Introduction To Experimental Particle Physics

Maurizio Pierini

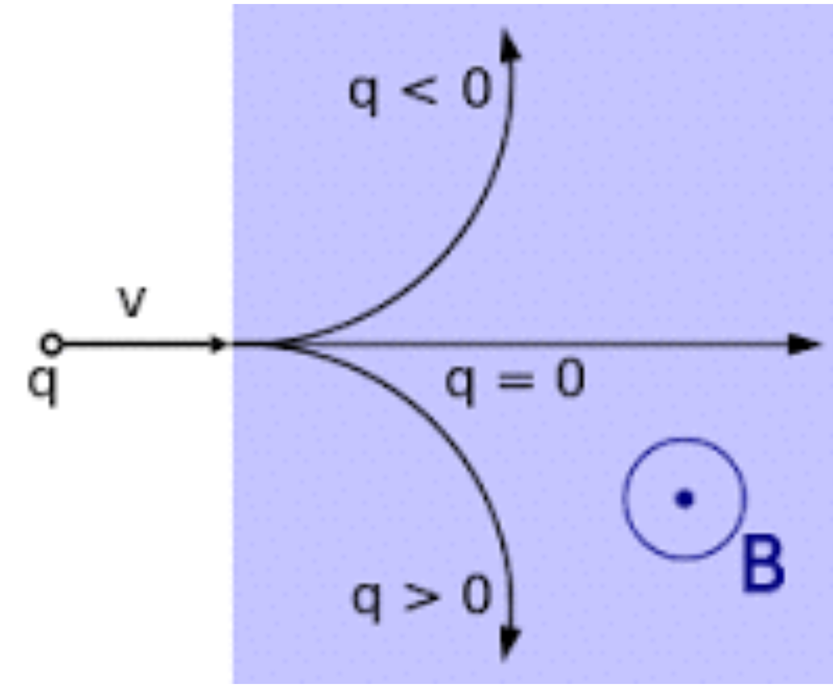
CERN

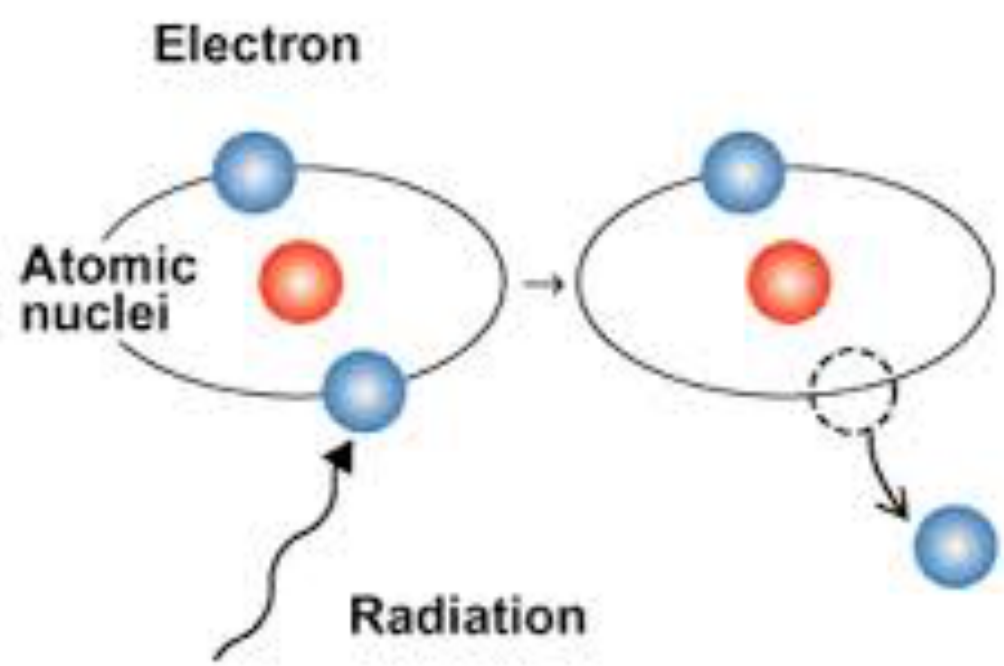
Outline

- Some History
- Particle Physics in a nutshell
- Colliders and Big Data
- The LHC and the LHC experiments
- How we take data
- What we do with our data
- An outlook to future

A little bit of History

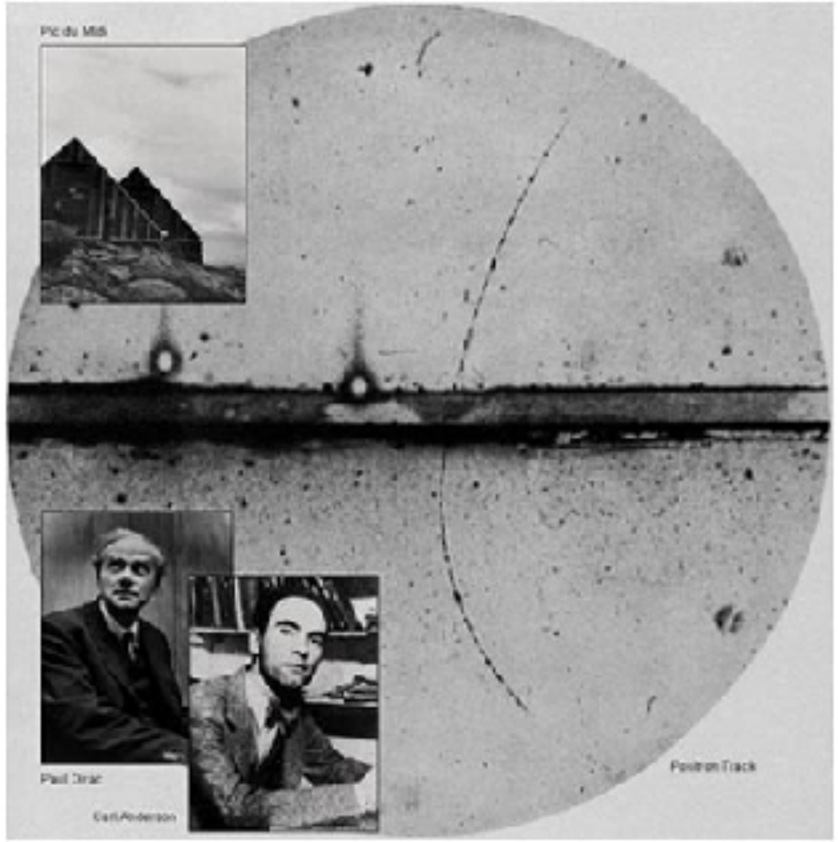
- Particle Physics started with detectors, exploiting interaction between particles and environment
  - Charged particles bend under magnetic field (Lorentz force)
  - Particles ionize material hitting atoms (e.g., the silicon in the camera of your phone)

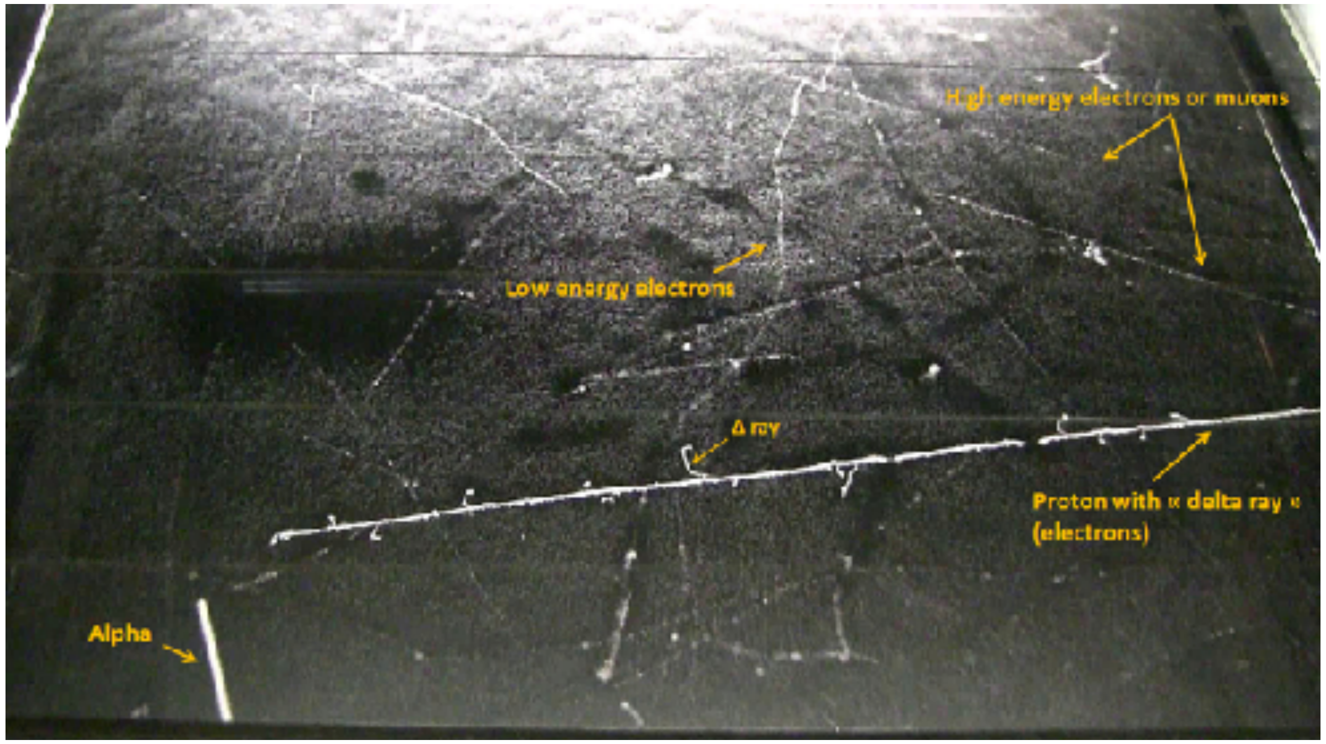




A little bit of History

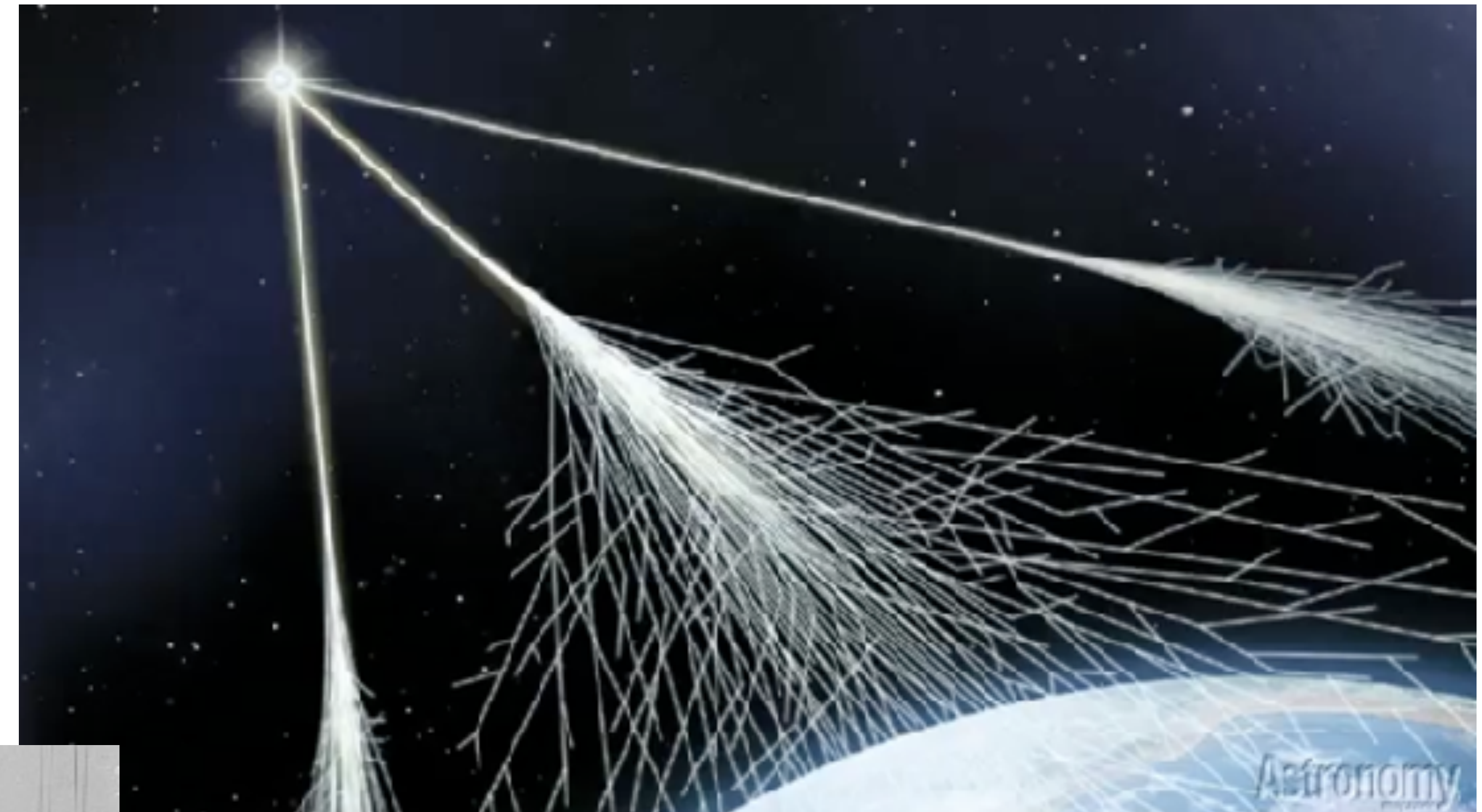
- The first detectors where chambers of low-pressure gas
  - Particles crossing ionize the gas and makes bubbles
  - Looking at pictures (one by one) and connecting the dots, humans could track particles





A little bit of History

- The first source of particles came from the sky: cosmic rays

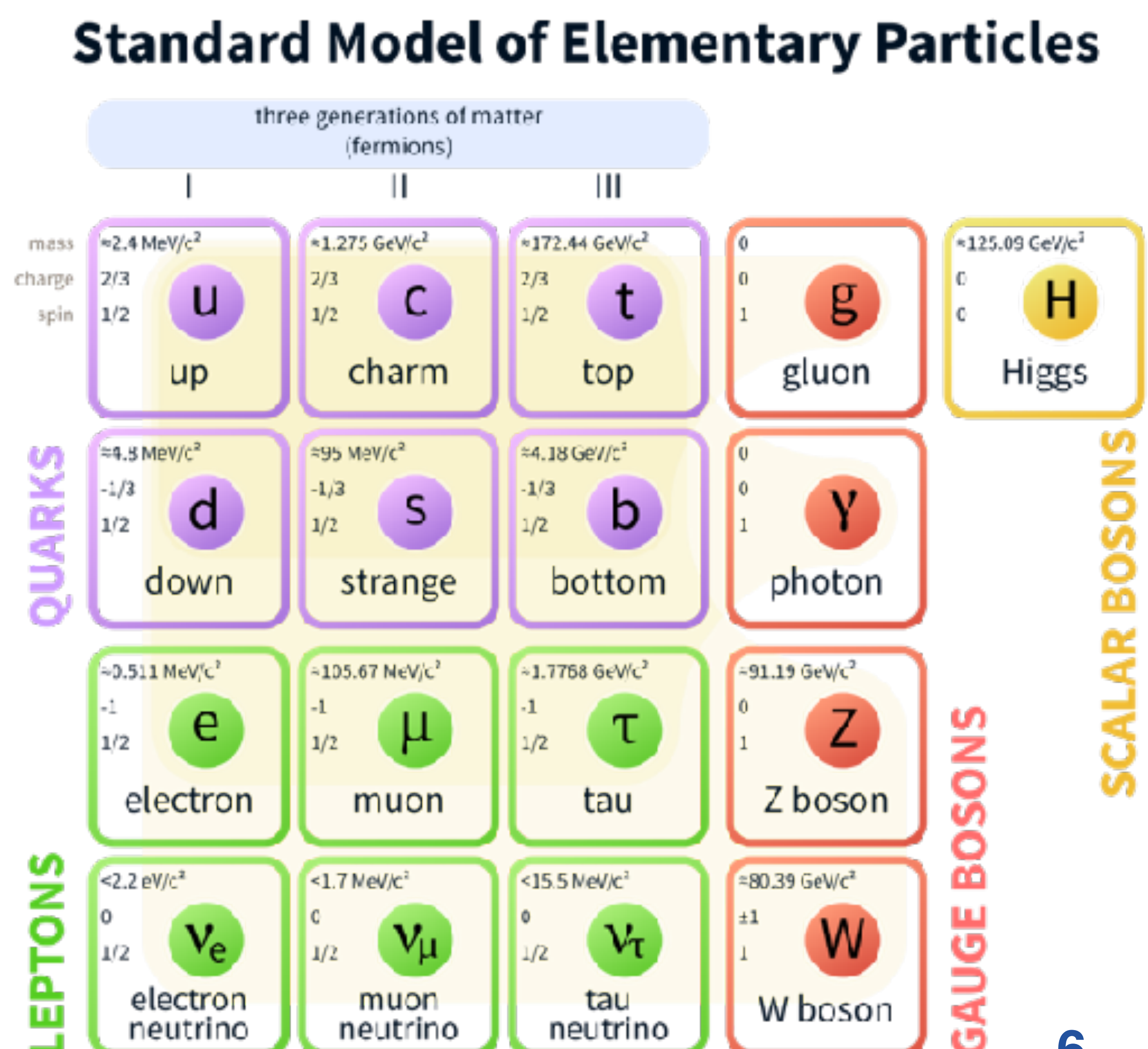


- More recently, images produced colliding a particle beam against a metal target
- Analysis performed by visual inspection

Modern Particle Physics

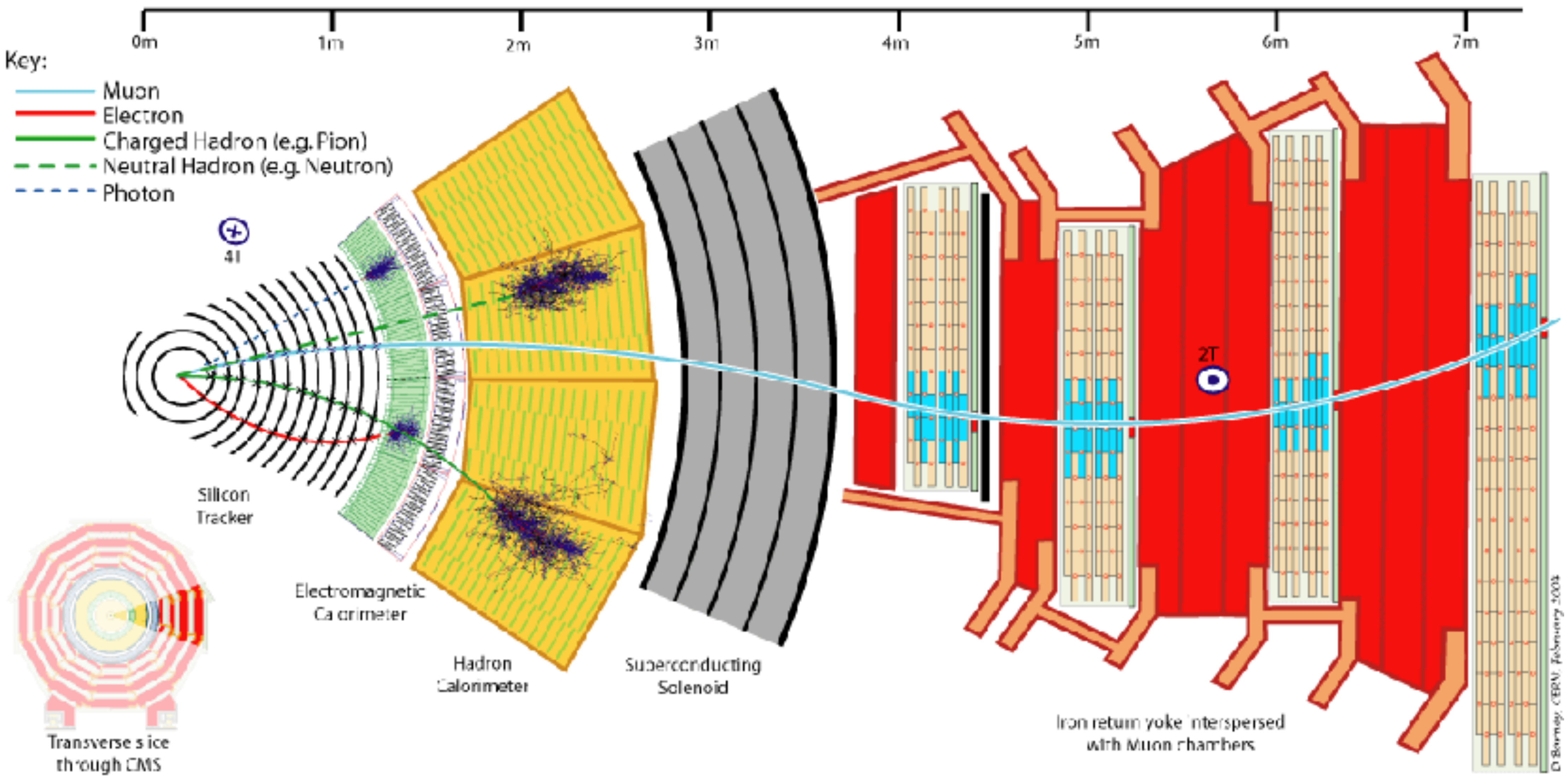
- These pioneer studies allowed many discoveries
  - antimatter
  - new heavier particles (muons, taus, etc)
  - new lighter particles (quarks inside the protons)

- More precise studies made possible with first particle colliders
  - controlled environment in laboratory
  - a lot of data, at tunable energy and intensity
- It was not possible anymore to perform single experiments by individuals



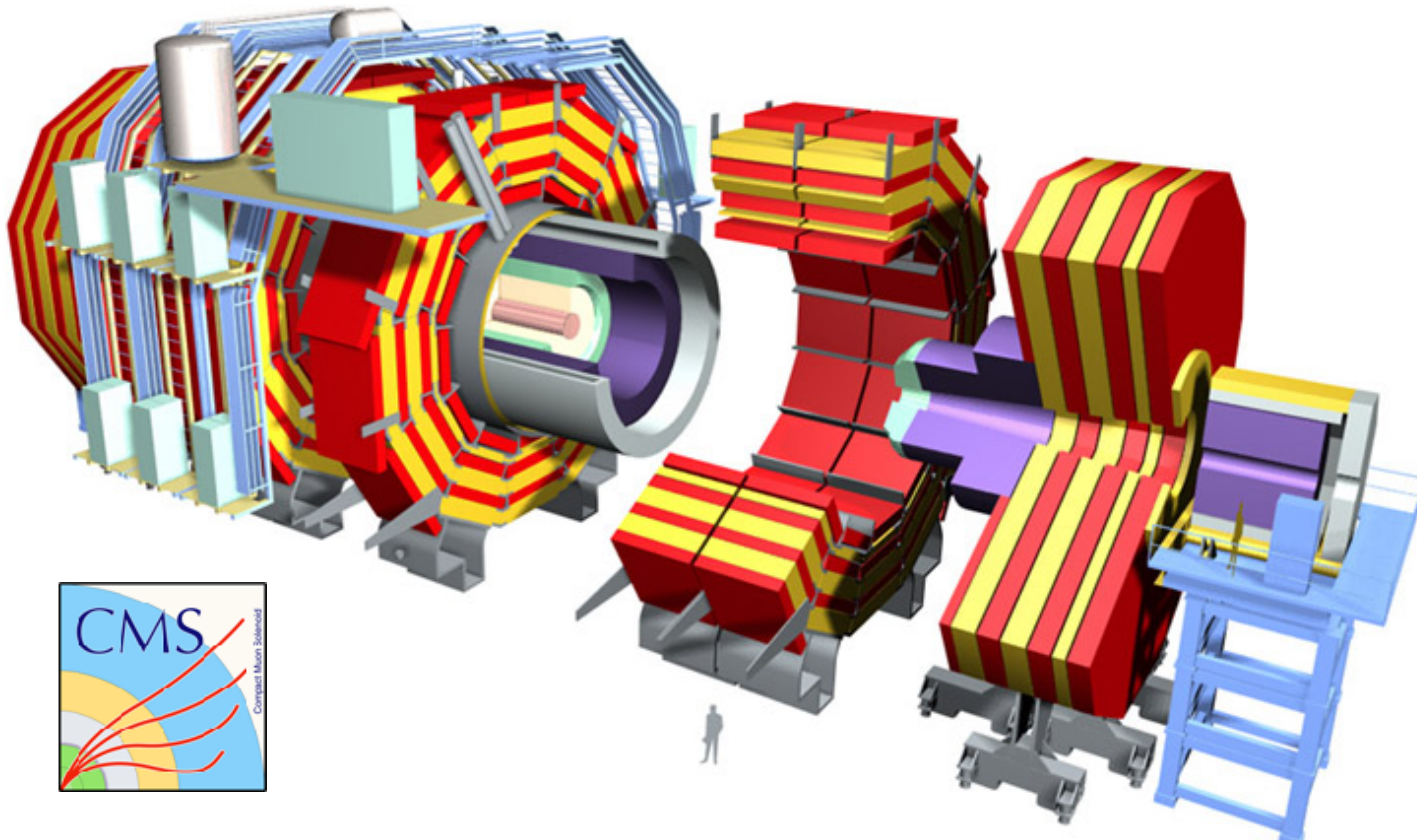
Multipurpose detectors

- With so many particles produced, one need a detector capable of seeing all of them
- Each particle has specific behaviour and needs a specific strategy to be detected



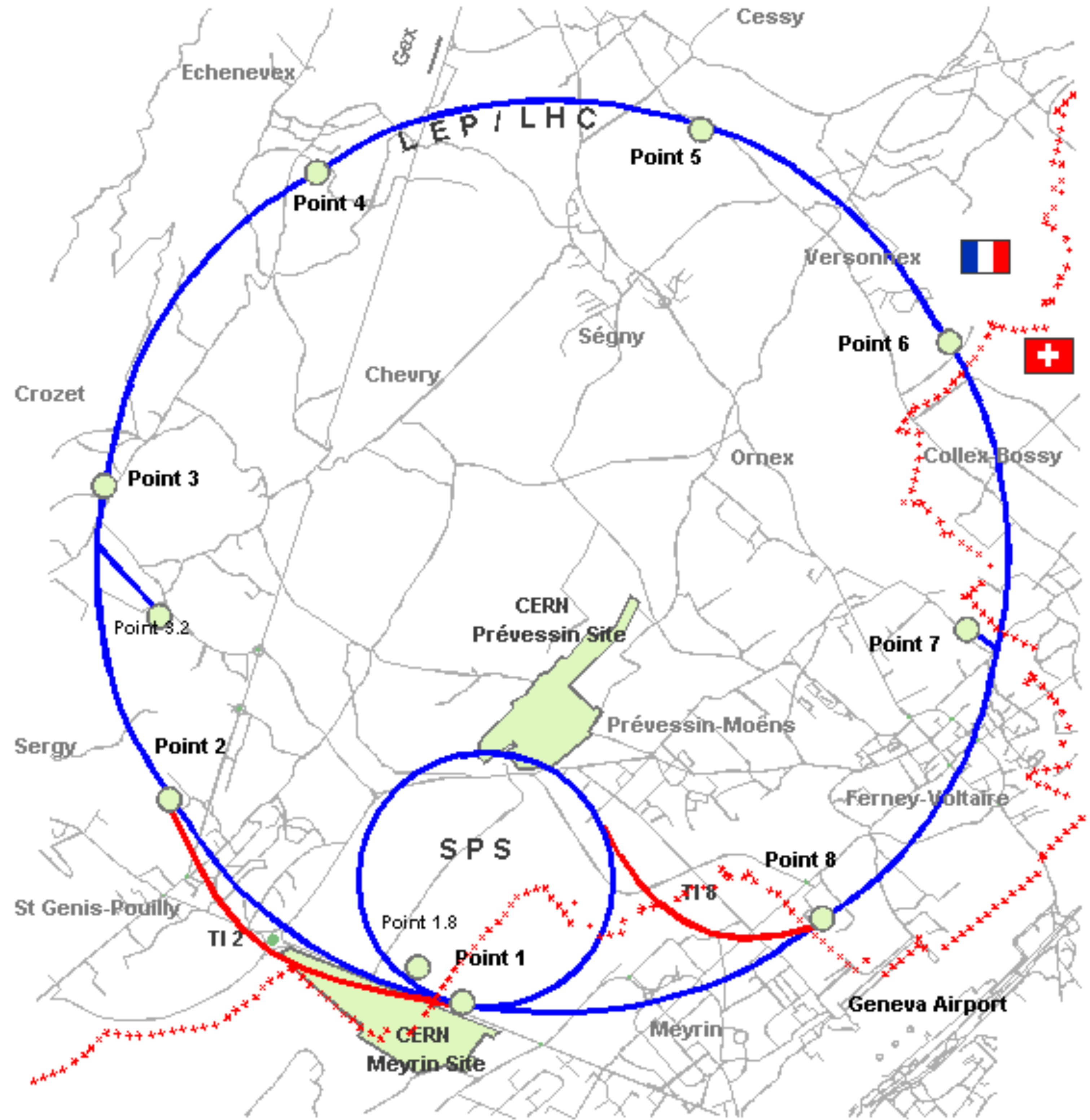
Hermetic detectors

- Particle colliders create collisions in one point
- Particles go everywhere. One needs a detector which covers as much as possible the space around the collision



The LHC

- The LHC is a proton collider
- Protons are accelerated to 13 TeV ~ 13000 the equivalent energy of their mass
- Collisions happen every 25 nsec
- Protons break in the collisions, creating heavier particles (proton energy turned into particle mass)



Map of CERN sites and LHC access points

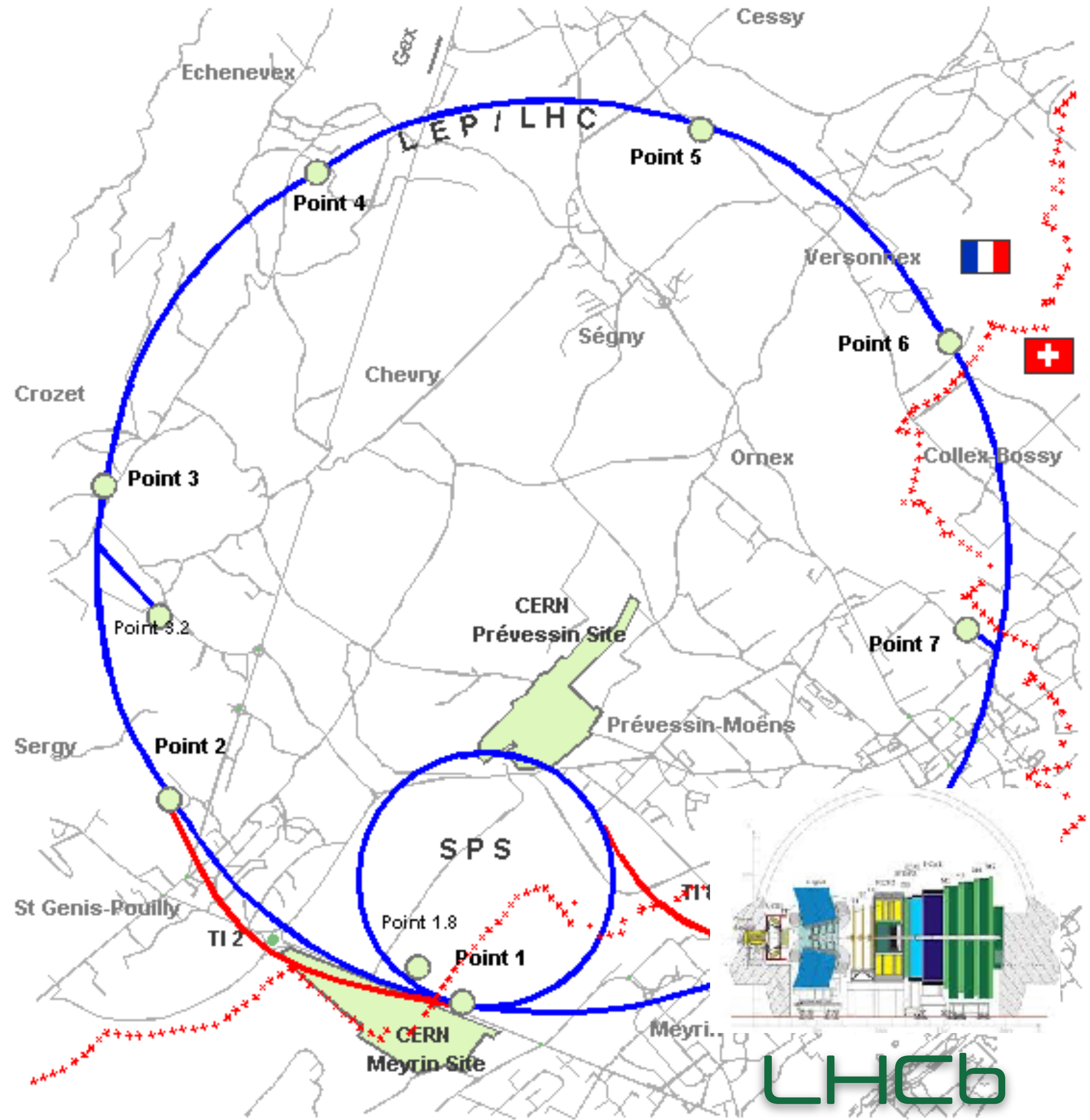
The LHC detectors

- Two multipurpose detectors study many different kinds of collision
  - Higgs boson
  - Weak and strong interactions
  - New physics (dark matter, supersymmetry, etc)

Map of CERN sites and LHC access points

The LHC detectors

- One detector studies the differences between matter and antimatter



Map of CERN sites and LHC access points

The LHC detectors

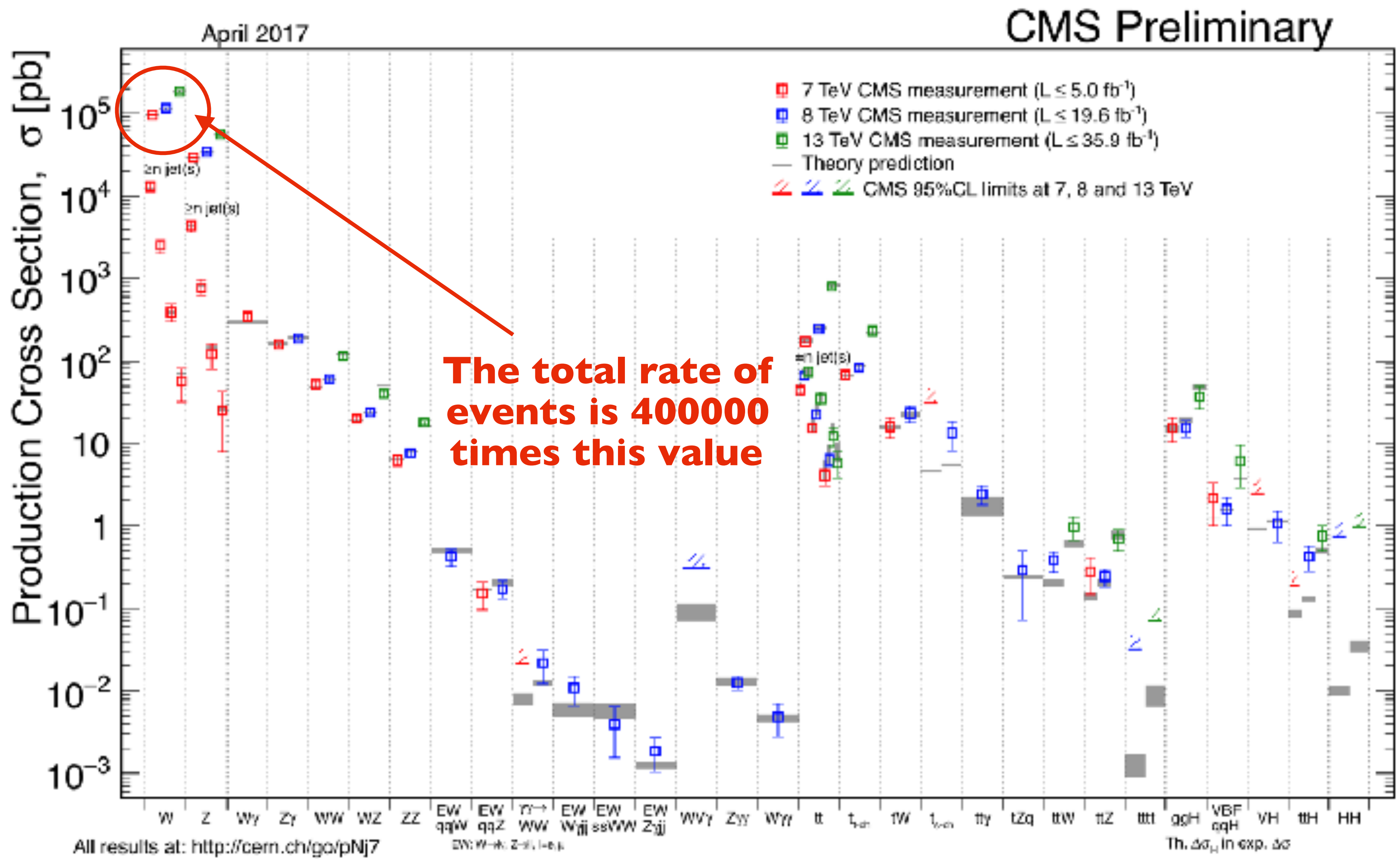
- One detector studies collisions between lead atoms (1 month/ year) and the soup of quark and gluons produced in these collisions

Map of CERN sites and LHC access points

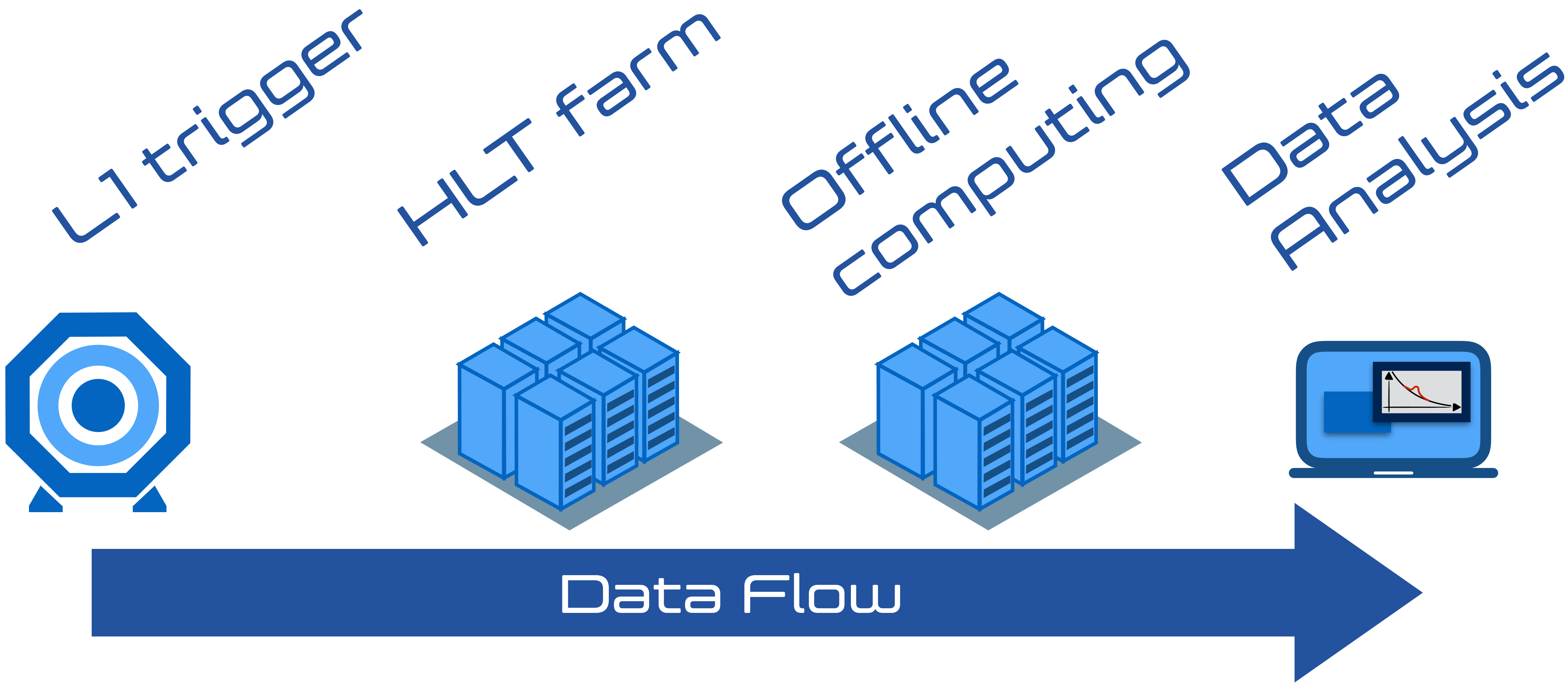
The LHC Big Data problem

- The LHC generates 40 million collisions every second
- The technology to store all these data doesn't exist
- Keeping data costs money (for disk, tape, and CPU for processing)
- We need to select in real time what we want to keep
- Some event is more interesting than other

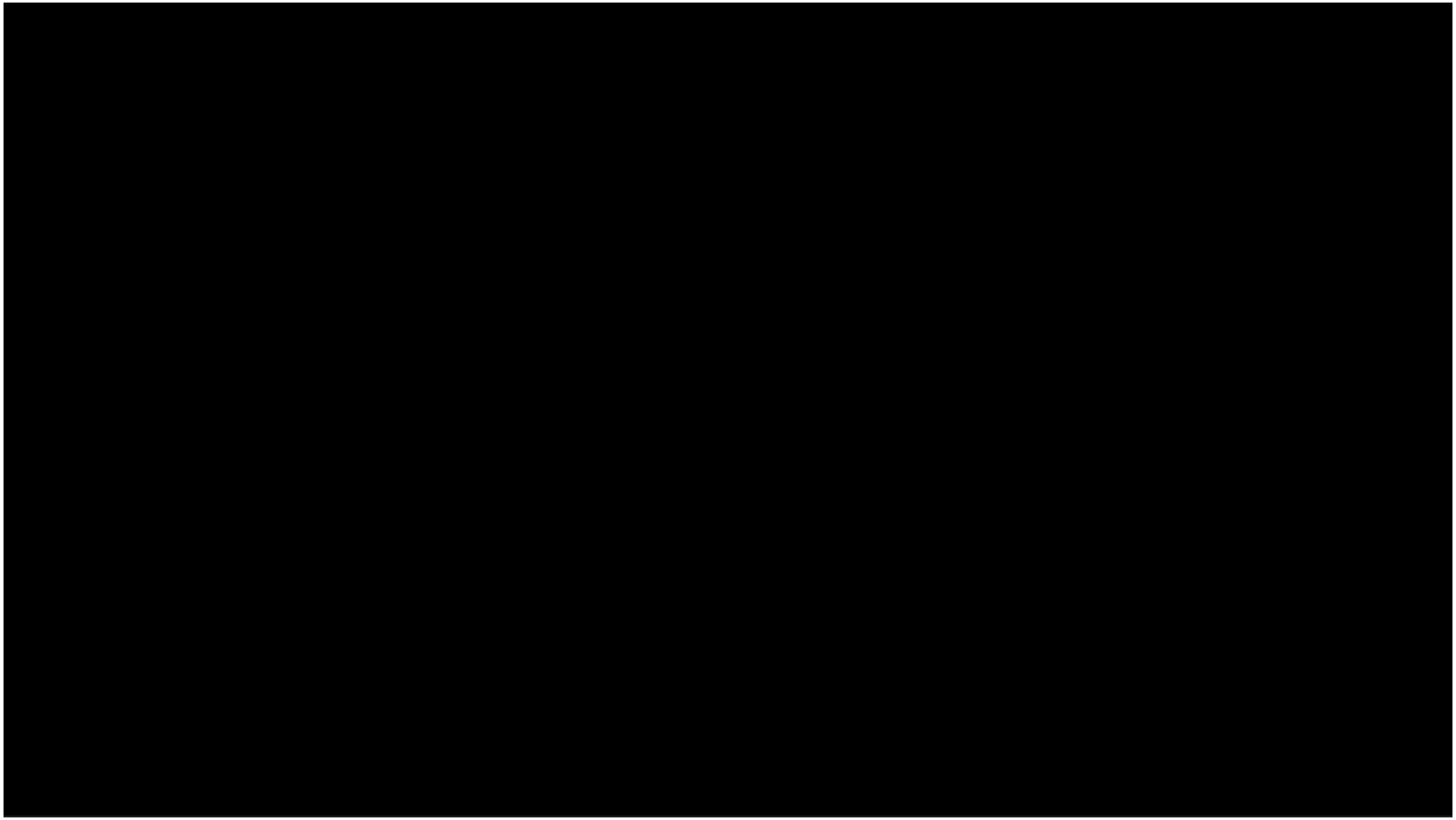
A BIG THEORETICAL BIAS that we have to pay



The LHC Big Data problem

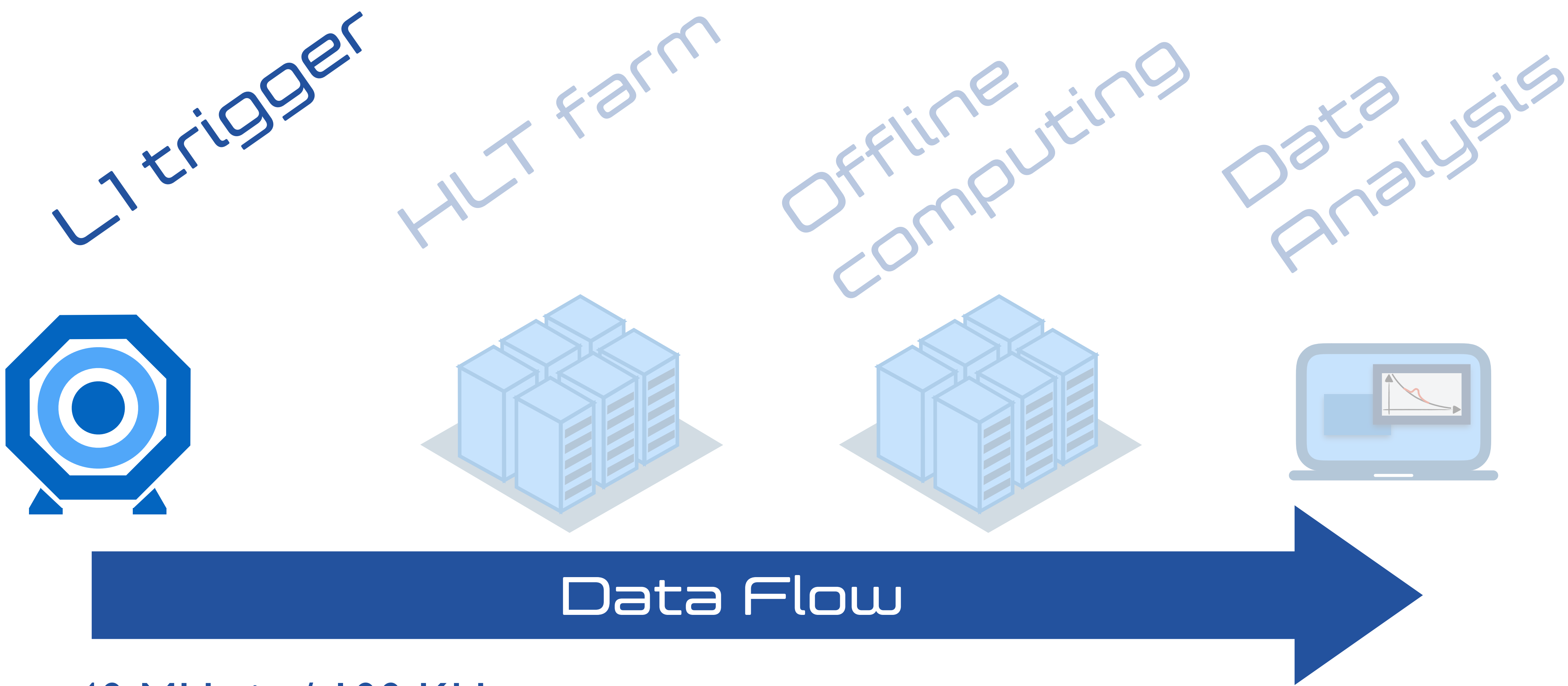


The LHC Big Data problem



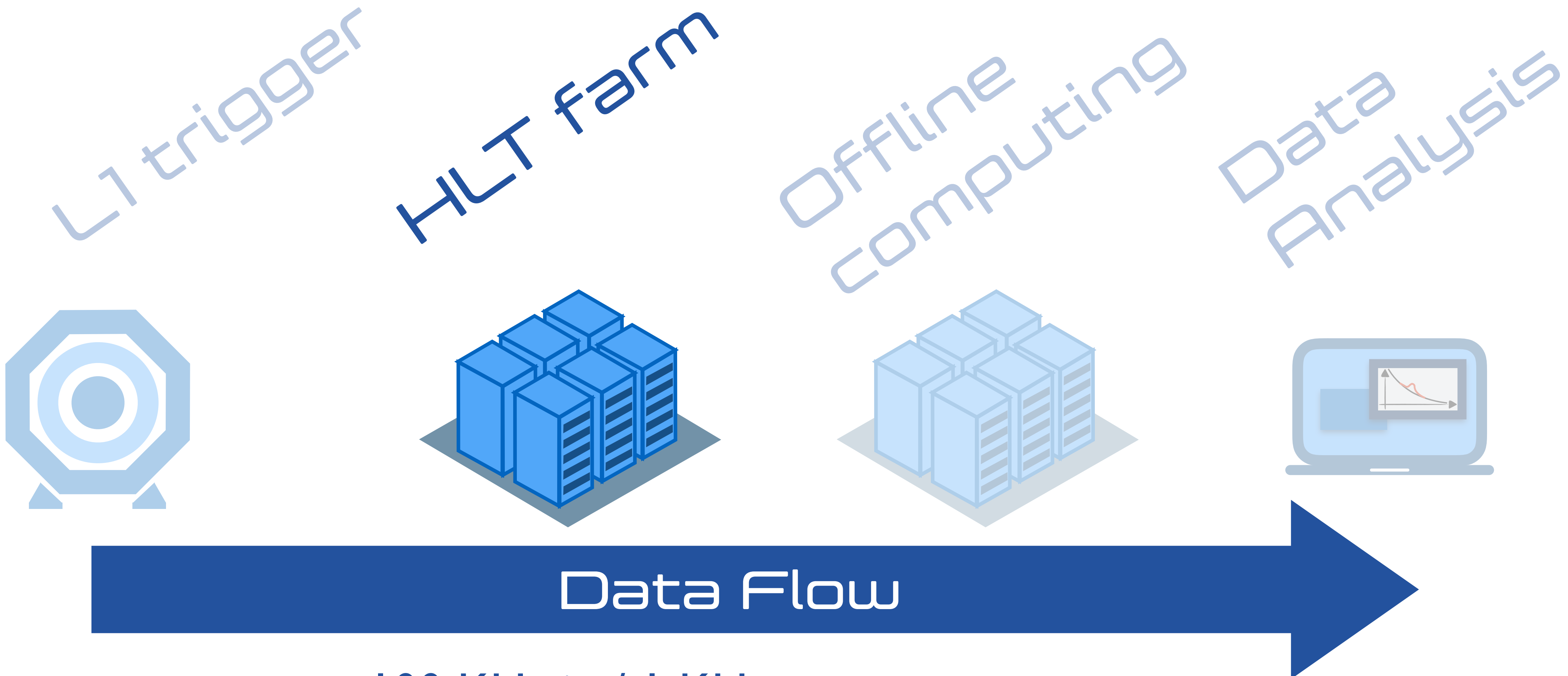
https://www.youtube.com/watch?v=jDC3-QSiLB4

The LHC Big Data problem



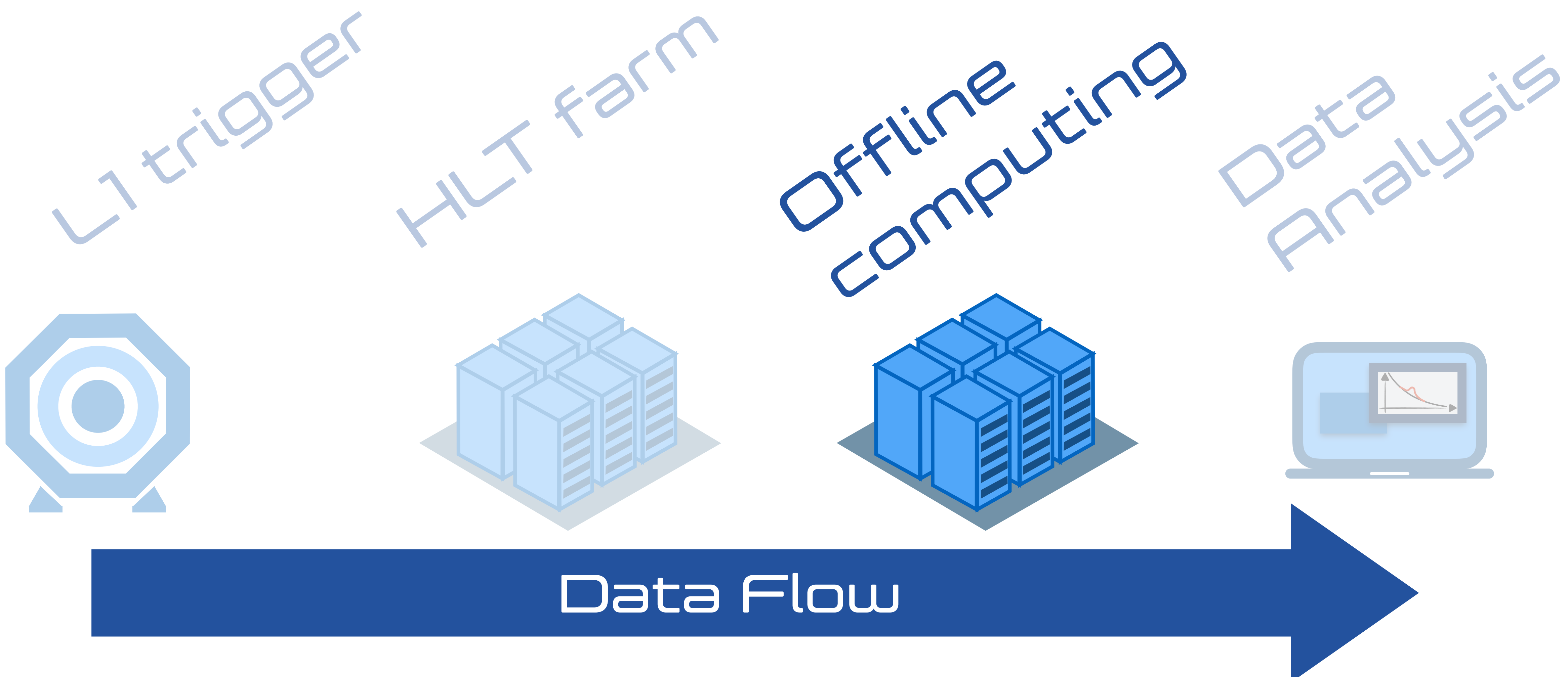
- 40 MHz in / 100 KHz out
- ~ 500 KB / event
- Processing time: ~10 µs
- Based on coarse local reconstructions
- FPGAs / Hardware implemented

The LHC Big Data problem



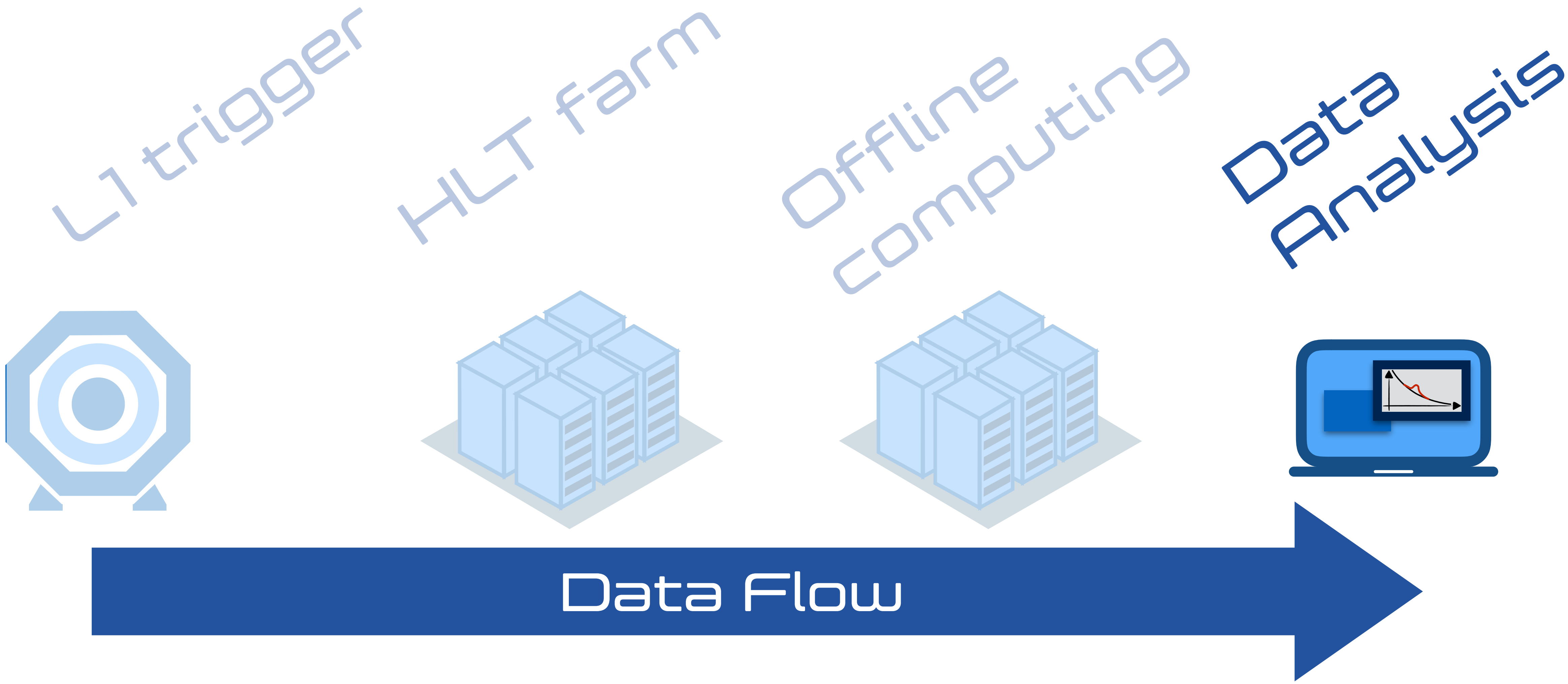
- 100 KHz in / 1 KHz out
- ~ 500 KB / event
- Processing time: ~30 ms
- Based on simplified global reconstructions
- Software implemented on CPUs

The LHC Big Data problem



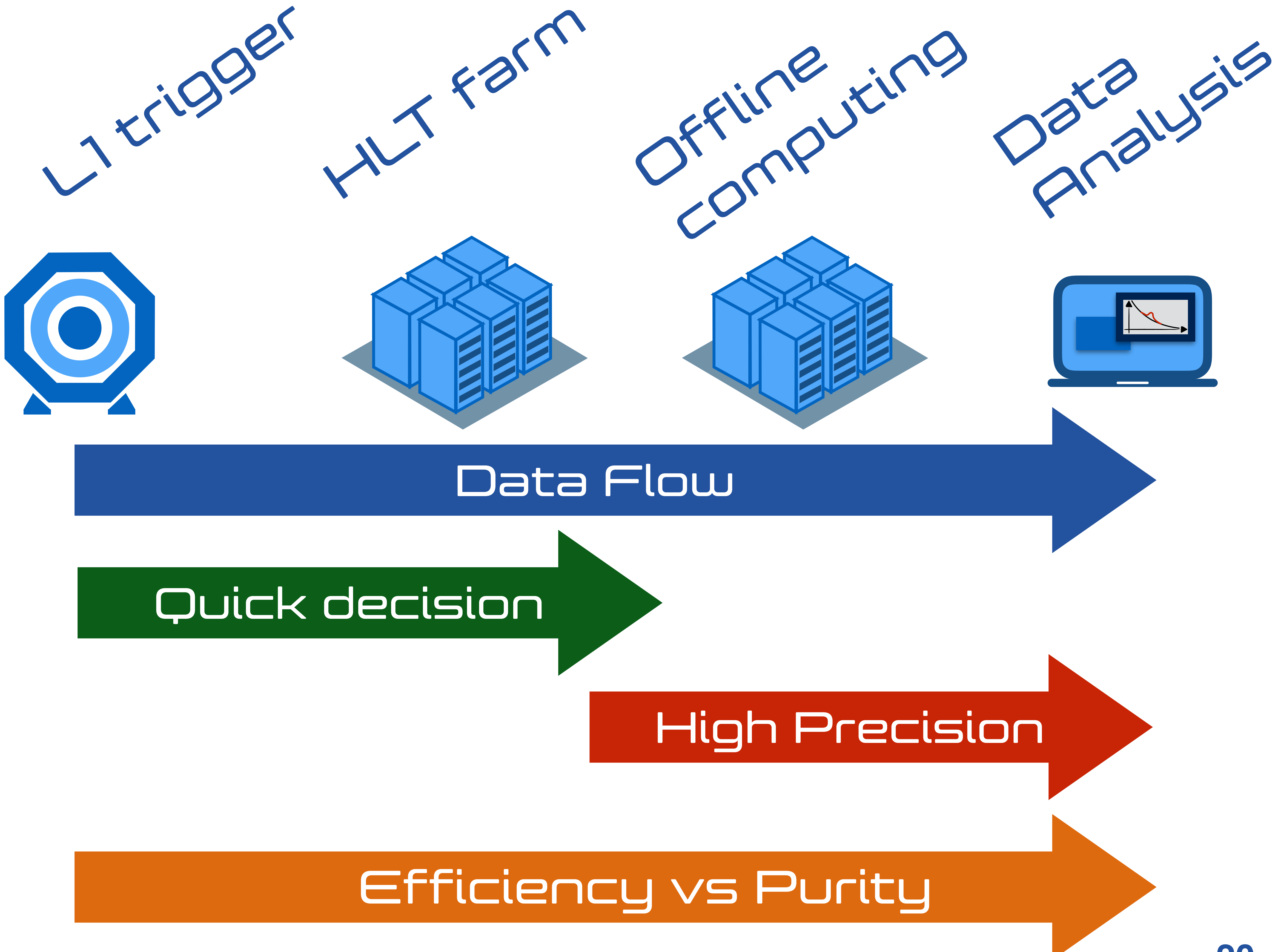
- 1 KHz in / 1.2 KHz out
- ~ 1 MB / 200 KB / 30 KB per event
- Processing time: ~20 s
- Based on accurate global reconstructions
- Software implemented on CPUs

The LHC Big Data problem



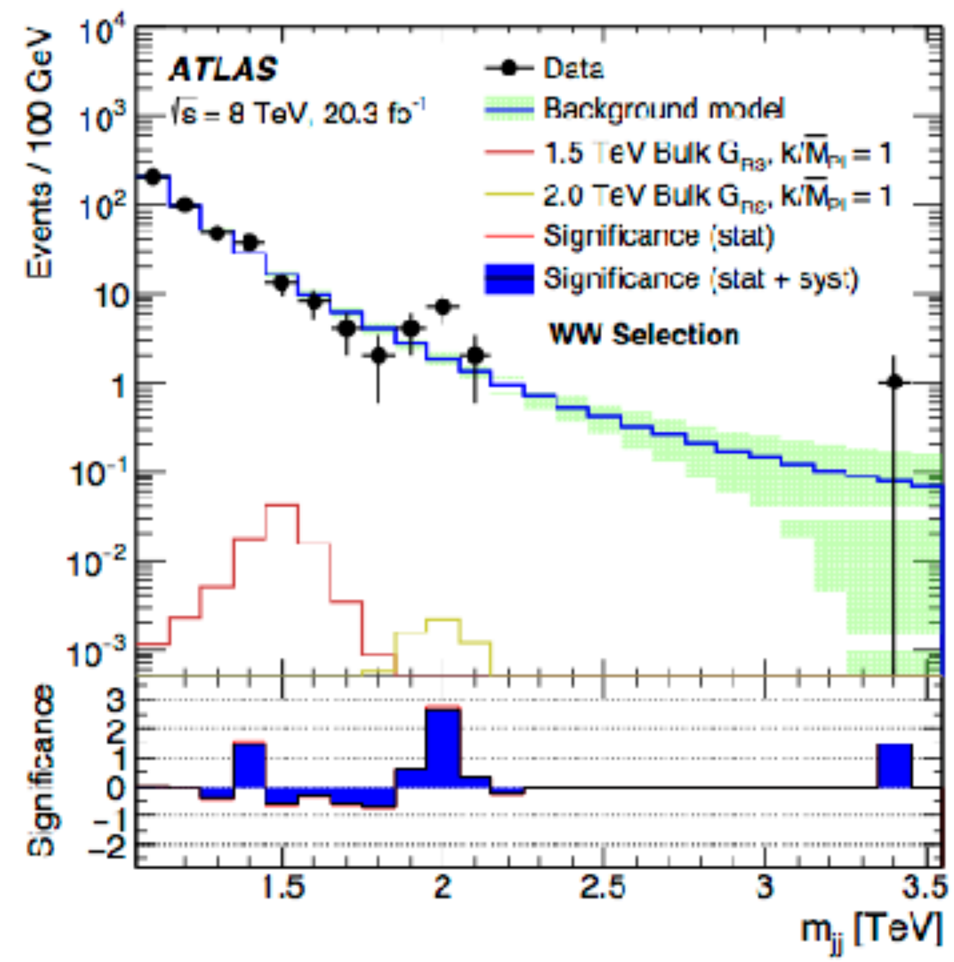
- Up to ~ 500 Hz In / 100-1000 events out
- <30 KB per event
- Processing time irrelevant
- User-written code + centrally produced selection algorithms

The LHC Big Data problem



What do we do with these data?

- We start from a question
  - Does the Higgs boson exist?
  - Is the LHC produce Dark Matter?
  - Are there heavier copies of the Standard Model particles?
- We work out the consequences of each test hypothesis
  - Higgs boson -> events with 2 photons and specific mass value
  - dark matter -> events with invisible particles
  - heavier SM copies -> pairs of SM particles with specific mass values



What do we do with these data?

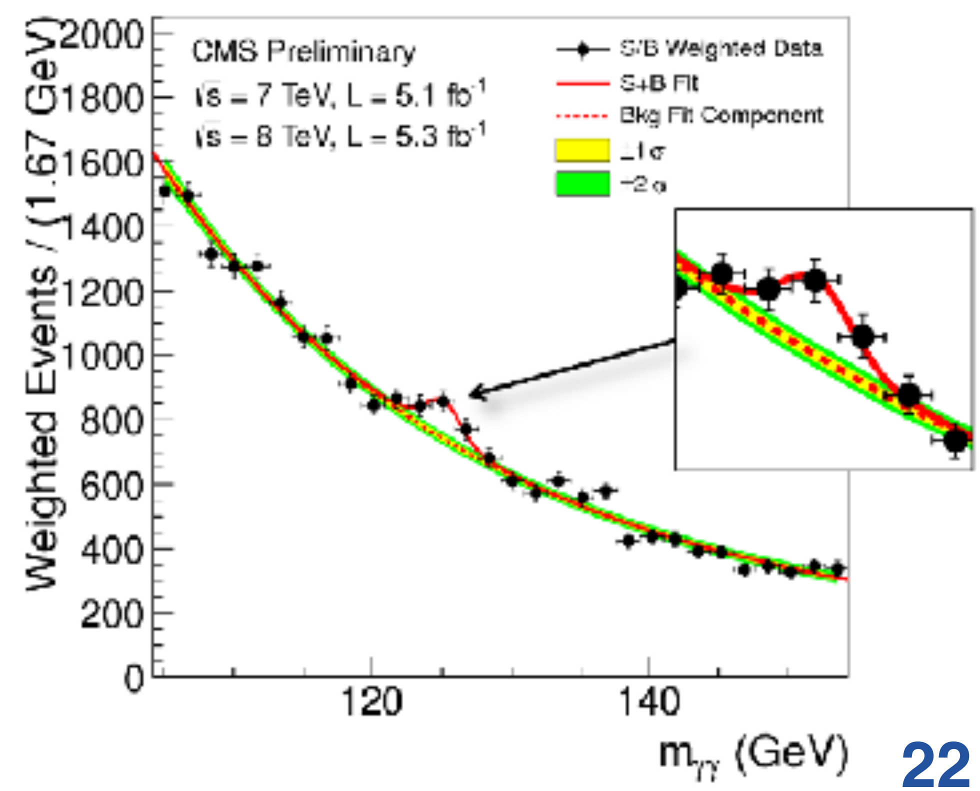
- We start from a question
  - Does the Higgs boson exist?
  - Is the LHC produce Dark Matter?
  - Are there heavier copies of the Standard Model particles?
- We work out the consequences of each test hypothesis
  - Higgs boson -> events with 2 photons and specific mass value
  - dark matter -> events with invisible particles
  - heavier SM copies -> pairs of SM particles with specific mass values

Start from all events with two photons

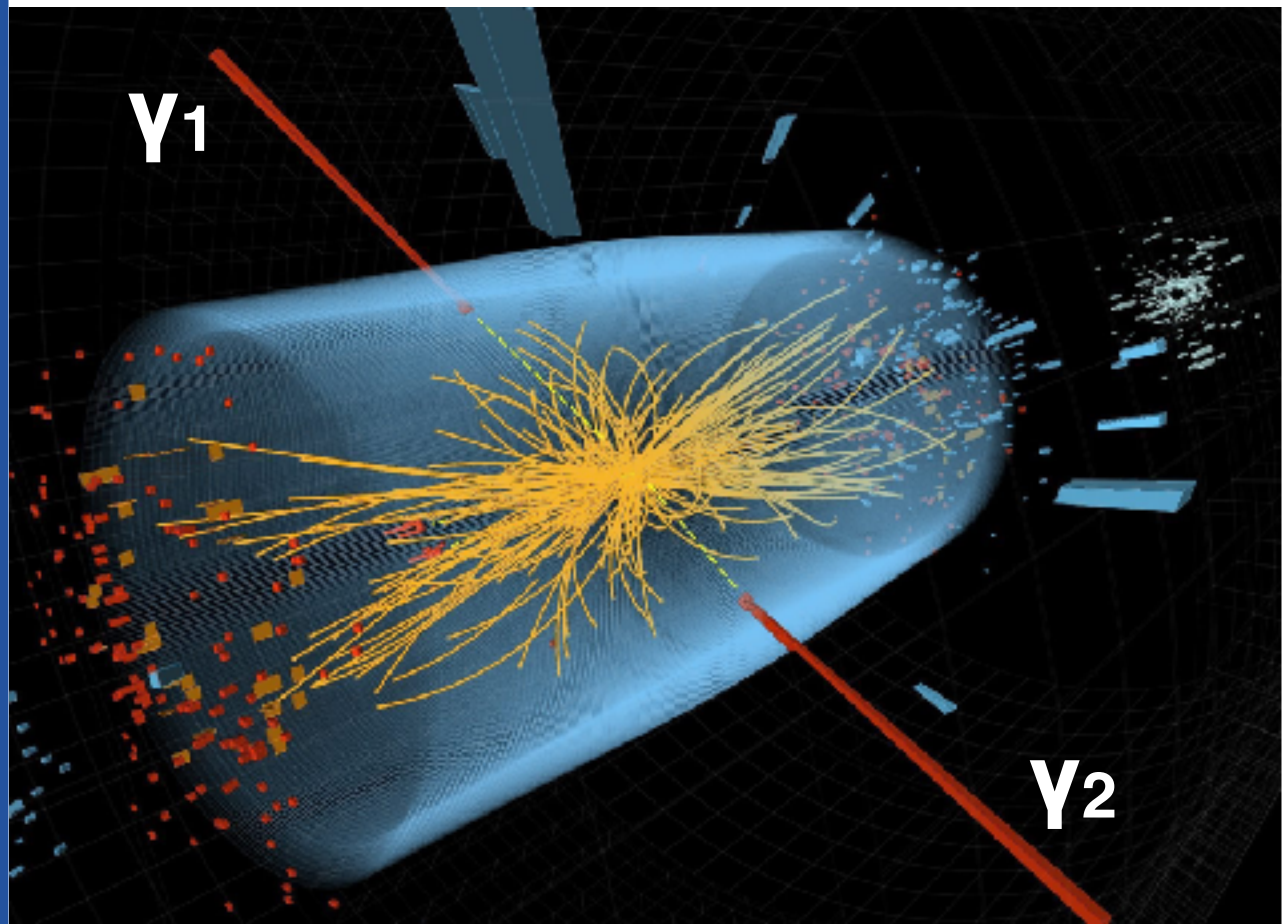
Filter the events applying cleanup algorithms

Compute the mass of the di-photon object

Run statistical analysis (search for signal over background)

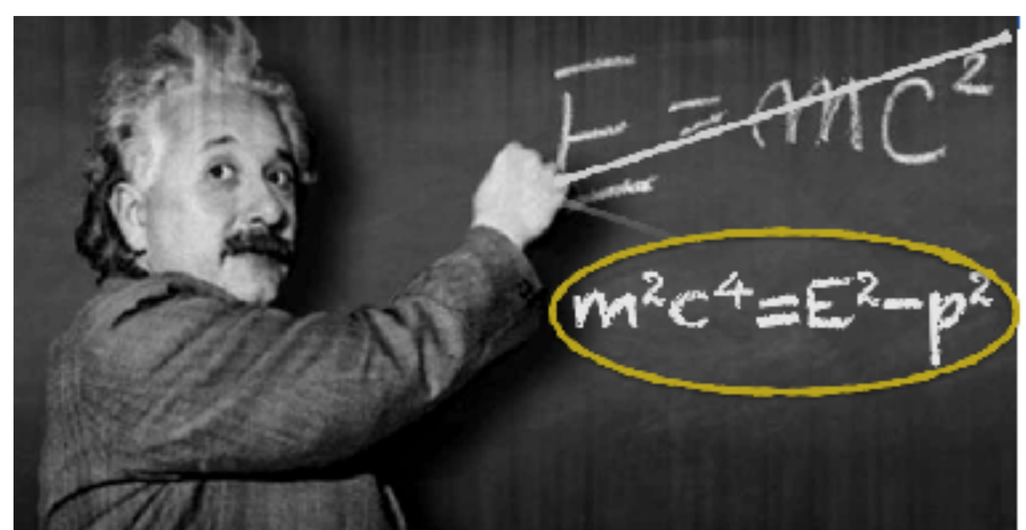


What do we do with these data?



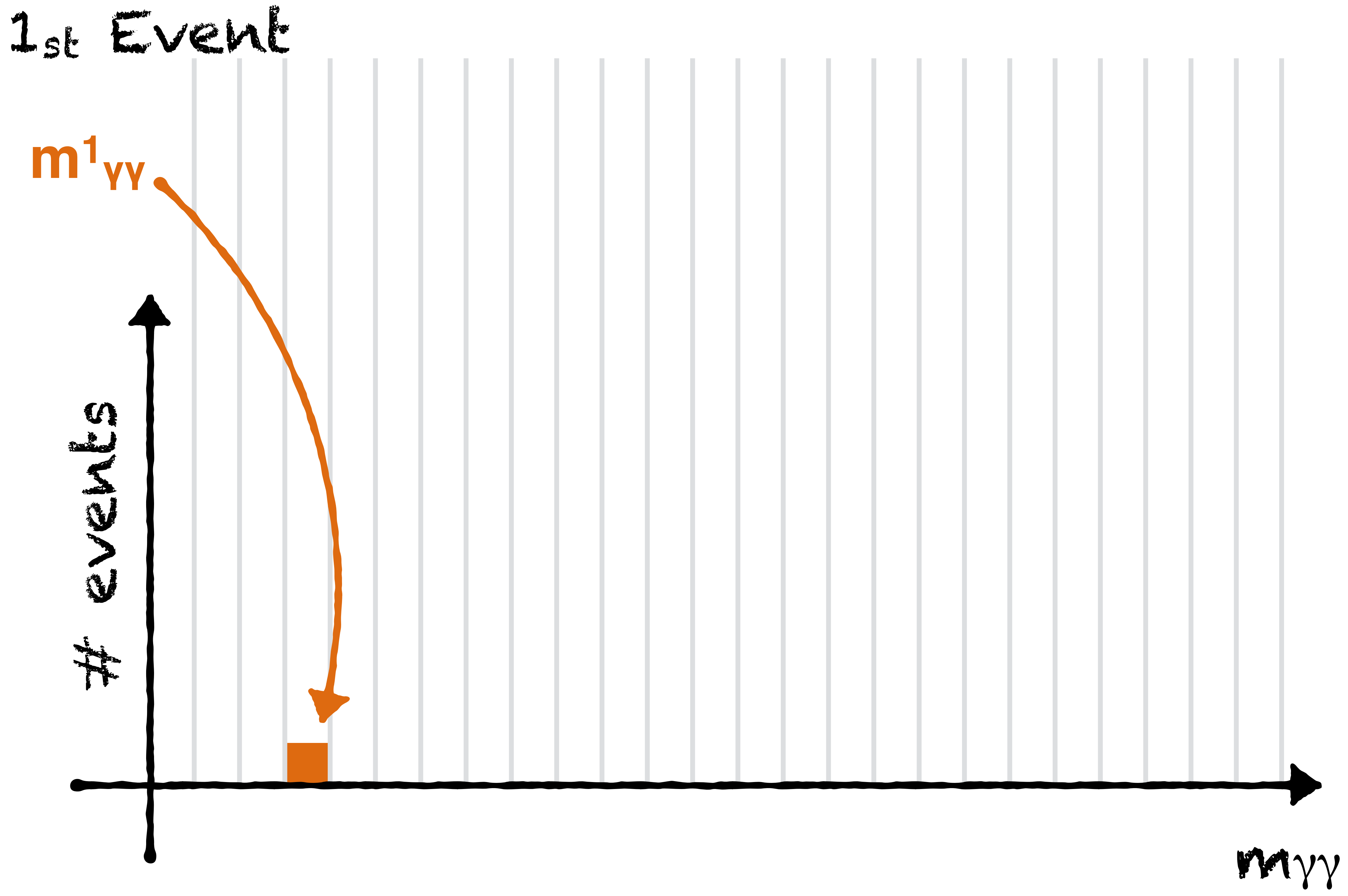
- Energy of the particle E measured by detector
- Location of the deposit gives the directions (v_x,v_y,v_z) and (w_x,w_y,w_z) for γ_1 and γ_2
- Photons have no mass

E_1	E_1v_x	E_1v_y	E_1v_z
+	+	+	+
E_2	E_2w_x	E_2w_y	E_2w_z
=	=	=	=
E	p_x	p_y	p_z

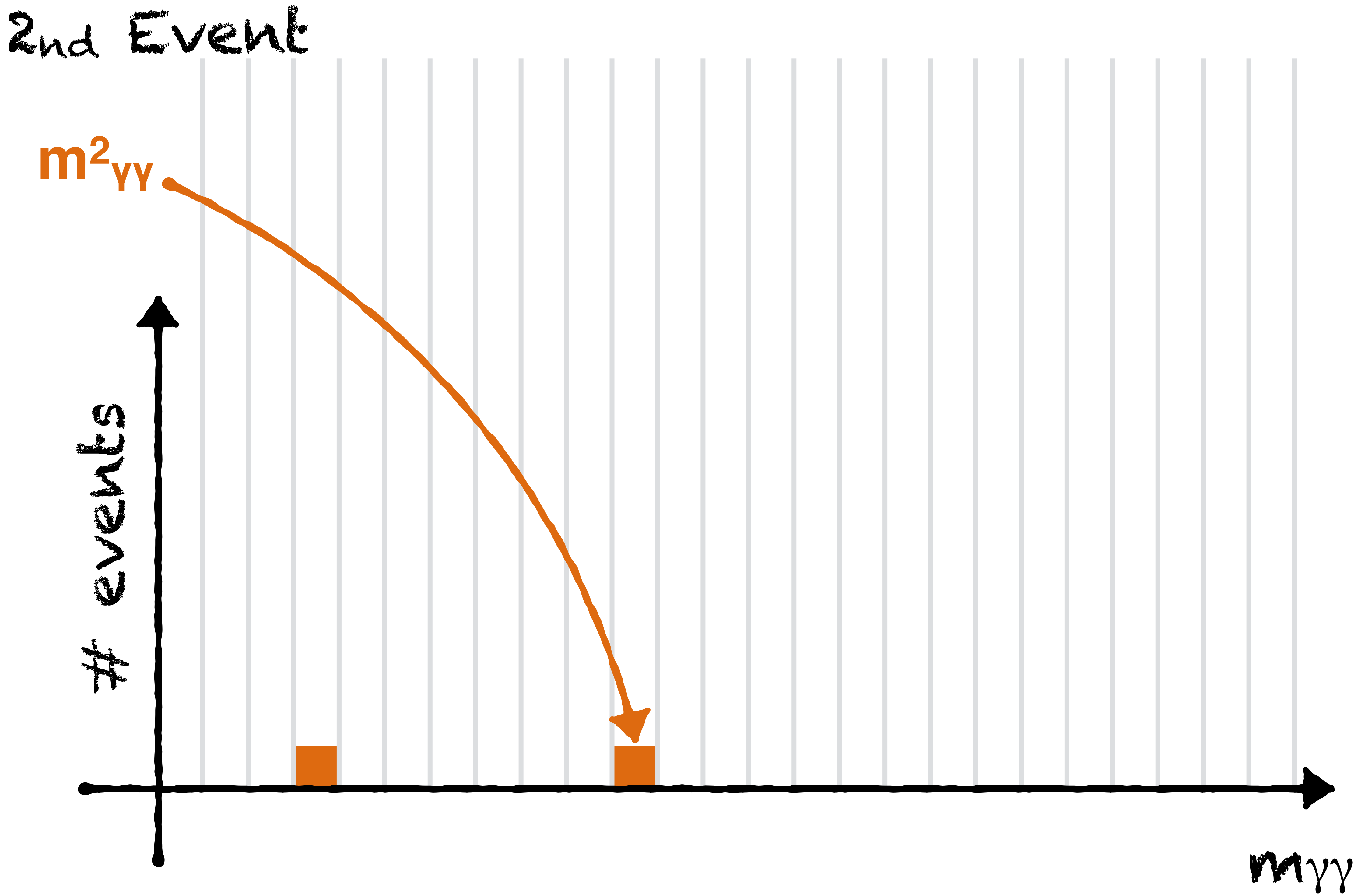


$m_{\gamma\gamma}$

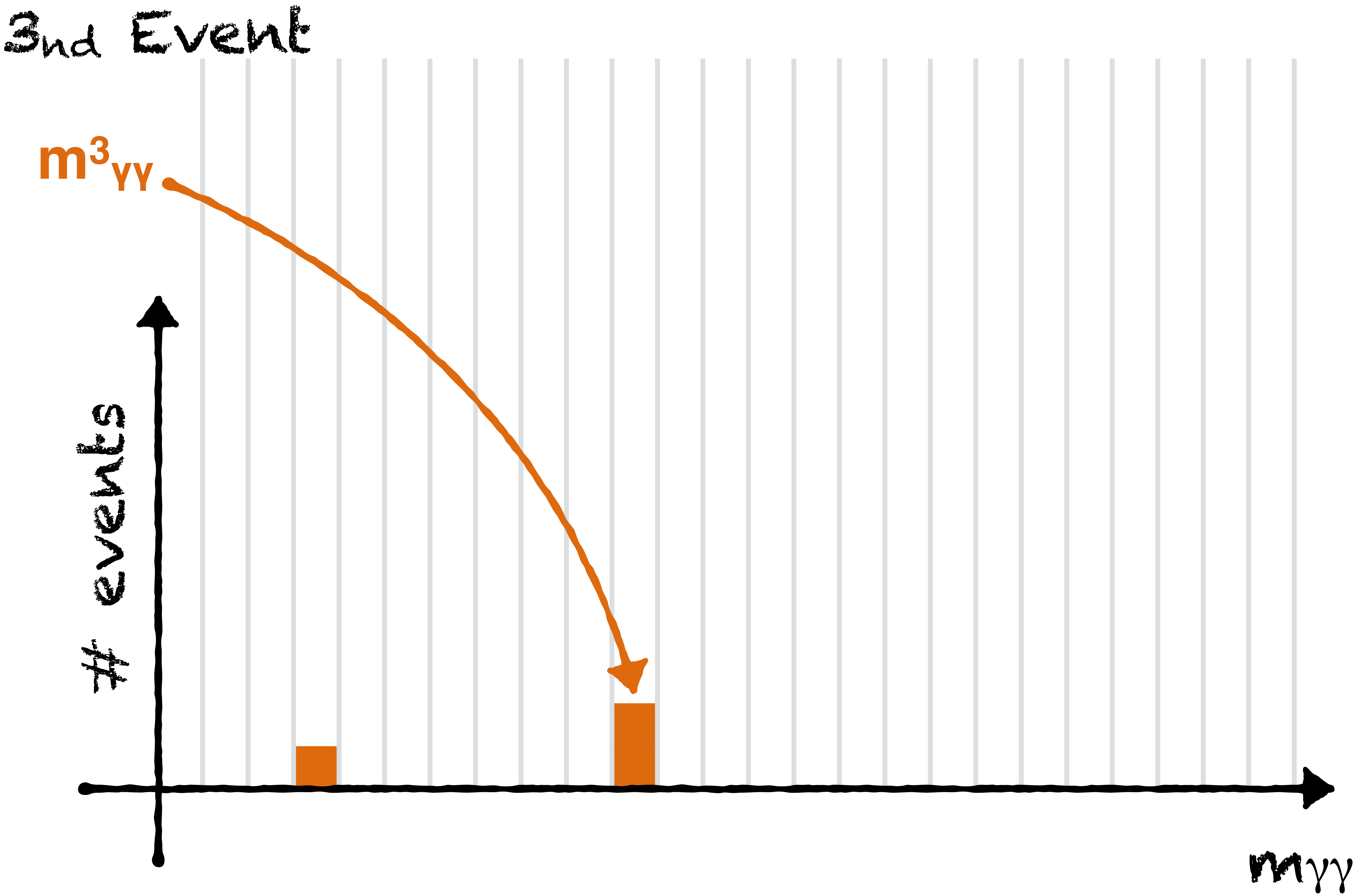
Filling the mass histogram



Filling the mass histogram

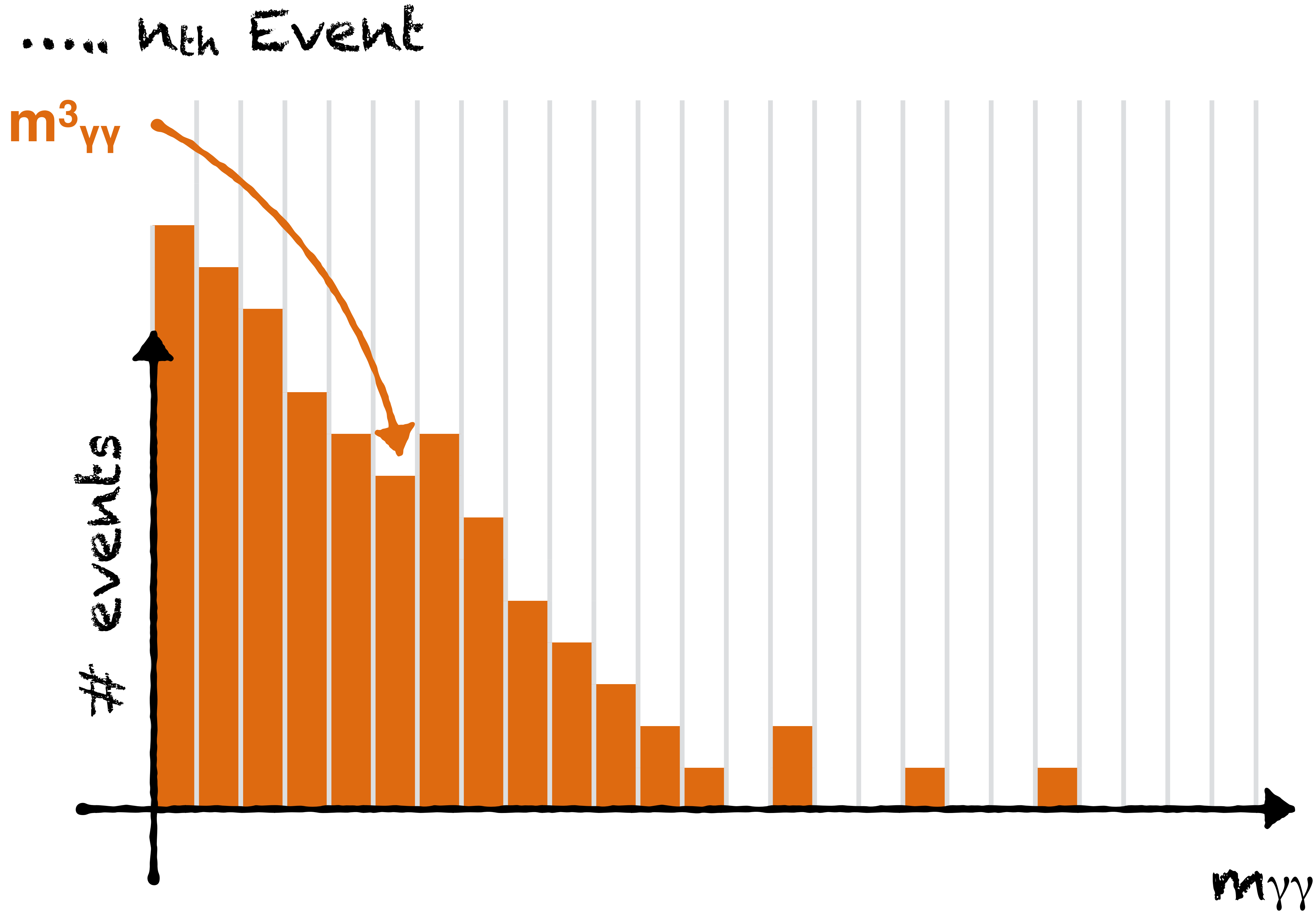


Filling the mass histogram

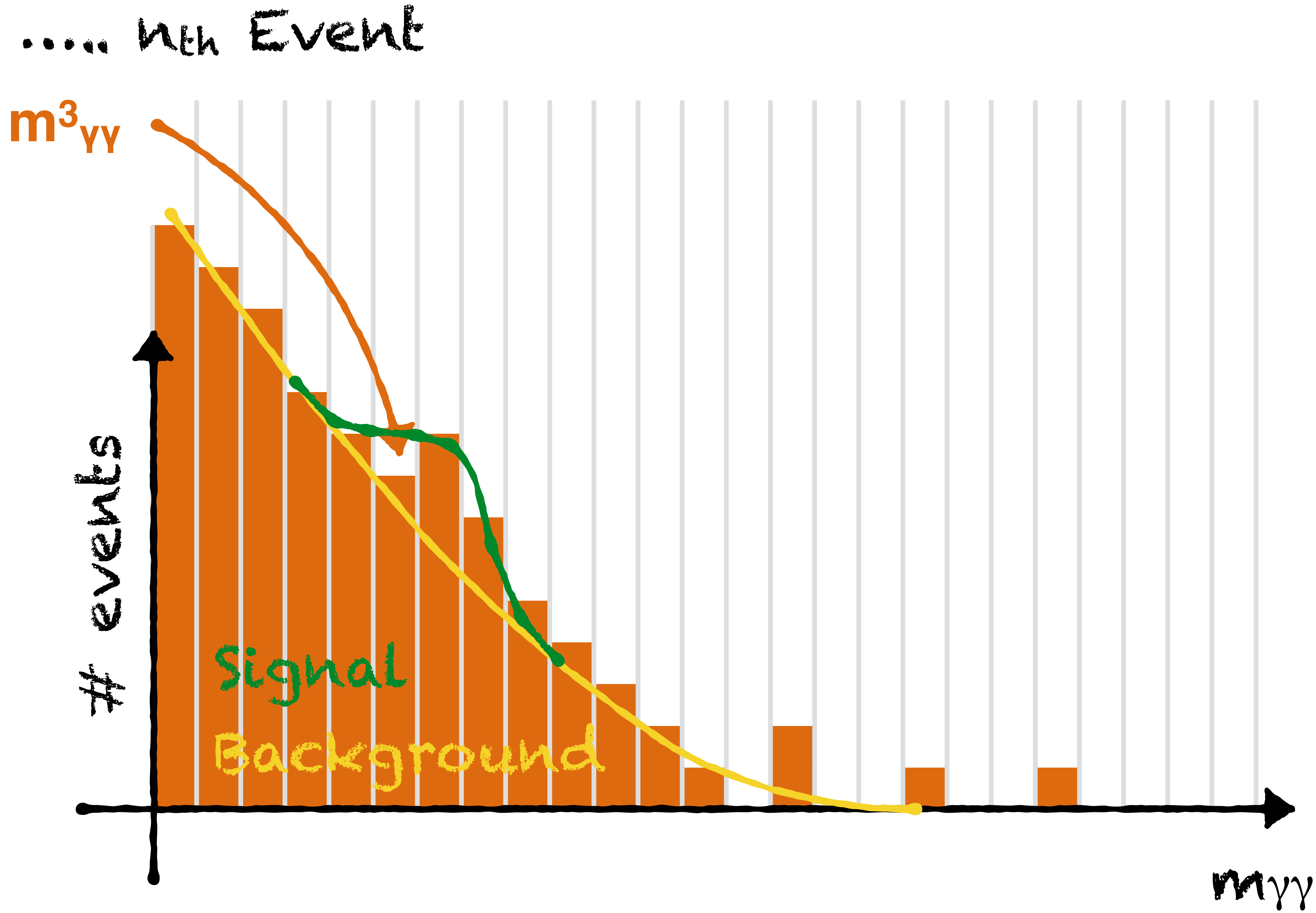


• • • •

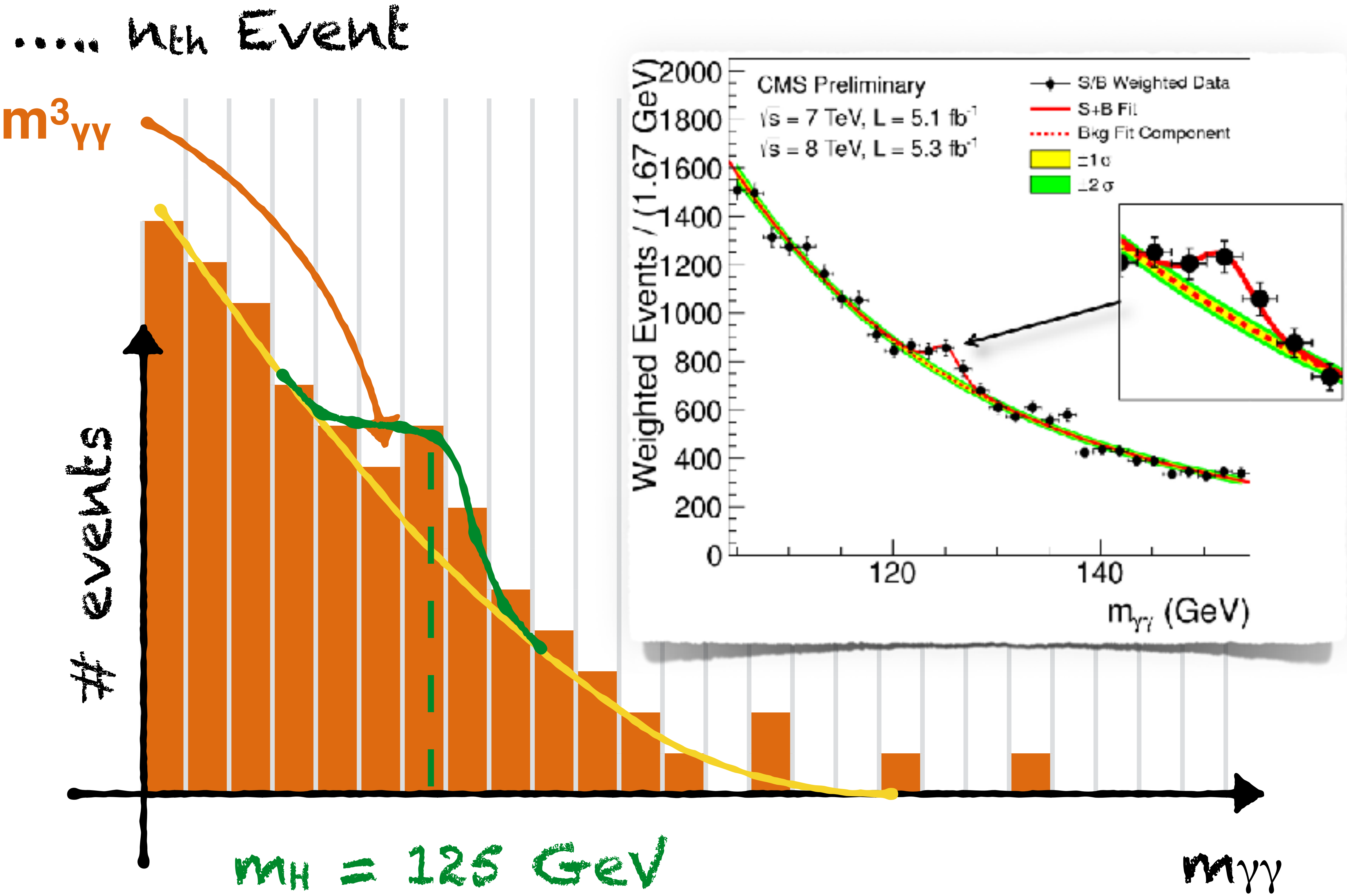
Filling the mass histogram



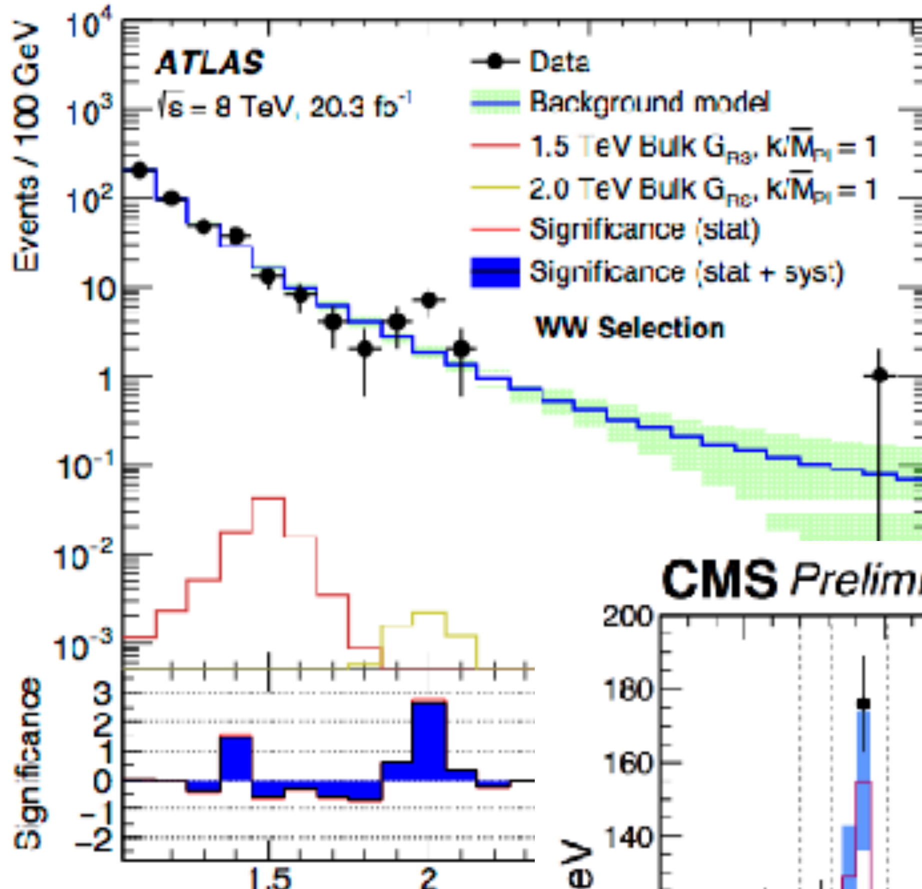
Fitting the mass histogram

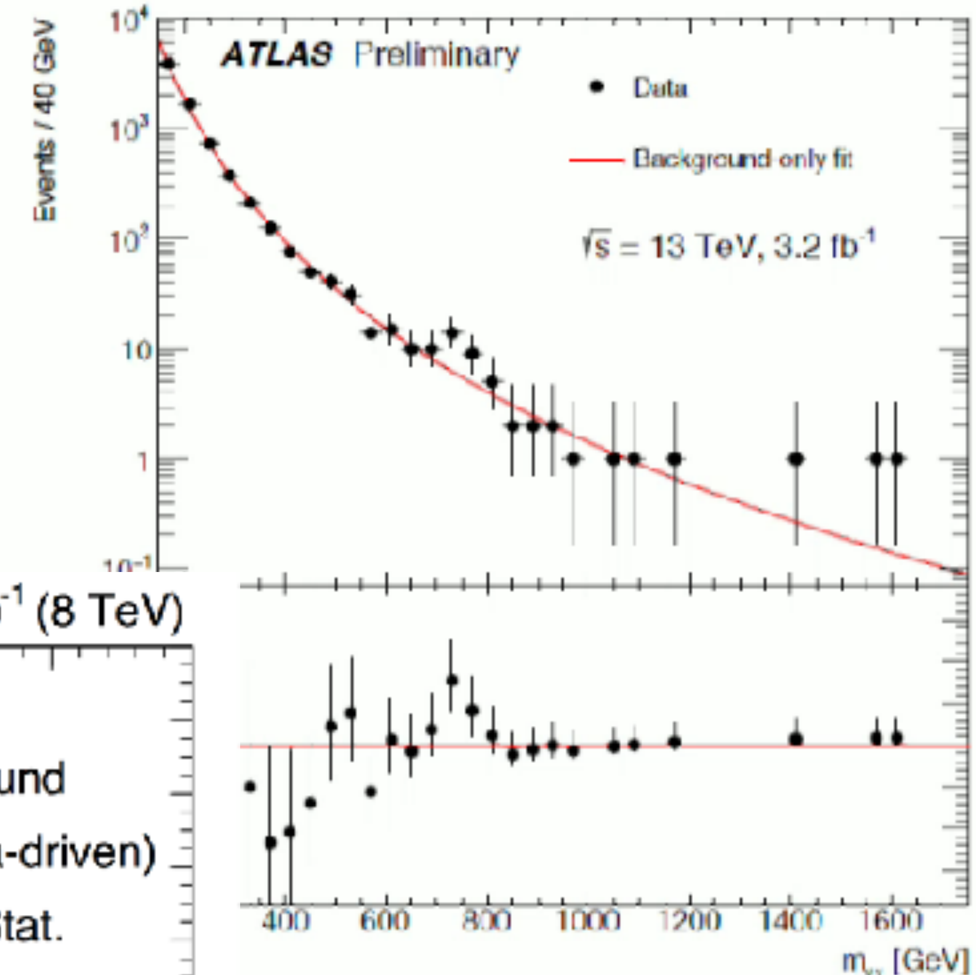


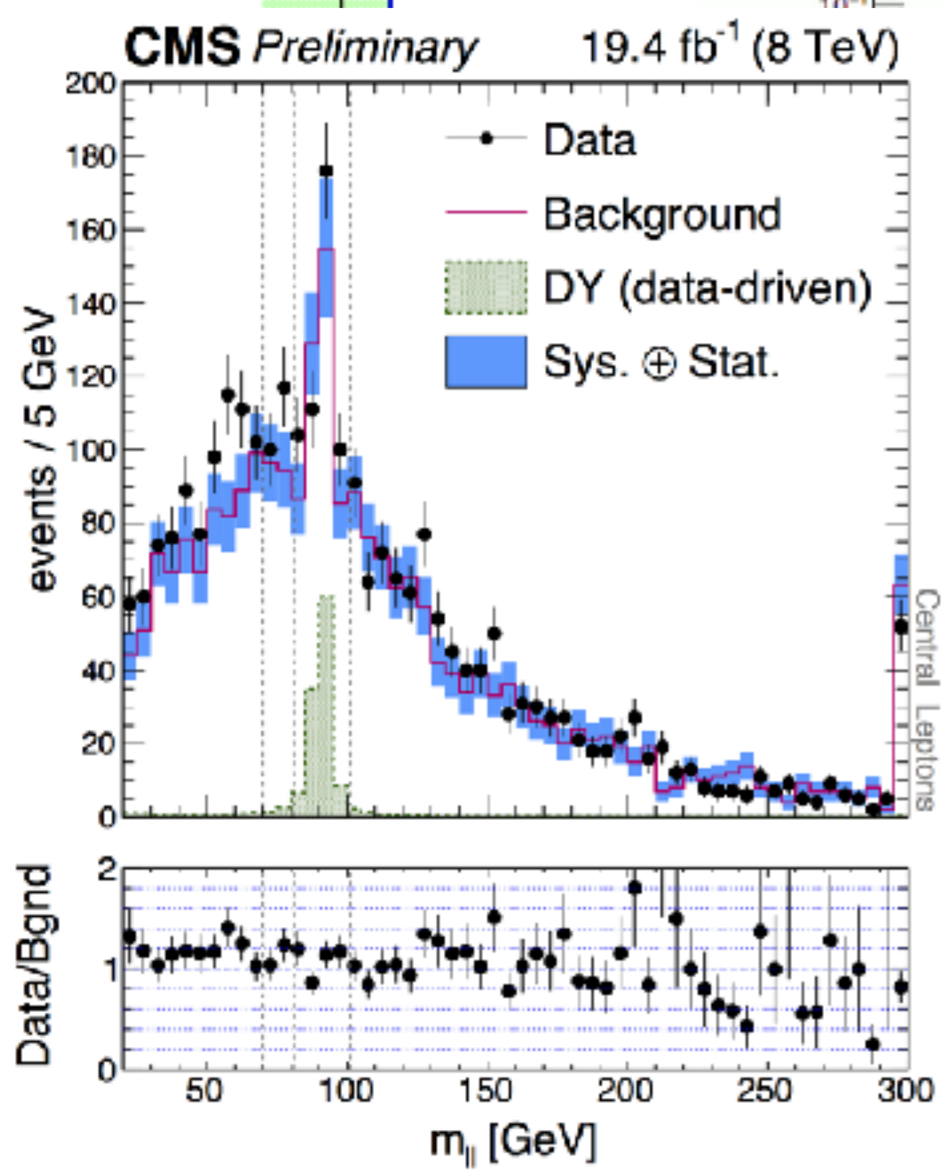
Fitting the mass histogram



Not all peaks are discoveries

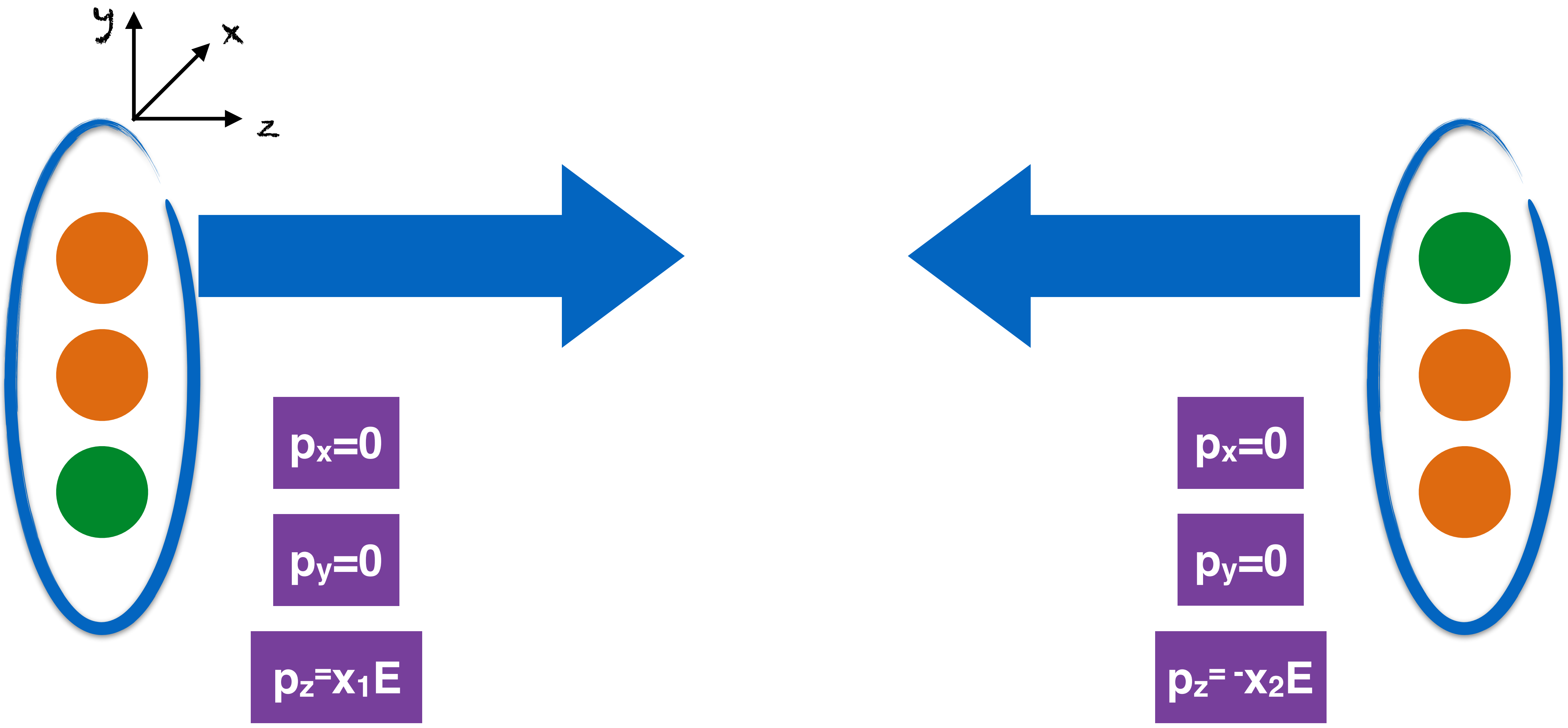






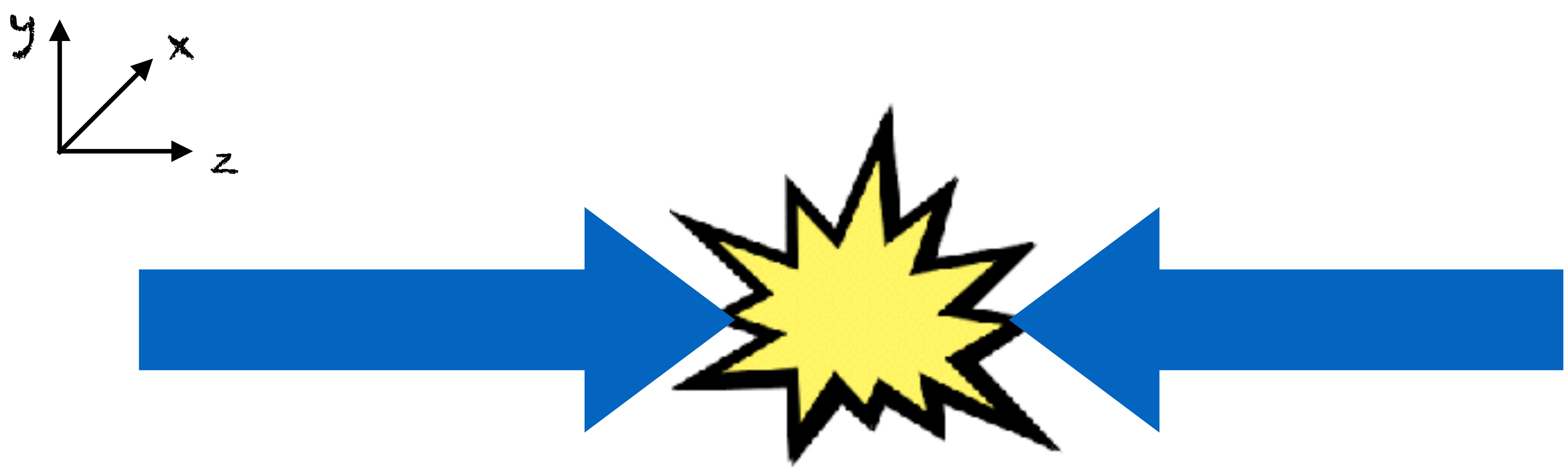
Search for Dark Matter

- We cannot see Dark Matter
- But we can make it
- We can observe Dark Matter indirectly, using energy/momentum conservation



Search for Dark Matter

- We cannot see Dark Matter
- But we can make it
- We can observe Dark Matter indirectly, using energy/momentum conservation



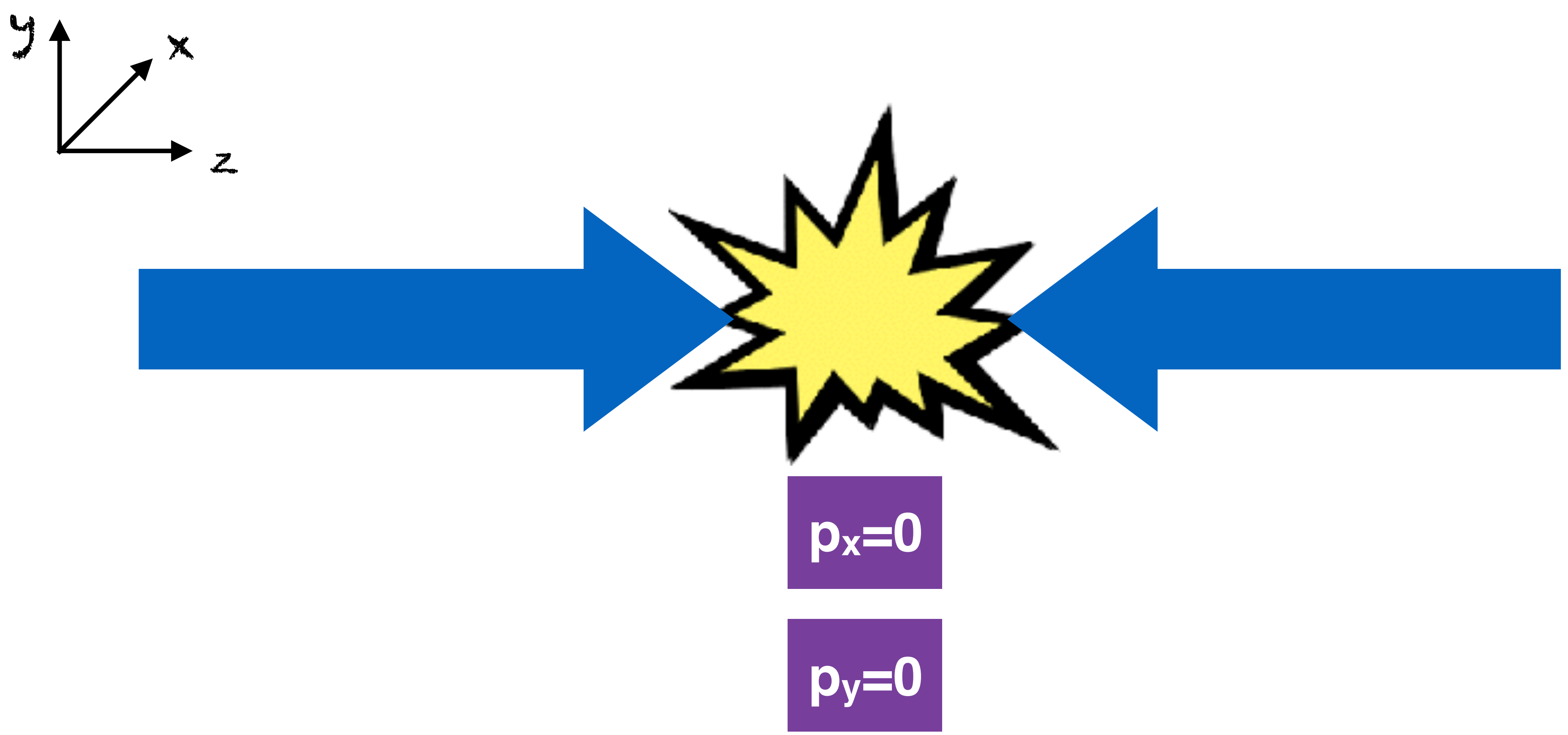
$p_x=0$

$p_y=0$

$p_z=(x_1-x_2)E$

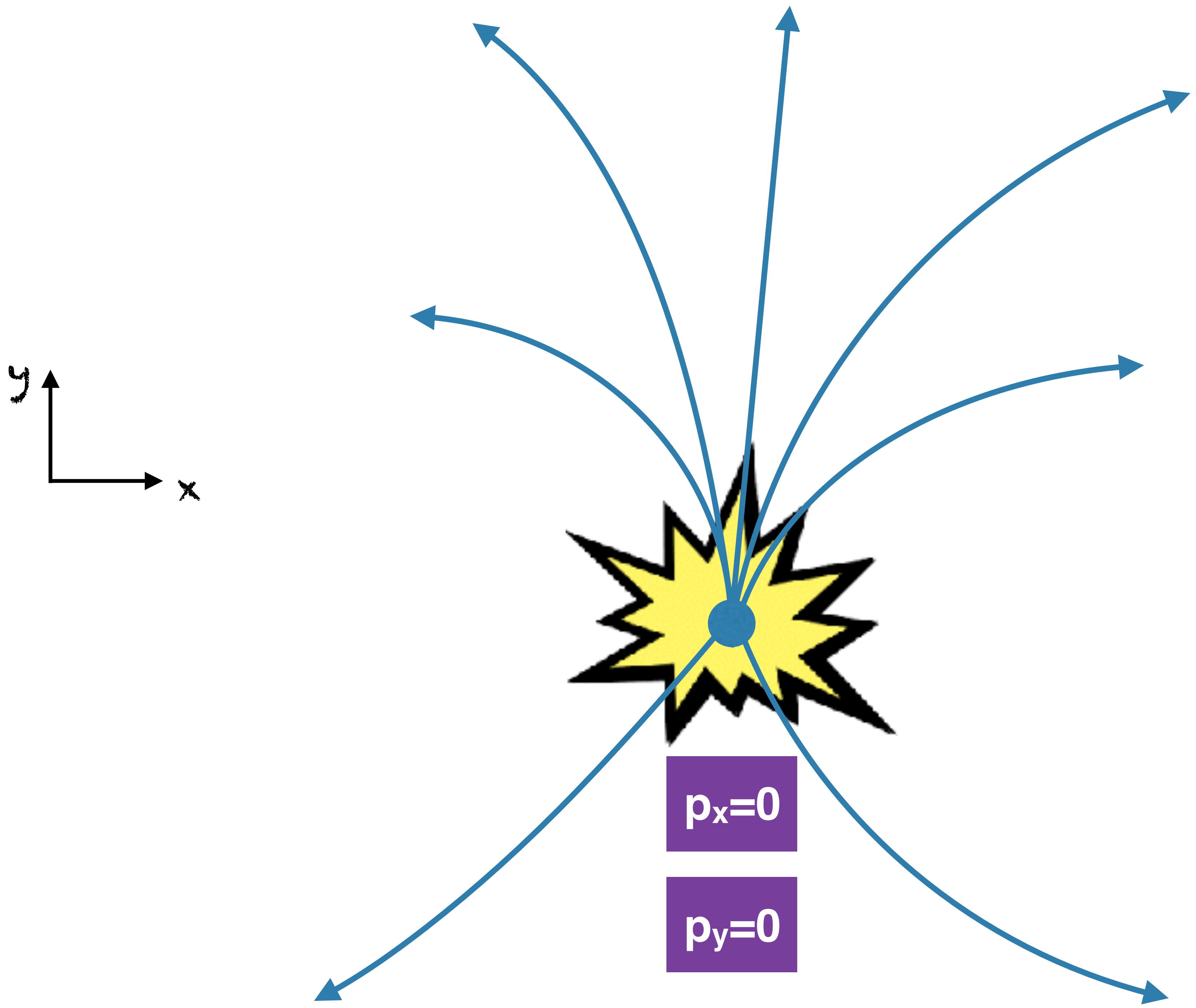
Search for Dark Matter

- We cannot see Dark Matter
- But we can make it
- We can observe Dark Matter indirectly, using energy/momentum conservation



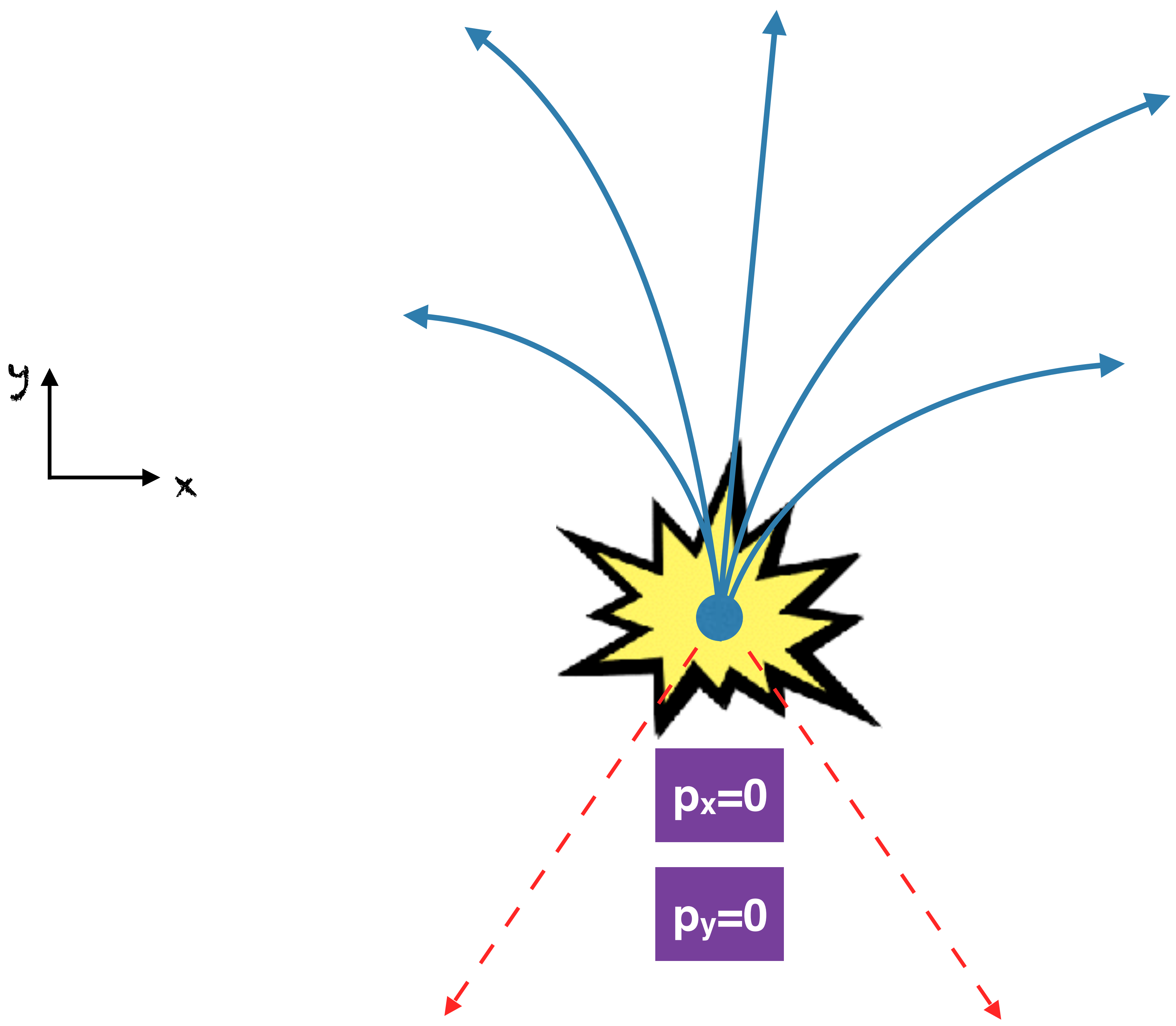
No initial momentum flows in the transverse plane

Search for Dark Matter



Total momentum flows of collision product should be 0

Search for Dark Matter



If not, something is escaping of the detector

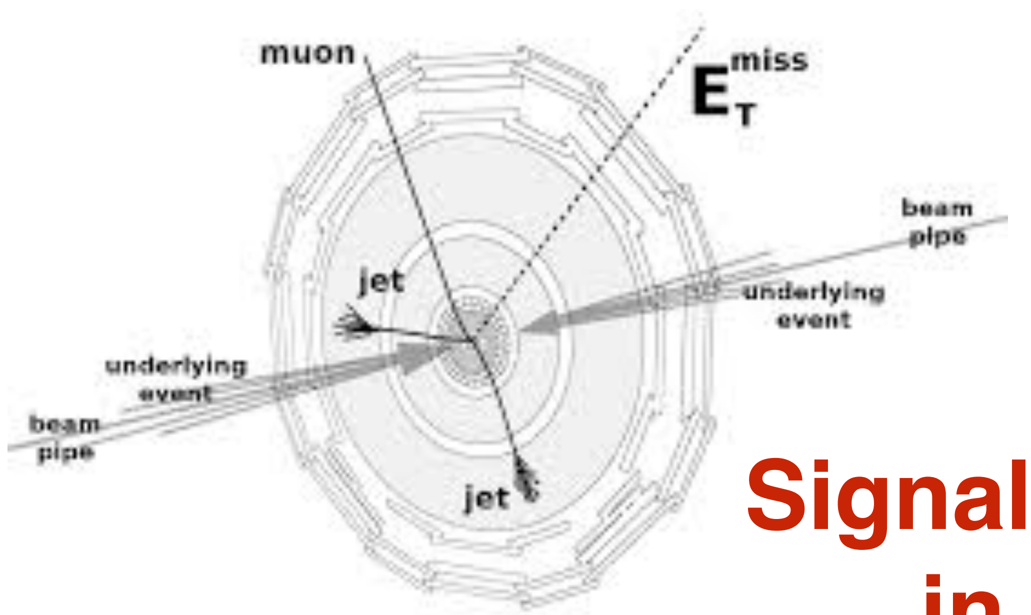
Missing Transverse Energy

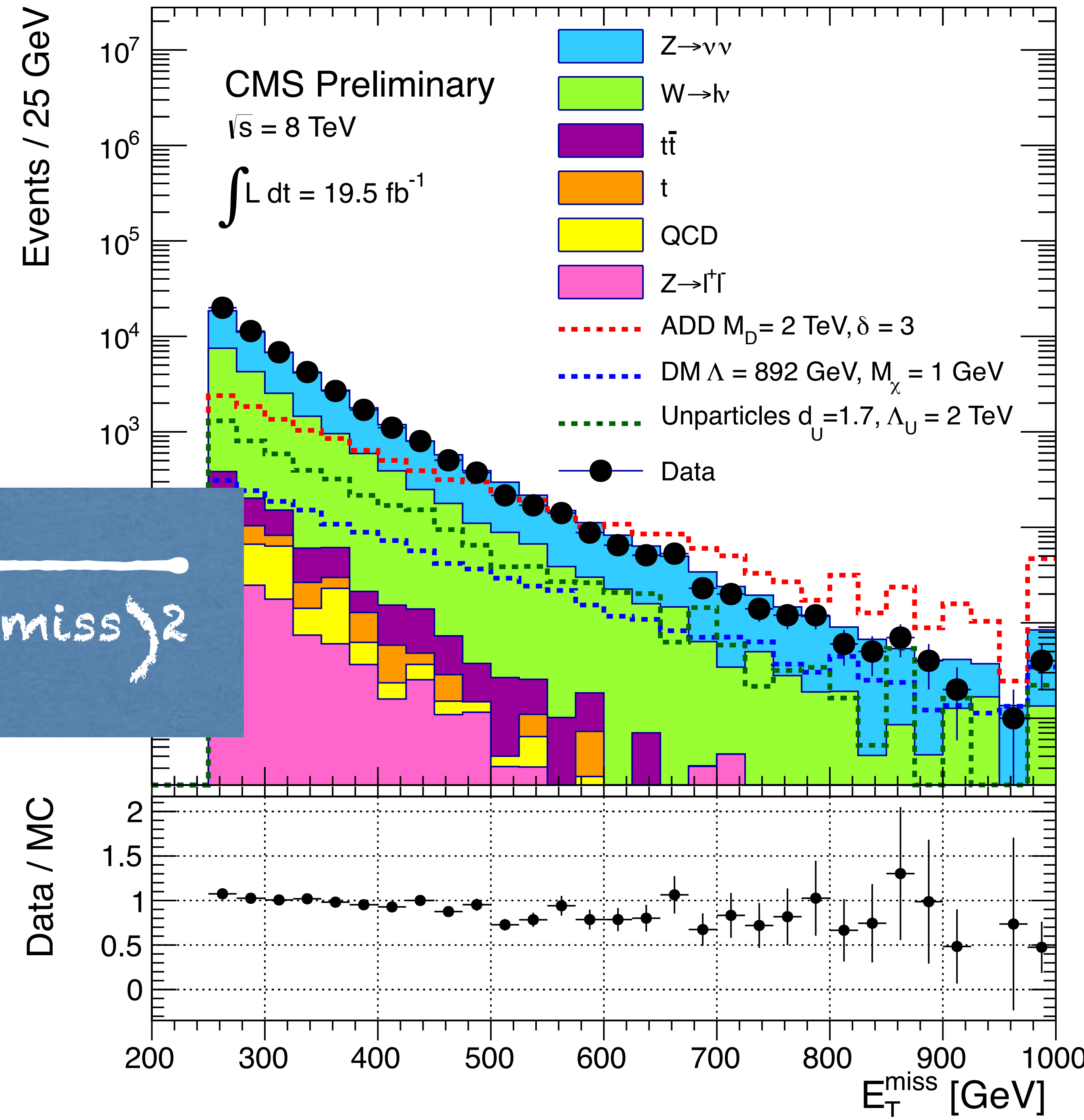
- Missing Momentum

$$p_x^{miss} = \Sigma_i \, p_x^i$$
$$p_y^{miss} = \Sigma_i \, p_y^i$$

- Missing Energy

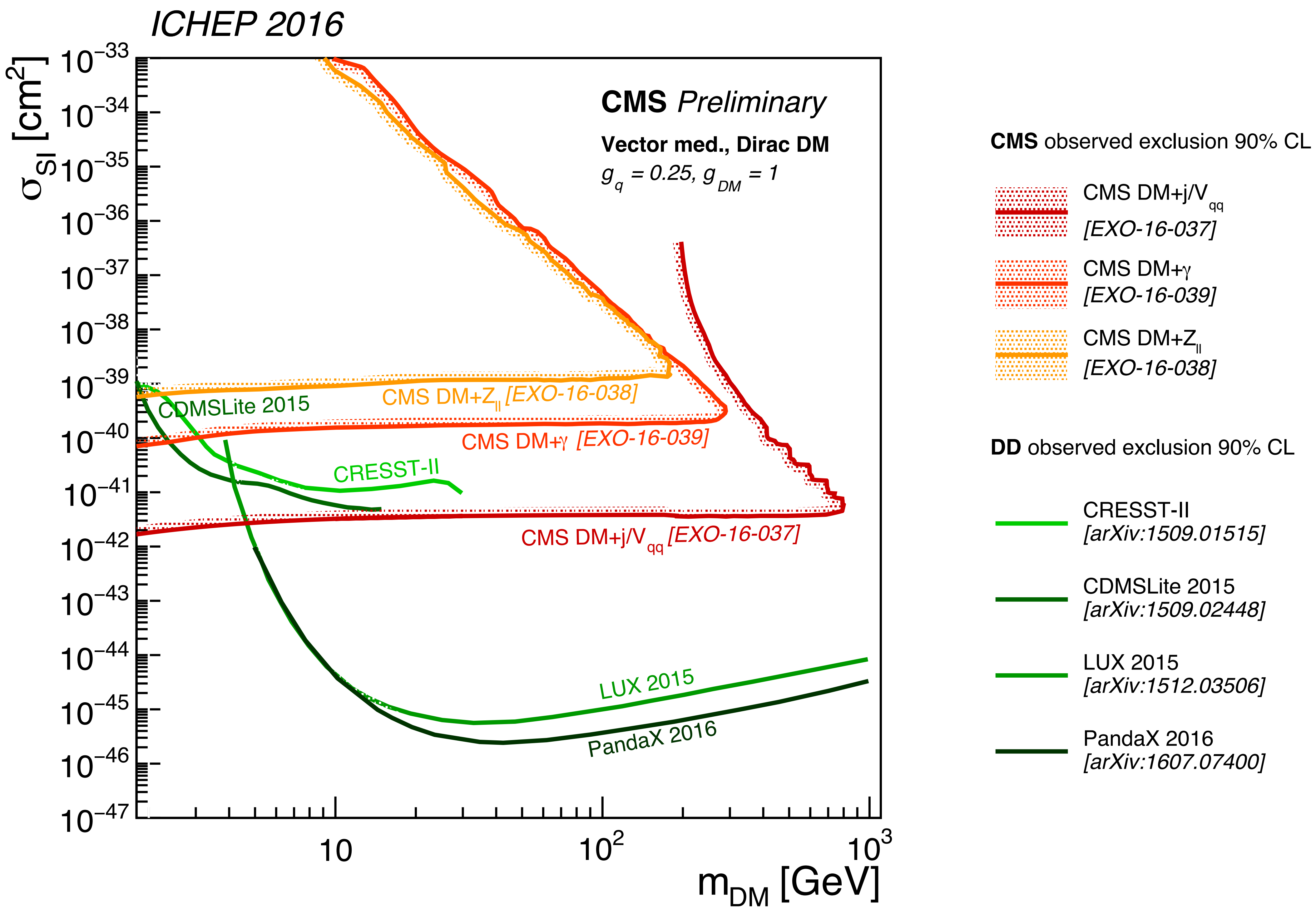
$$E_T^{miss} = \sqrt{(p_x^{miss})^2 + (p_x^{miss})^2}$$





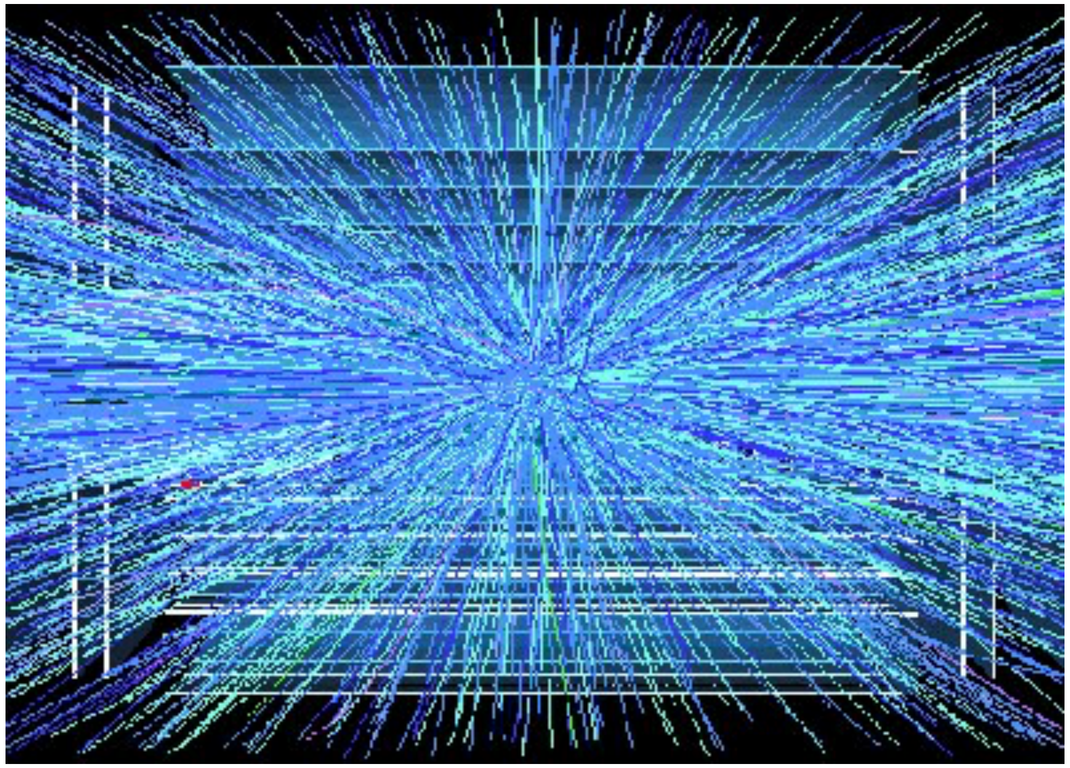
Signal should emerge as an access in the tail of the distribution

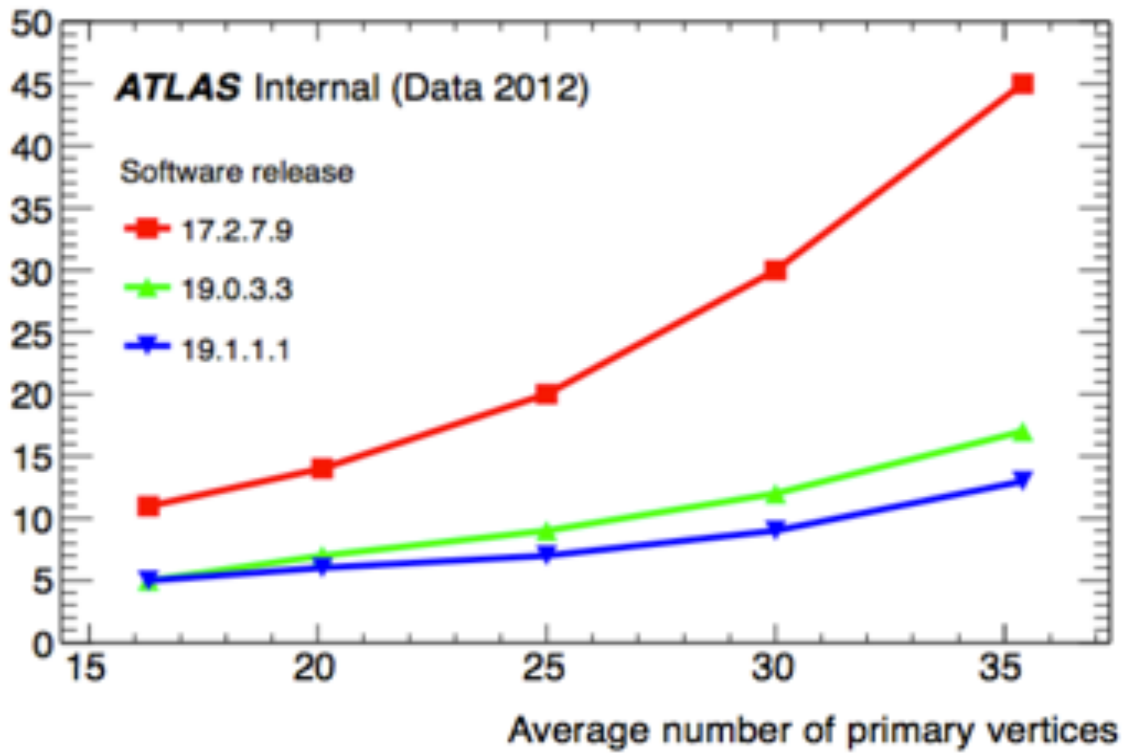
But it did not.. (so far)



The challenge ahead

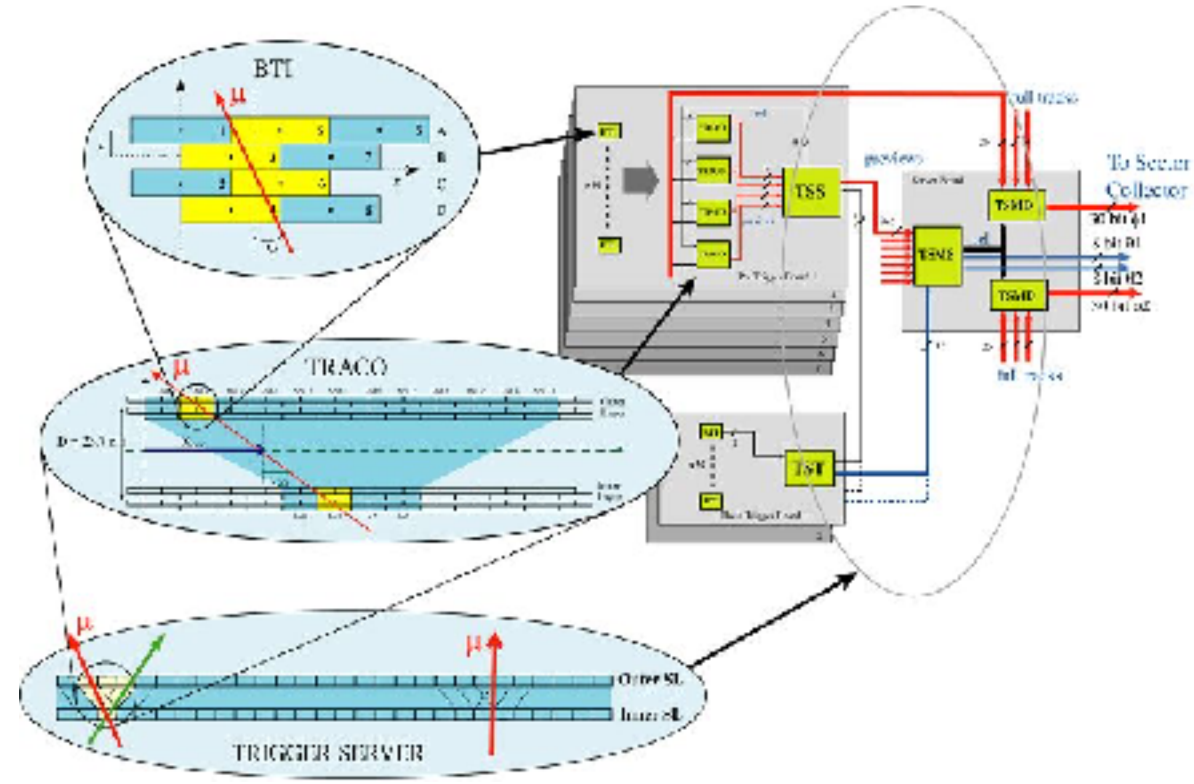
- The evolving conditions of the machine are drifting the experiments to more prohibitive environments (luminosity comes with a cost)
- More (& bigger) events to handle
- More noise from pileup interactions
- Increase in resources will not scale with needs
- Flat (or decreasing?) budget
- (Non linearly) increasing demand
- Need to find better ways to do things
- Problems can be formulated as image detection, where big progresses are happening (see ConvNNs)



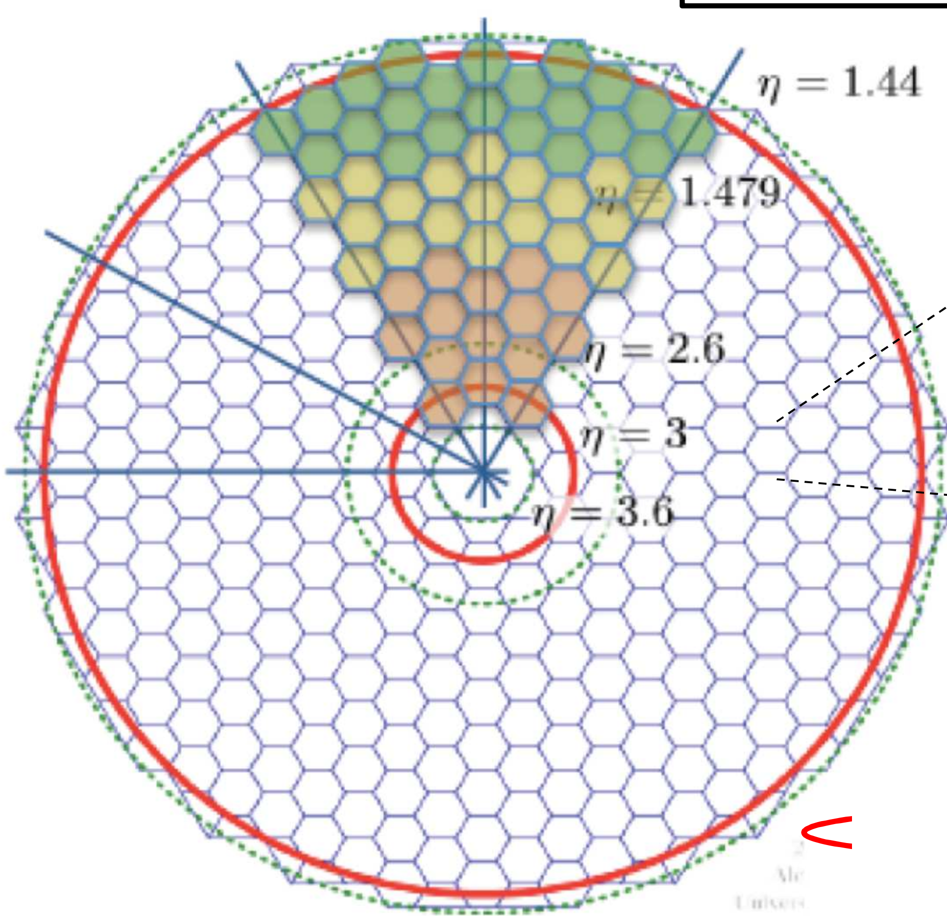


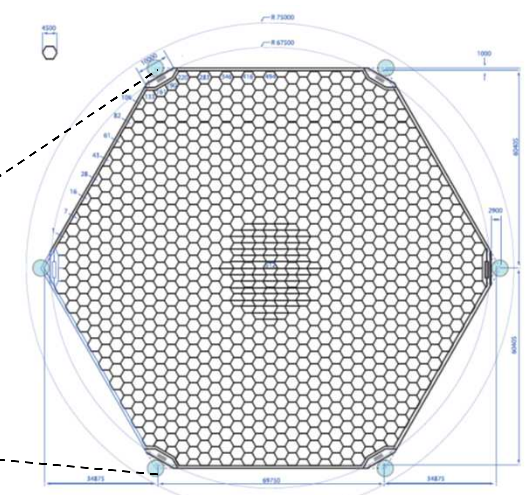
New instruments

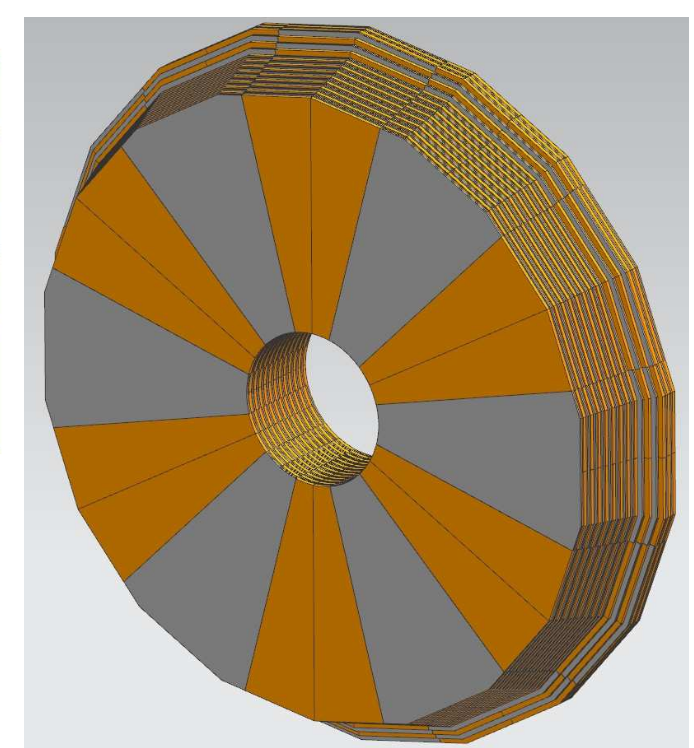
- The High-Luminosity challenges will be faced improving the detector
  - add tracking capability earlier in the game (@L1 trigger)
  - improve detector coverage
  - improve detector granularity



Hexagonal 6” Si wafer (256 or 512 channels

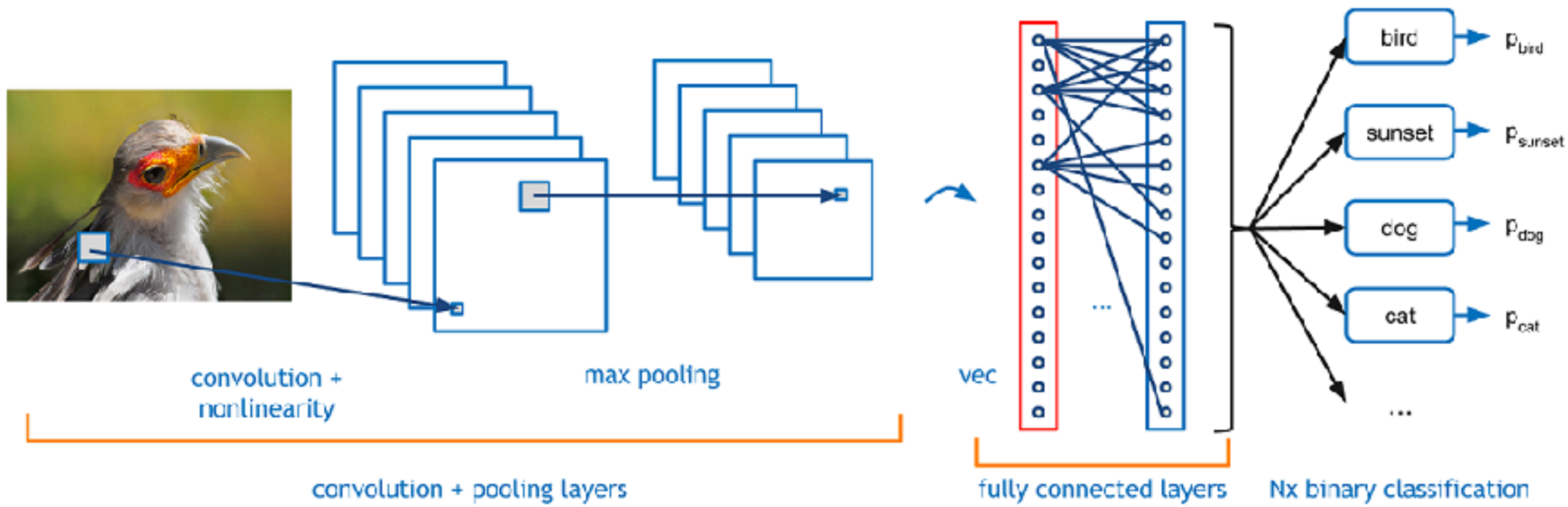






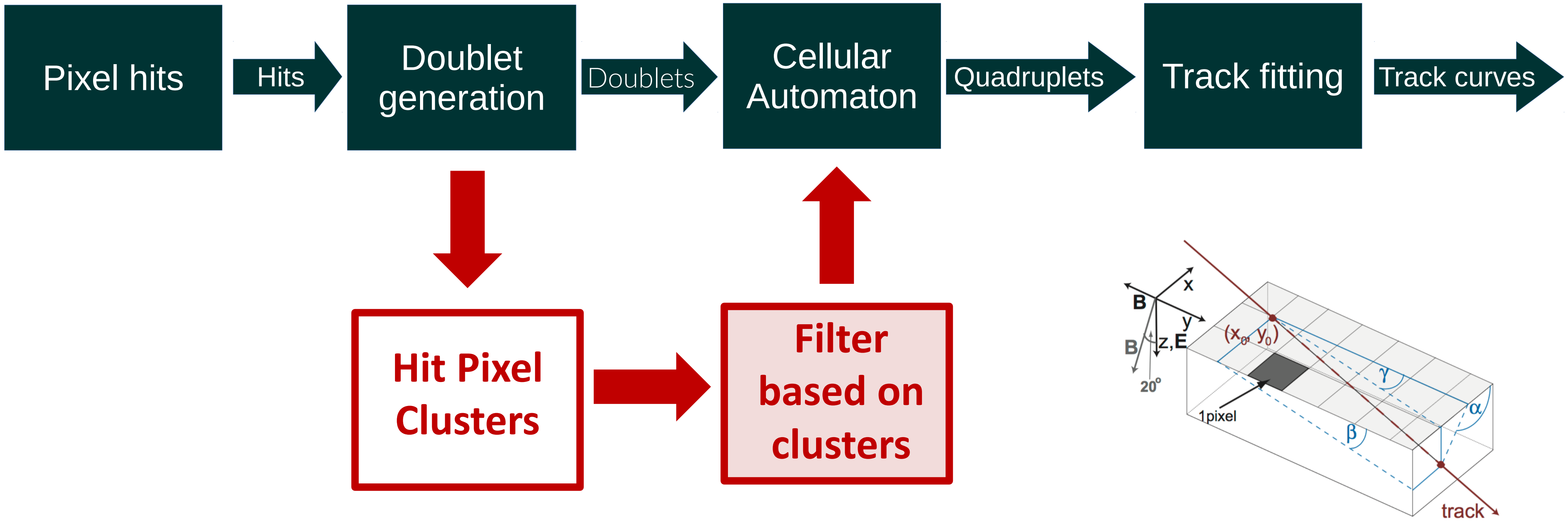
Convolutional NN

- New technique for computing vision & AI applications

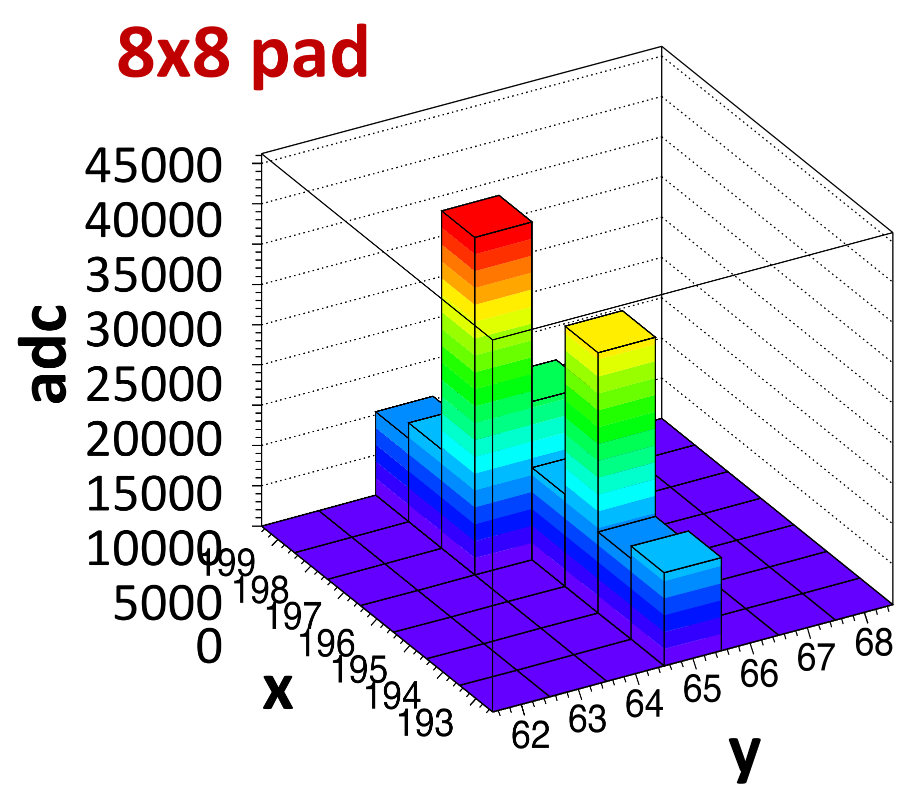


- Similar to human vision
  - process overlapping patches of image
  - combine them together
- Nowadays technology for deep learning (self-driving cars, etc)

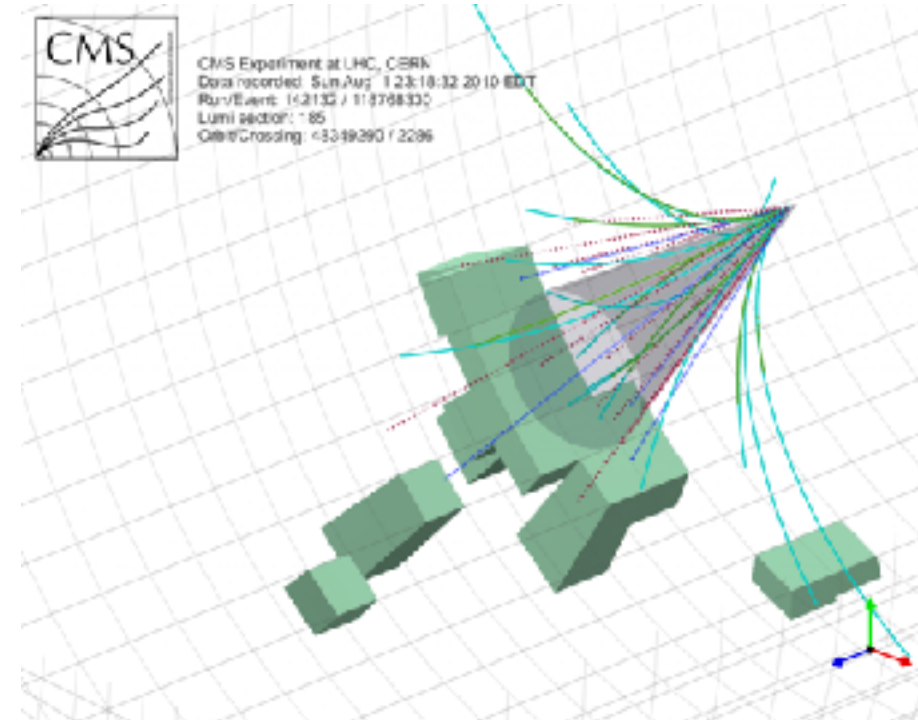
Examples: Tracking



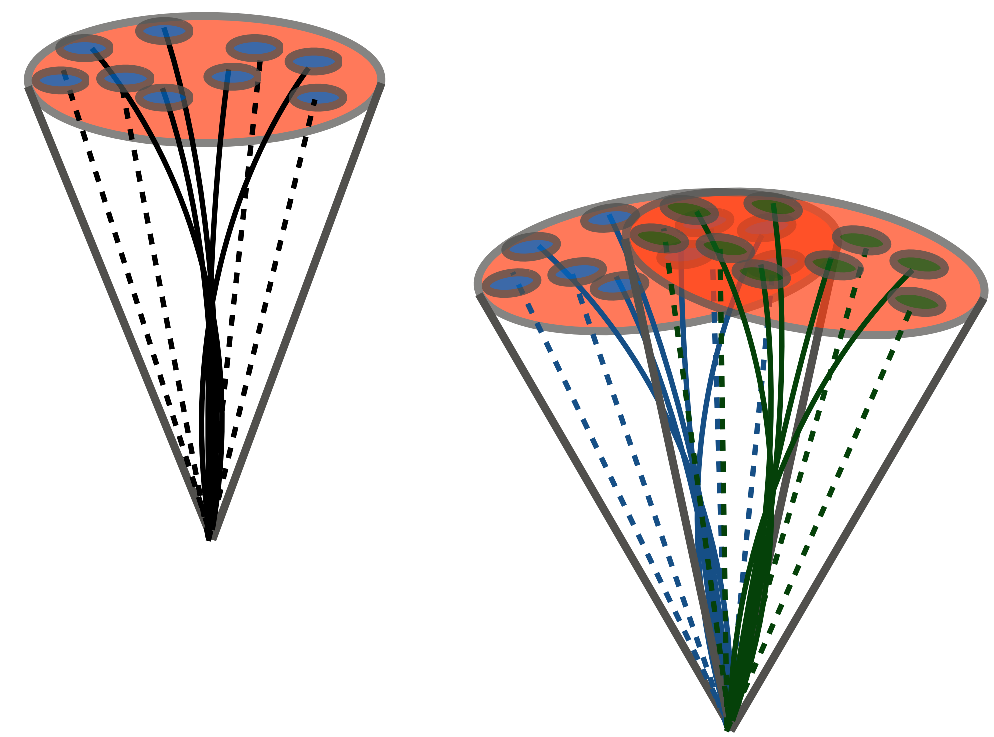
- Represent hits as 8x8 images
  - use the deposited energy (ADC counts) as temperature
- Use DNN to decide if a given pair of hits is a good match or a fake



Jet ID with ML

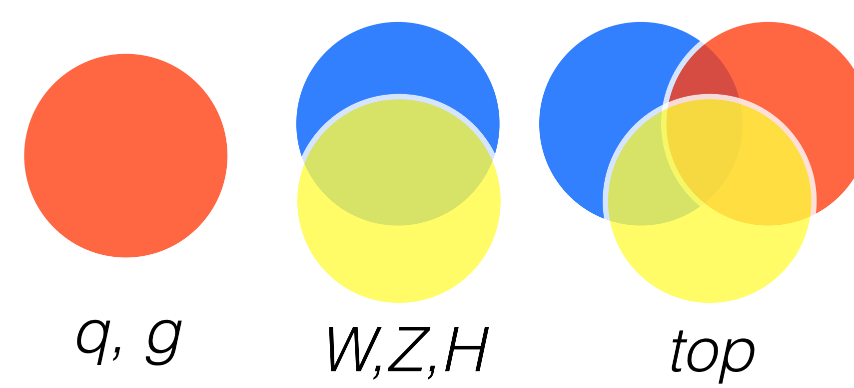


- Jets are cone-like showers of quarks and gluons that produce tens of particles, all close to each other



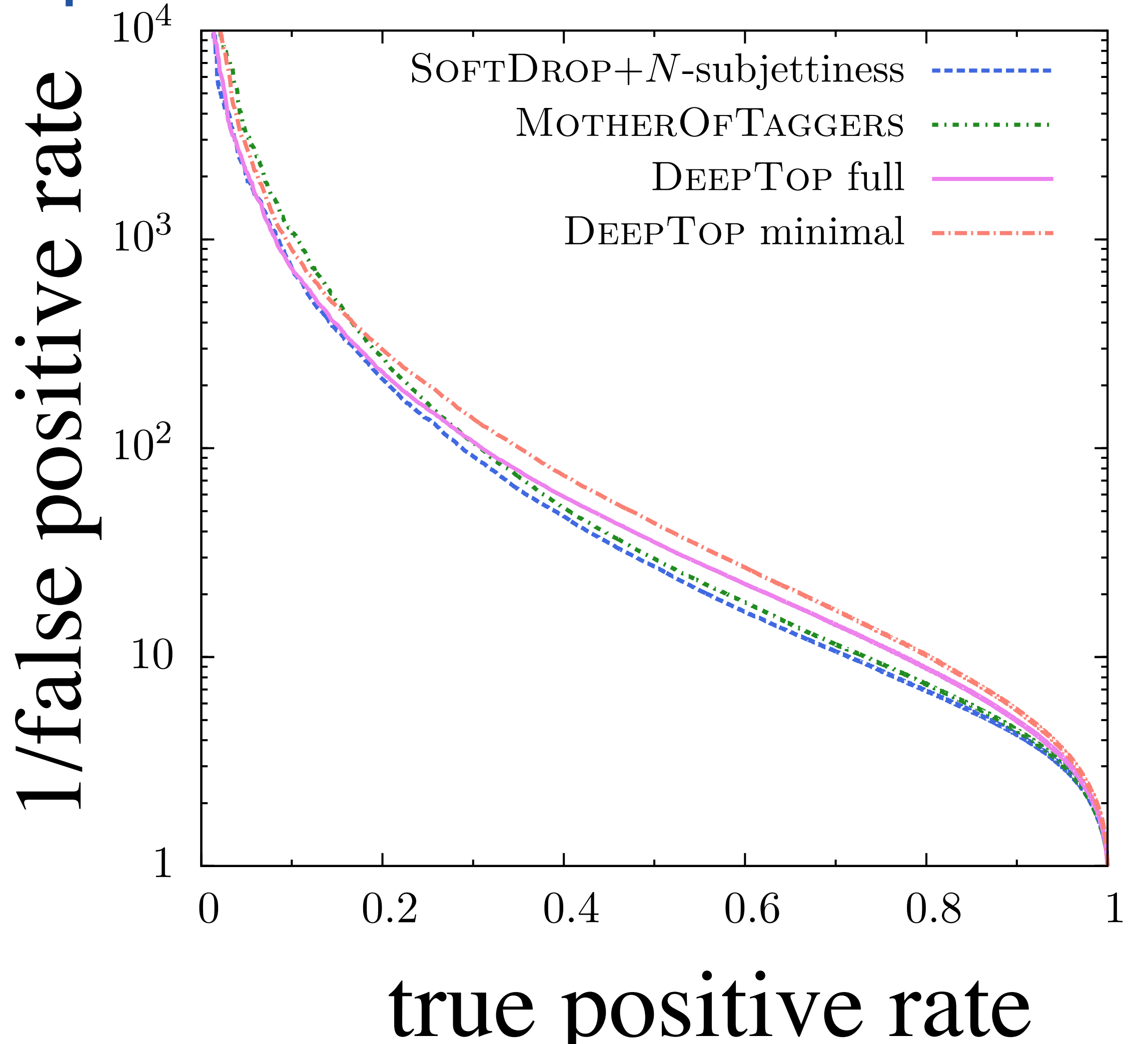
- With large energies (e.g., LHC), jets can also come from H, W, top particles (decaying to jets, which overlap)

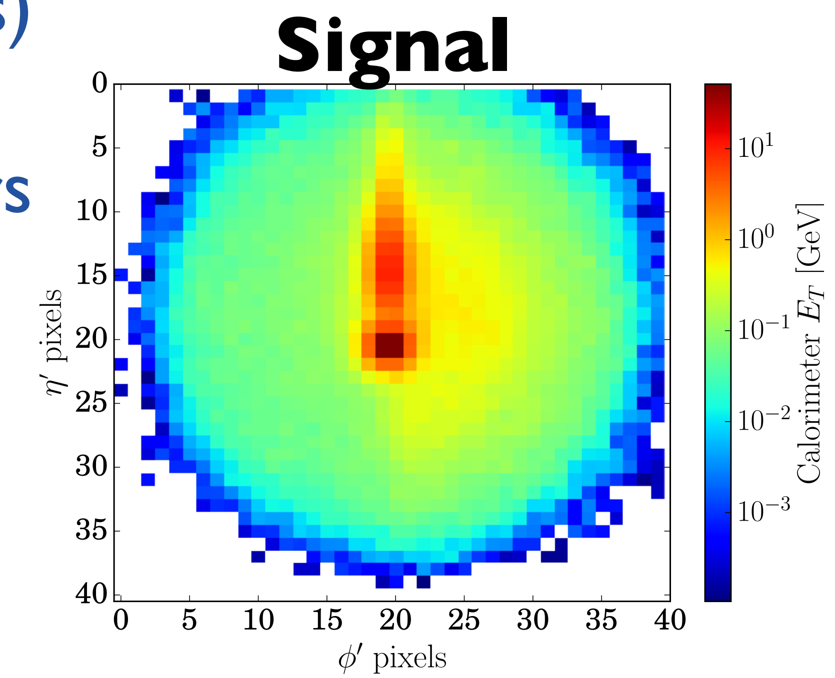
- Several papers in the last two years on DNN solutions to this problem

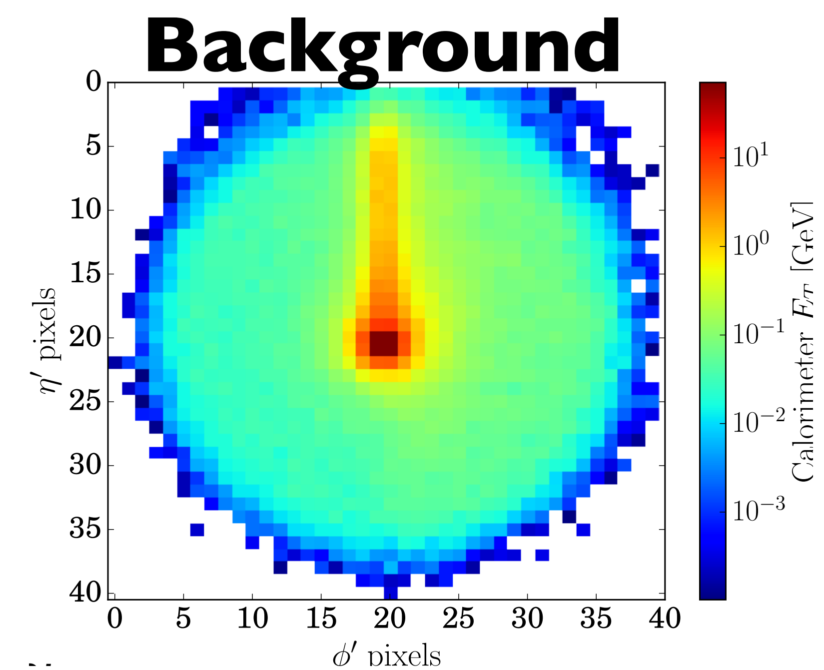


Jet ID with ConvNNs

- Major challenge: irregular detector geometry (vs "regular arrays" assumed in DL applications)
- Jet image processing to "regularise" jet showers and make DL work easier (cantering, rotating, flipping image)
- Good performances on simulated events



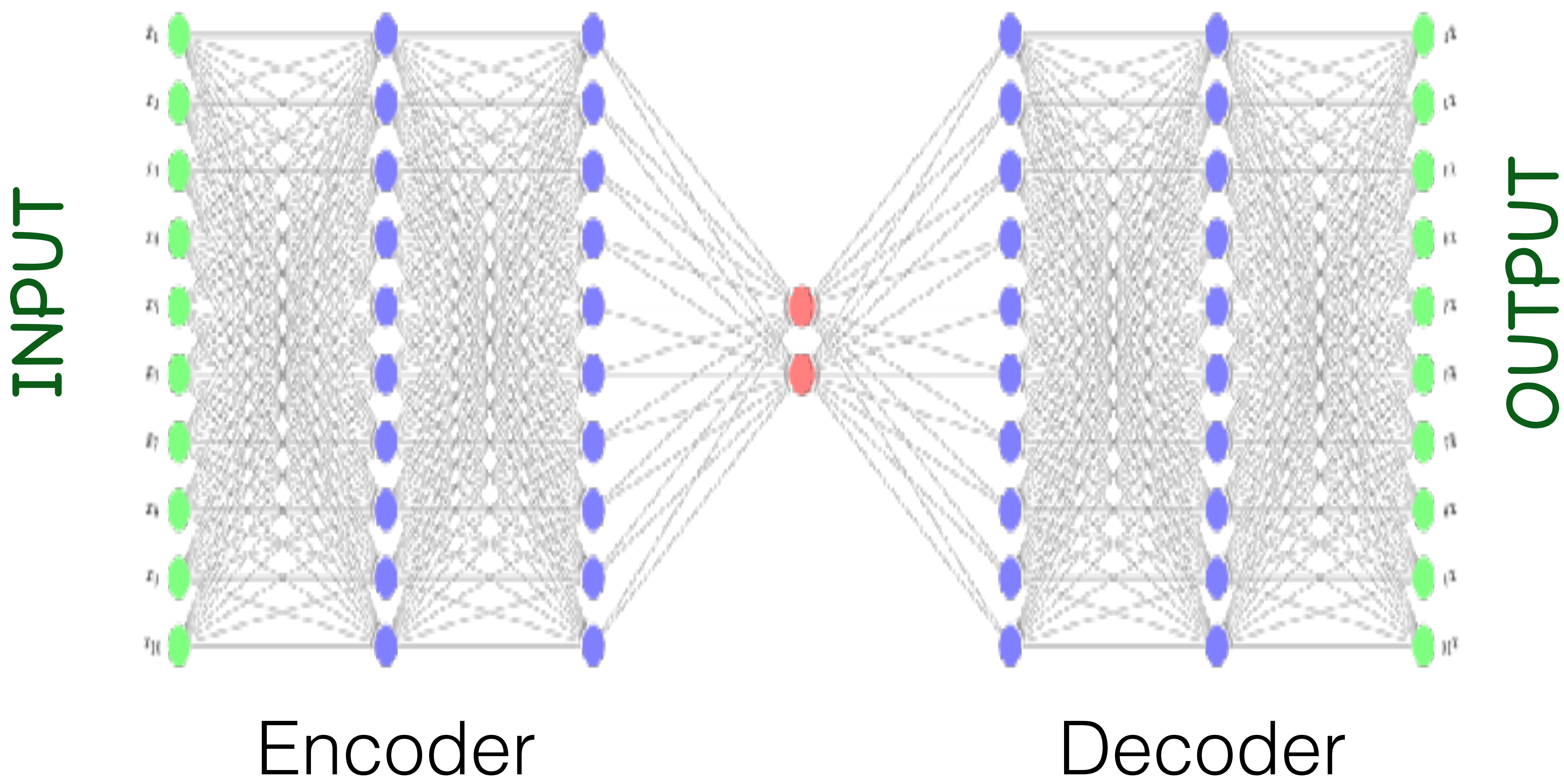




arXiv: 1701.08784

Generative Network

- Use NNs to generate new images from a sample of images
- Example: autoencoders



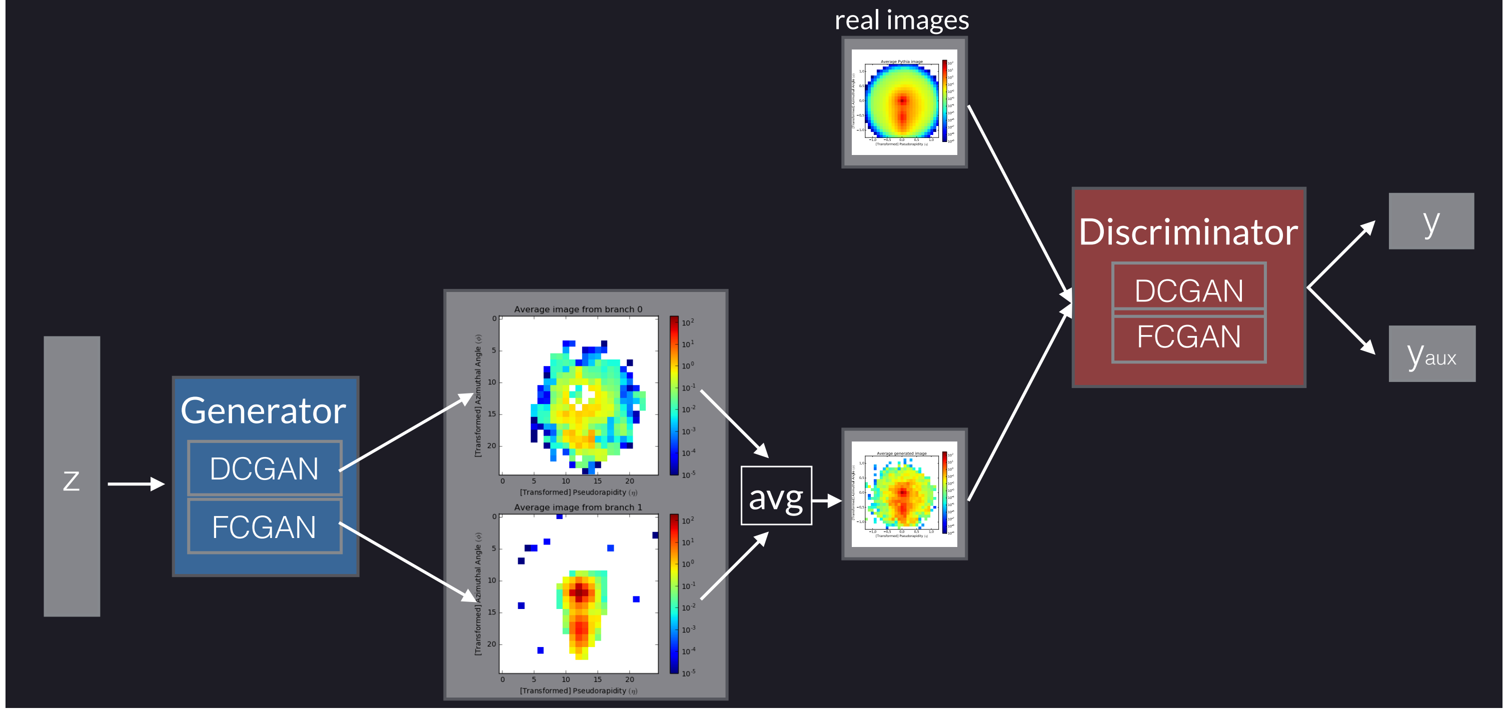
Project the input into an N-dim space
Sample from this N-dim space back into an output
Minimize output-input distance

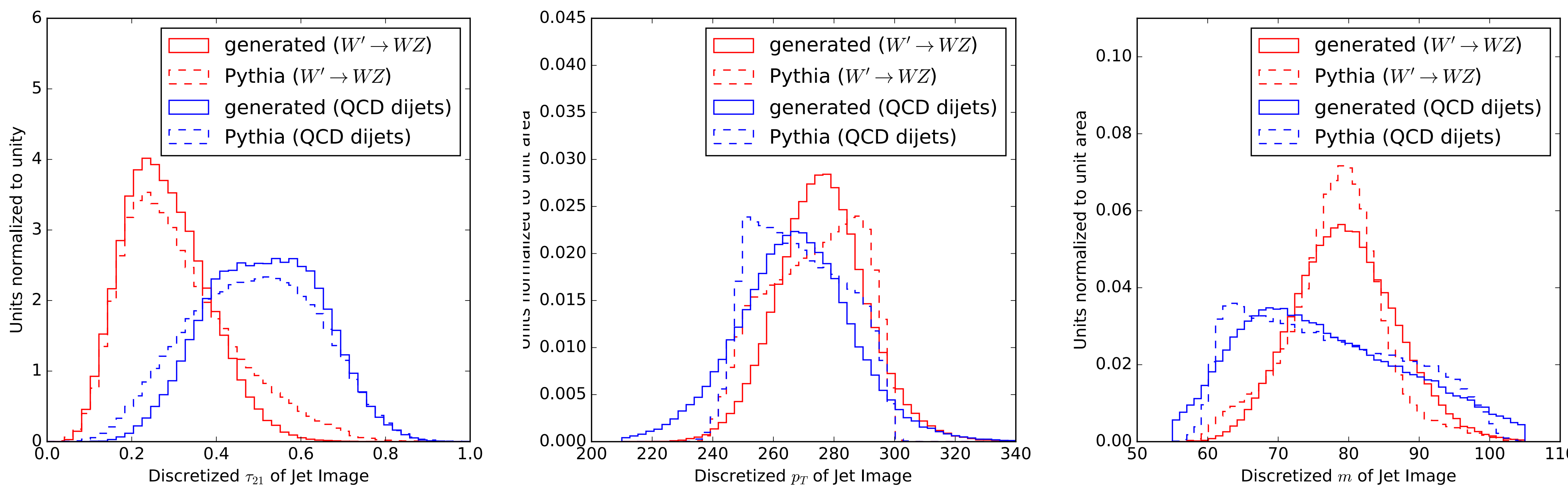
Adversarial Network

- Two networks in competition
  - One generates "fake" data
  - The other one tries to distinguish fake vs real data
- If the second fails, the first is good
- Can generate images

GANs for Jets

- DCGAN — convolutional layers in both *G* and *D*
- FCGAN — fully-connected layers in both *G* and *D*
- HYBRIDGAN — a combination of the two:

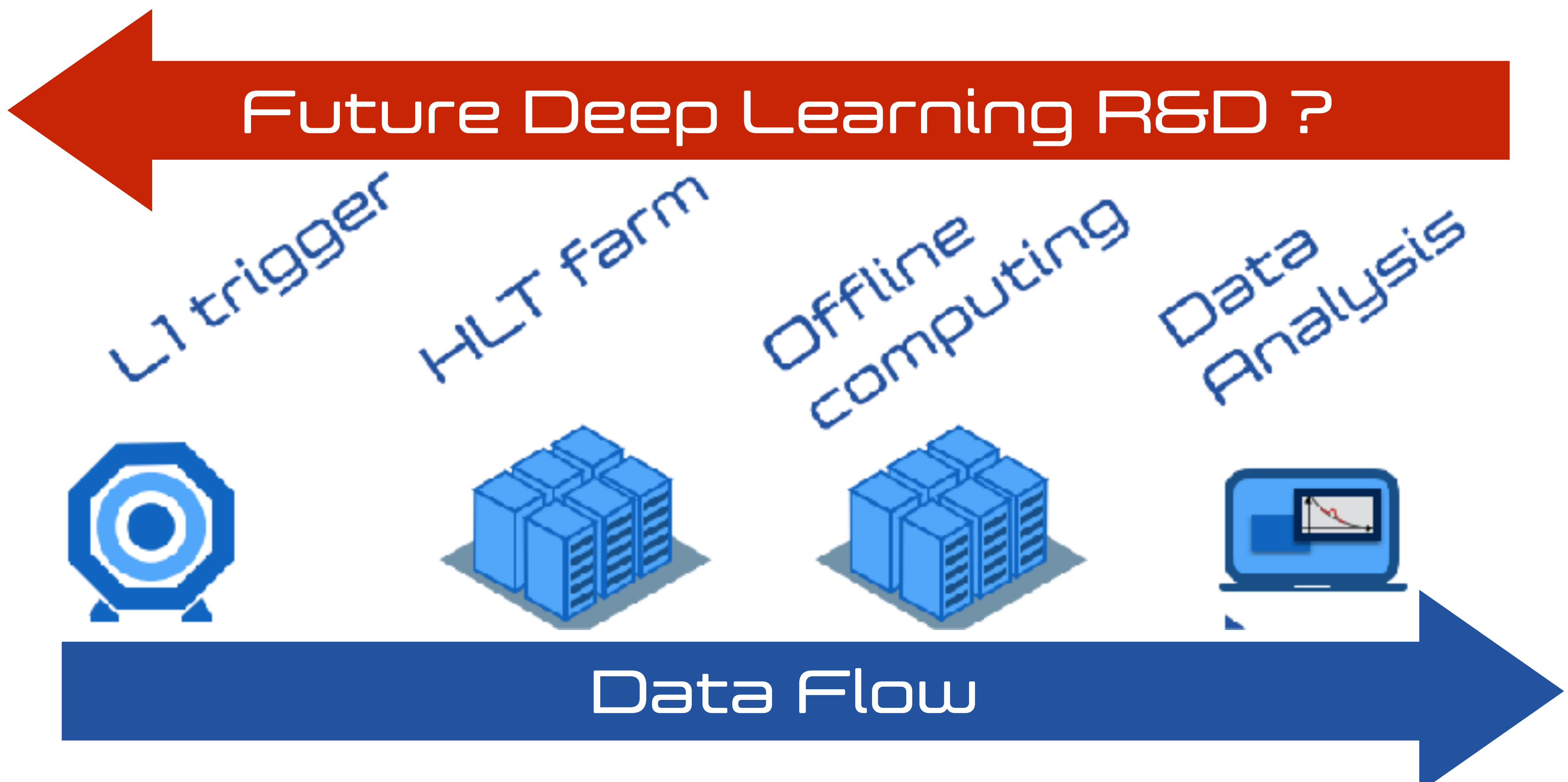




arXiv:1701.05927

Better? Faster? Both?

- We will sue Deep Learning to make reconstruction and selection faster
- We will move it to trigger layers
  - We will trigger faster
  - We will trigger better
  - We will save resources
  - We will automatise many tasks



Conclusions

- LHC experiments represent the ultimate technological advance in particle physics
  - very complicated conditions
  - very broad range of tasks to accomplish
- We are doing great (Higgs boson discovery) but this is not enough (no new physics yet)
- Future ahead challenges
  - More needs, because of more chaotic environment
  - Less resources (budget for science decreasing)
- We need to change approach
  - Looking fwd to Deep Learning as a way out