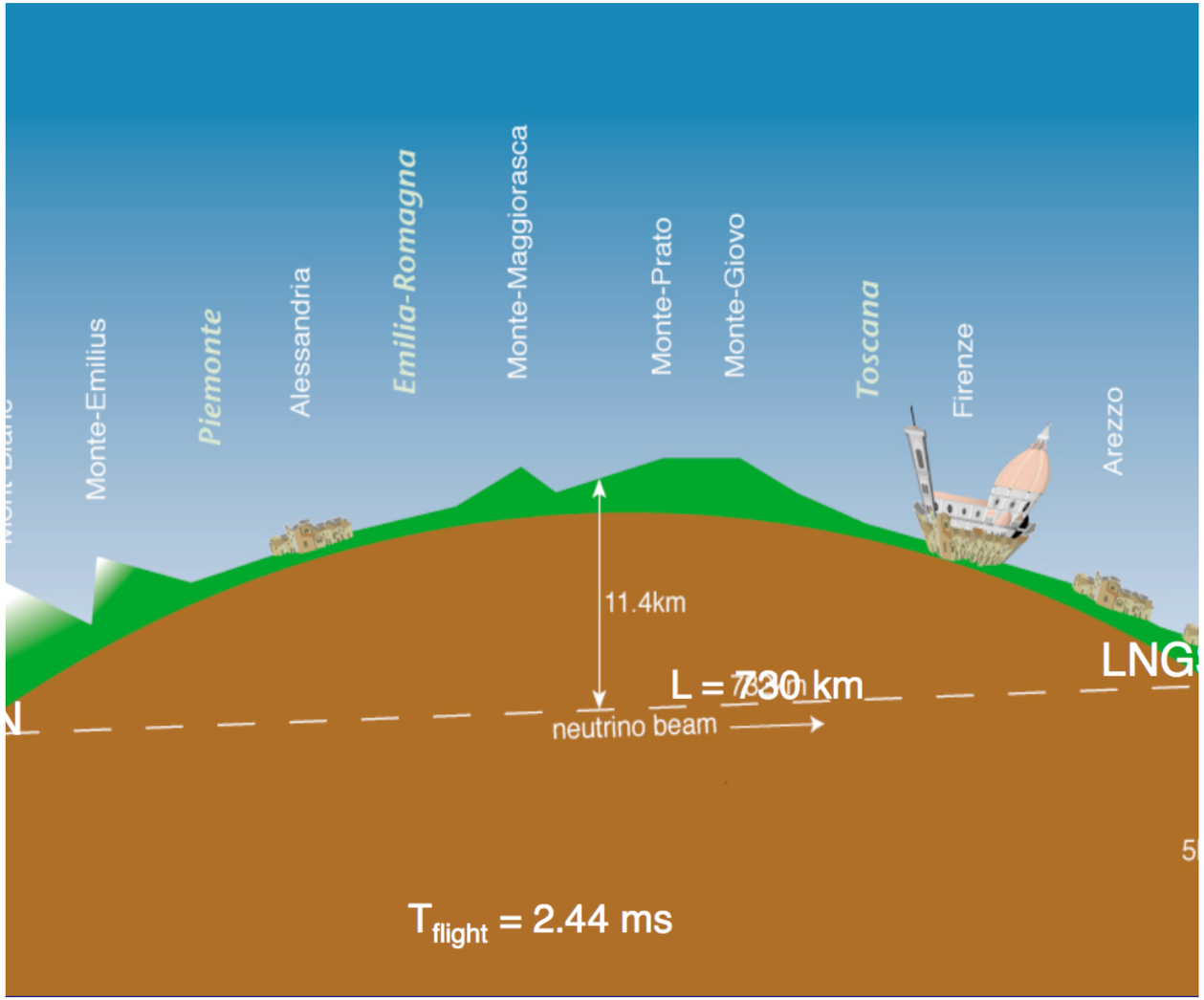
Dark Matter Signal Search.

Episode 1

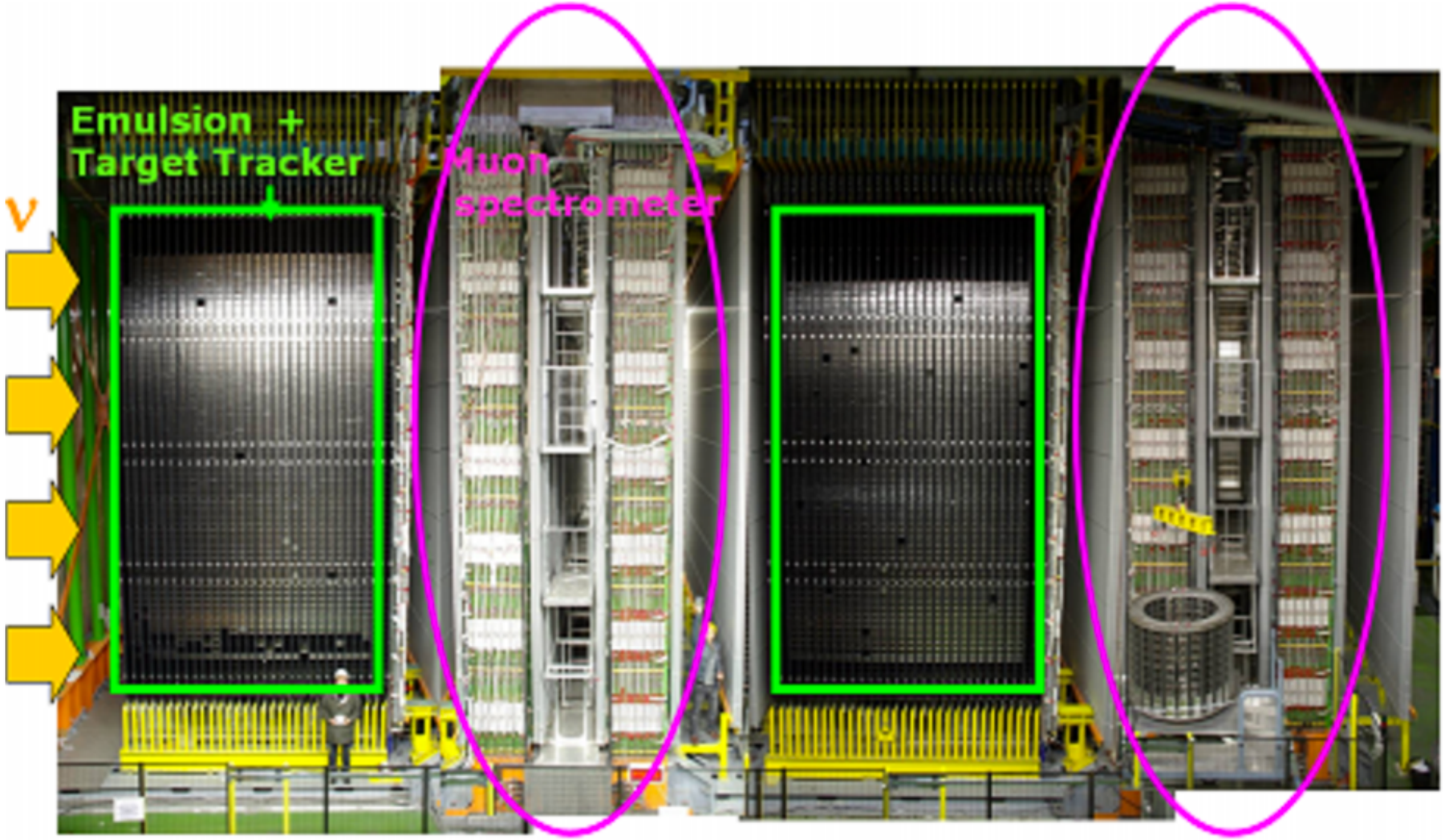
A. Filatov, S. Shirobokov, Andrey Ustyuzhanin
Yandex School of Data Analysis
NRU Higher School of Economics
INFN, Napoli

From CERN to OPERA, overview



- Goal: find neutrino oscillations
- Detector: photo emulsion
- Data taking: 2008-2012
- Results: 5 $\boldsymbol{\nu}_{\mu} \longrightarrow \boldsymbol{\nu}_{\tau}$ observed,
- 2015 – Nobel prize in Physics for discovery of neutrino oscillations
- http://operaweb.lngs.infn.it

OPERA detector



OPERA ECC brick

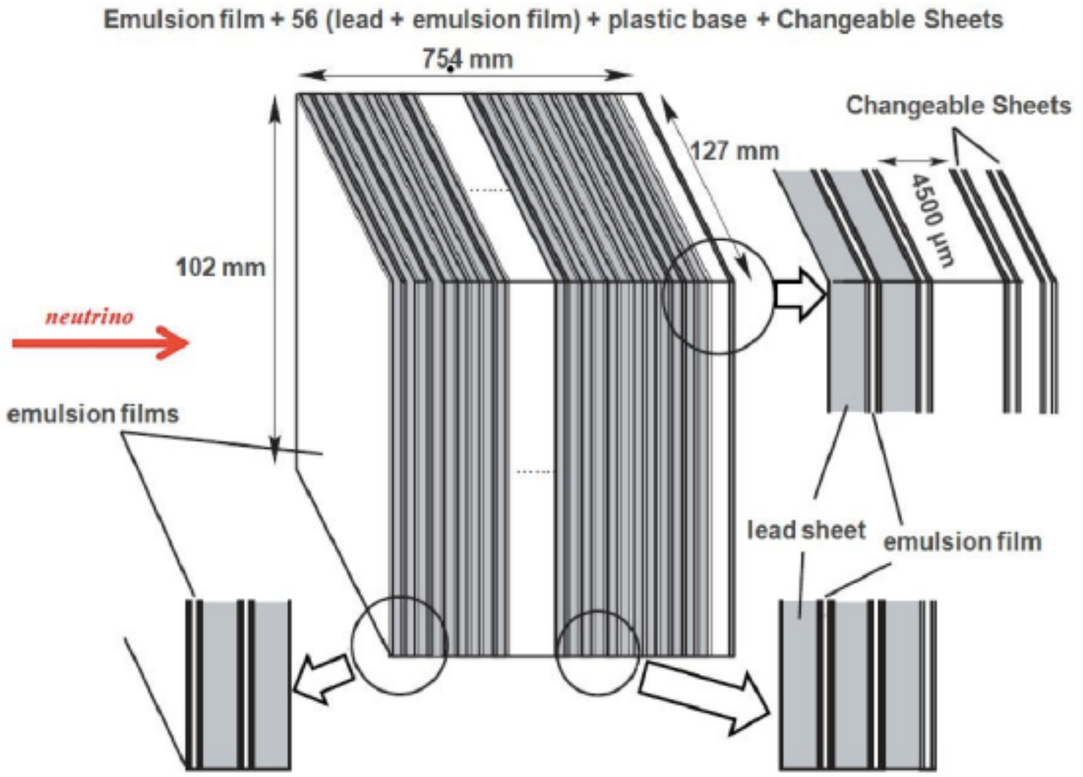


Figure 2.4 – Schematic structure of an ECC brick.

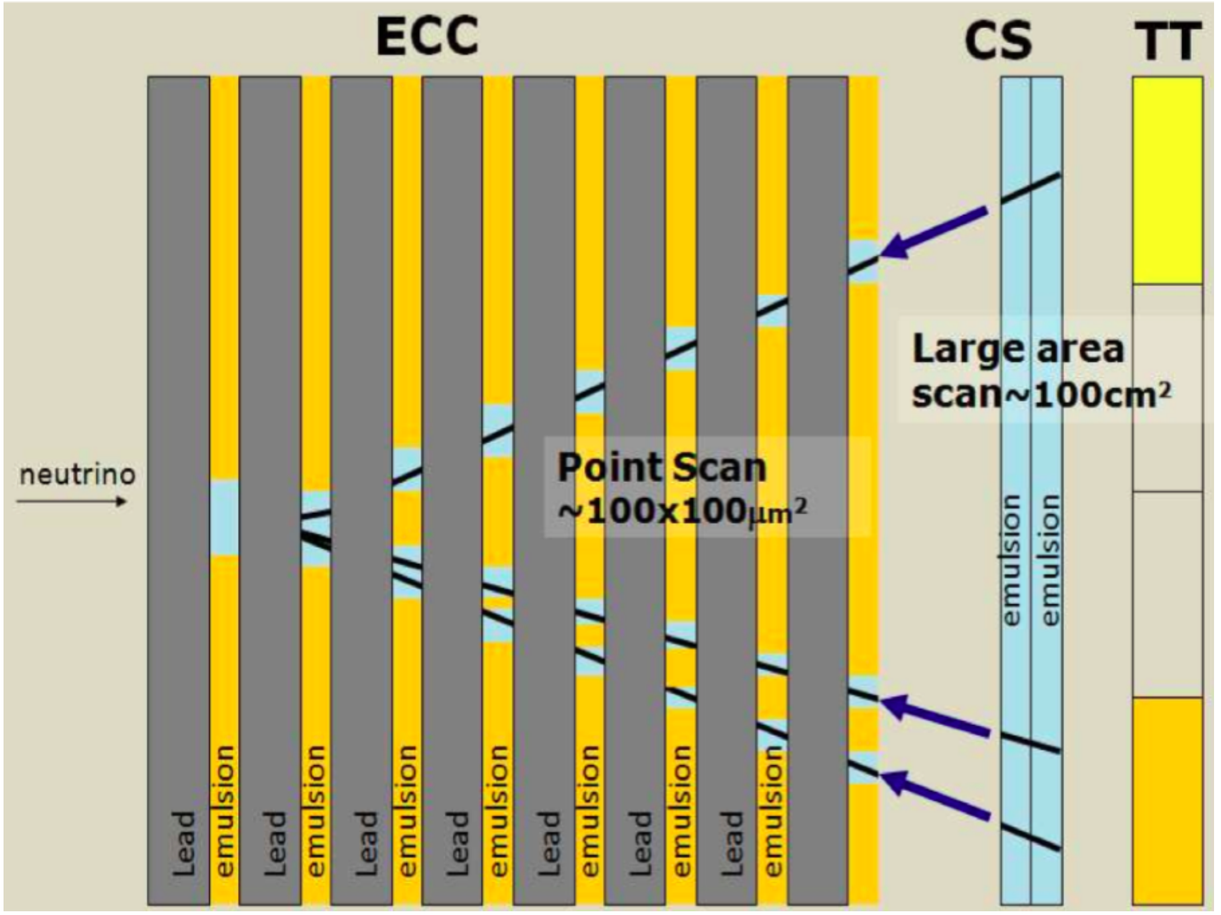
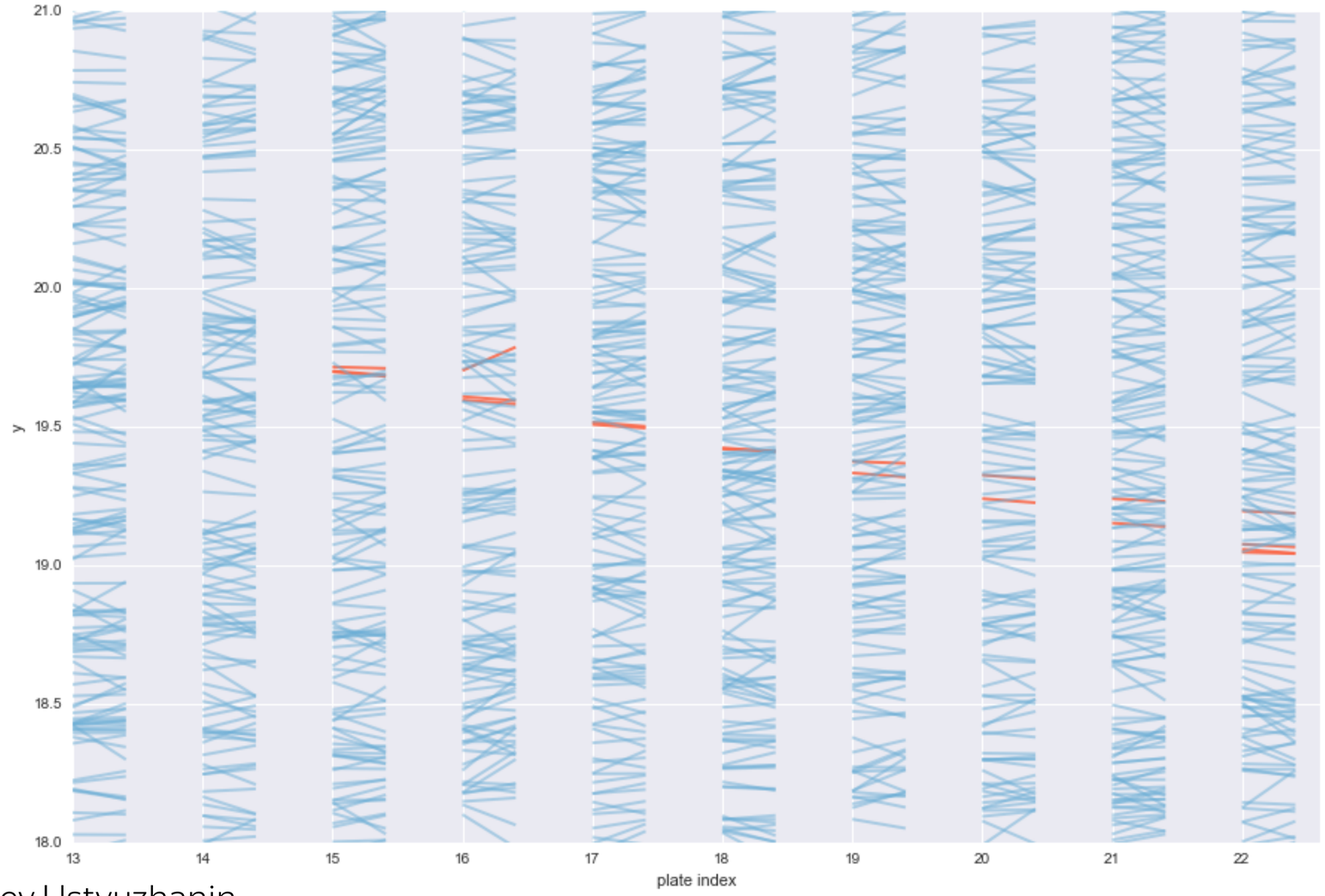


Plate of a brick

Dark matter clues

- While passing through lead material neutrino can scatter
    - on atomic nuclei – known behaviour (hadronic showers)
    - on atomic electron – similar to light-dark matter interaction. results in electro-magnetic showers
- SHiP experiment at CERN is
    - being design to search for particles no-one has ever seen (including light dark matter, http://ship.web.cern.ch/ship/)
    - going to use photo-emulsion detector similar to OPERA

Brick structure



Atomic track element: **basetrack**

- x,
- y,
- z,
- TX,
- TY,
- χ^2

Given

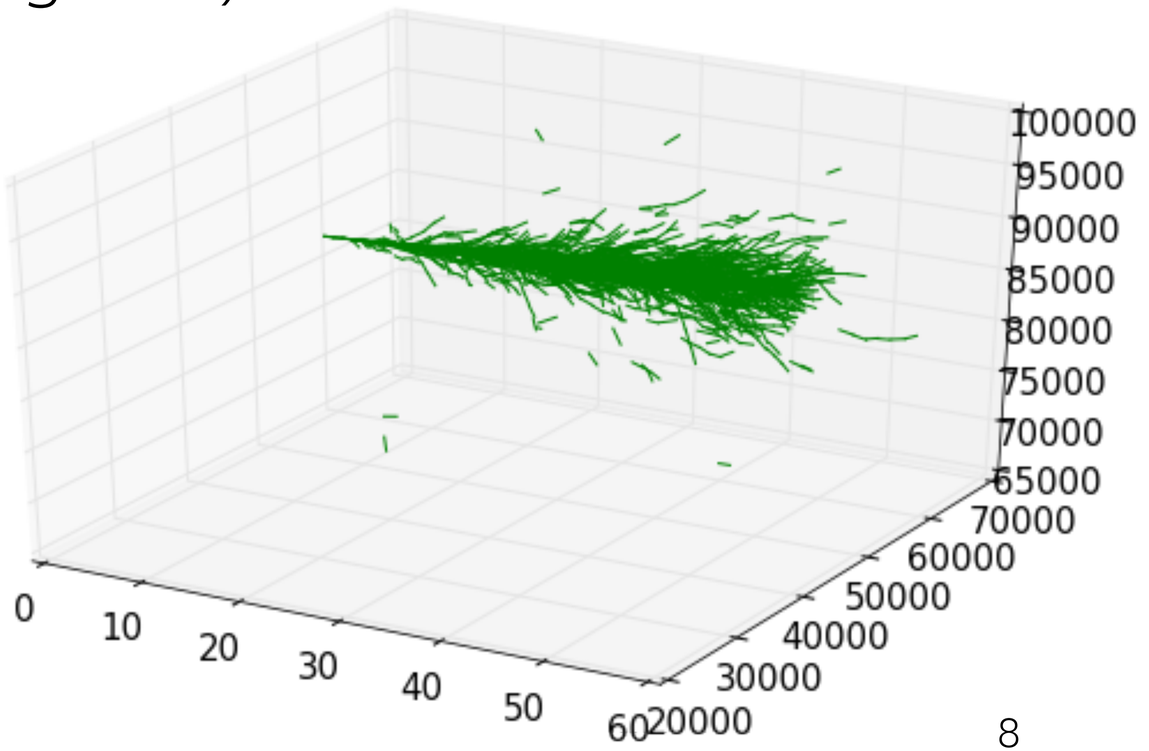
Data Background: 1 brick, ~ 10^6 base tracks (signal=0)

MC Signal: simulation of pure EM showers
(100 events, 10^2-10^3 basetracks per shower) (signal=1)

DS_1_train.csv, DS_1_test.csv,

Origin of the mother-particle is known (x, y, z, TX, TY, \chi2)

DS_1_electron_train.csv,
DS_1_electron_test.csv



Challenge

Develop algorithm that can

- detect electromagnetic shower basetracks within a brick basetracks (in test sample we have only description of the track (x, y, z, TX, TZ) for every track and set of mother-particles)
- Figure of Merit: ROC AUC

Hosted at: Kaggle, https://inclass.kaggle.com/c/dark-matter-signal-search-episode-1, requires valid account!

Competition deadline: 19-July-2017 23:59 UTC+0

Prize: memorable prizes + talk on Thursday

Very quick start

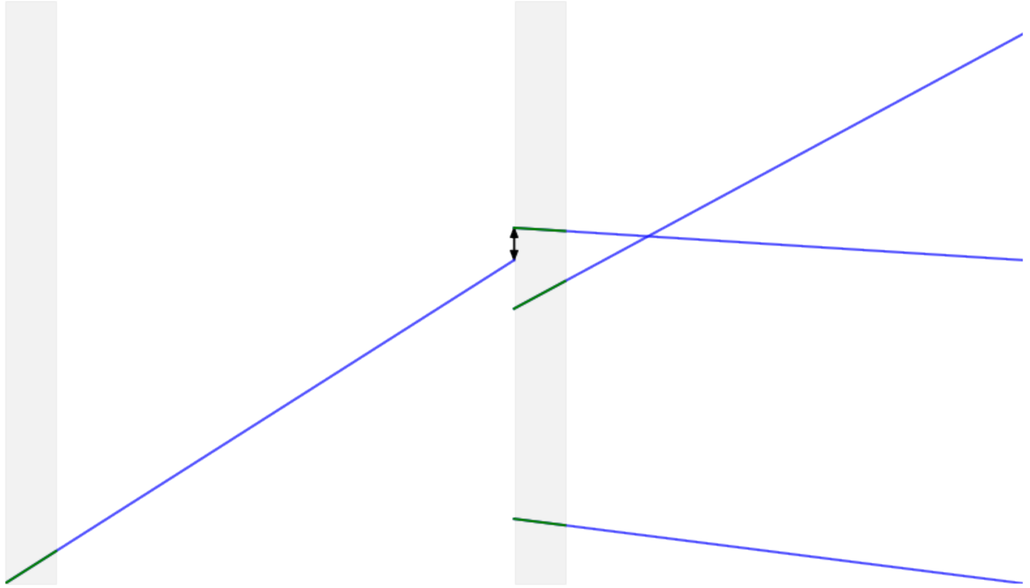
Baseline solution: https://github.com/yandexdataschool/mlhep2017
episode1_baseline folder

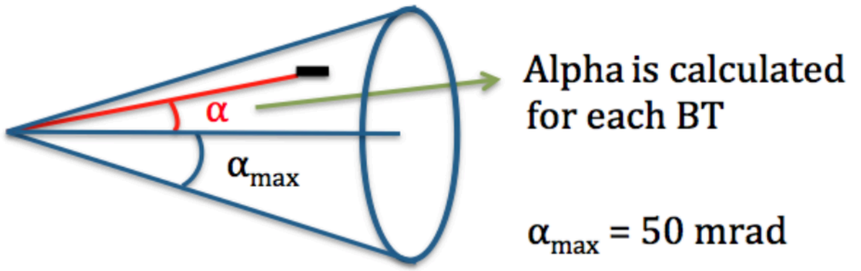
Can be run with everware (click ”run@everware” badge of the repository), data is at: **/data_mlhep2017**

Run chi2_baseline.ipynb, download baseline.csv, upload it to kaggle website

Ideas

- Use information about origin
- Consider tracks within certain angle from every known origin
- Play with new features:
  - distance between tracks
  - \Alpha (see figure on the right)
  - \Theta (angle between basetracks)
  - dTX, dTY (slope difference)
  - IP– Impact Parameter





Conclusion

- Intriguing problem, realistic dataset
- Non-realistic mixture:
    - In real life (SHiP) background sample is different (a lot of muon tracks + hadronic showers)
    - No origin is known
    - Number of electro-magnetic showers is going to be around 200
- End of challenge: 2017-07-19 23:59 UTC+0
- Use baseline!

Happy hacking!

References

- (thesis) Buonaura, Annarita **Study of nu-tau properties with the SHiP experiment,** http://cds.cern.ch/record/2268663

- (talk) Electron and proton beams for Dark Sector Searches at the CERN North Area, Andrey Golutvin, https://agenda.infn.it/getFile.py/access?contribId=20&sessionId=6&resId=0&materialId=slides&confId=12410

- (talk) Search for new physics with the SHiP experiment at CERN, https://indico.cern.ch/event/466934/contributions/2563221/attachments/1489290/2314198/talk.slides.pdf