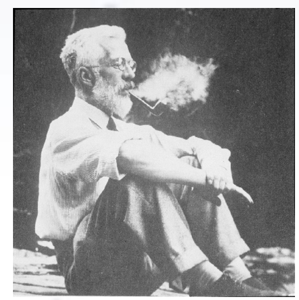
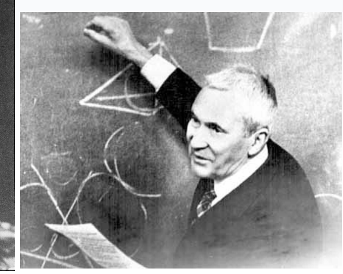
Applied Statistics

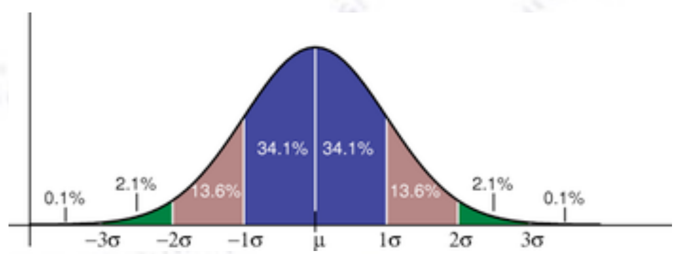
ATLAS test beam data analysis





Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

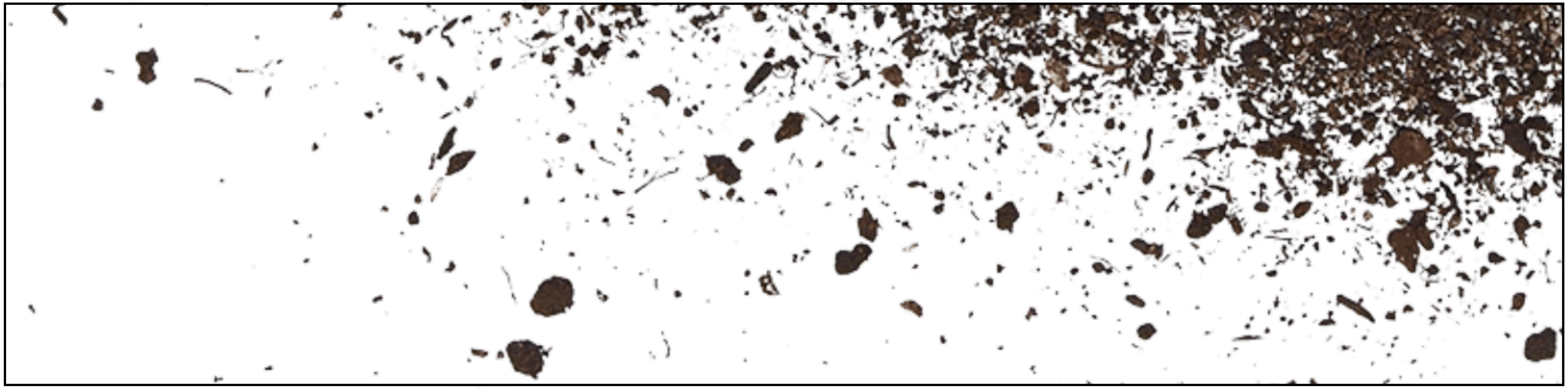
REAL (complicated) DATA

This exercise is the closest we'll come to "a real data analysis" in this course!

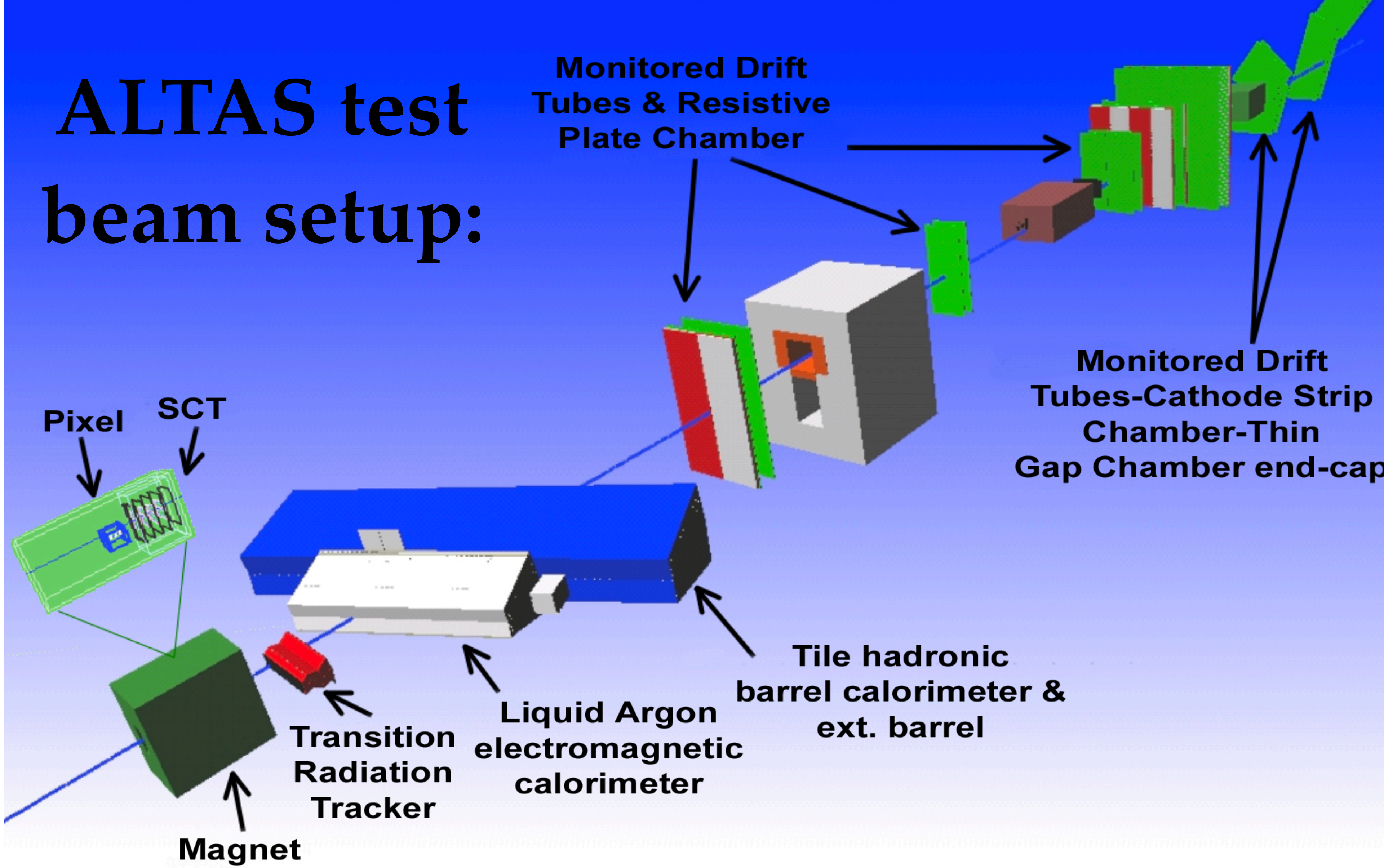
It is from the ATLAS test beam setup, used to test the different detectors and the associated software before assembling the full ATLAS detector.

However, it could in principle be "anything" - it doesn't really matter! In all essence, it is data with associated questions.

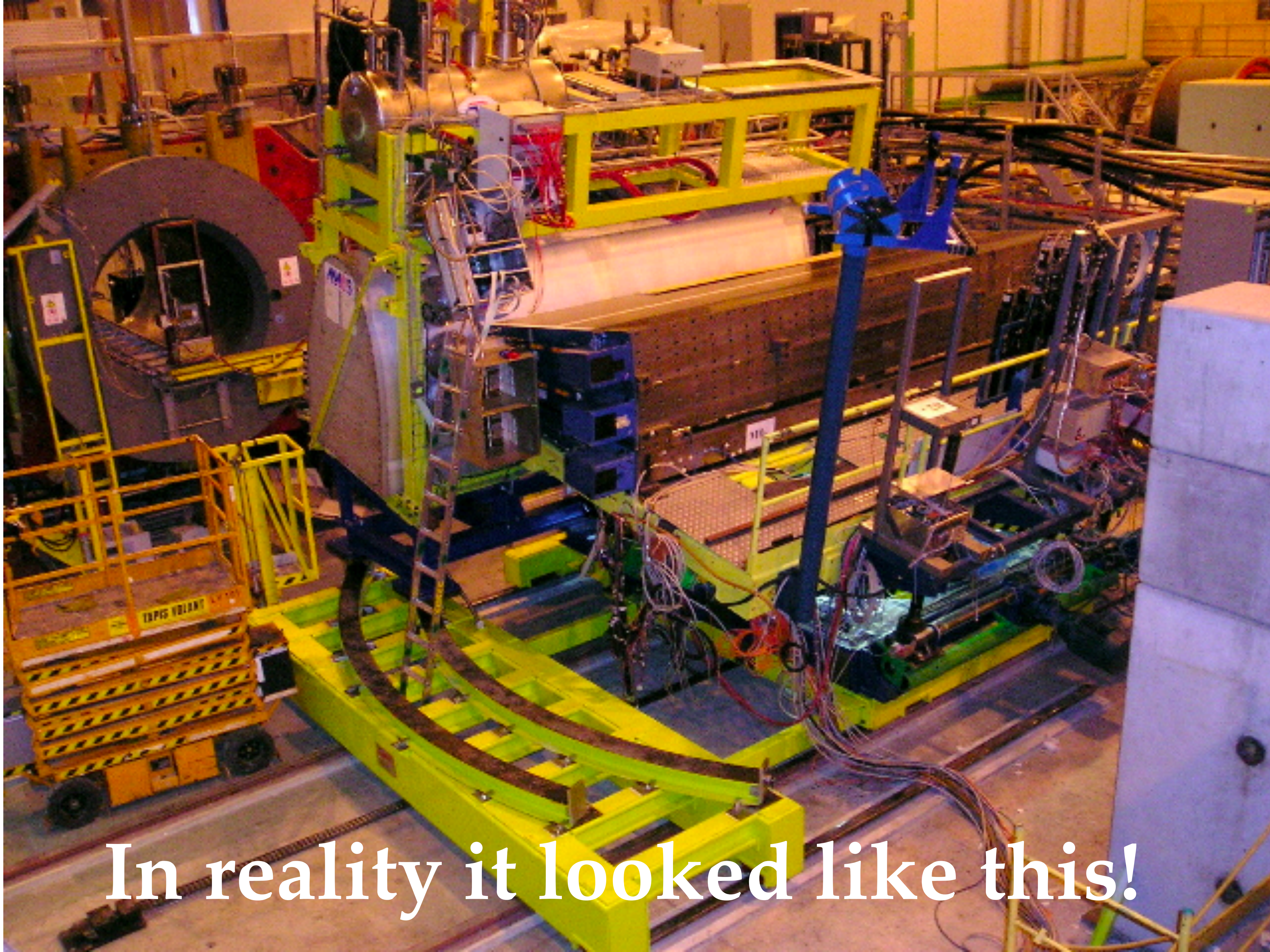
The aim is to classify the particles (or patients, products, promotions, etc.) and determine how well this can be done.



ALTAS test beam setup:

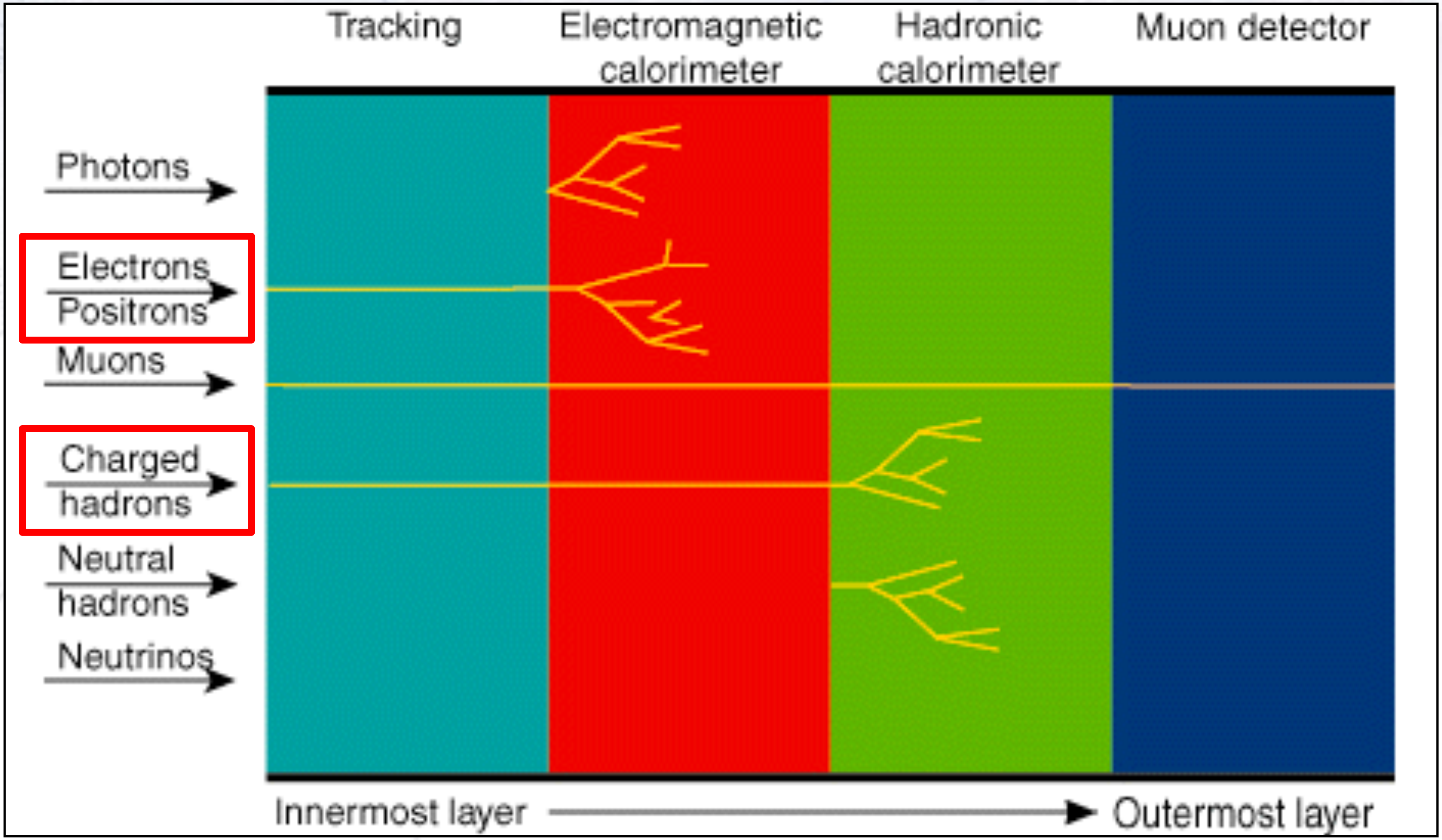


The ATLAS test beam setup aimed to test, measure, and optimise the performance of the various detector parts and the associated reconstruction software.

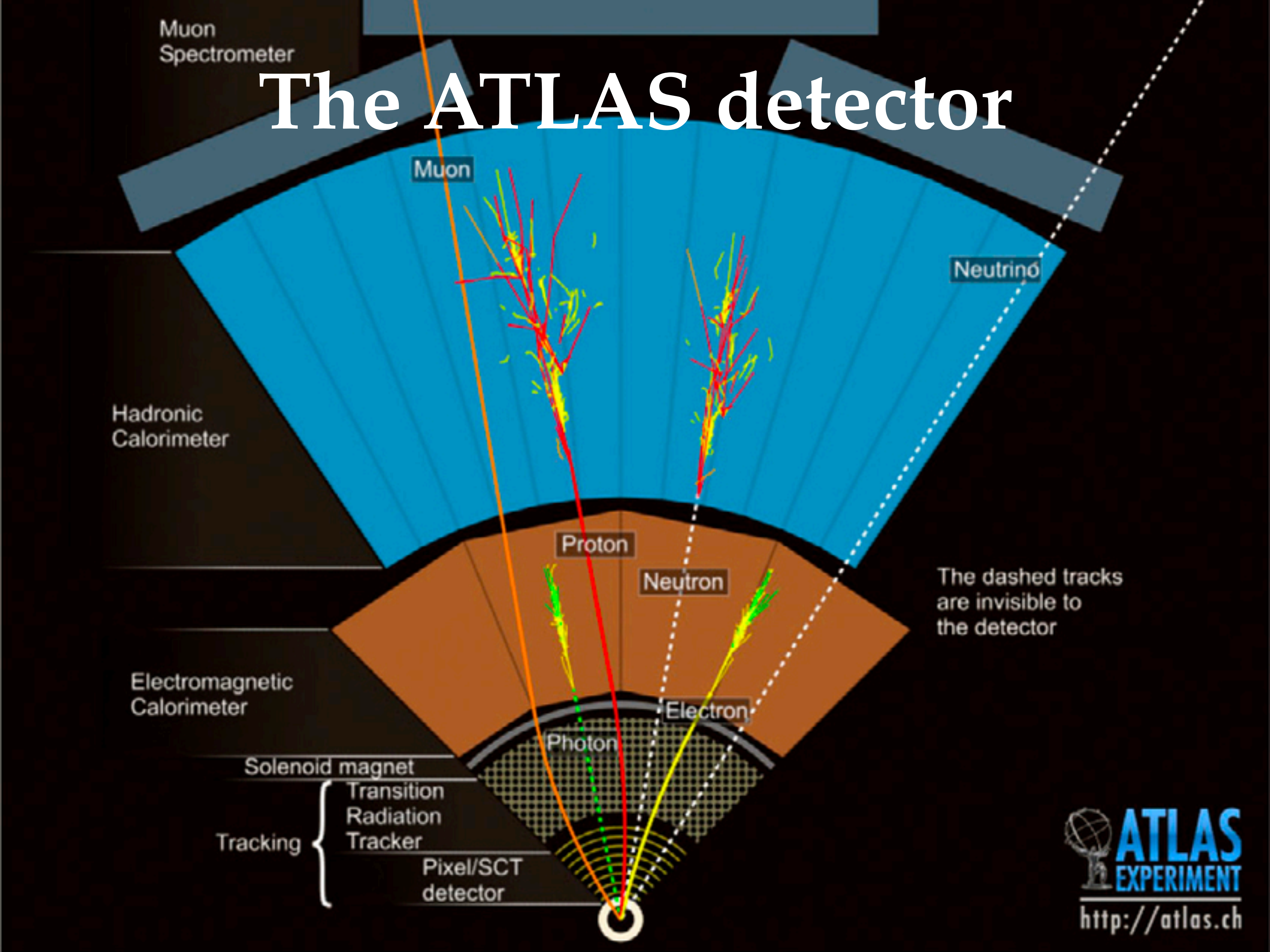

TAPIS VOLANT
In reality it looked like this!

How particles interact

In our case, there are only two types of particles: **Electrons and pions**
Electrons react significantly and thus stop early. **Pions** do the opposite!



The ATLAS detector



The ATLAS test beam data

The data for this analysis was taken at a beam energy of 2 GeV. Single particles (**electrons or pions**) were sent through the detector layers, and for each particle passing into the detector slice, the following numbers (among 100s of others!) were recorded:

- **Cherenkov counts** (1 number). The Cherenkov counter reacts on particles, which travels faster than the speed of light IN THAT MEDIUM, thus mostly to electrons.

- **Transition Radiation Tracker** (2 numbers, both integer). Number of hits for tracking in the TRT (first) and number of High Threshold (HT) hits (second) for identifying electrons, since these have a higher chance of creating such a HT hit.

- **ElectroMagnetic Calorimeter** (4 numbers). This type of detector stops electrons, which thus tend to deposit their energy AS EARLY as possible. The four numbers correspond to the energy deposit (in GeV) in each of the 4 layers of the ATLAS EM calorimeter.

- **Hadronic Calorimeter** (3 numbers). Much thicker than the EM calorimeter, this detector stops all particles except muons with more than 3-4 GeV in energy. Gives no signal, if no particle reaches the detector. Electrons hardly ever reach this detector.

- **Muon detector** (1 number). If a muons passes through, this detector gives a higher signal than when not. However, the coverage of the muon detector used was very small, and thus only few muons gives a signal. Not really used in the exercise.

Checking the data

If someone tells you that the raw data "is fine", then **distrust that person**!!!

The more "raw" the data, the more inspection and cleaning is needed (recall the Table Measurement exercise).

For this reason, but also to get a feel for the data, a sample is printed:

```
Cher:  634.0    nLT, nHT: 32,  8    EM:  0.12  0.47  0.87 -0.04   Had:  0.00  0.00  0.00    Muon: 417.0
Cher:  816.0    nLT, nHT: 42, 12    EM:  0.05  0.87  0.77 -0.01   Had:  0.00  0.00  0.00    Muon: 388.0
Cher:  943.0    nLT, nHT: 46,  8    EM:  0.19  0.53  0.96  0.01   Had:  0.00  0.00  0.00    Muon: 377.0
Cher:  907.0    nLT, nHT: 33,  3    EM:  0.20  0.74  0.72  0.02   Had:  0.00  0.00  0.00    Muon: 372.0
Cher:  775.0    nLT, nHT: 35,  4    EM: -0.01  0.82  0.57  0.02   Had:  0.00  0.00  0.00    Muon: 392.0
Cher:  773.0    nLT, nHT: 39,  9    EM:  0.10  1.24  0.27 -0.03   Had:  0.00  0.00  0.00    Muon: 398.0
Cher:  782.0    nLT, nHT: 30,  7    EM:  0.35  0.89  0.60 -0.04   Had:  0.00  0.00  0.00    Muon: 369.0
Cher:  700.0    nLT, nHT: 39,  9    EM:  0.03  1.14  0.60  0.02   Had:  0.00  0.03  0.76    Muon: 408.0
Cher:  542.0    nLT, nHT: 42,  1    EM: -0.03 -0.01  0.13 -0.05   Had:  0.00  0.00  0.00    Muon: 412.0
Cher:  752.0    nLT, nHT: 40,  7    EM:  0.09  0.66  0.57  0.10   Had:  0.00  0.00  0.00    Muon: 387.0
Cher:  576.0    nLT, nHT: 29,  1    EM:  0.08  0.02  0.04  0.03   Had:  0.00  0.00  0.00    Muon: 392.0
Cher:  772.0    nLT, nHT: 46,  6    EM:  0.15  1.11  0.49  0.09   Had:  0.03  0.00  0.00    Muon: 393.0
Cher:  735.0    nLT, nHT: 39,  8    EM: -0.02  0.72  1.04  0.09   Had:  0.00  0.00  0.00    Muon: 413.0
Cher:  751.0    nLT, nHT: 44,  7    EM:  0.43  1.11  0.33  0.03   Had:  0.00  0.00  0.00    Muon: 359.0
Cher:  903.0    nLT, nHT: 37,  6    EM:  0.26  0.58  0.67 -0.04   Had:  0.00  0.00  0.00    Muon: 394.0
Cher:  557.0    nLT, nHT: 44,  3    EM: -0.06 -0.01  0.38 -0.04   Had:  0.30  1.32  0.81    Muon: 505.0
Cher:  635.0    nLT, nHT: 45,  7    EM:  0.08 -0.07  0.04  0.01   Had:  0.00  0.00  0.00    Muon: 378.0
Cher:  801.0    nLT, nHT: 36,  3    EM:  0.45  0.93  0.47 -0.00   Had:  0.00  0.00  0.00    Muon: 425.0
Cher:  940.0    nLT, nHT: 27,  3    EM:  0.27  1.05  0.56 -0.10   Had:  0.00  0.00  0.00    Muon: 373.0
Cher:  760.0    nLT, nHT: 39, 10    EM:  0.23  1.11  0.32 -0.04   Had:  0.00  0.00  0.00    Muon: 421.0
```

ATLAS test beam data

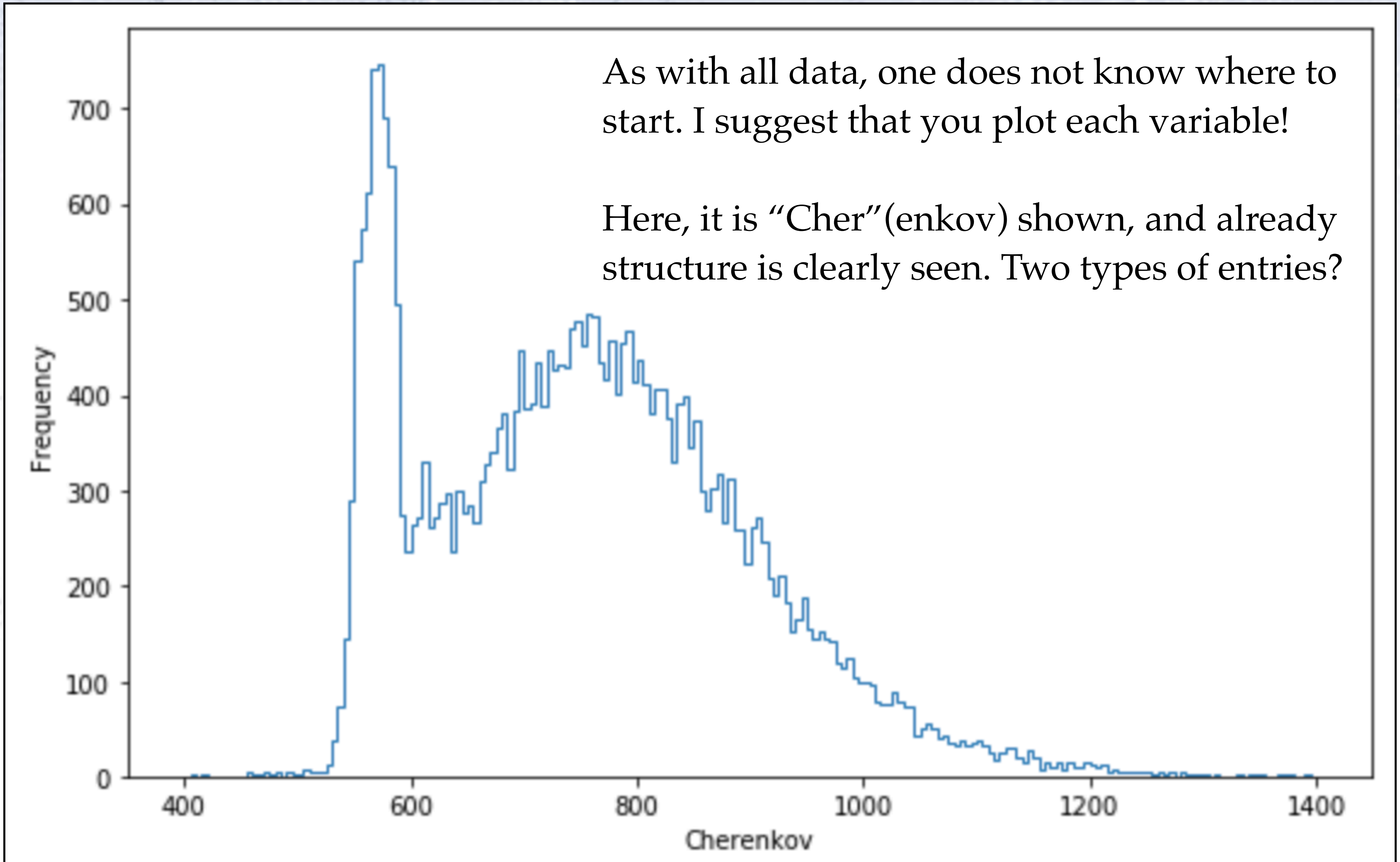
Cher	nLT	nHT	EM0	EM1	EM2	EM3	Had1	Had2	Had3	Muon
634.0	32	8	0.123	0.468	0.868	-0.037	0.000	0.000	0.000	417.0
816.0	42	12	0.049	0.870	0.768	-0.006	0.000	0.000	0.000	388.0
943.0	46	8	0.188	0.534	0.965	0.012	0.000	0.000	0.000	377.0
907.0	33	3	0.198	0.740	0.724	0.022	0.000	0.000	0.000	372.0
775.0	35	4	-0.010	0.822	0.574	0.023	0.000	0.000	0.000	392.0
773.0	39	9	0.100	1.236	0.272	-0.033	0.000	0.000	0.000	398.0
782.0	30	7	0.355	0.885	0.596	-0.042	0.000	0.000	0.000	369.0
700.0	39	9	0.031	1.141	0.602	0.023	0.000	0.033	0.757	408.0
542.0	42	1	-0.028	-0.014	0.133	-0.054	0.000	0.000	0.000	412.0
752.0	40	7	0.092	0.659	0.574	0.098	0.000	0.000	0.000	387.0
576.0	29	1	0.076	0.018	0.044	0.025	0.000	0.000	0.000	392.0
772.0	46	6	0.151	1.107	0.489	0.090	0.027	0.000	0.000	393.0
735.0	39	8	-0.018	0.723	1.044	0.088	0.000	0.000	0.000	413.0
751.0	44	7	0.431	1.110	0.328	0.029	0.000	0.000	0.000	359.0
903.0	37	6	0.261	0.580	0.667	-0.044	0.000	0.000	0.000	394.0
557.0	44	3	-0.059	-0.007	0.378	-0.039	0.302	1.325	0.810	505.0
635.0	45	7	0.079	-0.072	0.040	0.010	0.000	0.000	0.000	378.0
801.0	36	3	0.452	0.926	0.474	-0.004	0.000	0.000	0.000	425.0
940.0	27	3	0.267	1.049	0.559	-0.099	0.000	0.000	0.000	373.0
760.0	39	10	0.229	1.107	0.316	-0.041	0.000	0.000	0.000	421.0
1003.0	48	3	-0.070	0.585	0.600	-0.089	0.000	0.000	0.000	412.0
894.0	44	10	0.095	0.911	0.623	-0.055	0.000	0.000	0.000	393.0

ATLAS test beam data

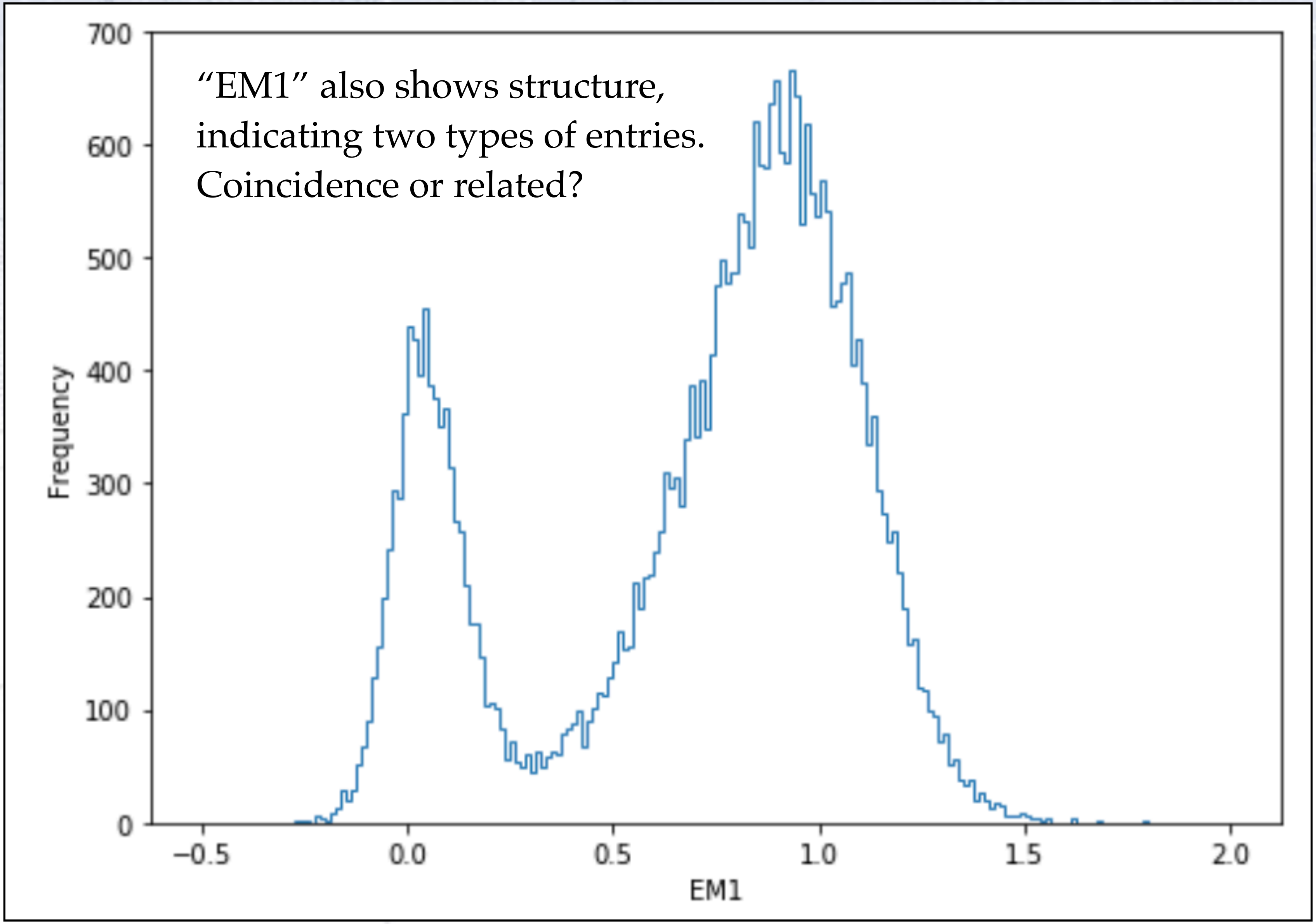
Cher	nLT	nHT	EM0	EM1	EM2	EM3	Had1	Had2	Had3	Muon
634.0	32	8	0.123	0.468	0.868	-0.037	0.000	0.000	0.000	417.0
816.0	42	12	0.049	0.870	0.768	-0.006	0.000	0.000	0.000	388.0
943.0	46	8	0.188	0.534	0.965	0.012	0.000	0.000	0.000	377.0
907.0	33	3	0.198	0.740	0.724	0.022	0.000	0.000	0.000	372.0
775.0	35	4	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.000	0.000	392.0
773.0	39	9	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.000	0.000	398.0
782.0	30	7	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.000	0.000	369.0
700.0	39	9	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.033	0.757	408.0
542.0	42	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.000	0.000	412.0
752.0	40	7	0.092	0.659	0.574	0.098	0.000	0.000	0.000	387.0
576.0	29	1	0.076	0.018	0.044	0.025	0.000	0.000	0.000	392.0
772.0	46	6	0.151	1.107	0.489	0.090	0.027	0.000	0.000	393.0
735.0	39	8	-0.018	0.723	1.044	0.088	0.000	0.000	0.000	413.0
751.0	44	7	0.431	1.110	0.328	0.029	0.000	0.000	0.000	359.0
903.0	37	6	0.261	0.580	0.667	-0.044	0.000	0.000	0.000	394.0
557.0	44	3	-0.059	-0.007	0.378	-0.039	0.302	1.325	0.810	505.0
635.0	45	7	0.079	-0.072	0.040	0.010	0.000	0.000	0.000	378.0
801.0	36	3	0.452	0.926	0.474	-0.004	0.000	0.000	0.000	425.0
940.0	27	3	0.267	1.049	0.559	-0.099	0.000	0.000	0.000	373.0
760.0	39	10	0.229	1.107	0.316	-0.041	0.000	0.000	0.000	421.0
1003.0	48	3	-0.070	0.585	0.600	-0.089	0.000	0.000	0.000	412.0
894.0	44	10	0.095	0.911	0.623	-0.055	0.000	0.000	0.000	393.0

Variables to be used!
They are from three
INDEPENDENT detectors…

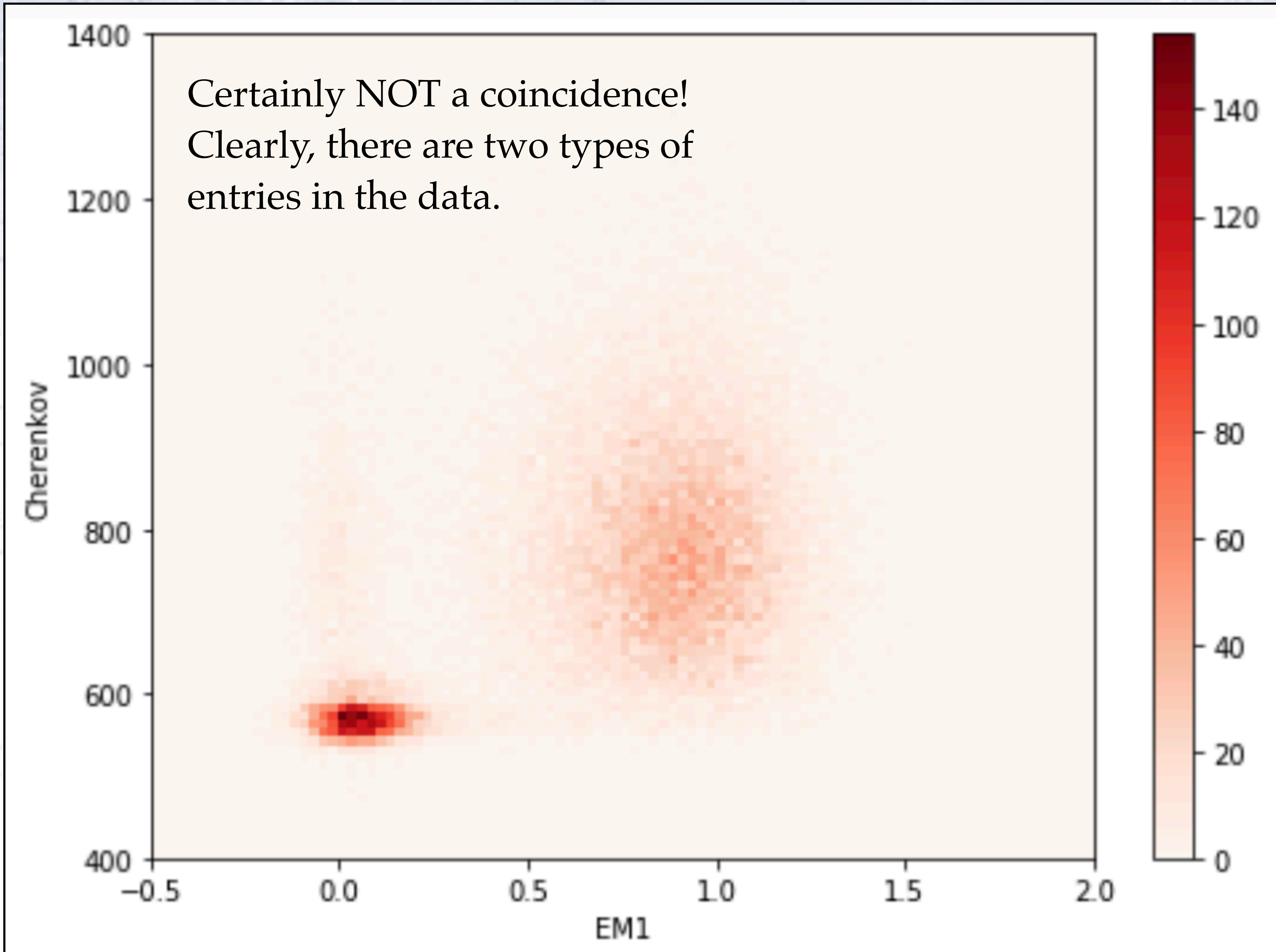
Looking at the data



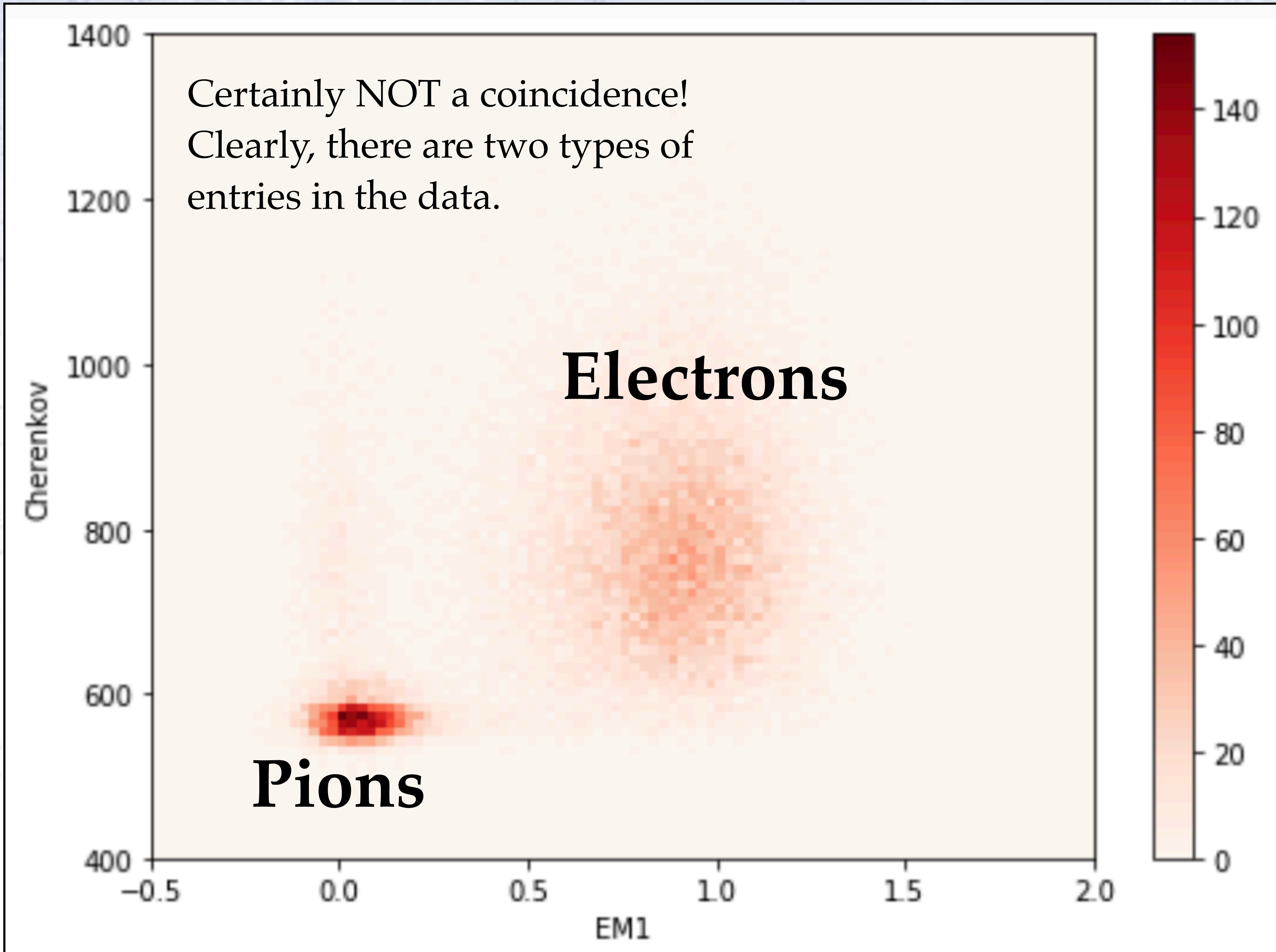
Looking at the data



Looking at the data

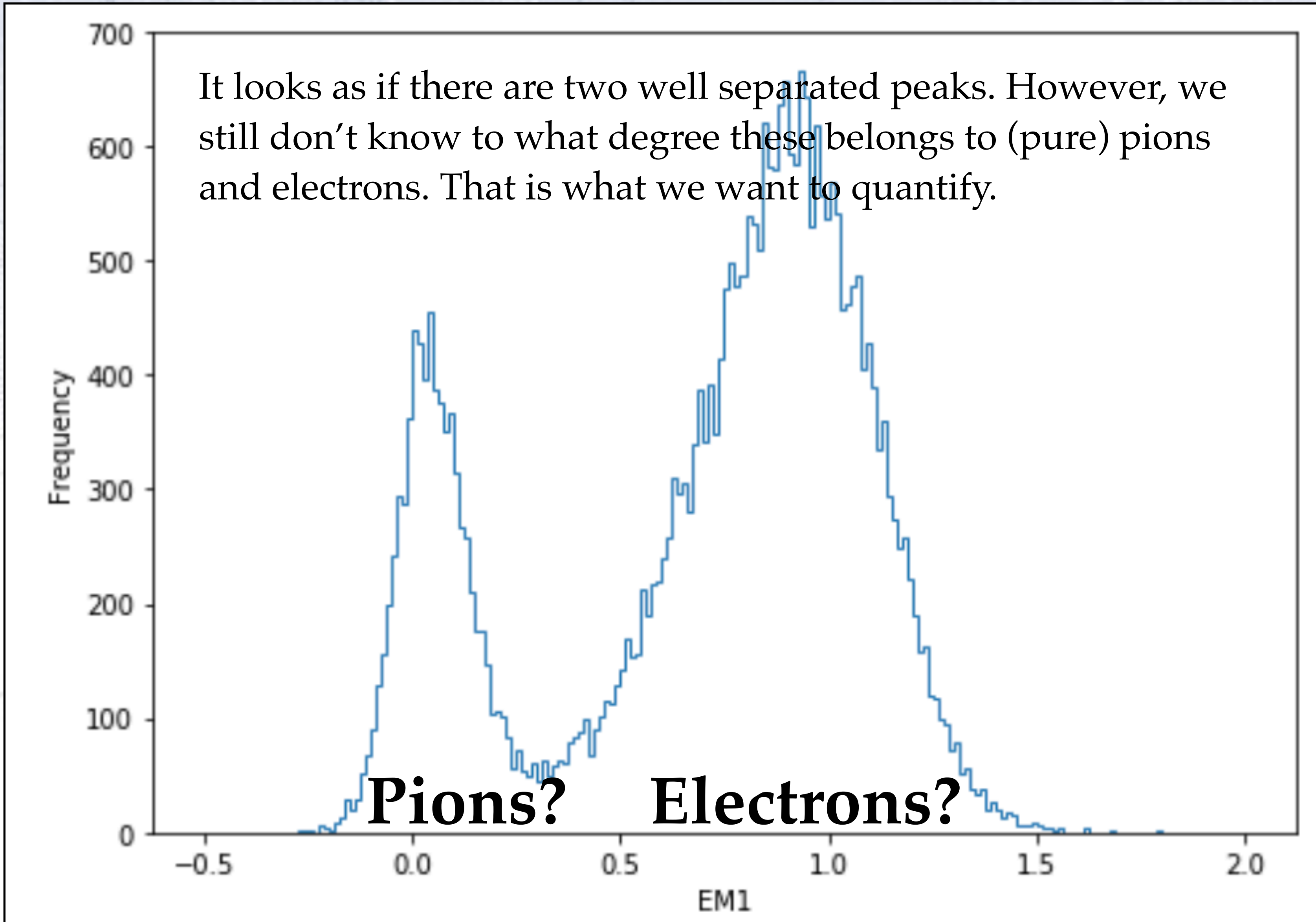


Looking at the data



Looking at the data

It looks as if there are two well separated peaks. However, we still don’t know to what degree these belongs to (pure) pions and electrons. That is what we want to quantify.

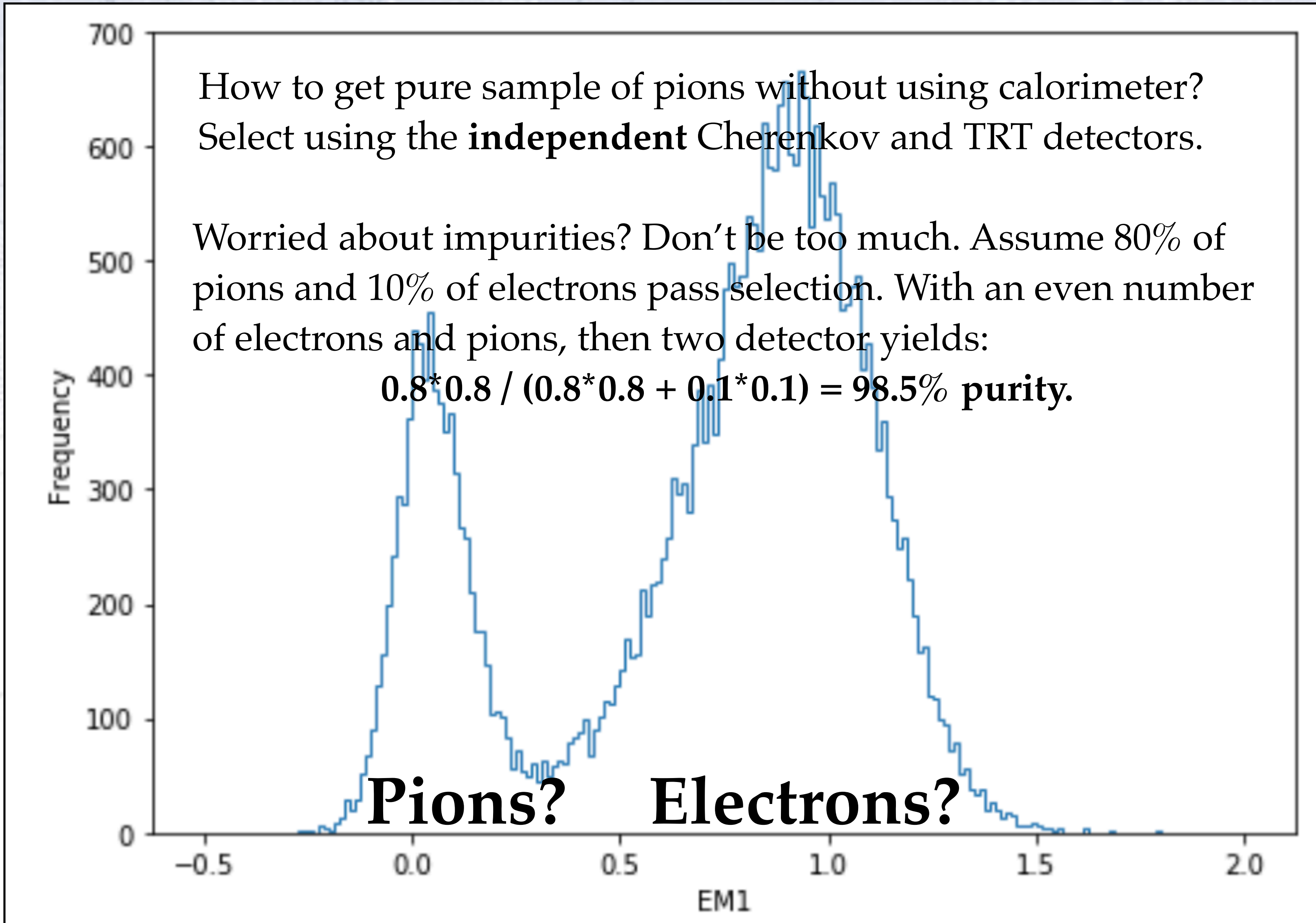


Looking at the data

How to get pure sample of pions without using calorimeter?
Select using the **independent** Cherenkov and TRT detectors.

Worried about impurities? Don't be too much. Assume 80% of pions and 10% of electrons pass selection. With an even number of electrons and pions, then two detector yields:

0.8*0.8 / (0.8*0.8 + 0.1*0.1) = 98.5% purity.



Questions to the data

Main question: How good are each of the three detectors in question (Cherenkov, TRT, and Calorimeter) to distinguish between electrons and pions?

Suggested method: Using purified sample (based on cuts on the two other detectors), ask what fraction of electrons and pions passes your electron selection. The fraction of electrons, that are not selected as electrons will be your TYPE I errors, denoted alpha, while the fraction of pions, that do pass the cut will be your TYPE II errors, denoted beta.
Optimise your selection criteria (cuts), and measure the two error rates for each of the three detector types and particle types. Each detector should get six numbers:

The electron cut value above which you accept an electron.

- The signal efficiency (i.e. 1-alpha) for electrons of this cut.
- The background fake rate (i.e. beta) for pions of this cut.

The pion cut value below which you accept a pion.

- The efficiency (i.e. 1-alpha) for pions of this cut.
- The fake rate (i.e. beta) for electrons of this cut.

Finally, draw ROC curves for each detector's selection variable.

Questions to the data

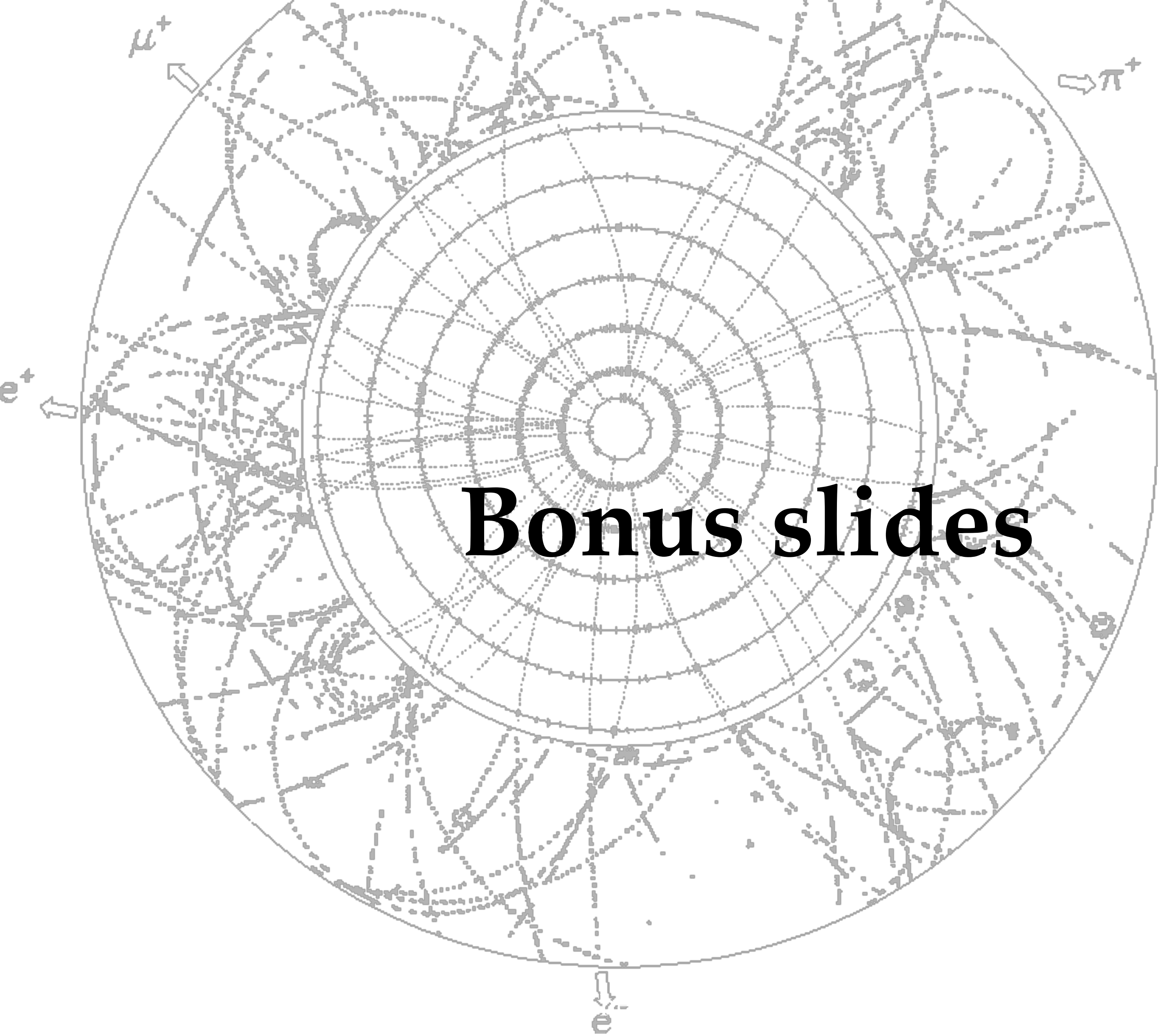
What does the sample consist of, i.e. number of electrons and pions in the data?

What are your sample purities, and are you able to get a 99.9% pure electron sample with 90% selection efficiency based on the TRT and calorimeter alone? (these are the detectors in ATLAS, while the Cherenkov was for test beam only).

Are the probabilities of HT hits approximately constant and independent for electrons and pions? If so, what distribution should the number of HT hits follow (for a fixed number of LT hits), and is this really the case both for electrons and pions?

Can you combine the four calorimeter energies into a Fisher discriminant? I.e. can you boost the performance of the calorimeter by using a linear combination of the inputs? You might consider providing some ratios instead of the raw measurement.

More questions can of course be asked, but this should suffice.

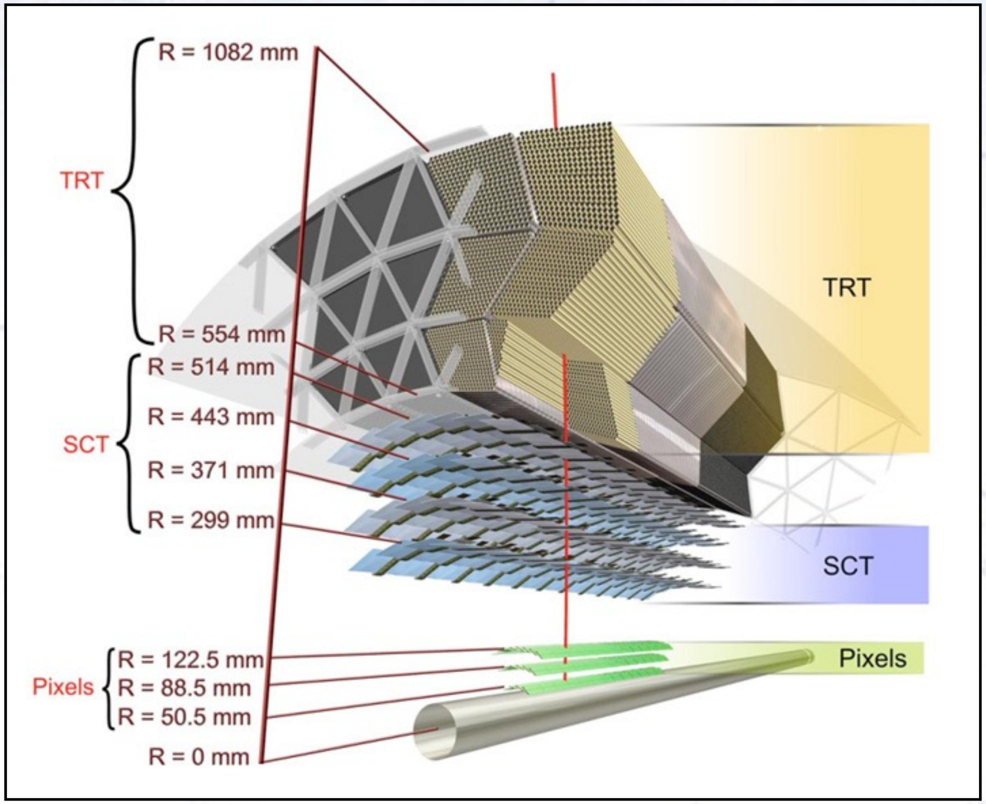


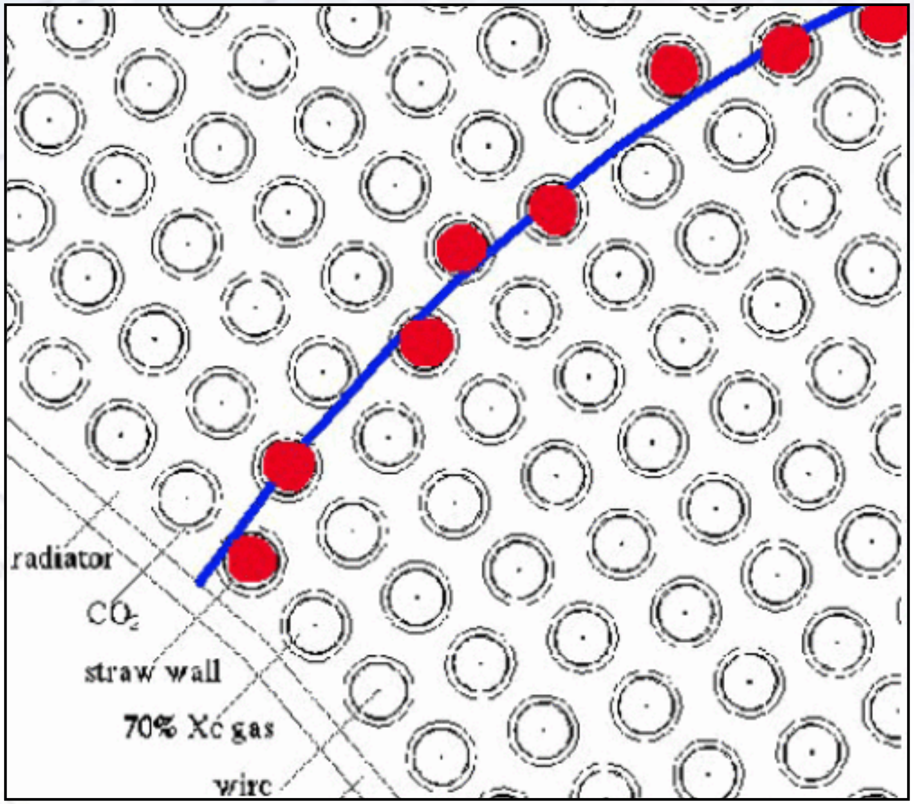
Bonus slides

ATLAS TRT detector

ATLAS TRT detector

The ATLAS TRT detector works by measuring the ionisation of charged particles passing through. In addition, it can measure Transition Radiation from electrons!

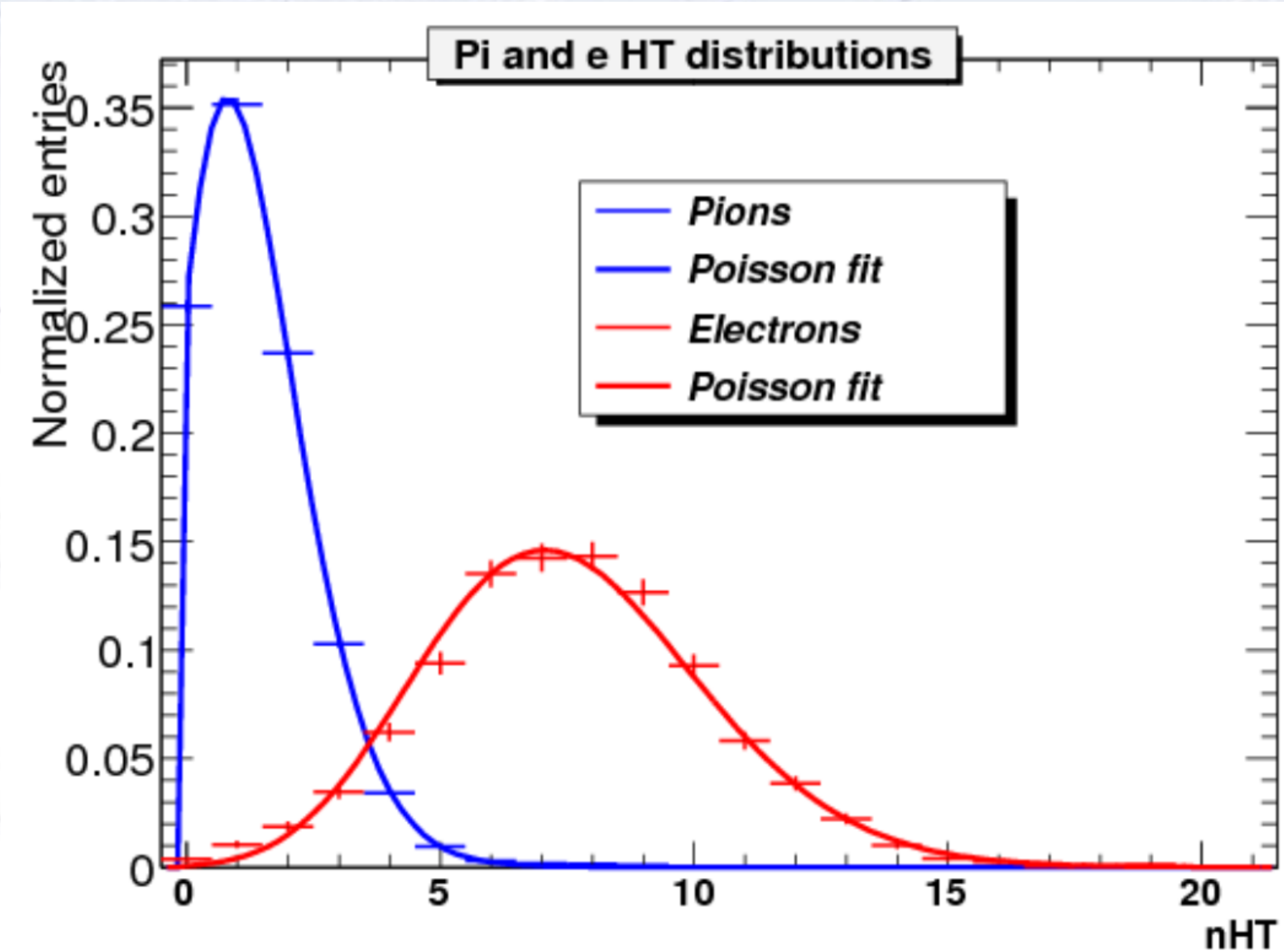




The probability of TR hit (called High-Threshold hit) is about 20-25% for electrons, while it is 4-6% for pions. Given 30-35 hits, this gives a difference.

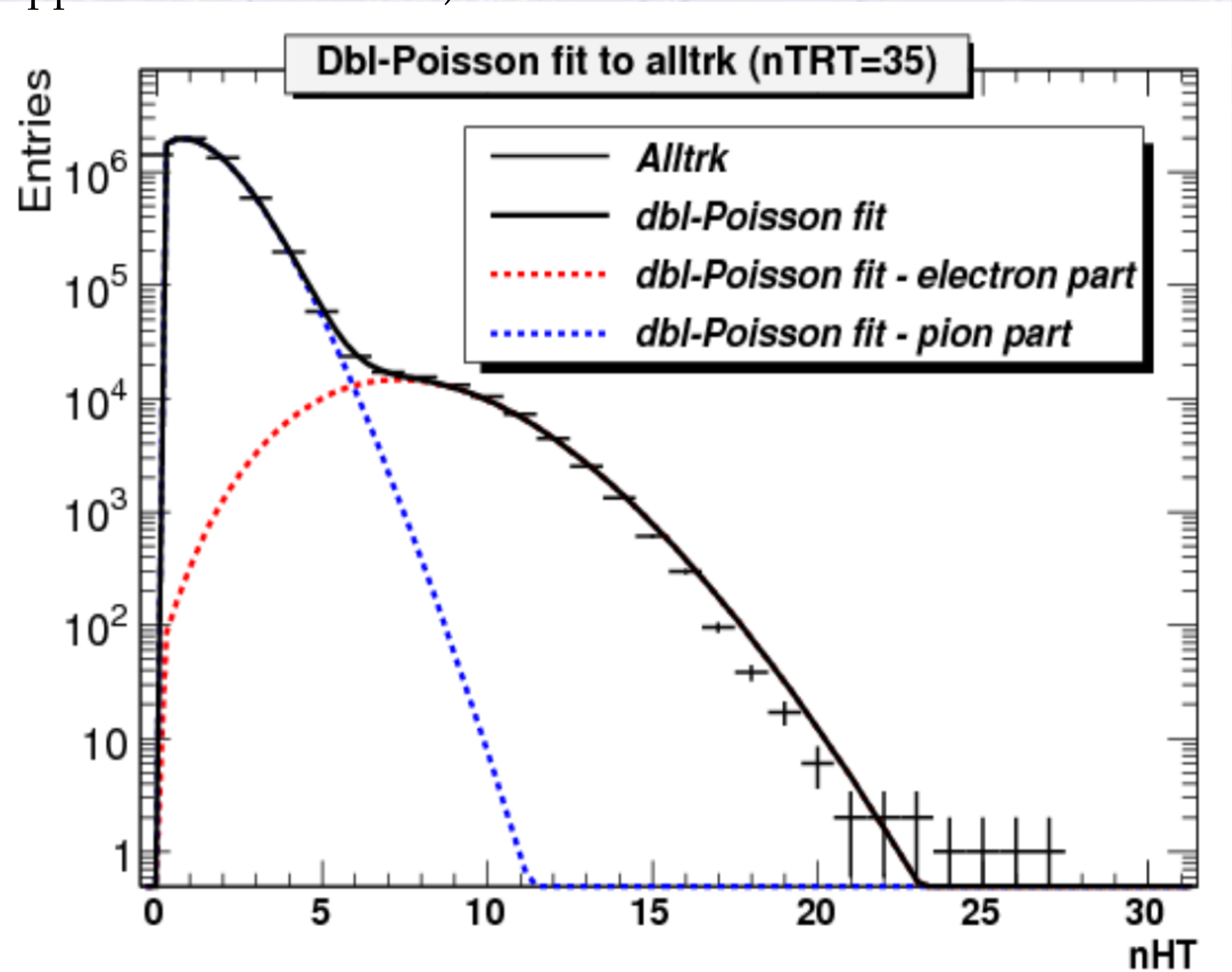
ATLAS TRT detector

Given a FIXED number of TRT hits, the distribution of HT hits is Binomial (and to some approximation Poisson).



ATLAS TRT detector

Given a FIXED number of TRT hits, the distribution of HT hits is Binomial (and to some approximation Poisson).



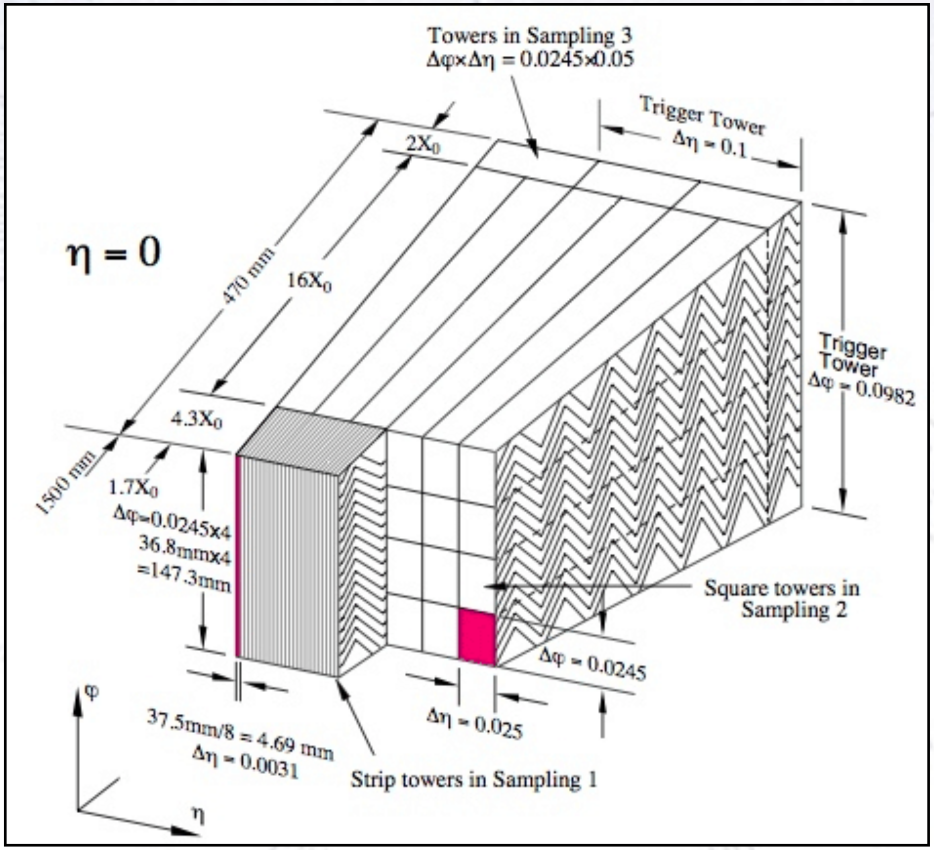
The NBI ColliderScope

If you feel, that you've seen this light pattern before, then it might be from the lights on the facade of NBI (ColliderScope by Skeel, Skrive, Ellegaard & Petersen).

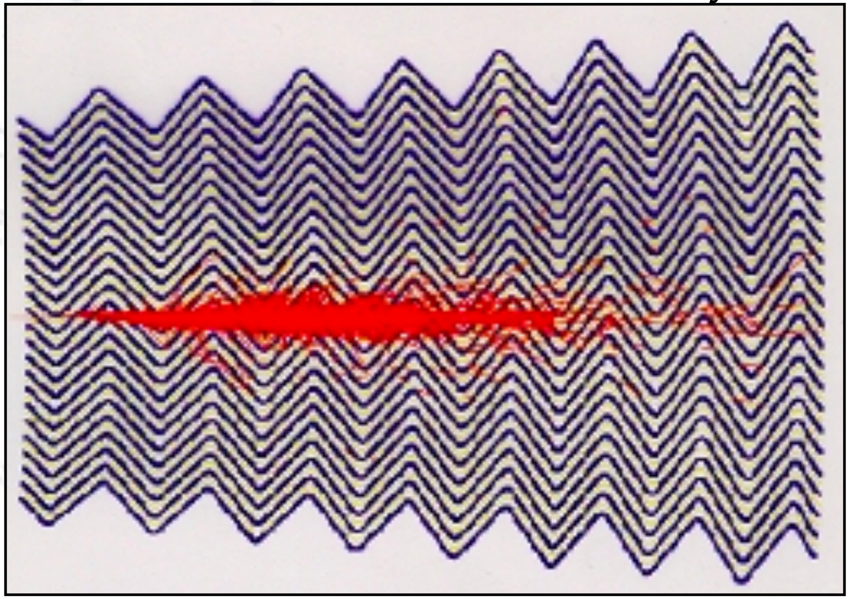
ATLAS EM Calorimeter

ATLAS EM Calorimeter

The ATLAS Electro-Magnetic (EM) Calorimeter consists of four non-equal layers (left figure), in which electrons produce an EM shower (right figure). This leaves energy in each layer, which is characteristic of electrons.



Shower across three main calo layers.



Compared to other particles, electrons tend to leave more energy early on.