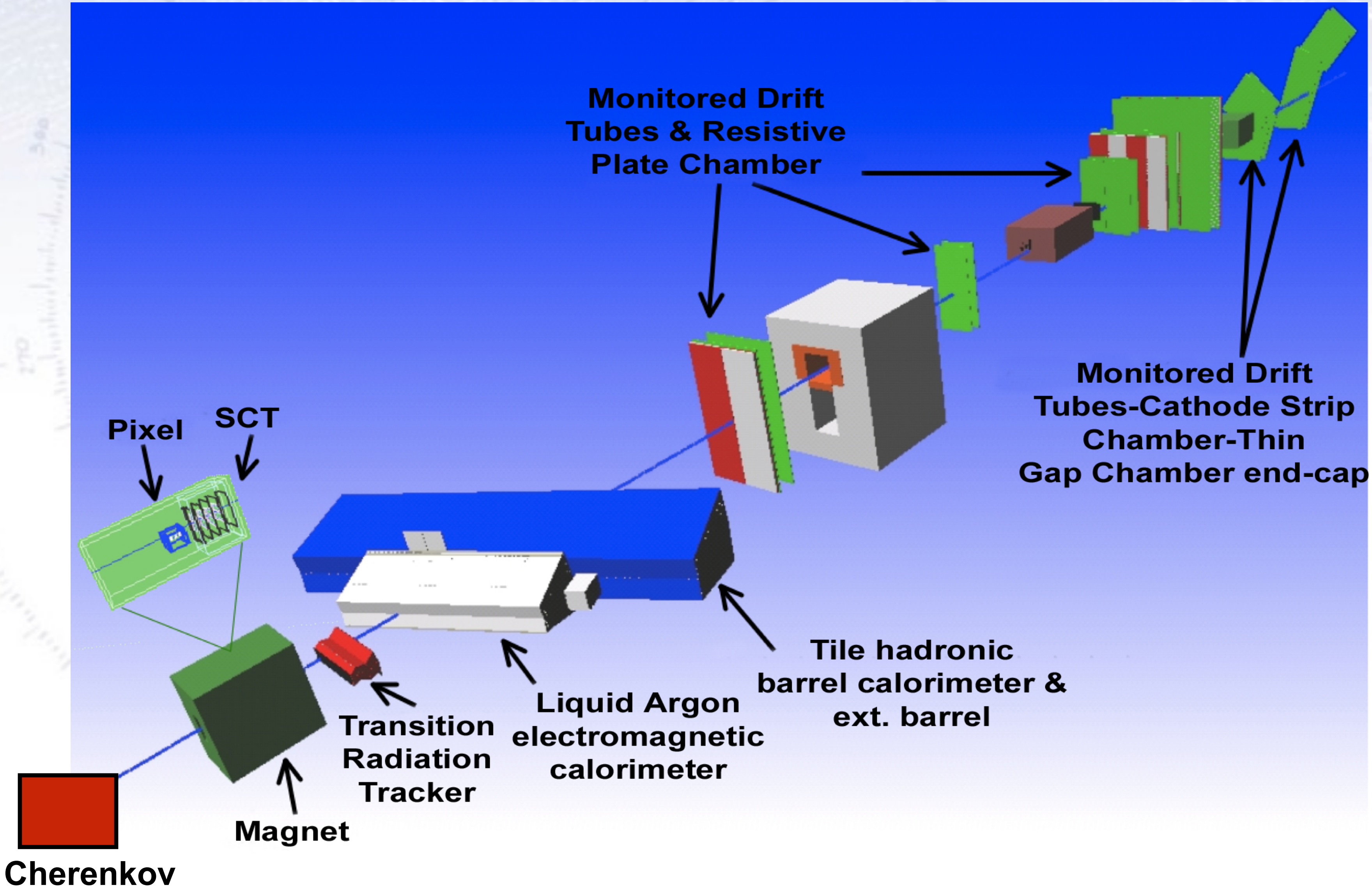
YOU are able to understand the data, can analyse it, and extract information from it.

With this in mind, I hope that you - if possible in groups - work your way through this data. The essential element in this analysis is, that there are **independent** measurements of the same particle, which is the one feature, which allows one (i.e. you) to determine the distribution of each variable.

ATLAS TestBeam setup

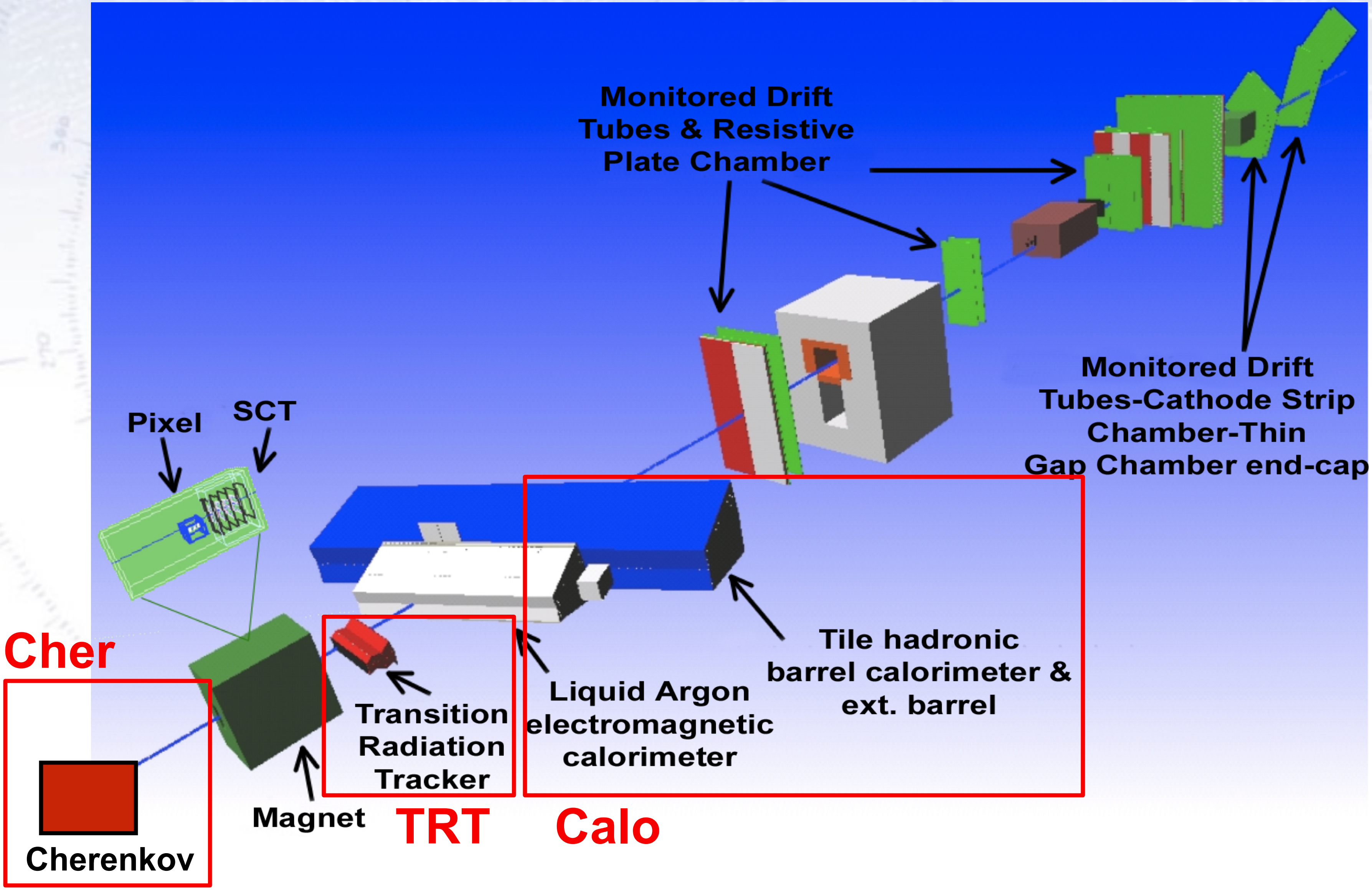
Setup of ATLAS TestBeam

The setup is shown below, with detector parts from all parts of ATLAS.

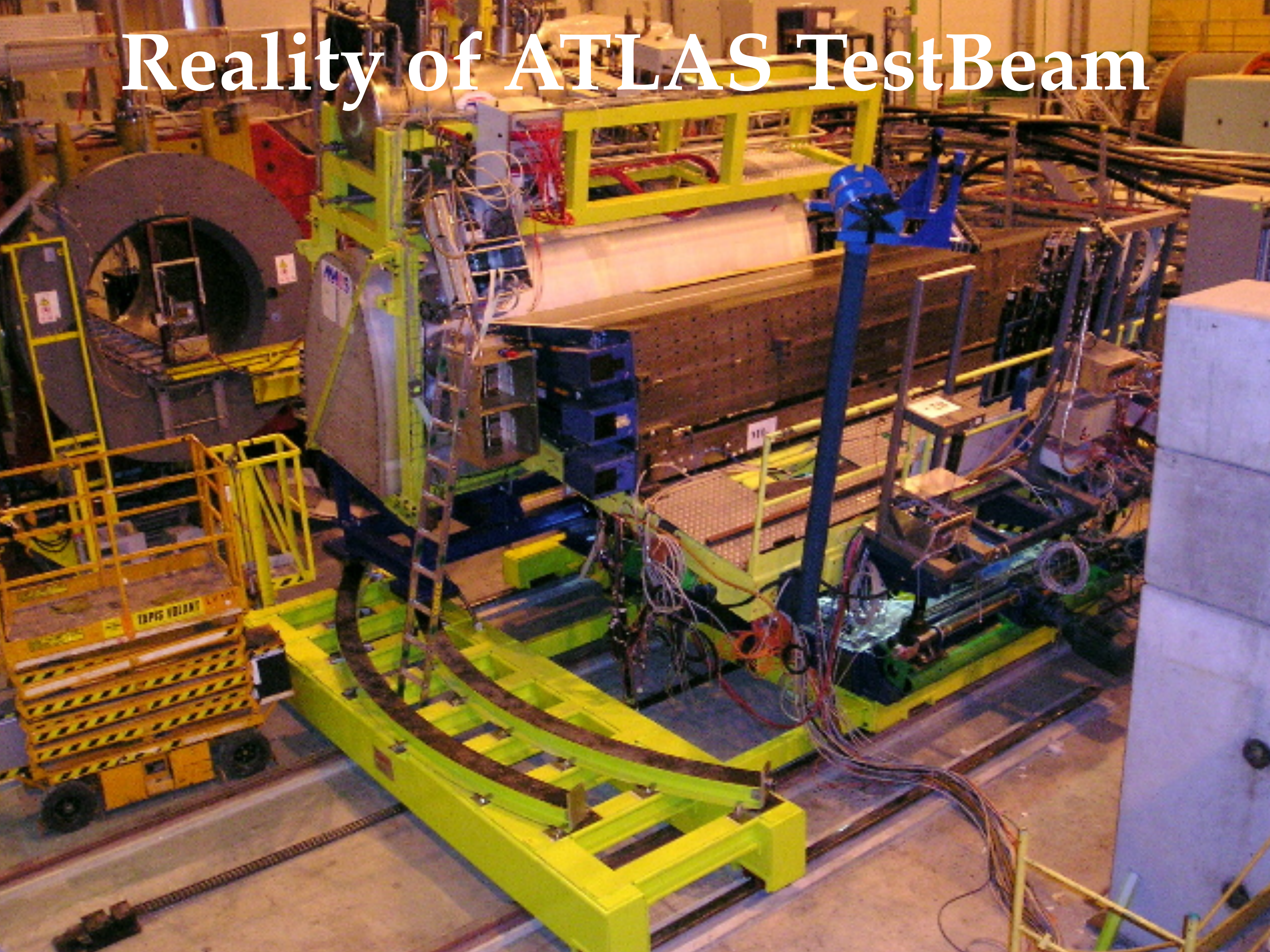


Setup of ATLAS TestBeam

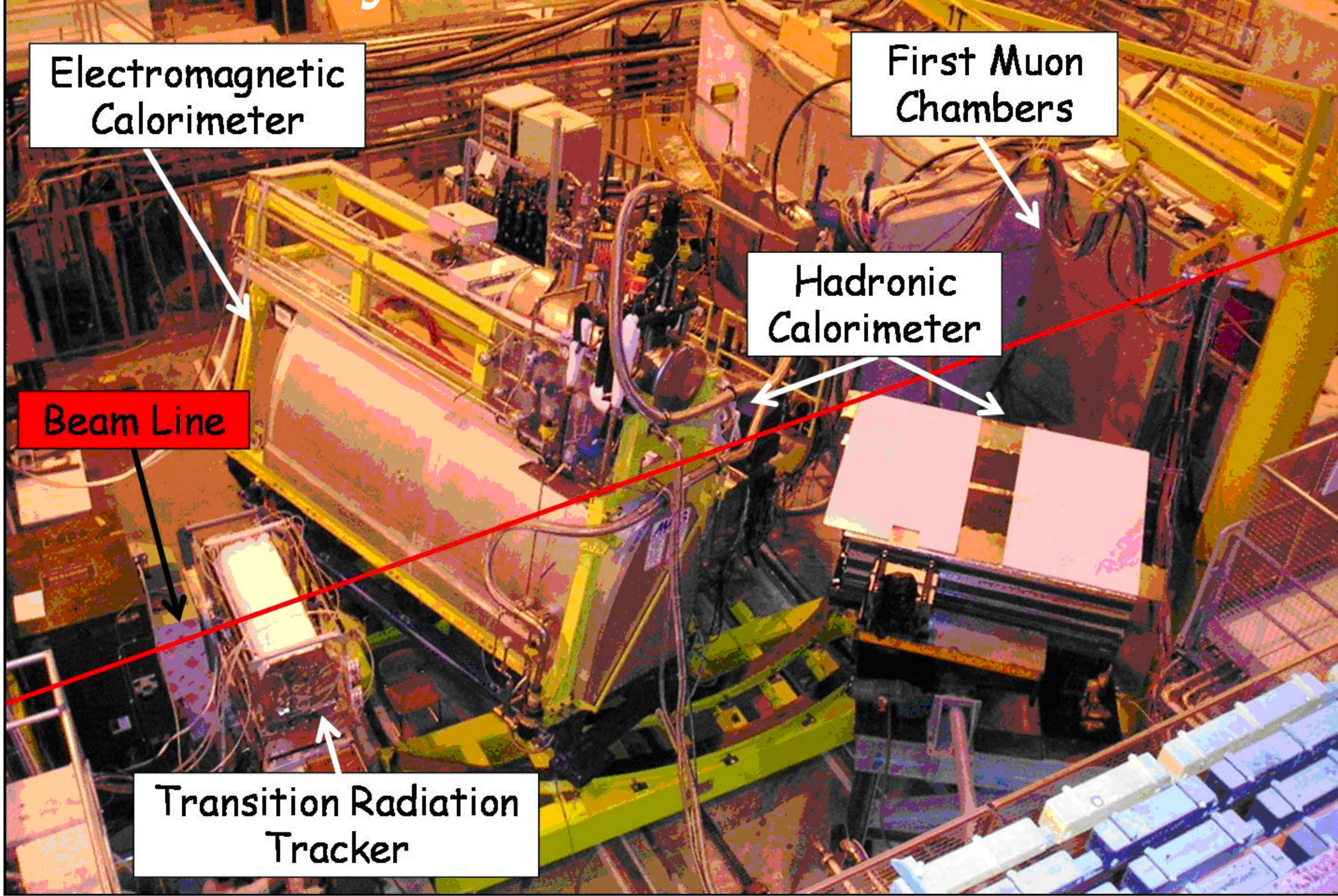
The setup is shown below, with detector parts from all parts of ATLAS.



Reality of ATLAS TestBeam



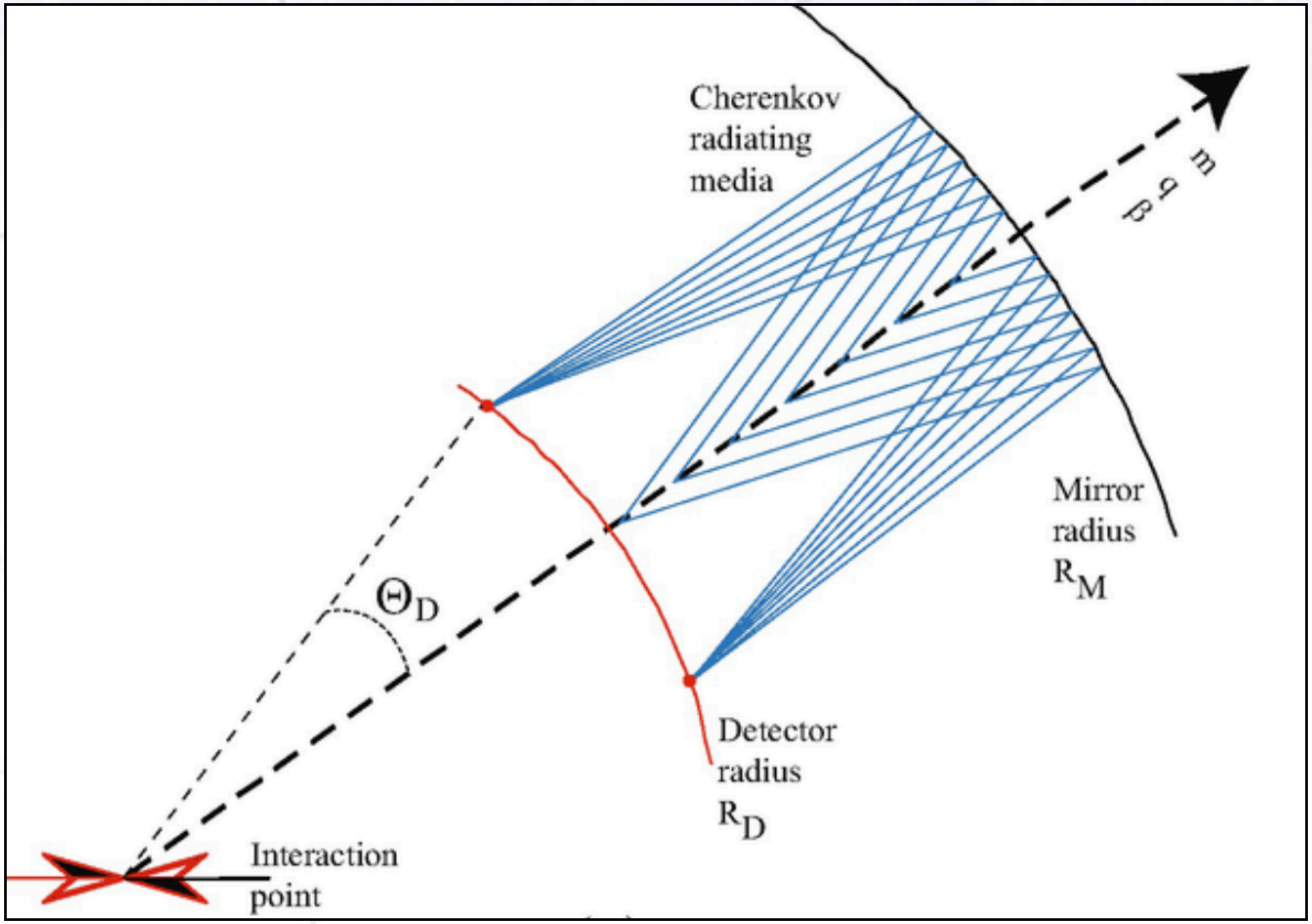
Reality of ATLAS TestBeam



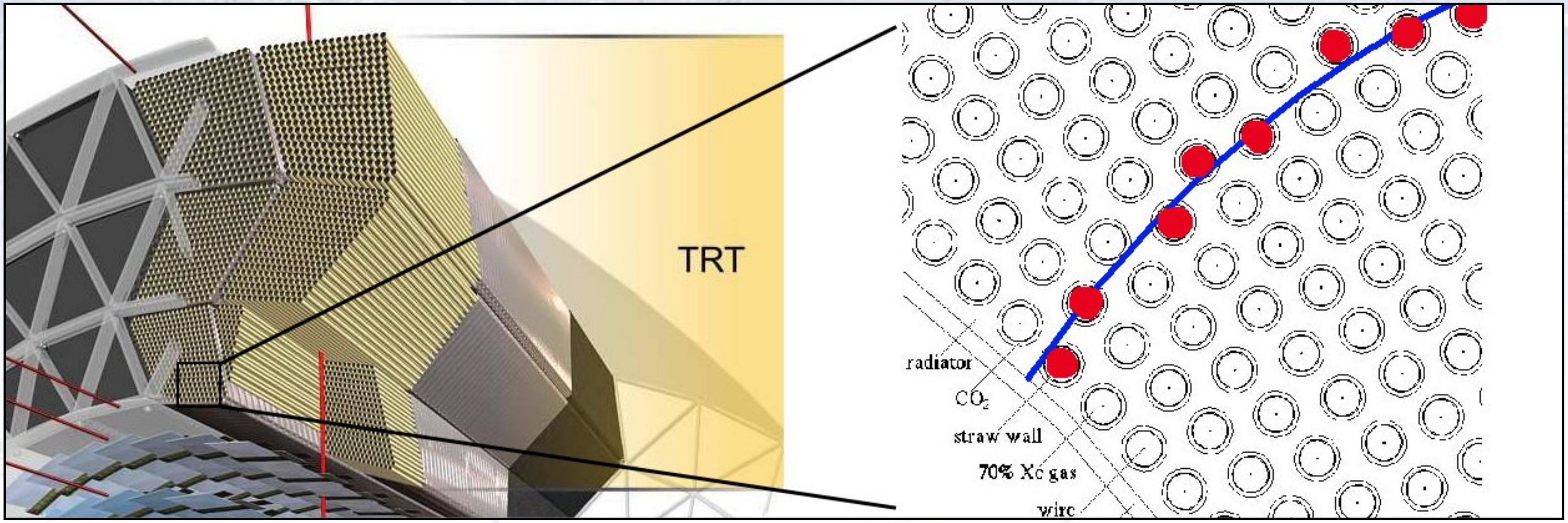
The individual detectors…

Cherenkov Detector

A charged particle emits Cherenkov light, when it travels faster than the speed of light *in that medium*. A Cherenkov detector gives a large signal, when this happens, otherwise a lower (dark current) signal. The medium (i.e. gas) can be chosen to only let this happen for some (light) particles, in our case electrons.

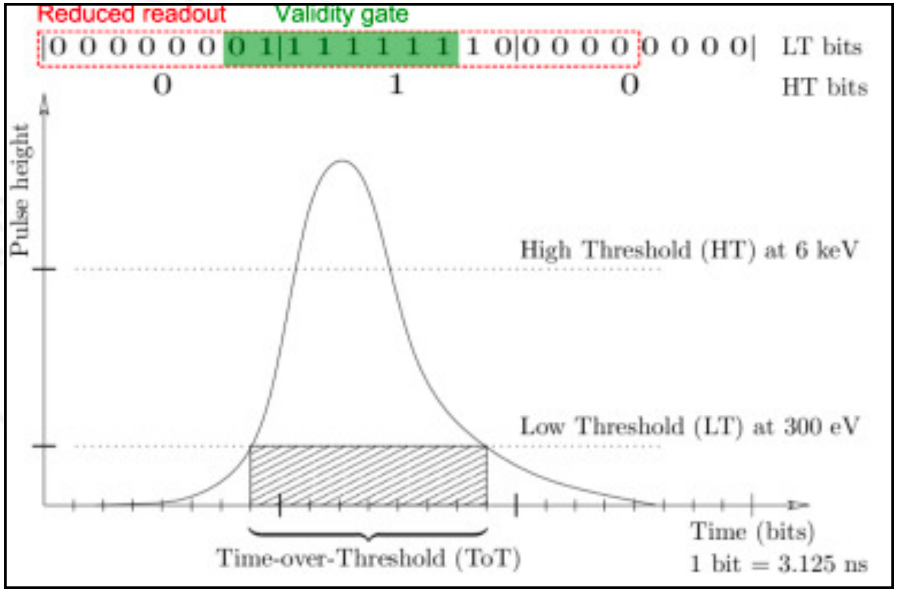


Transition Radiation Tracker



The Transition Radiation Tracker (TRT) records the passage of charged particles. The particles typically pass through 30-35 straws, each giving two signal:

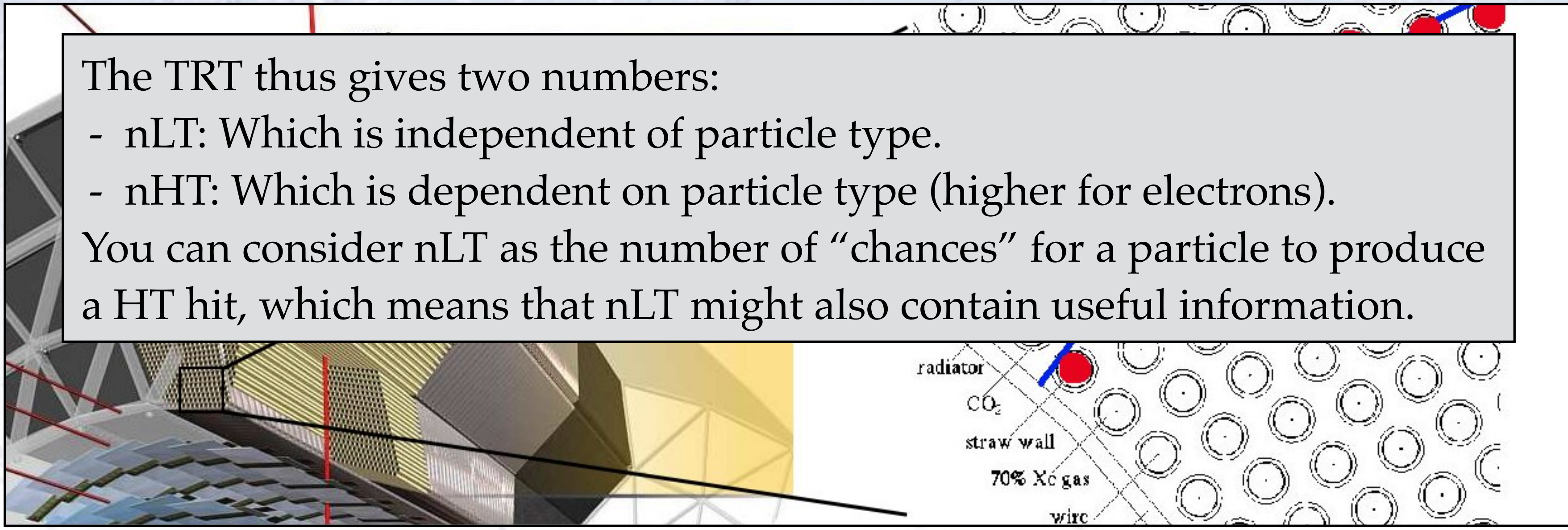
- Low-Threshold (LT), which is triggered by both types of particles with very high efficiency (> 98%).
- High-Threshold (HT), which is triggered mainly by electrons (~20%) and less so by pions (~6%).



Transition Radiation Tracker

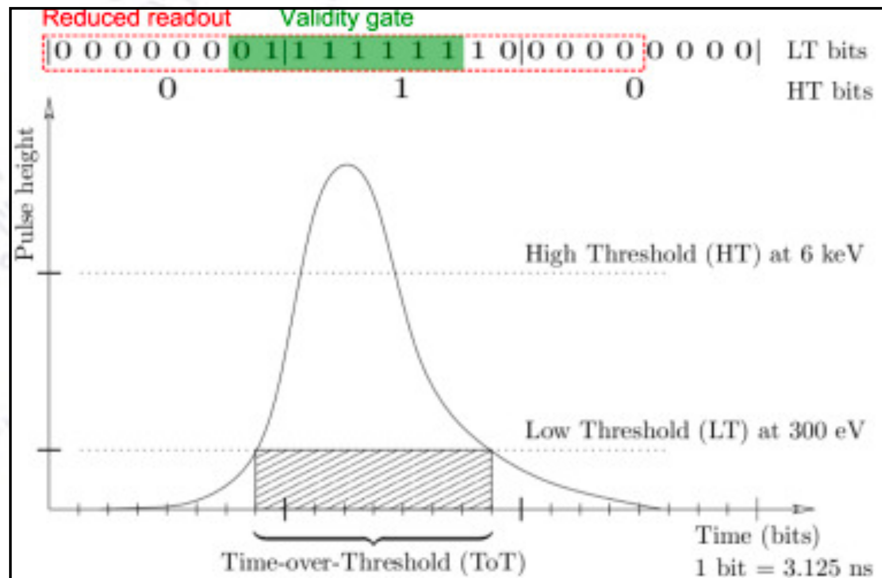
The TRT thus gives two numbers:
- nLT: Which is independent of particle type.
- nHT: Which is dependent on particle type (higher for electrons).

You can consider nLT as the number of "chances" for a particle to produce a HT hit, which means that nLT might also contain useful information.

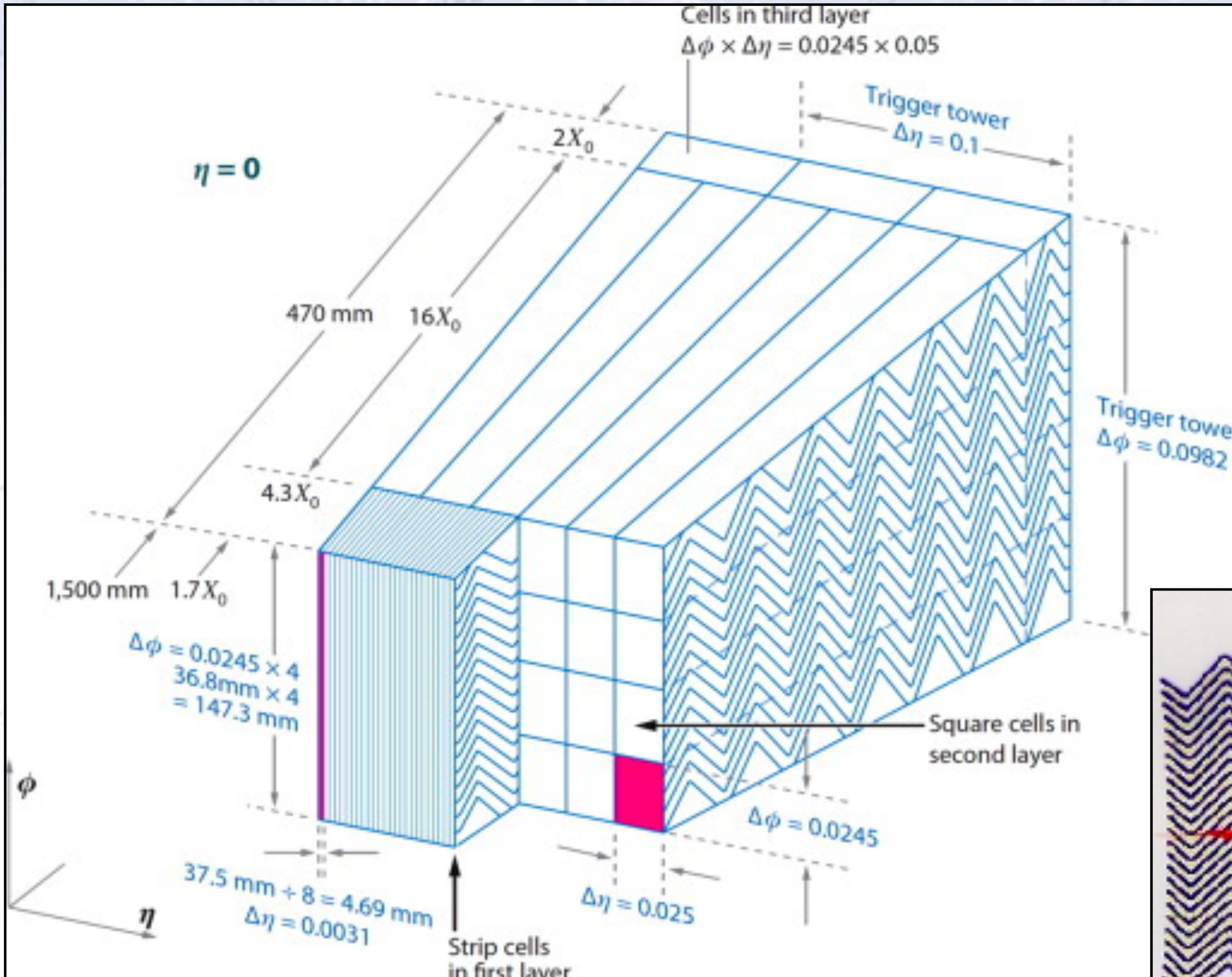


The Transition Radiation Tracker (TRT) records the passage of charged particles. The particles typically pass through 30-35 straws, each giving two signal:

- Low-Threshold (LT), which is triggered by both types of particles with very high efficiency (> 98%).
- High-Threshold (HT), which is triggered mainly by electrons (~20%) and less so by pions (~6%).

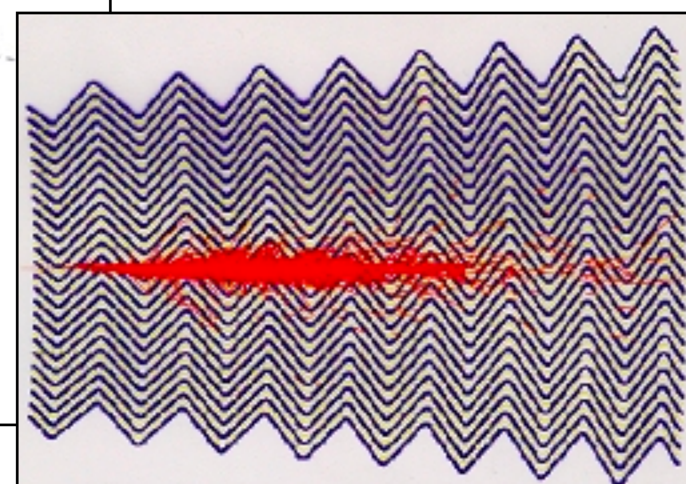


EM Calorimeter

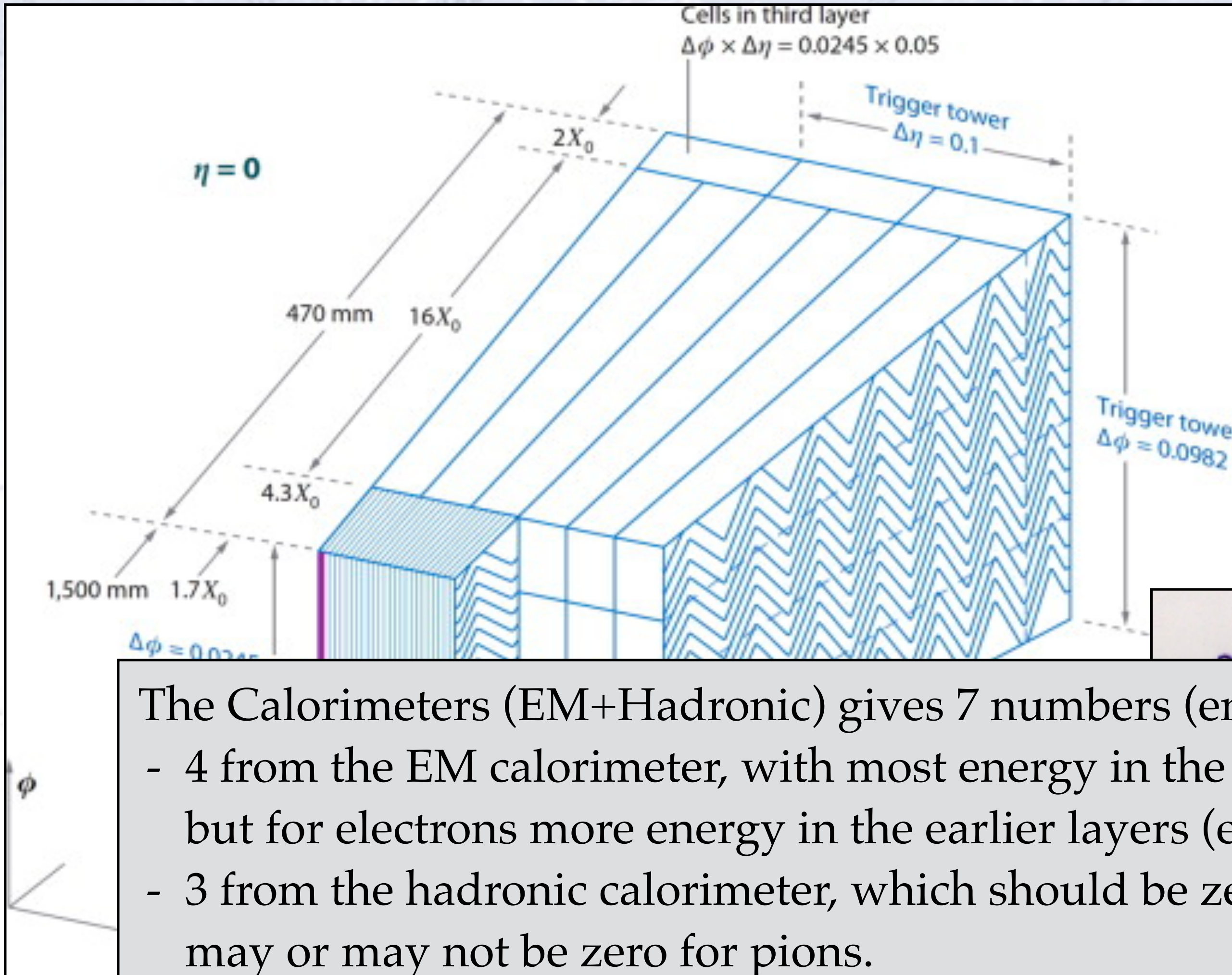


The ElectroMagnetic (EM) Calorimeter records the energy of a shower.

- Electrons shower fast and leaves all their energy.
- Pions shower slower and more irregularly, sometimes going through the calorimeter.

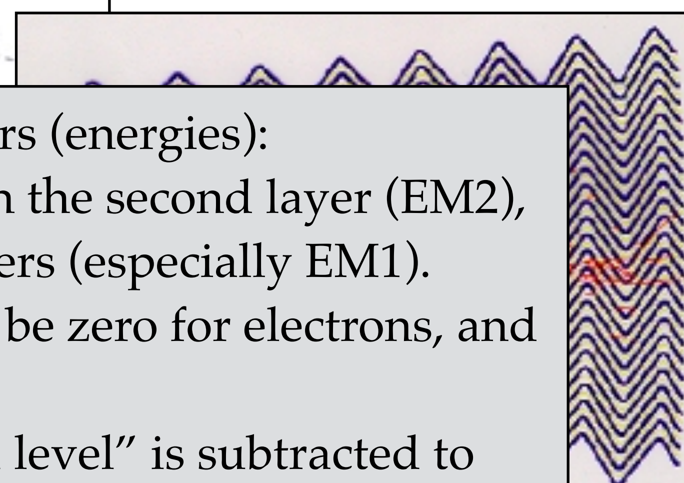


EM Calorimeter



The ElectroMagnetic (EM) Calorimeter records the energy of a shower.

- Electrons shower fast and leaves all their energy.
- Pions shower slower and more irregularly, sometimes going through the calorimeter.



The Calorimeters (EM+Hadronic) gives 7 numbers (energies):
- 4 from the EM calorimeter, with most energy in the second layer (EM2), but for electrons more energy in the earlier layers (especially EM1).
- 3 from the hadronic calorimeter, which should be zero for electrons, and may or may not be zero for pions.

Negative energy may occur, since a “background level” is subtracted to yield the correct average energy.

Your analysis

Your analysis

Generally, this analysis is about separating electrons and pions (and determining how well this can be done), followed by a few questions characterising the detector response to each type of particle.

Thus, you should imagine, that your new detector/equipment/questionnaire gave you this output, and now it is up to you to find out, what this tells you about your experiment, and how to extract information from it in the best possible way.

The questions asked should be considered as guiding to you, some/most of which your analysis should cover, but you do **not** have to follow them blindly (I've put "Optional" on those that are not essential).

Start by considering the data, and get a feel for the typical range of each variable. Plot the variables, both in 1D and also 2D! From considering these plots, guess/estimate an approximate knowledge of how electrons and pions distribute themselves in the variables above, and how to make a selection of these.

Your analysis

As described on the webpage introducing the data, there are three (relevant) detectors:

- Cherenkov,
- TRT (Transition Radiation Tracker) and
- Calorimeters

They are each capable of separating electrons and pions. As they are *INDEPENDENT* (three separate detectors), they may be used to cross check each other, and this is what you should use, in fact the essential part of this (and many other) analysis!

Note that this exercise has **little fitting in it**, as the distributions are typically **not very simple**. There is however one exception, namely the hits in the TRT. Given a fixed number of LT hits (e.g. N=35), the number of HT hits should follow a certain distribution… for you to fit.

Your analysis

As described on the webpage introducing the data, there are three (relevant) detectors:

- Cherenkov,
- TRT (Transition Radiation Tracker) and
- Calorimeters

They are each capable of separating electrons and pions. As they are *INDEPENDENT* (three separate detectors), they may be used to cross check each other, and this is what you should use, in fact the essential part of this (and many other) analysis!

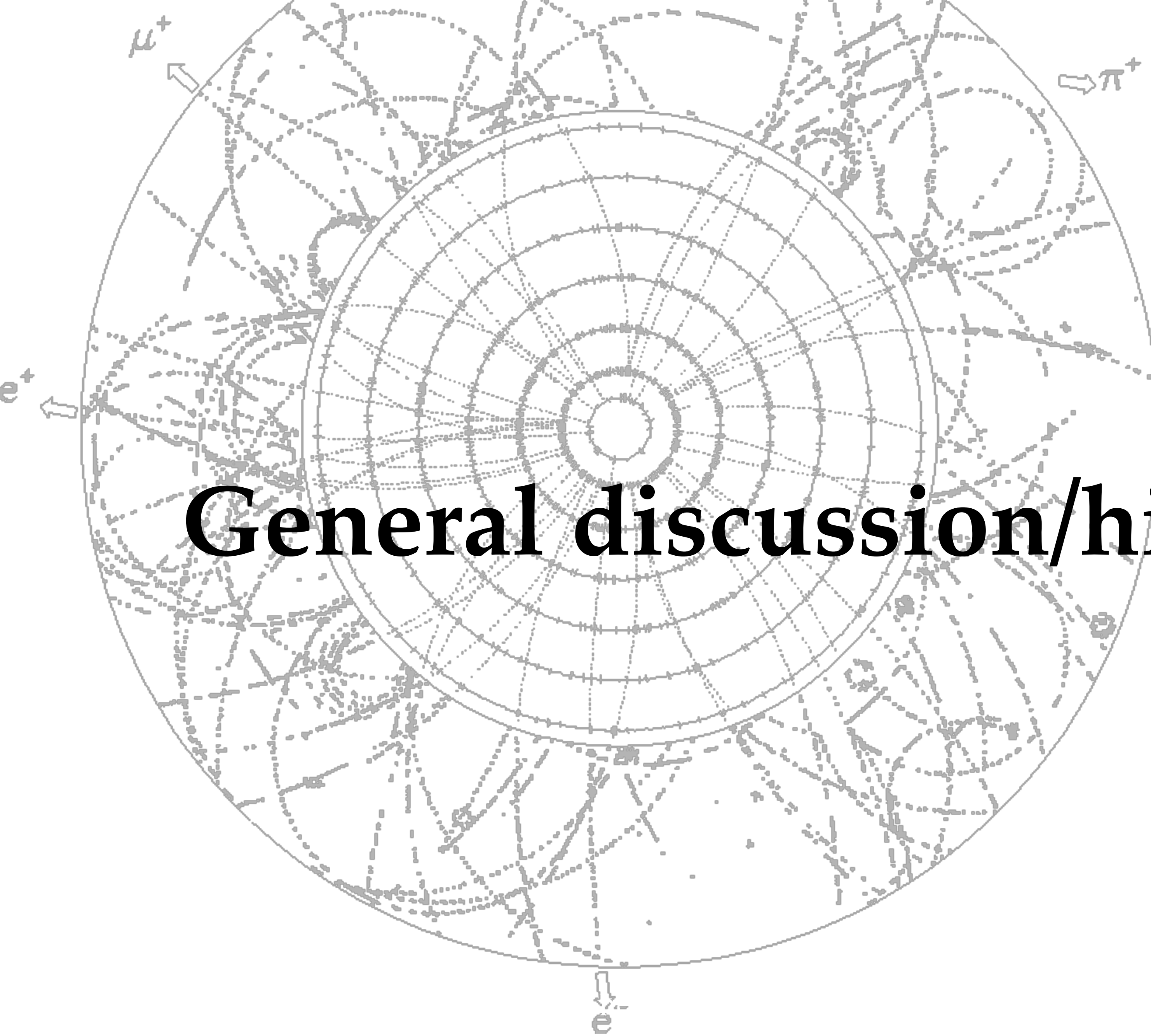
Note that this exercise has **little fitting in it**, as the distributions are typically **not very simple**. There is however one exception, namely the hits in the TRT. Given a fixed number of LT hits (e.g. N=35), the number of HT hits should follow a certain distribution… for you to fit.

Full Ph.D. thesises have been written about this data.

<u>It is real data in its perfect form:</u>

Simple, well planned, cleaned, yet with challenges, flaws, and strange correlations!

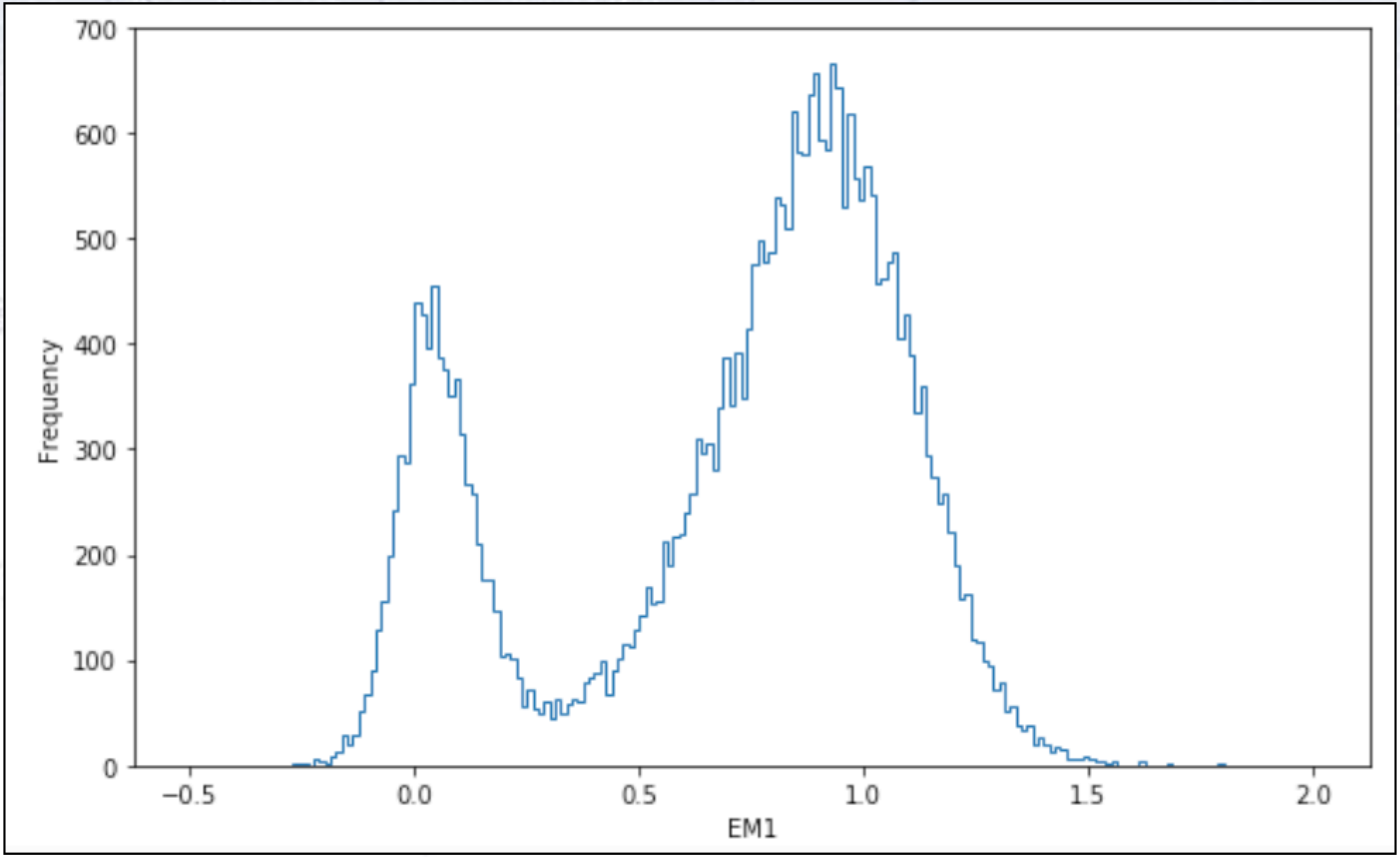
Throw yourself boldly at the data, and have "fun" with it.



General discussion/hints

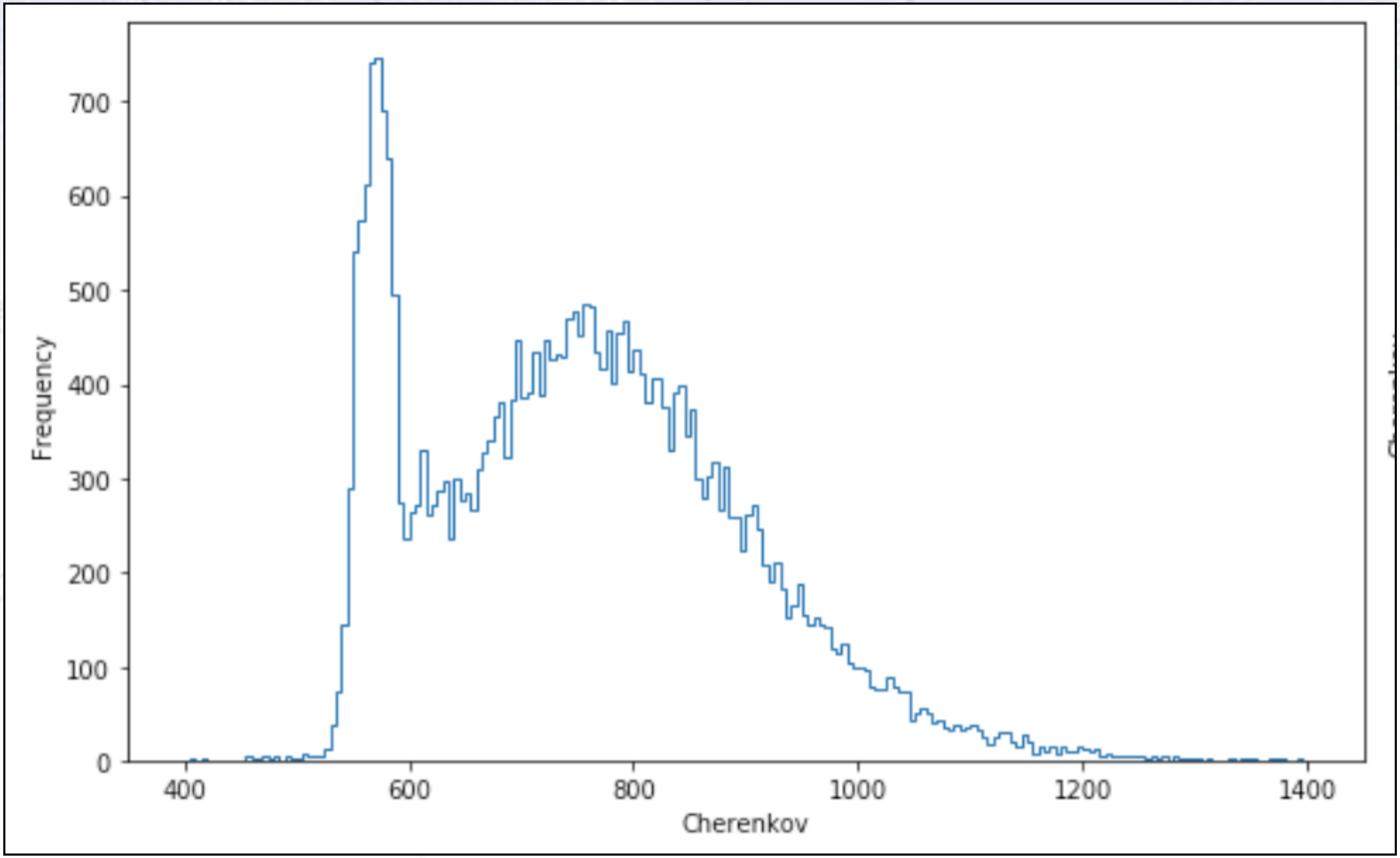
Consider this distribution

What does the following distribution of values of "EM1" look like?



Consider this distribution

How about this distribution, with the values of “Cher”?



Questions for discussion

How many "populations" (i.e. types of cases) do you think this data contains?

How could you check this line of thinking?

Are the three detectors independent? How could you check this? Does it matter, if they are not absolutely perfectly independent?

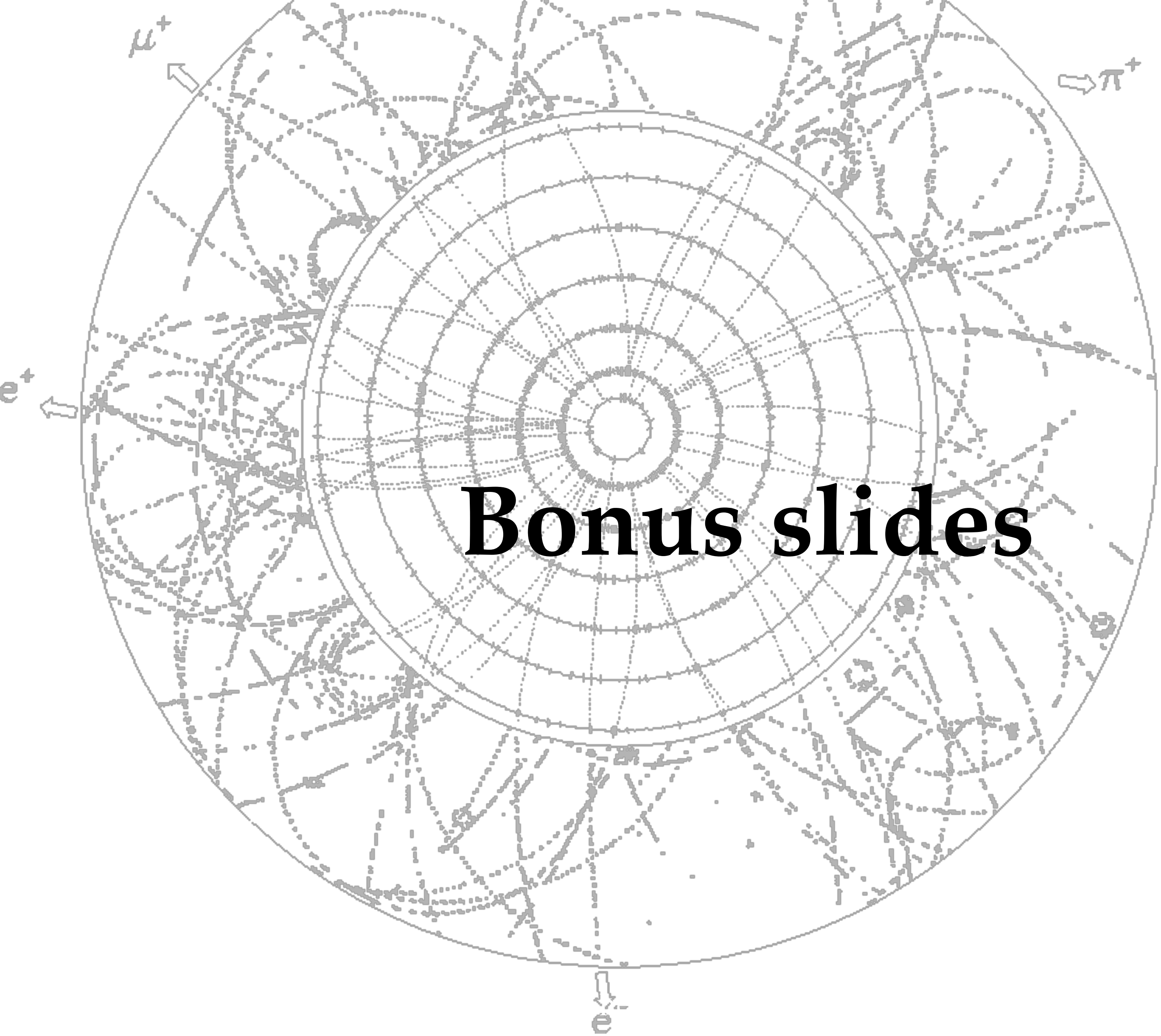
Do you think that distributions are generally Gaussian? Do you need to fit them?

Can you measure the performance of each detector? Are there systematic errors?

Can you improve the calorimeter's ability to separate particle types? If so, how?

Overall note:

This data is NOT meant for "fine tuned analysis", but rather "crude inspection". Try to get simple approximate answers out… I'm sure that you will afterwards be able to fine tune them.



Bonus slides

X