From Quarks to Cosmos with AI, Carnegie-Mellon University July 12-16 2021

Muon Energy Regression Hackathon Challenge: Introduction

Tommaso Dorigo[1,2], Lukas Layer[1,3], Giles Strong[1,2]

1 – INFN

2 - Università di Padova

3 - Università di Napoli

The Standard Model

A misnomer – it is not a model but a full-blown theory which allows us to compute the result of subatomic processes with high precision

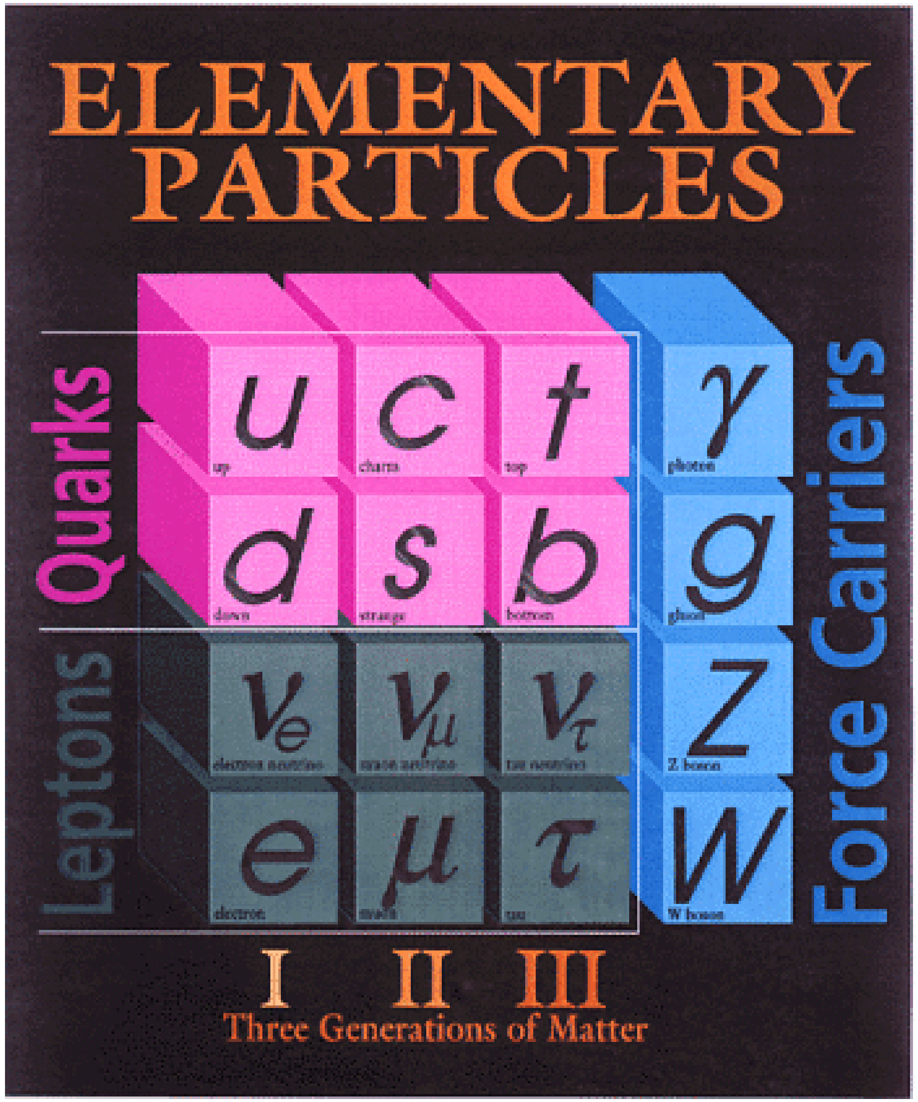
Three families of **quarks**, and three families of **leptons**, are the matter constituents

Strong interactions between quarks are mediated by 8 **gluons**, g

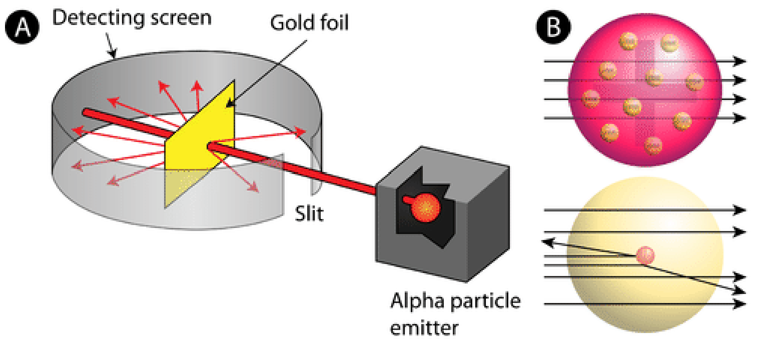
Electromagnetic interactions between charged particles are mediated by the photon, **γ**

The weak force is mediated by **W and Z** bosons

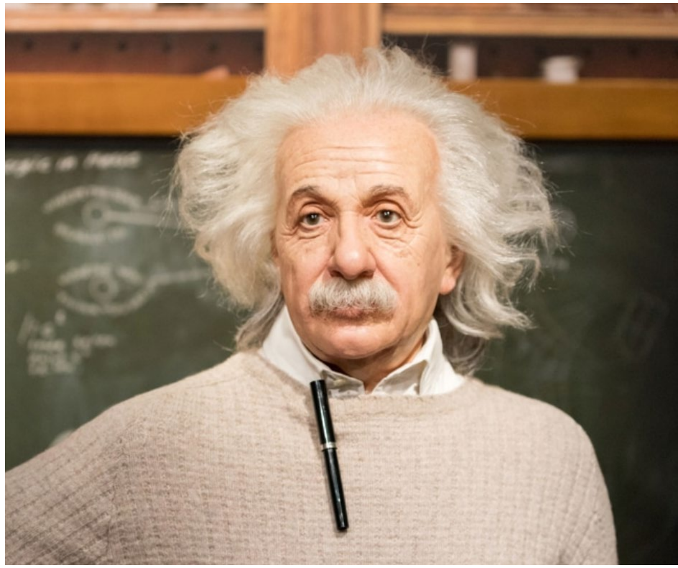
The **Higgs boson** is an additional peculiar particle that gives mathematical consistency to the whole construction



How do we study matter constituents?

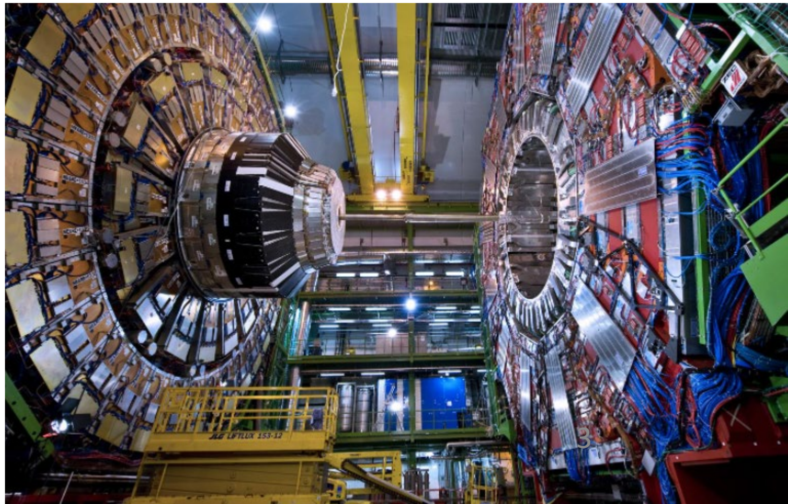


Rutherford, Geiger, Marsden in 1909 discovered the atomic nucleus by studying the deflection of charged alpha particles off a thin gold foil



From there on, the paradigm has always been the same: if you can't see inside something with imaging techinques, throw something at it!

A further important notion: **$E=mc^2$** → if you give your projectiles LOTS of kinetic energy, the latter may transform into the mass of new particles, which we can then study

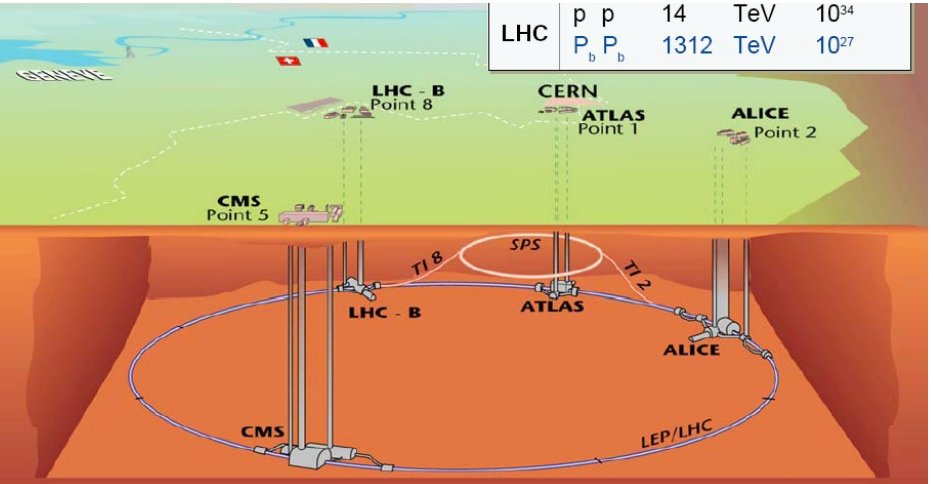


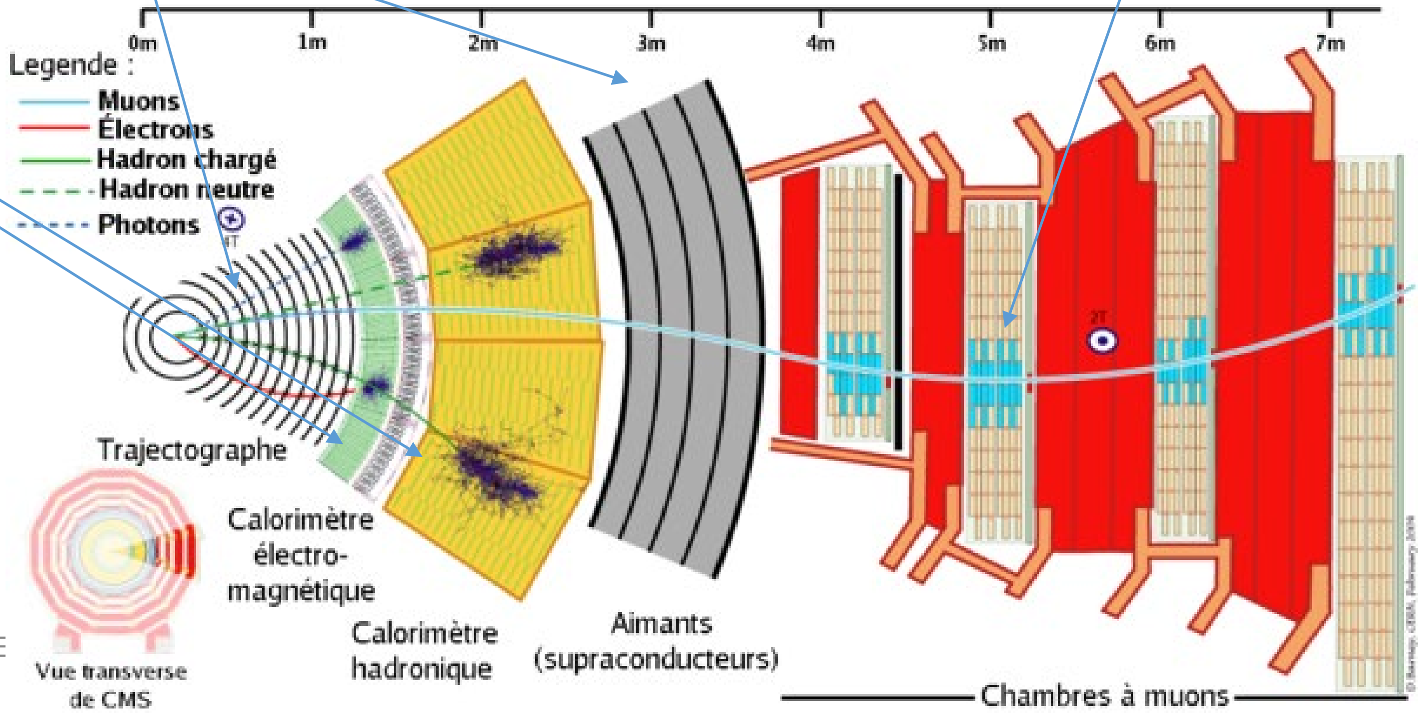
How to do that? By pushing them in particle accelerators, and by detecting them in sensitive devices - particle detectors.

How we detect particles in a collider detector

Charged particles are detected in the **tracker**, through the ionization they leave in silicon; a powerful **magnet** bends their trajectories, allowing a measurement of their momentum Then **calorimeters** destroy both charged and neutral ones, measuring their energy. Muons are the only particles that can traverse the dense material and get tracked outside

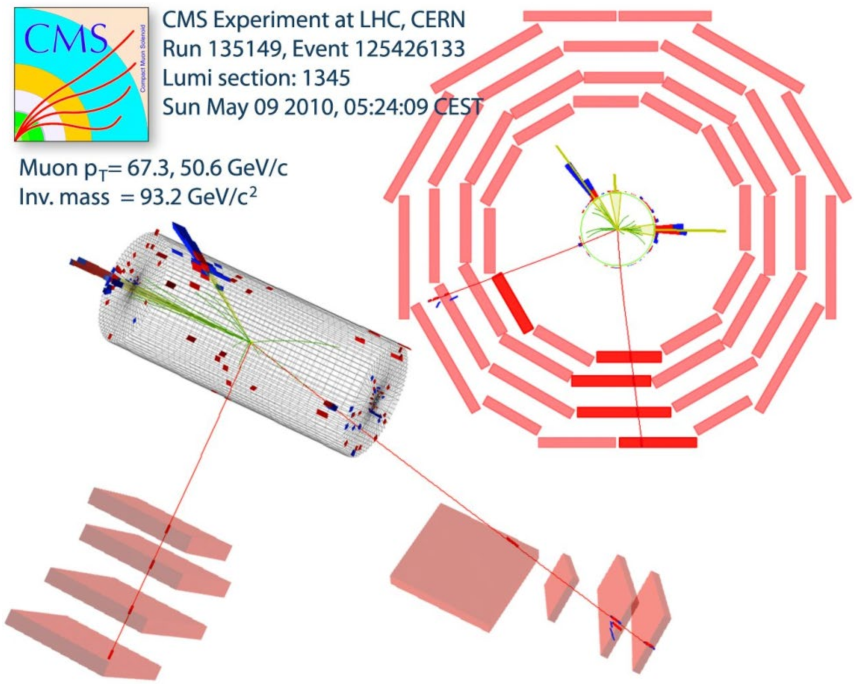
Right: a slice of the CMS detector, as seen from the beam line, shows the different interaction of particles with detector elements
Below: the LHC tunnel runs for 27km below the France-Switzerland border, 100m underground



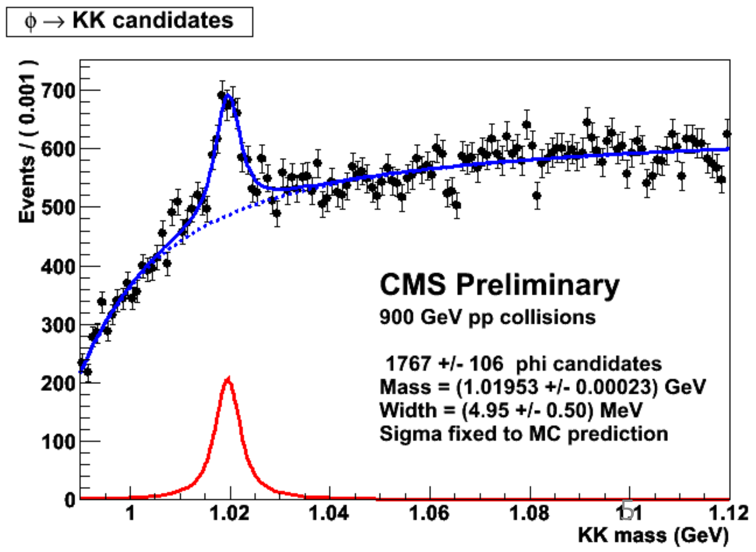


How we see a collision

A reconstruction of the O(10M) electronic signals provides us a «view» of the created objects: using their characteristics we build O(100) high-level variables which we compare to theoretical models after a further compression (usually into a 1-dim test statistic) → then we do measurements and inference



This is a huge dimensionality reduction…

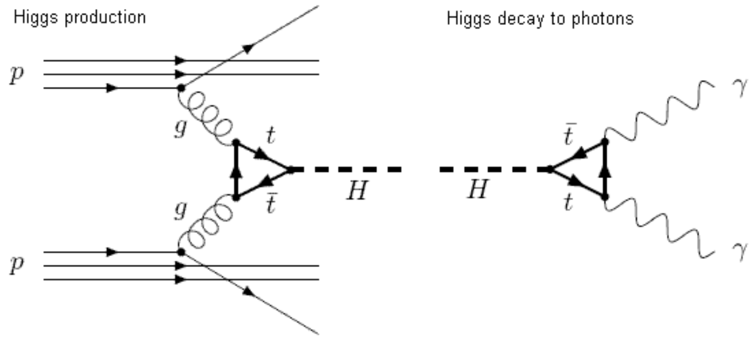


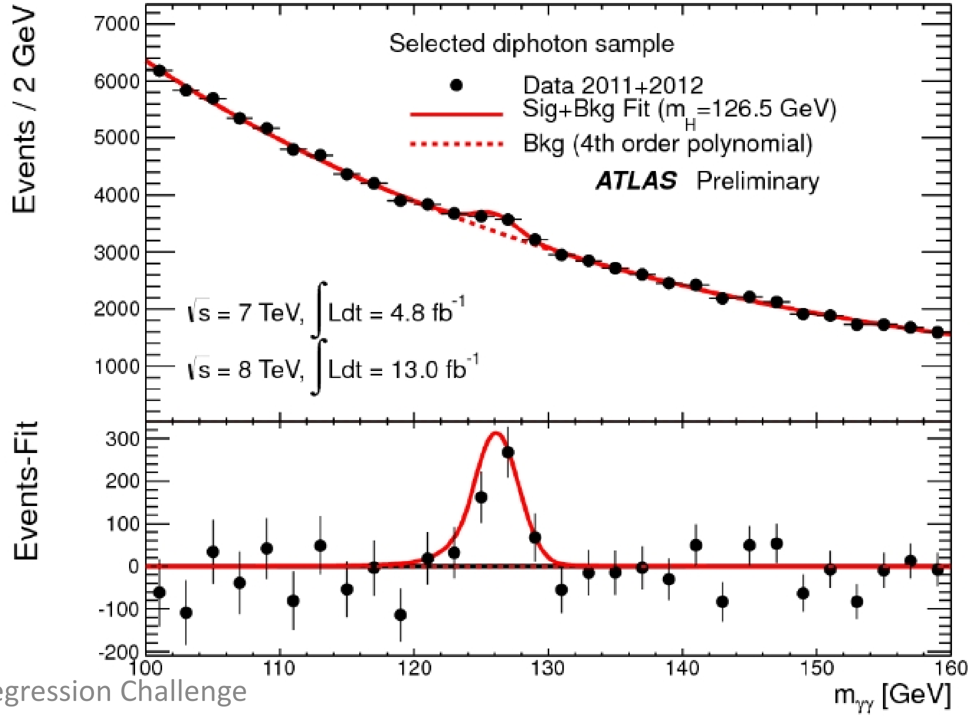
Searches for new particles

The typical search for a new particle involves a model which predicts its existence

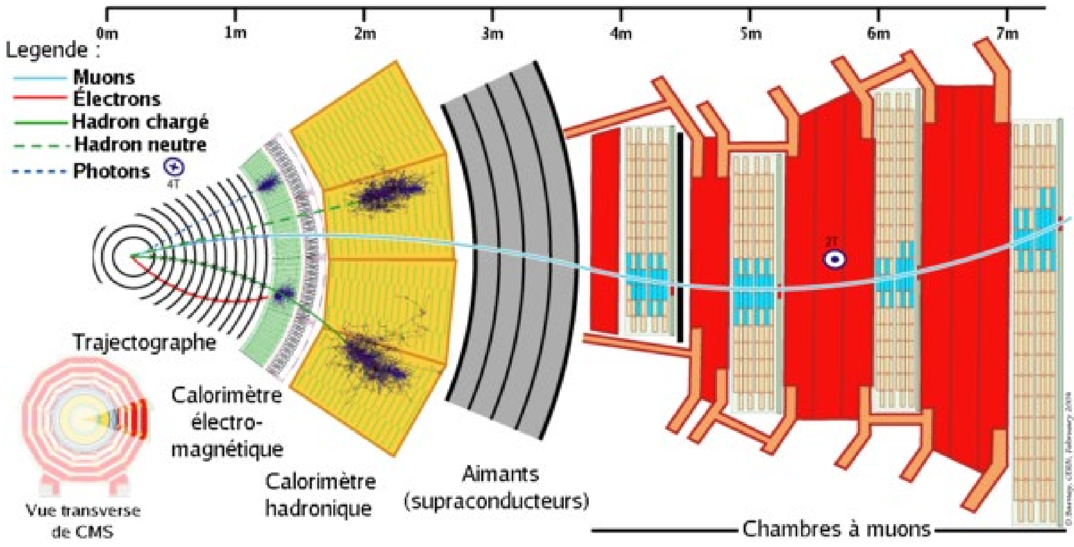
- Monte Carlo generators use the model to produce **simulated datasets** that **teach us how the signal looks like**
- A data selection isolates a sample where we try to evidence the particle (which usually is highly unstable → so we don't see it but its decay products)
- Typically we attempt to reconstruct the **particle mass** from the measured features of its decay products. As mass is a fixed attribute, a histogram may betray the particle via a narrow bump on a smooth background

- **The resolution on the mass, itself a function of the resolution on the energies of the disintegration products, is the crucial ingredient**





Meet the Muon



The muon is a heavier partner (207 times) of the electron. It does not exist in ordinary matter, but can be produced in the decay of heavier unstable particles

Its peculiar properties make it highly distinguishable in a particle detector; it is an excellent, sensitive probe of new particle decays.
We measure muon energy from the curvature of their track in a magnetic field B: if B is in Tesla, R in meters, and E (>>m) is measured in GeV, we have

E = 0.3 B R

However, B cannot exceed a few Tesla, and the particle path a few meters.
For a E=1000 GeV muon, the radius of curvature in a B=2T field is R=1666m, so it is hard to distinguish from a straight line!
→ Curvature-based energy measurement breaks down at high energy

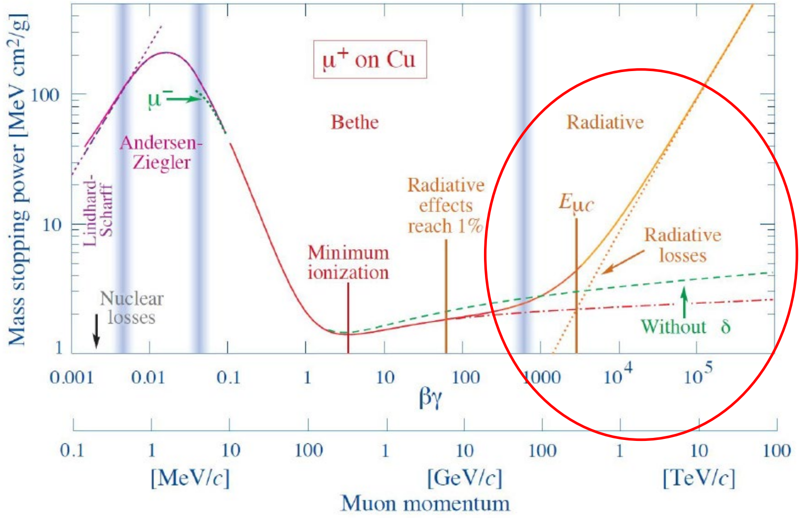
How to measure multi-TeV muons?

In future colliders, whose energy will possibly create new particles of multi-TeV mass, retaining sensitivity to muon energy is a critical ingredient.

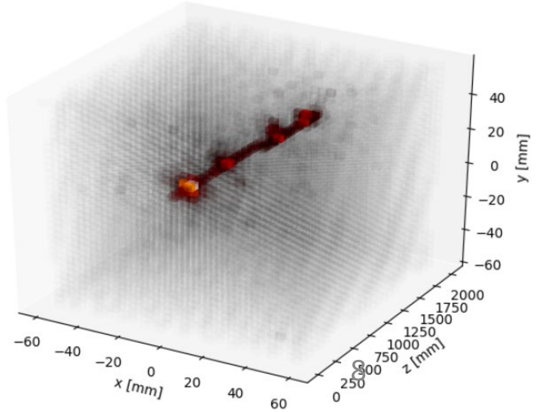
We propose to measure the energy of muons by looking at their radiative energy losses in a dense, granular, homogeneous device: a highly-segmented calorimeter

Muons lose **little** energy in thick media, and the process is stochastic → not easy to infer total energy from radiated energy.

But... if we combine the total energy loss with **spatial information** from the granular picture of their deposits, maybe we can?



Above: energy loss of muons versus their energy; below: energy deposition in a high-granularity calorimeter



Why should spatial information help?

Besides ionizing the medium, muons lose energy mainly through the process of bremsstrahlung: energetic photons are released when muons cross intense electric fields of the atoms

Photons travel in the calorimeter cells and, if they have energy above a few MeV, produce electron-positron pairs, which in turn radiate further photons → this may give rise to localized «clusters» of energy depositions

Softer photons undergo other processes (e.g. Compton scattering) which also allow their energy to be detected in the calorimeter cells

There is no map between the energy of the primary photon and the muon energy, but constructing variables sensitive to these photon energies may inform the regressor

Inferring the curvature from calorimeter cells

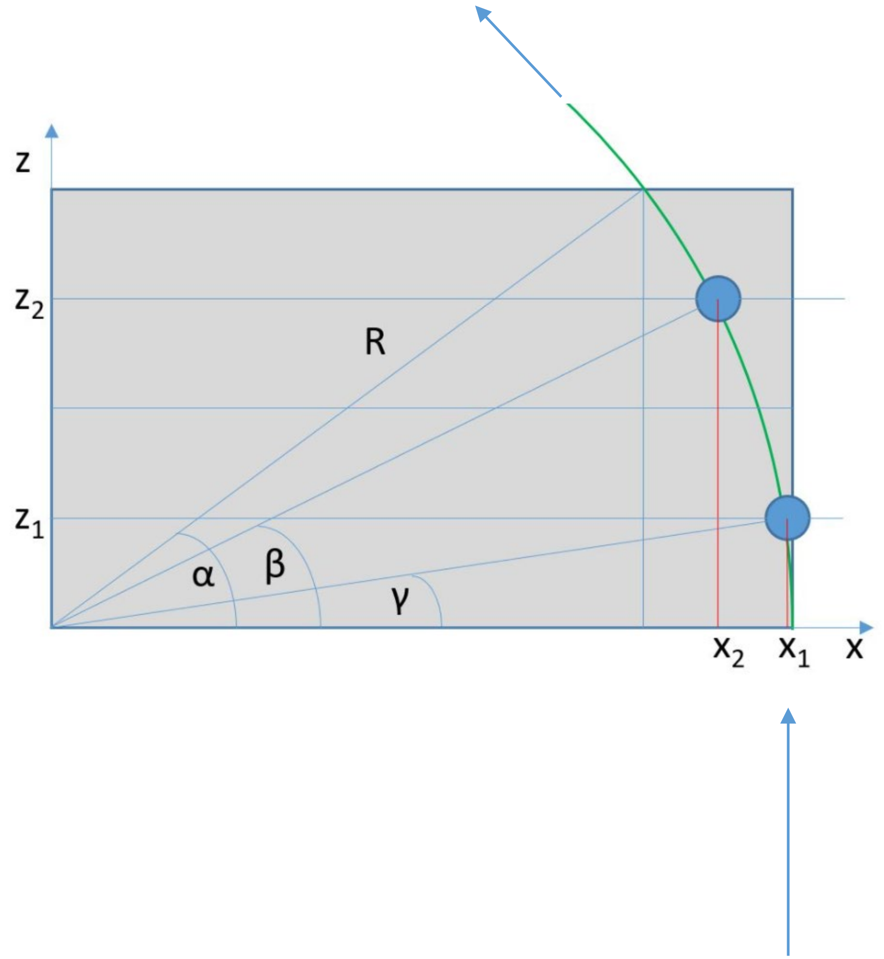
One further idea is to do what the tracking measurement does (with higher precision) using the energy deposits

If the muon energy is <1 TeV, its trajectory «bends» by a few millimeters while traversing the 2 meters → also energy deposits (from ionization, along the muon track, and from bremsstrahlung) should exibit the same pattern

Due to the sparsity of cells with recorded energy, one can aggregate the information, by e.g. constructing two centers of gravity of energy deposits at different calorimeter depths, and then solving the equation of a curve with tangent orthogonal to the front face of the calorimeter

An explanation is provided in this blog post:

https://www.science20.com/tommaso_dorigo/challenge_measure_muons_energy_with_highschool_math_and_win_a_mug-253021



The data

The data used for the hackathon are produced by an accurate simulation (GEANT4) of muons traversing a 2.032m-long, finely segmented calorimeter, subjected to a 2 Tesla magnetic field oriented along y. Muons are generated at the front face of the calorimeter (z=0) and they traverse the 2.032m, exiting on the back end.

In each of 51,200 cells (in a grid of 32 in x, 32 in y, and 50 in z) is recorded, for each muon, the energy deposited by radiative processes, in GeV. For each muon of the training dataset is also given the true energy, which we must regress to using the (sparse) 51,200 (x,y,z,E) vectors.

Training-set muons have energies in the [50,8000] GeV range,

The figure of merit

How well can we regress the energy of muons? This is not a sufficiently quantitative question. We must specify what will be the metric of evaluation.

Since the resolution on the energy given by a curvature measurement in a tracking detector can be shown to be of the form $\boldsymbol{\sigma_{E,tracker} = k\, E}$ (linear, due to linearity of curvature with energy), we wish to combine our determination with an assumed tracking measurement in a 2T field, given by

$$\sigma_{E,tracker}=0.2E \qquad (== 200\ \mathrm{GeV\ at\ 1\ TeV\ Energy})$$

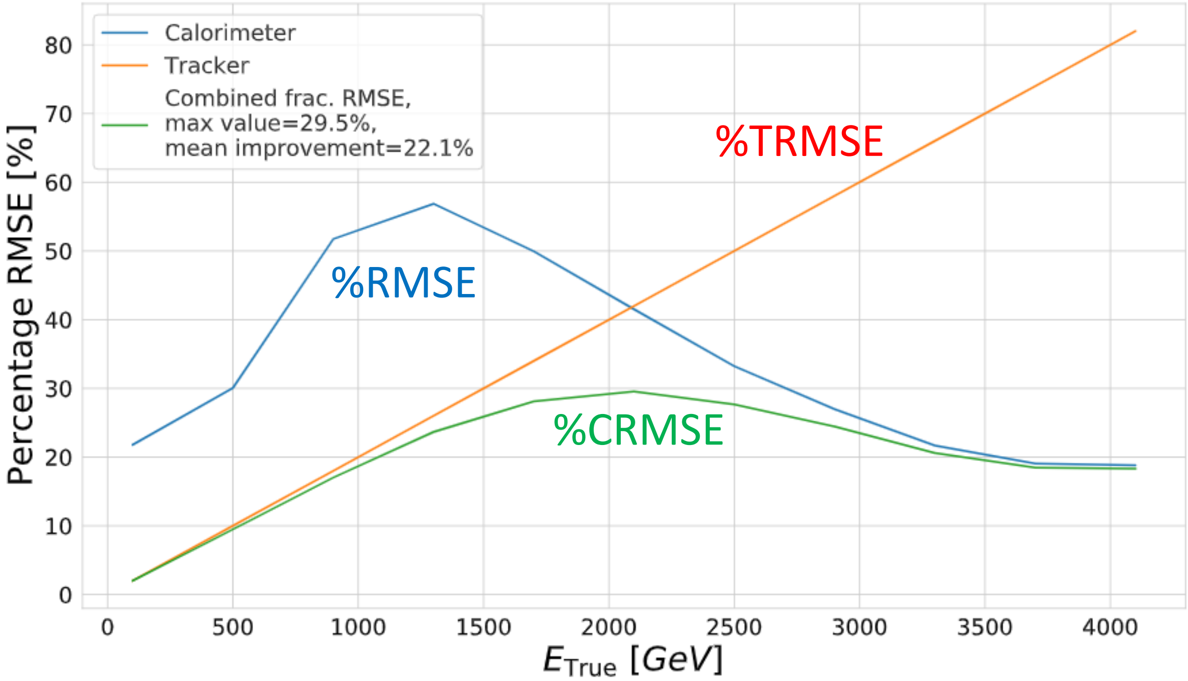
If you obtain an independent measurement of energy with a resolution $\boldsymbol{\sigma_{E,calo}}$, the combination (simple weighted average) has a global resolution of

$$\sigma_{comb} = [1/(\sigma_{E,tracker})^2+1/(\sigma_{E,calo})^2]^{-1/2}$$

We will determine the resulting resolution $\boldsymbol{\sigma_{E,calo}}$ of our regressor, and then combine it with the fixed tracker resolution. An example is shown in the next slide, where our final figure of merit is clarified.

Area between line and curve

Since our task is to improve over $\sigma_{E,tracker}$, our figure of merit is defined as the integral of the area between the tracker resolution curve and the combined resolution, in the [50,8000] GeV interval of true muon energies



On the y axis is shown the relative resolution on muon energy, defined as the mean squared error divided by true energy,

$$\textbf{\%RMSE} = (\text{bias}(E_{true})^2 + \text{variance}(E_{true})]^{1/2} / E_{true}$$

Area metric / 2

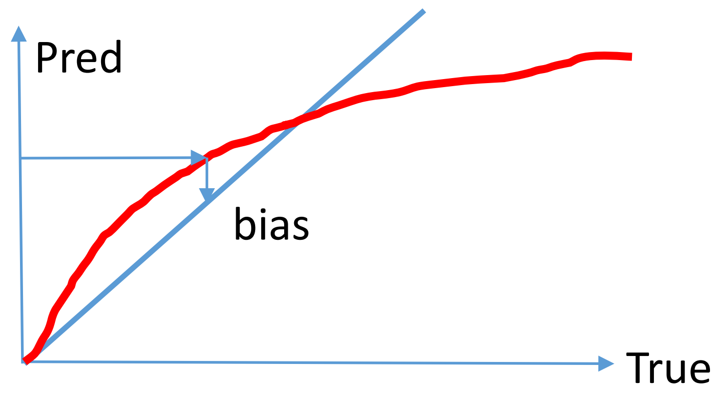
So we need to calculate the %RMSE for our model, then combine with the tracker resolution %TRMSE, obtaining a %CRMSE as a function of energy in TeV, and finally integrate the difference over the tracker-only %TRMSE=0.2 E /TeV (=20% at 1 TeV):

$$FOM = \frac{\int_{0.05}^{8}[0.2E - \%CRMSE]dE}{\int_{0.05}^{8} dE}$$

- As this is a figure of merit, your goal should be to **max**imize it!

We provide code to compute the metric given the predictions and true energy

One detail you will probably find useful: as the energy deposition grows logarithmically with energy, all regression outputs tend to display a curved shape of Pred(True) vs True energy
→ you may want to apply a final post-regression bias correction, by extracting on training data the dependence of True(Pred) and use it as a bias term



Submitting your solution

You should process the test data (which do not have the true energy per each muon) and produce a .csv file with one line per muon, containing two numbers:

event number, predicted energy

The file, named after you (e.g. kuusela.csv) should be emailed to

giles.strong@outlook.com

by 10.30 ET on Friday,

giving us time to score and rank the solutions for the presentations.

Example notebooks

→ go to Jupyter and load Lukas' examples

Supplementary material

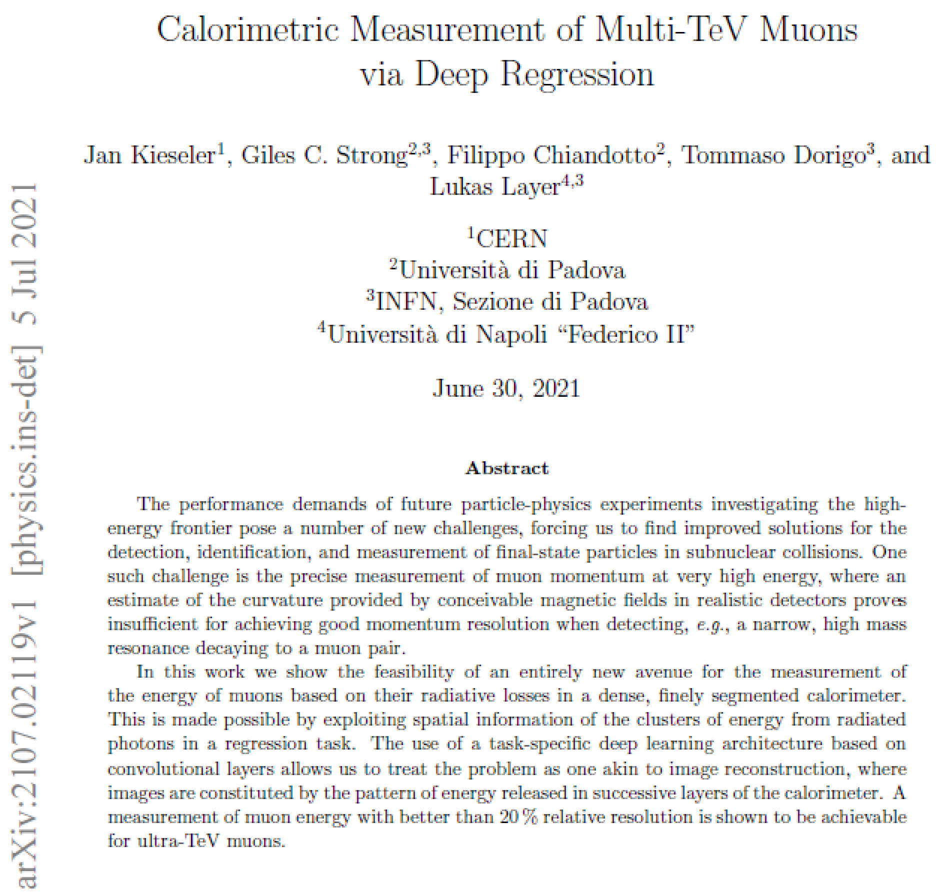
A preprint describing a CNN-based regression of the data you are provided with is available in https://arxiv.org/abs/2107.02119

The article also describes how one could construct some high-level features from the (x,y,z,E) vectors of each muon. However, you may be able to construct better ones than those suggested there!

The figure of merit we have constructed in the previous pages is complicated, so it is not completely straightforward to encode it in a loss function. You may start with the %MFSE,

$$L = \frac{1}{N}\sum_{i=1}^{N}\frac{(T_i - P_i)^2}{T_i}$$

where T is the true and P is the predicted energy.



Calorimetric Measurement of Multi-TeV Muons via Deep Regression

Jan Kieseler[1], Giles C. Strong[2,3], Filippo Chiandotto[2], Tommaso Dorigo[3], and Lukas Layer[4,3]

[1]CERN
[2]Università di Padova
[3]INFN, Sezione di Padova
[4]Università di Napoli "Federico II"

June 30, 2021

Abstract

The performance demands of future particle-physics experiments investigating the high-energy frontier pose a number of new challenges, forcing us to find improved solutions for the detection, identification, and measurement of final-state particles in subnuclear collisions. One such challenge is the precise measurement of muon momentum at very high energy, where an estimate of the curvature provided by conceivable magnetic fields in realistic detectors proves insufficient for achieving good momentum resolution when detecting, *e.g.*, a narrow, high mass resonance decaying to a muon pair.

In this work we show the feasibility of an entirely new avenue for the measurement of the energy of muons based on their radiative losses in a dense, finely segmented calorimeter. This is made possible by exploiting spatial information of the clusters of energy from radiated photons in a regression task. The use of a task-specific deep learning architecture based on convolutional layers allows us to treat the problem as one akin to image reconstruction, where images are constituted by the pattern of energy released in successive layers of the calorimeter. A measurement of muon energy with better than 20 % relative resolution is shown to be achievable for ultra-TeV muons.