PART 3

SMARTHEP
REAL-TIME ANALYSIS FOR
SCIENCE AND INDUSTRY

What will this lecture be about?

Introduction

- Definitions and basic concepts

Input to the physics

- The data: trigger, data preparation
- The theory: Monte carlo simulations
- **Reconstruction, or how to translate detector signals to particles**

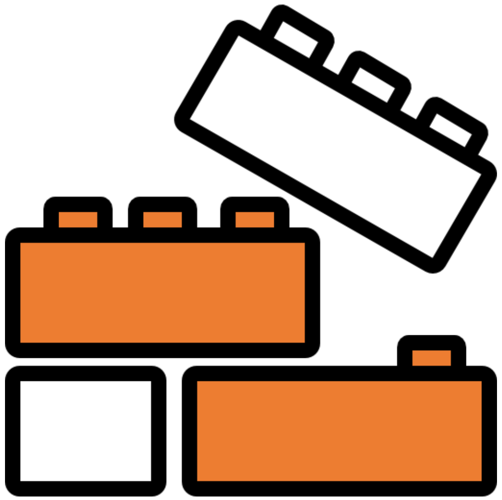
Physics analyses

- Through example, step-by-step
- Discussion of analysis methods

SMARTHEP
REAL-TIME ANALYSIS FOR SCIENCE AND INDUSTRY

Is there a topic you would like to add to this material?
If so: please let me know at the end of this lecture and I will see if I can add it!

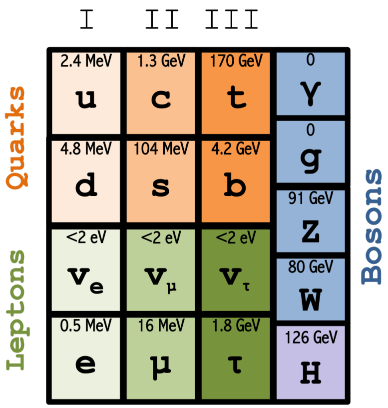
Reconstruction

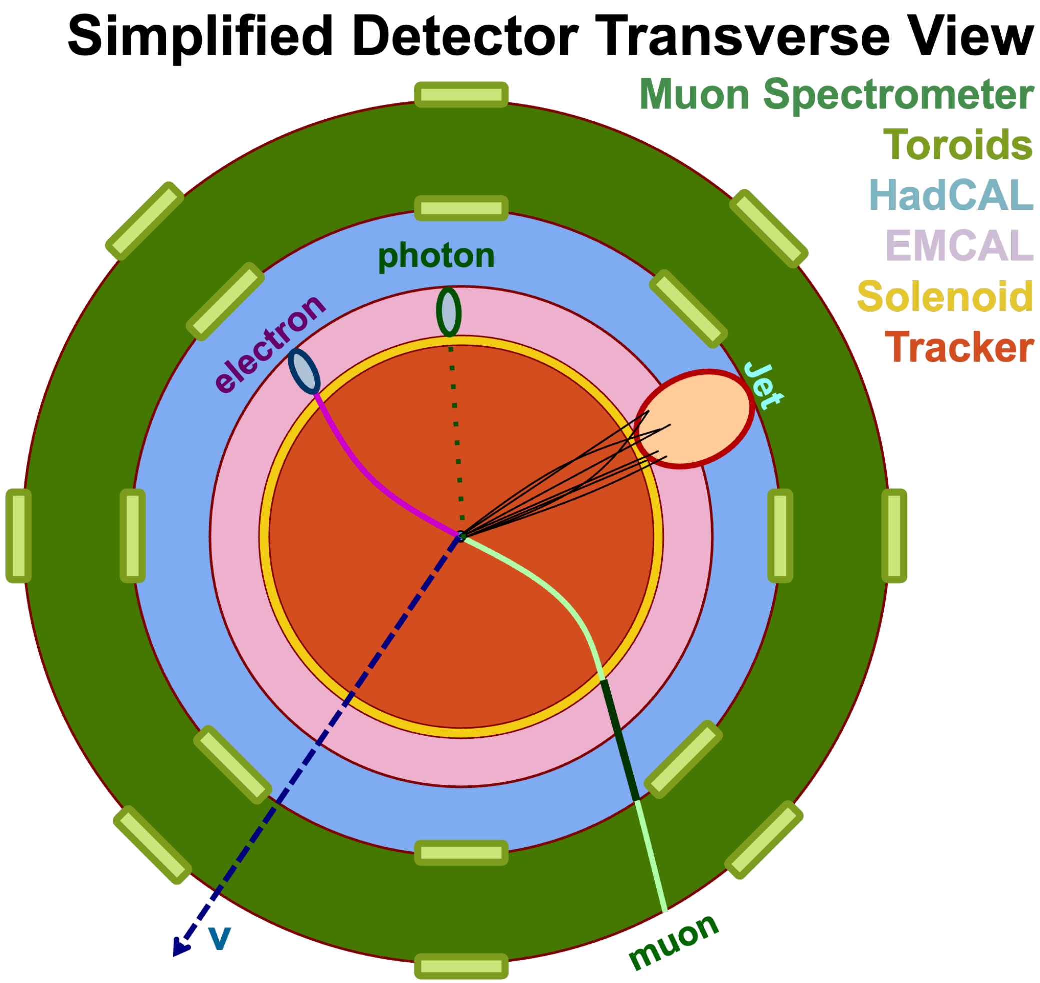


What do we reconstruct?

- Tracks and clusters

- Combining those:
  - "objects", i.e. "particles"

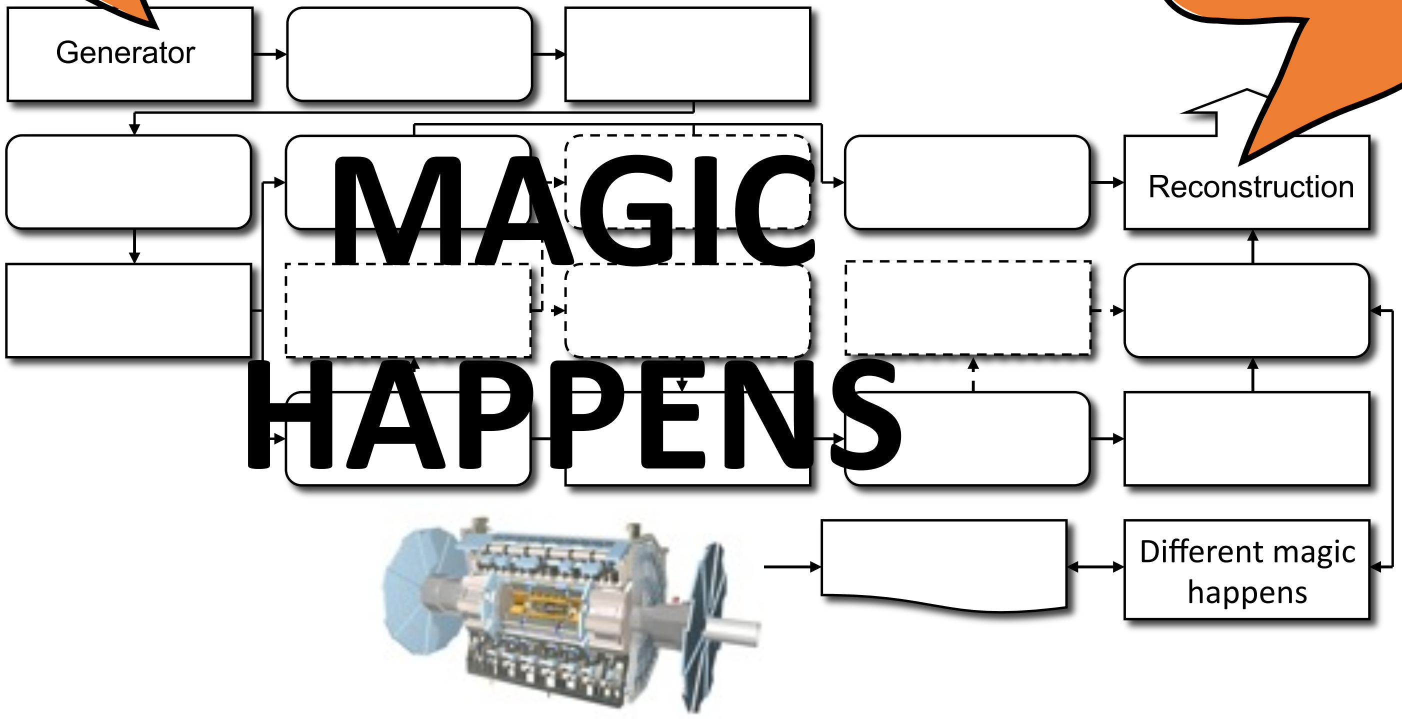




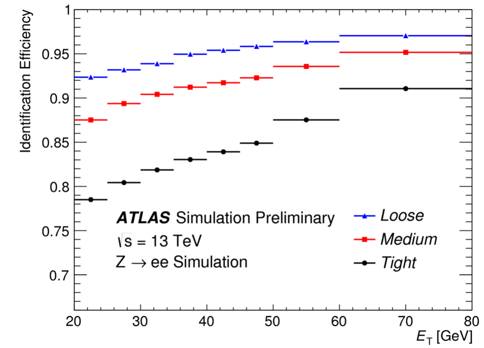
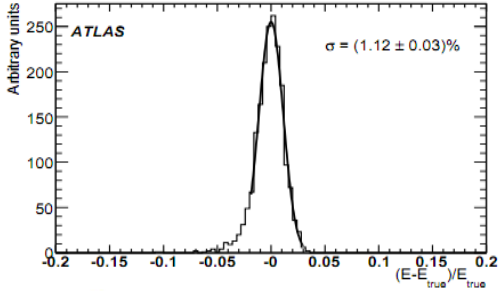
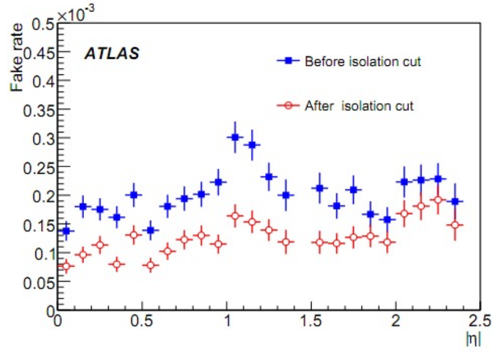
Reconstruction – Figures of Merit

“True” quantities
i.e. quantities at MC generator level

“Reconstructed” quantities
i.e. quantities after having run detector simulation, digitization and reconstruction

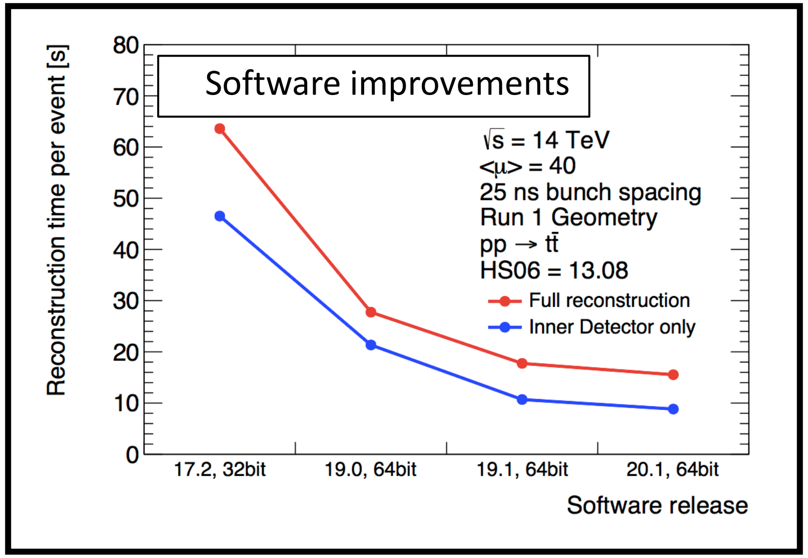


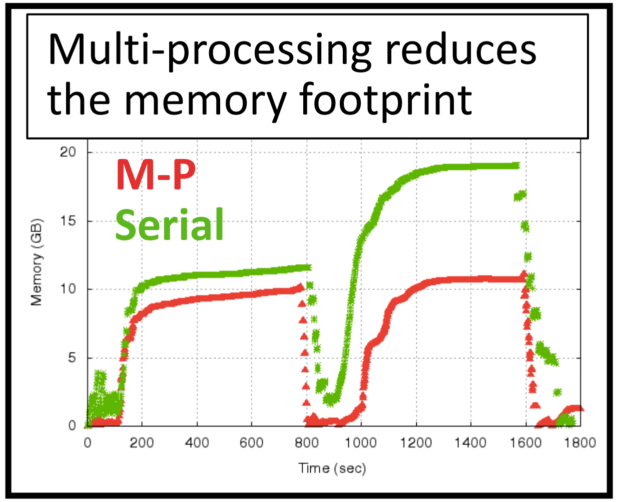
Reconstruction – Figures of Merit

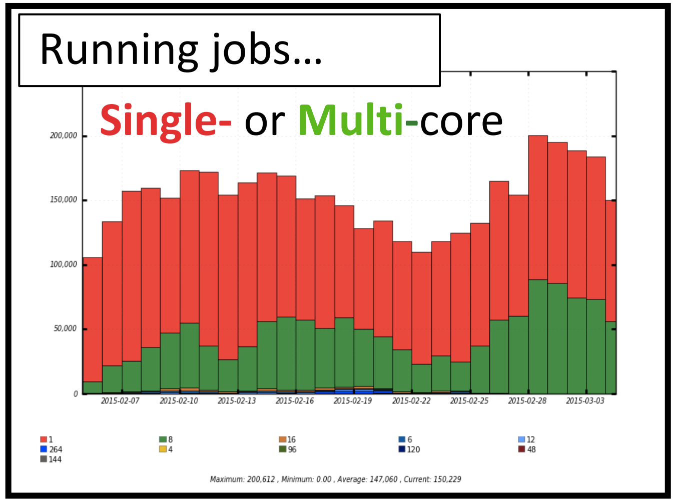
	Definition	Example		Needs be:
Efficiency	how often do we reconstruct the object we are interested in	**electron identification efficiency** = (number of reconstructed electrons) / (number of true electrons) in bins of transverse momentum		**High**
Resolution	how accurately do we reconstruct the quantity	**energy resolution** = (measured energy – true energy) / (true energy)		**Good** *(a small number)*
Fake Rate	how often we reconstruct a different object as the object we are interested in	**a jet faking an electron, fake rate** = (Number of jets reconstructed as an electron) / (Number of jets) in bins of pseudorapidity		**Low**

Reconstruction – goals

- **High efficiency**
- **Good resolution**
- **Low fake rate**
- **Robust against detector problems and data-taking conditions:**
  - Noise
  - Dead regions of the detector
  - Increased pile-up
- **Computing-friendly** →
  - CPU time per event
  - Memory use



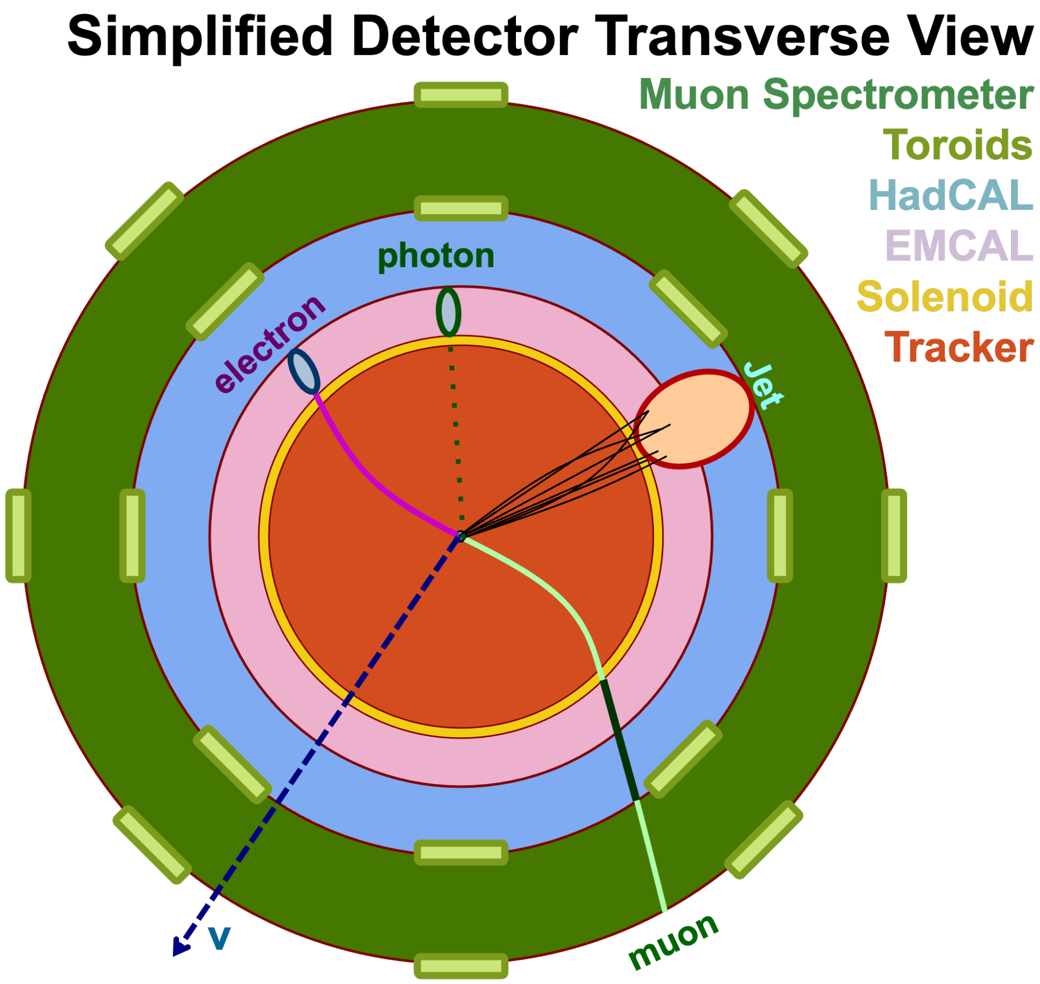




What do we reconstruct?

Tracks and clusters

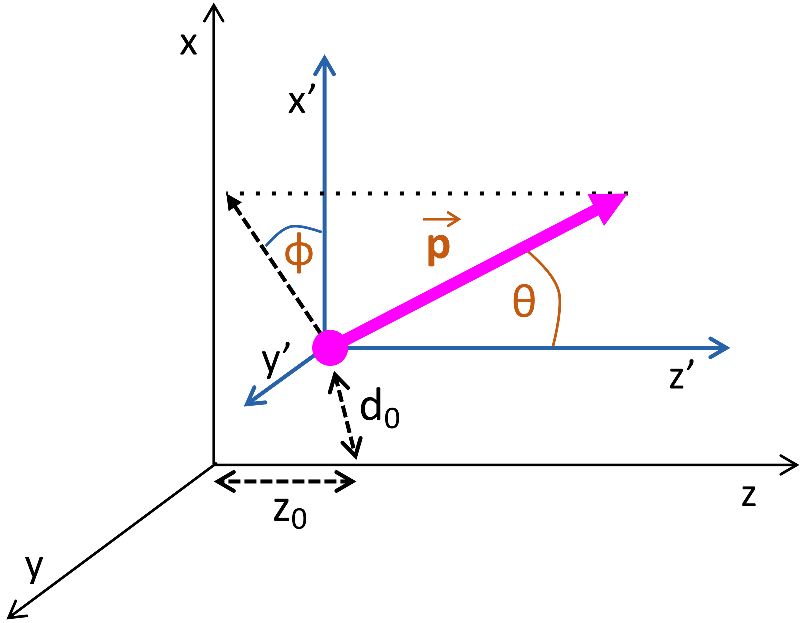
- Combining those:
  - "objects", i.e. "particles"



Tracking in a Nutshell

- **A track represents a measurement of a charged particle that leaves a trajectory as it passes through the detector.**

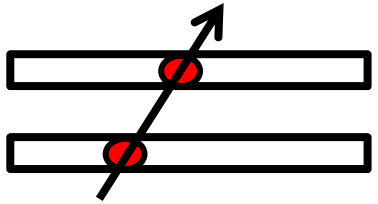
- **For a track we measure:**
  - Its momentum;
  - Its direction;
  - Its charge;
  - Its "perigee": the closest point to a reference line, transverse (d_0) or longitudinal (z_0).



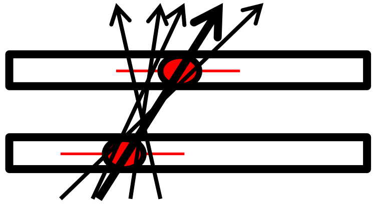
- Tracks are key ingredients of most of particle reconstruction.

Tracking in a Nutshell – Track Fitting

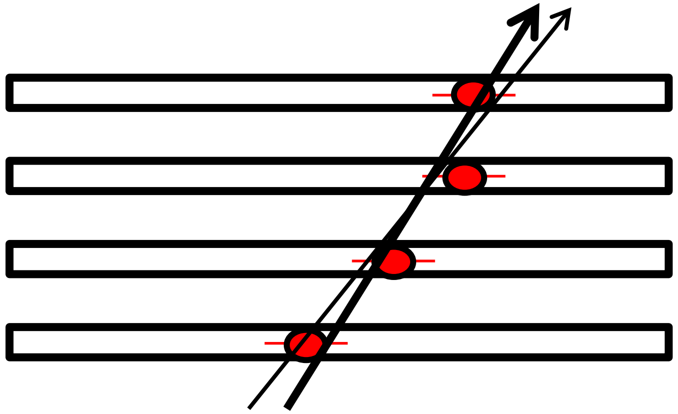
- **Perfect measurement – ideal**



- **Imperfect measurement – reality**



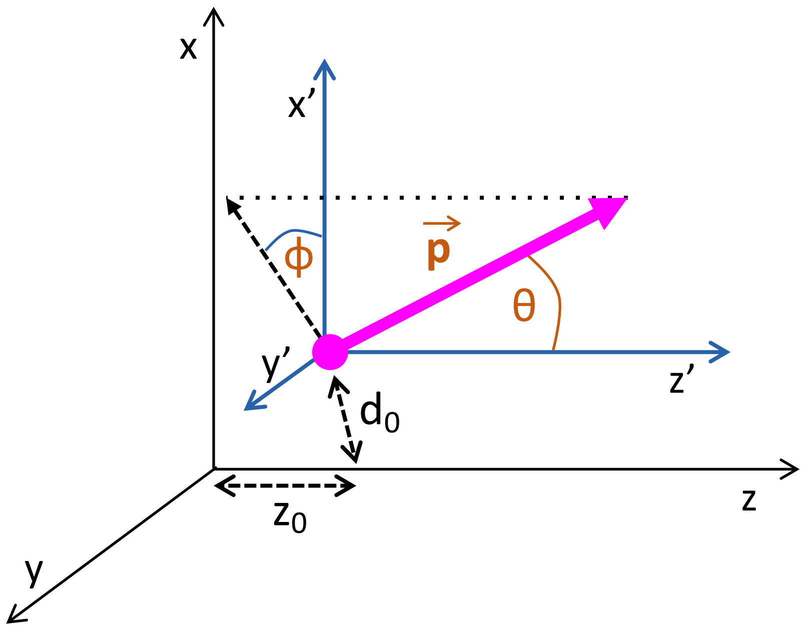
- **Small errors and more points help to constrain the possibilities**



- **Quantitatively:**
  - Parameterize the track;
  - Find parameters by Least-Squares-Minimization;
  - Obtain also uncertainties on the track parameters.

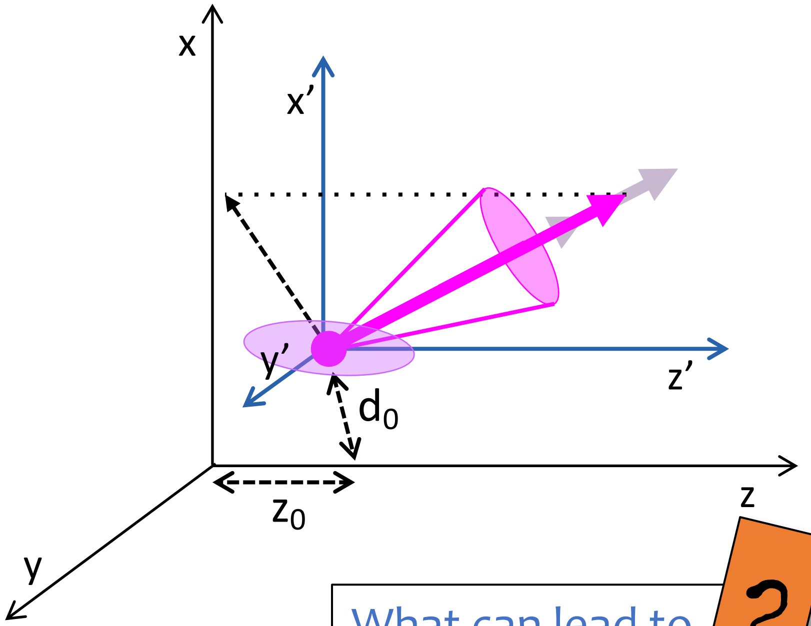
Tracking in a Nutshell – Track Fitting

- For a track we measure:
  - Its momentum;
  - Its direction;
  - Its charge;
  - Its "perigee": the closest point to a reference line, transverse (d_0) or longitudinal (z_0).



Tracking in a Nutshell – Track Fitting

- **For a track we measure:**
  - Its momentum;
  - Its direction;
  - Its charge;
  - Its "perigee": the closest point to a reference line, transverse (d_0) or longitudinal (z_0).
  - **And their uncertainty**



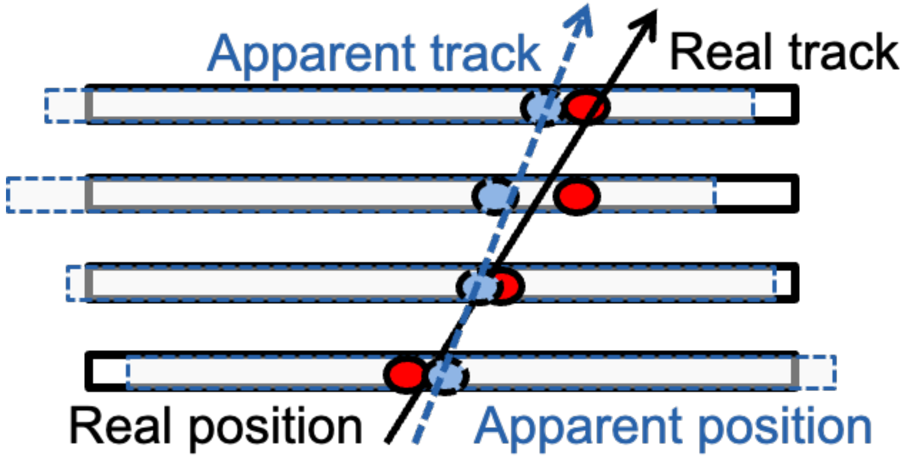
- **Small uncertainties are required.**
  - δdo is < O(10µm) and δθ < O(0.1mrad).
  - Allows separation of tracks that come from different particle decays (which can be separated at the order of mm).

What can lead to uncertainties?

Tracking in a Nutshell – the Uncertainties

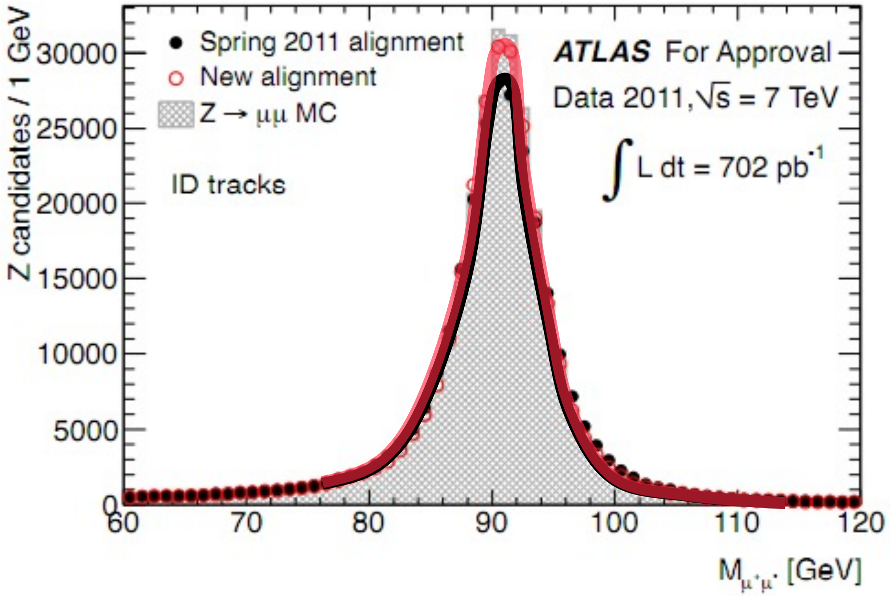
- Presence of Material
  - Coulomb scattering off the core of atoms
  - Energy loss due to ionization
  - Bremsstrahlung
  - Hadronic interaction

- Misalignment
  - Detector elements not positioned in space with perfect accuracy.
  - Alignment corrections derived from data and applied in track reconstruction.



Impact of good alignment

- Improving the tracker alignment description in the reconstruction gives better track momentum resolution which leads to better mass resolution.



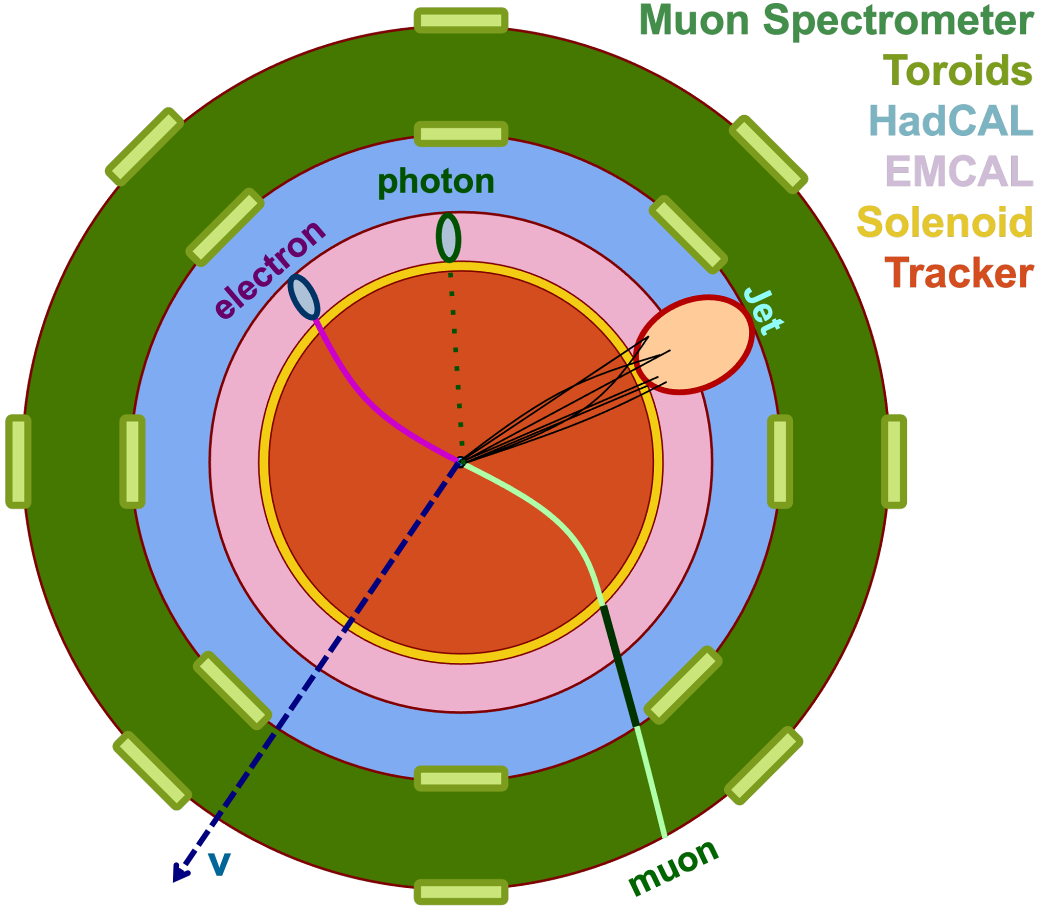
- Can see the reconstructed Z width gets narrower if we use better alignment constants. Very important for physics analysis to have good alignment.
- Alignment of detector elements can change with time, for example when the detector is opened for repair, or when the magnetic field is turned on and off.

What do we reconstruct?

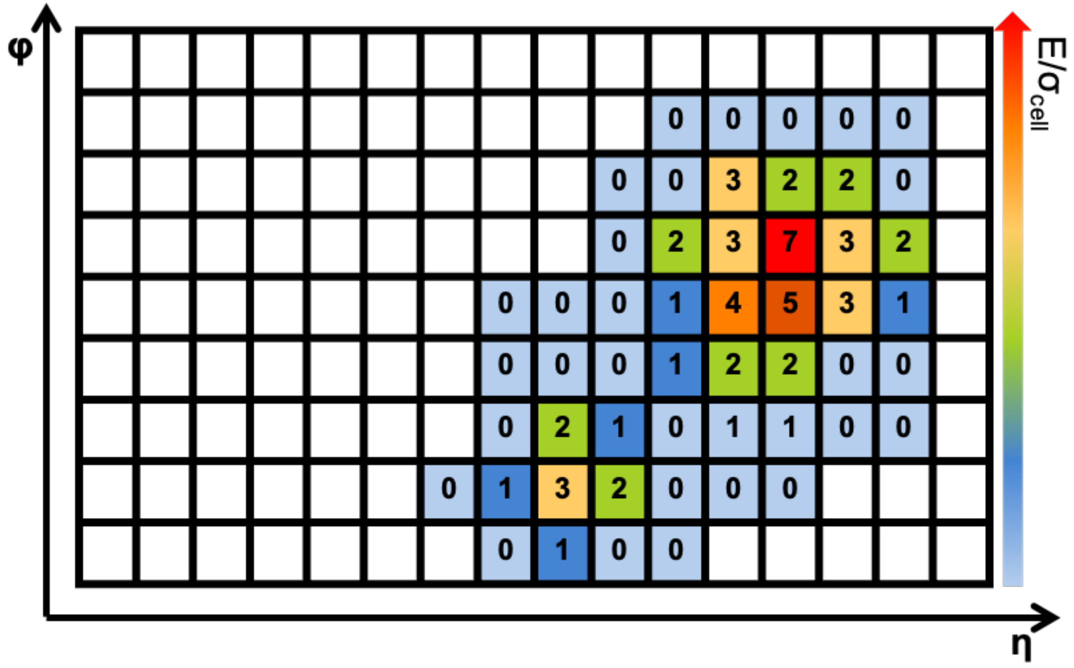
Tracks and clusters

- Combining those:
  - "objects", i.e. "particles"

Simplified Detector Transverse View



A calorimeter view

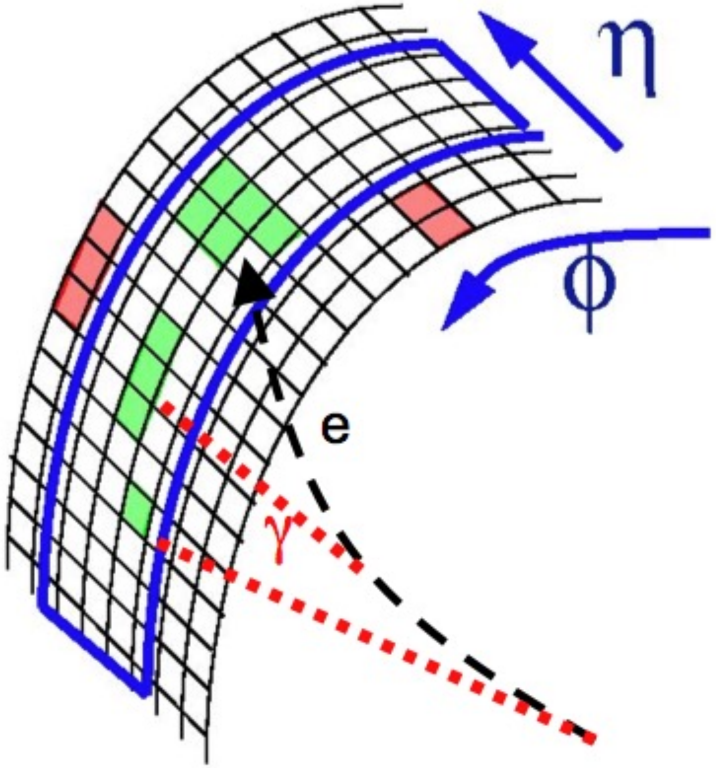


Clustering in a Nutshell

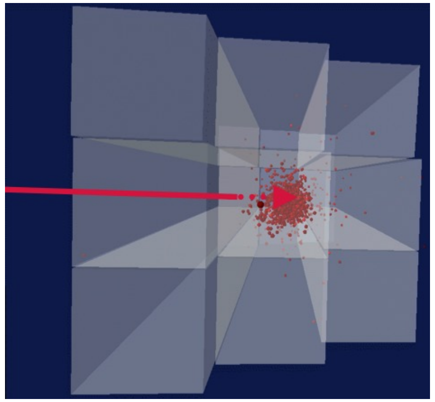
- **Reconstruct energy deposited in the calorimeter by charged or neutral particles; electrons, photons and jets.**
- **For a cluster we measure:**
  - The energy;
  - The position of the deposit;
  - The direction of the incident particles;
- **Calorimeters are segmented in cells.**
  - Typically, a shower created by a particle interacting with the matter extends over several cells.
- **Various clustering algorithms, e.g.:**
  - **Sliding window**. Sum cells within a fixed-size rectangular window.
  - **Topo-clustering**. Start with a seed cell and iteratively add to the cluster the neighbor of a cell already in the cluster.

Cluster Finding – an Example

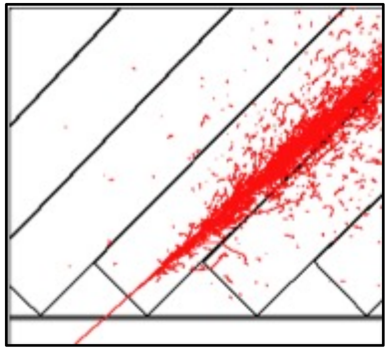
- CMS crystal calorimeter – ECAL clusters
  - electron energy in central crystal ~80%, in 5x5 matrix around it ~96%.



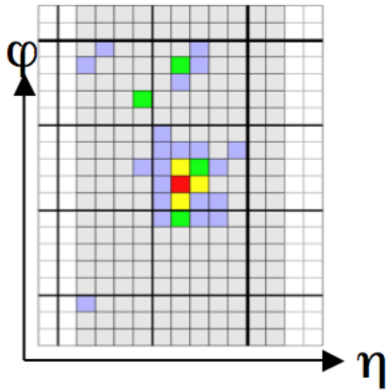
Front view



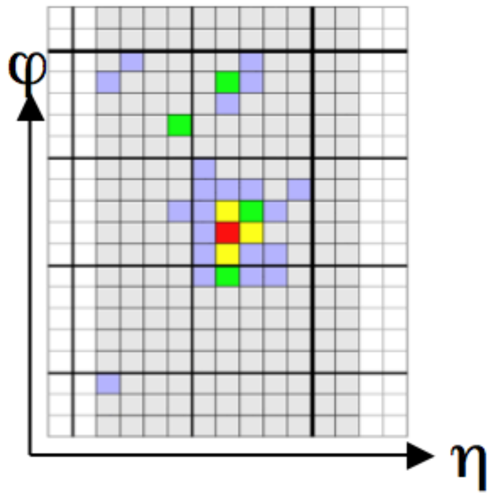
Side view

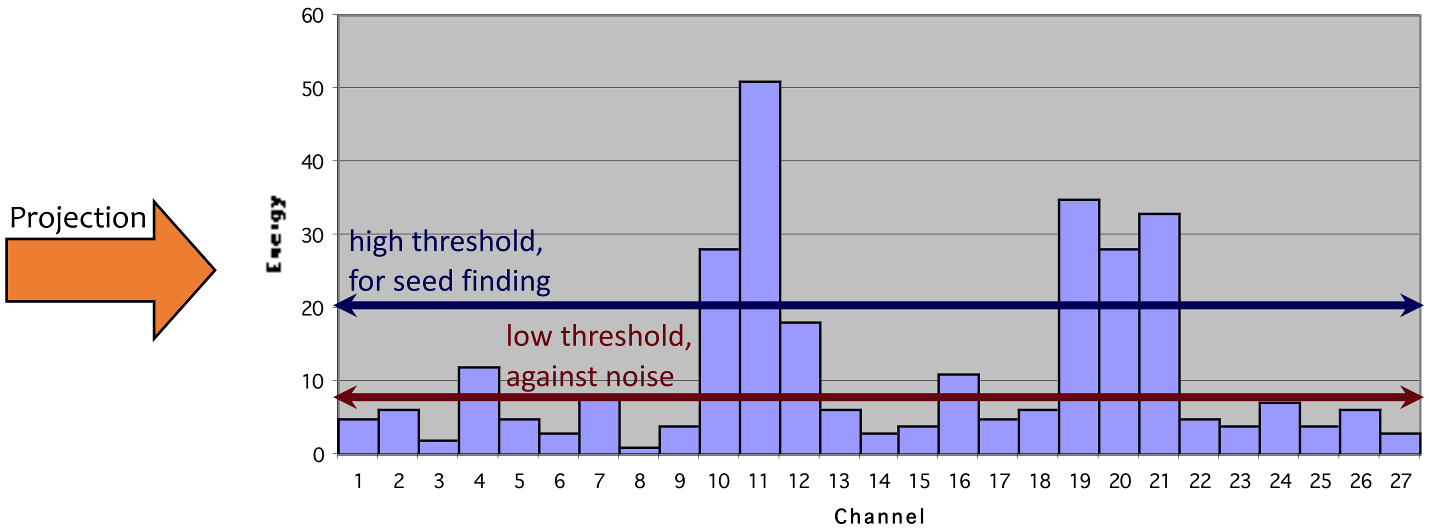


η-ϕ view



Cluster finding – an example



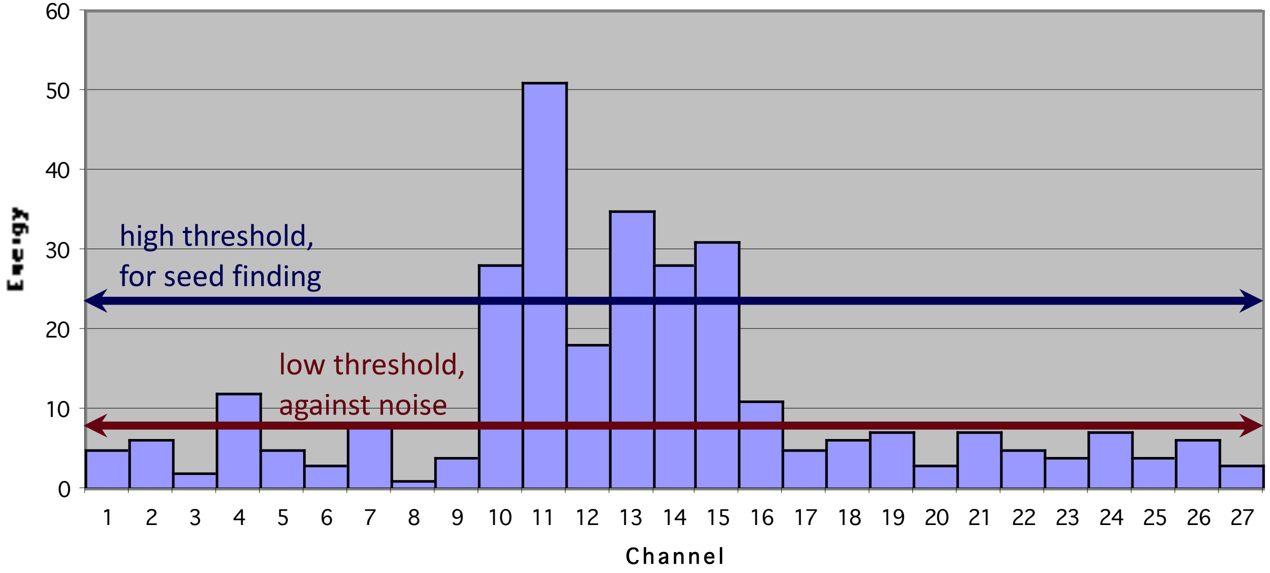


- **Simple example of an algorithm**
  - Scan for seed crystals = local energy maximum above a defined seed threshold
  - Starting from the seed position, adjacent crystals are examined, scanning first in φ and then in η
  - Along each scan line, crystals are added to the cluster if
    - The crystal's energy is above the noise level (lower threshold)
    - The crystal has not been assigned to another cluster already

Cluster finding – an example: Difficulties

- Careful tuning of thresholds needed.
  - needs usually learning phase;
  - adapt to noise conditions;
  - too low : pick up too much unwanted energy;
  - too high : loose too much of “real” energy. Corrections/Calibrations will be larger.

example : one lump or two?

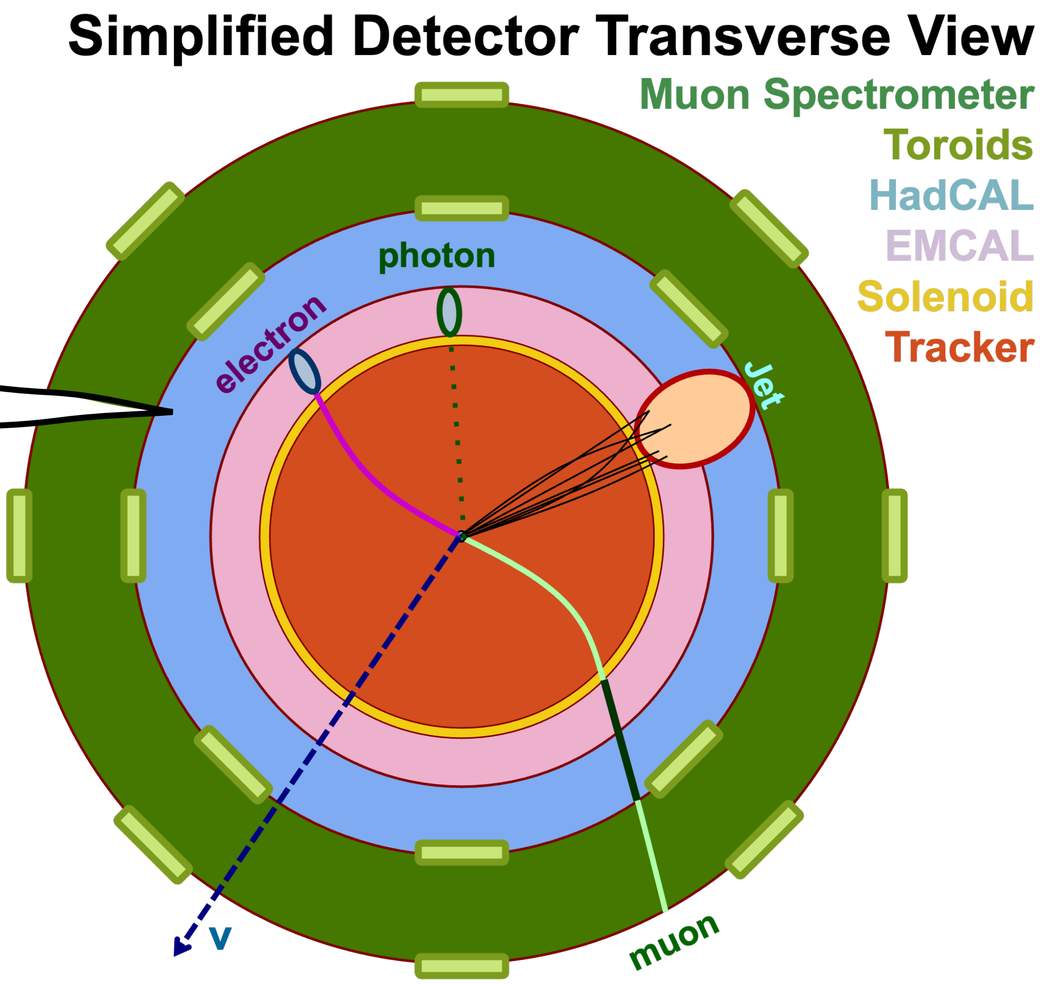


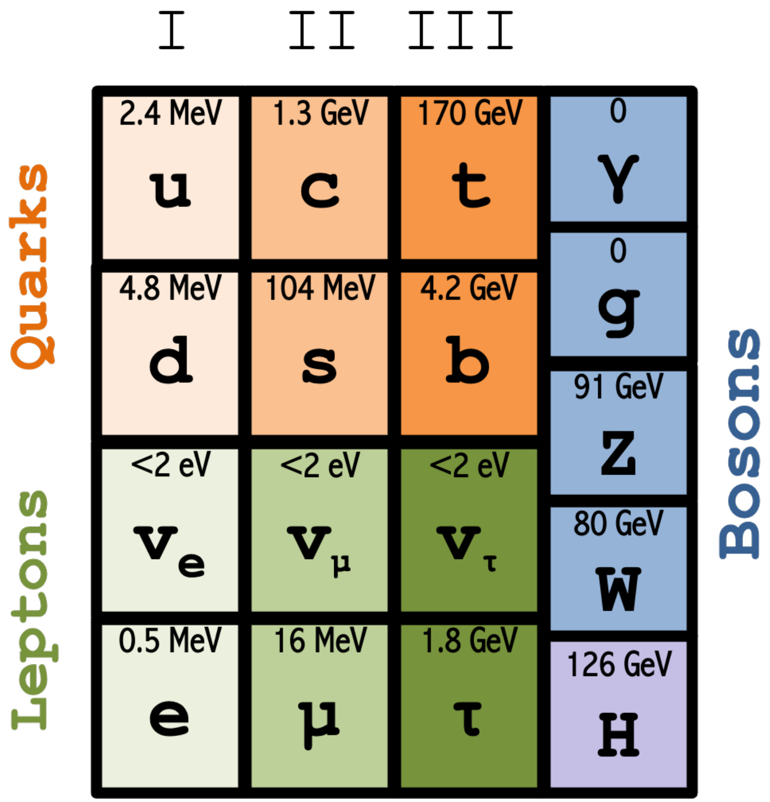
What do we reconstruct?

Tracks and clusters

Combining those:

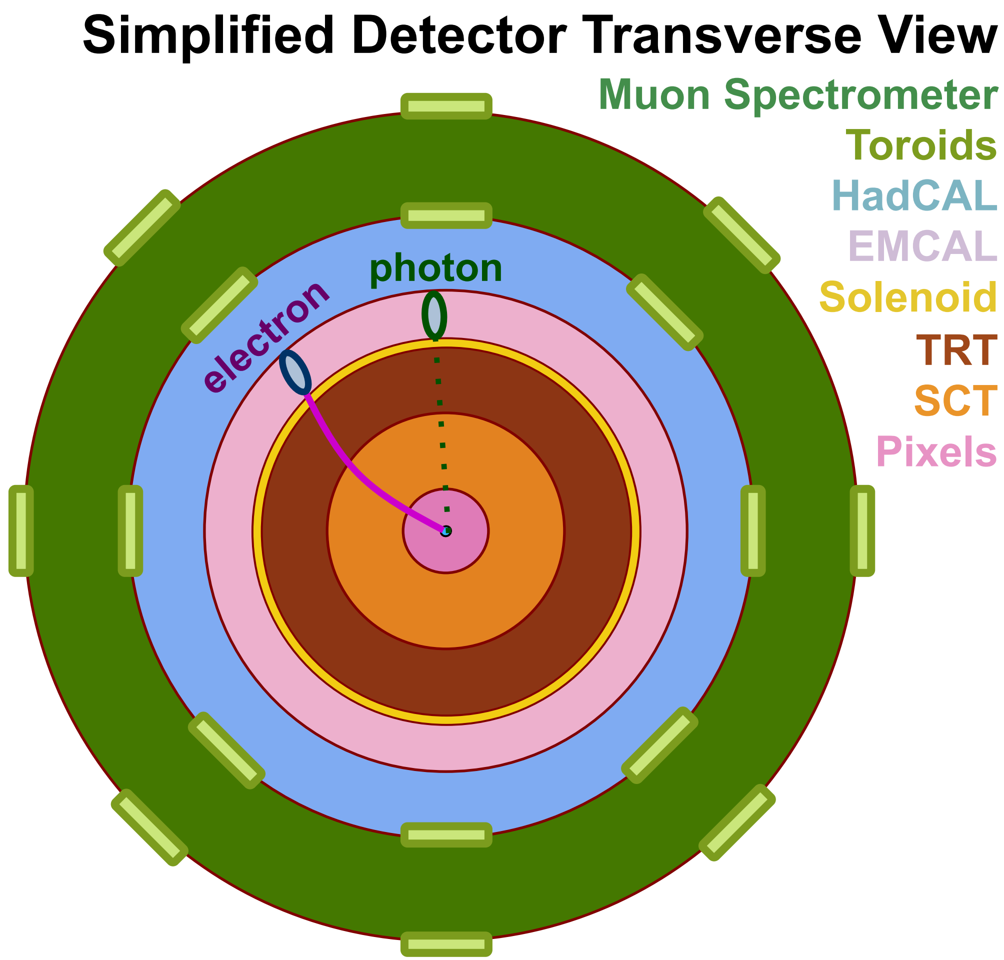
- "objects", i.e. "particles"




I
II
III
Quarks
Leptons
Bosons
2.4 MeV
u
1.3 GeV
c
170 GeV
t
0
γ
4.8 MeV
d
104 MeV
s
4.2 GeV
b
0
g
91 GeV
Z
<2 eV
ν_e
<2 eV
ν_μ
<2 eV
ν_τ
80 GeV
W
0.5 MeV
e
16 MeV
μ
1.8 GeV
τ
126 GeV
H

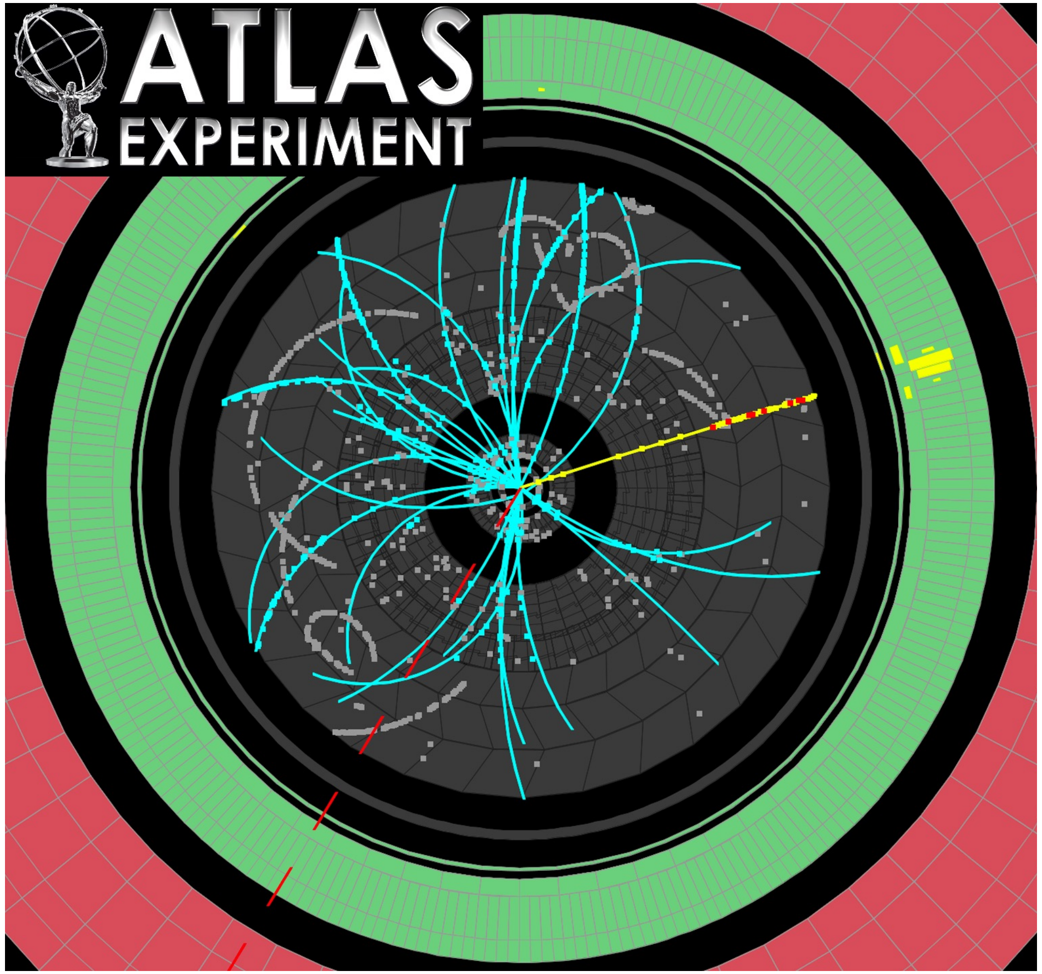
Electrons / Photons

- Final **Electron momentum measurement** can come from tracking or calorimeter information (or a combination of both)
  - Often have a final calibration to give the best electron energy
- Working points define categories
  - E.g. loose, medium, tight
  - **Trade-off**: Efficiency vs Fakes
- Often want "**isolated electrons**"
  - Require little calorimeter energy or tracks in the region around the electron



Electrons / Photons

- Final **Electron momentum measurement** can come from tracking or calorimeter information (or a combination of both)
  - Often have a final calibration to give the best electron energy
- Working points define categories
  - E.g. loose, medium, tight
  - **Trade-off**: Efficiency vs Fakes
- Often want "**isolated electrons**"
  - Require little calorimeter energy or tracks in the region around the electron



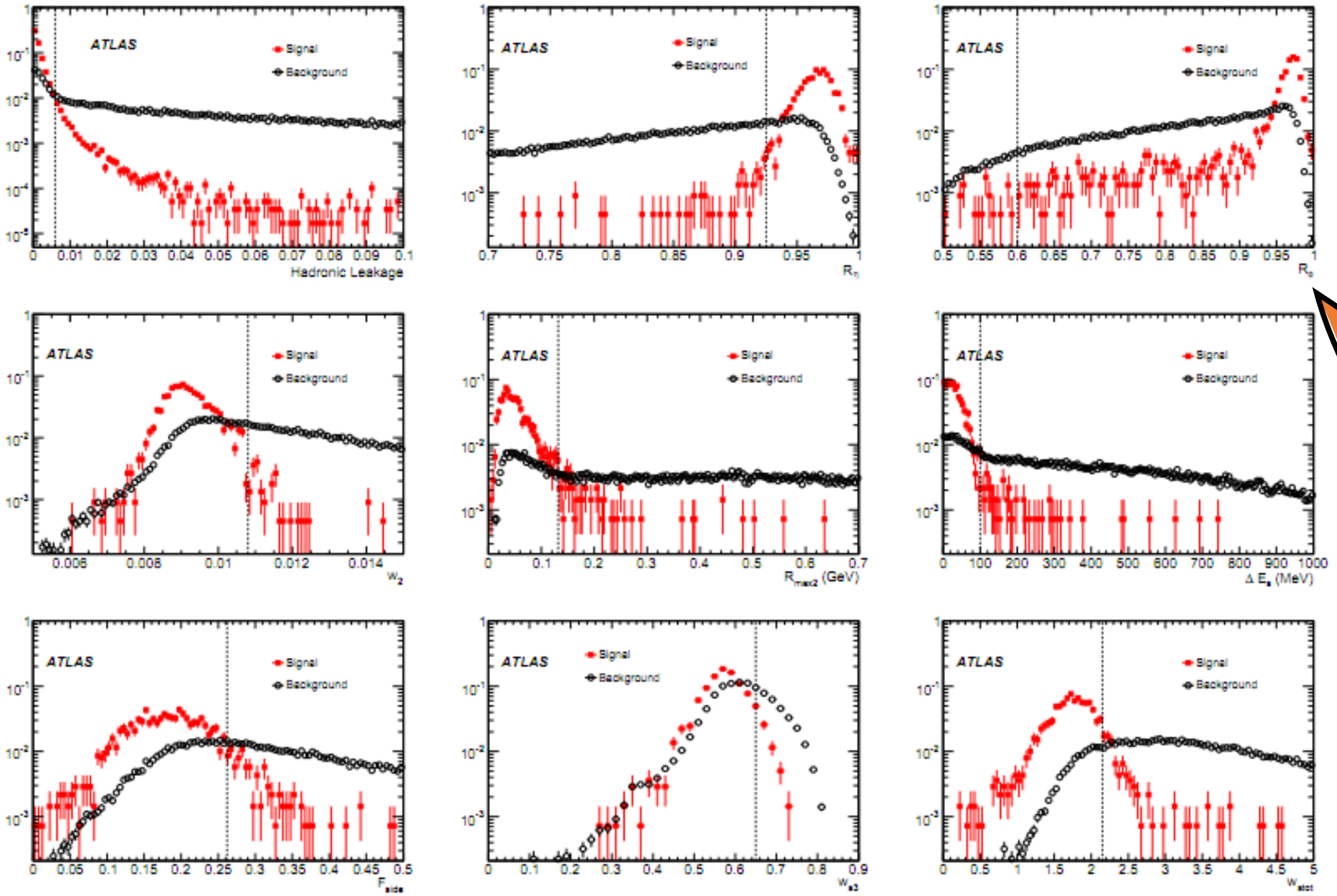
Electrons / Photons – Backgrounds

- **Sources of backgrounds:**
  - Hadronic jets leaving energy in calorimeter

- While calorimeter clusters are much wider for jets than for electrons/photons – there are many thousands more jets than electrons
  - rate of jets faking an electron needs to be very small ($\sim 10^{-4}$)

- **Complex identification algorithms** are required to give the rejection whilst keeping a high efficiency

Electrons – Identification Algos

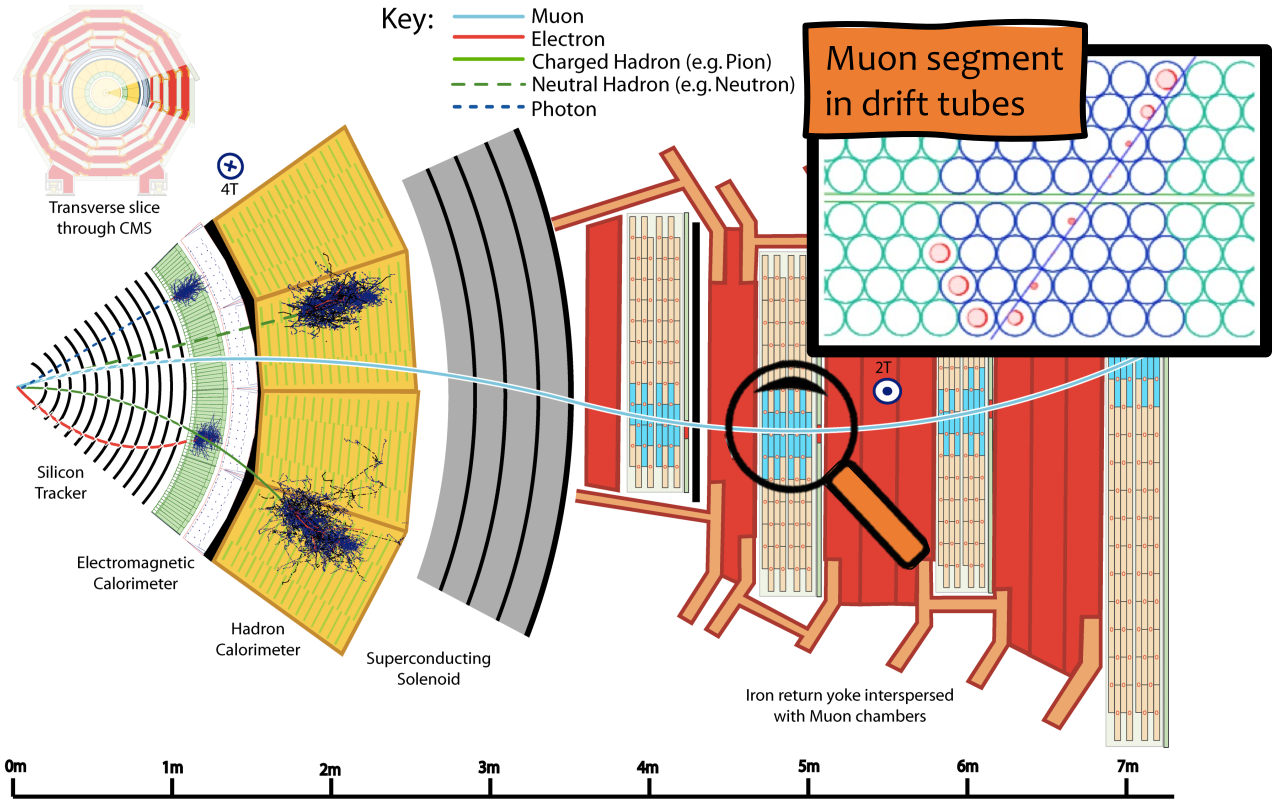


Example of different calorimeter shower shape variables used to distinguish electron showers from jets in ATLAS

Information can be exploited using multi-variate techniques such as **likelihood discriminants** or **boosted decision trees** or **other machine learning methods**.

Muons

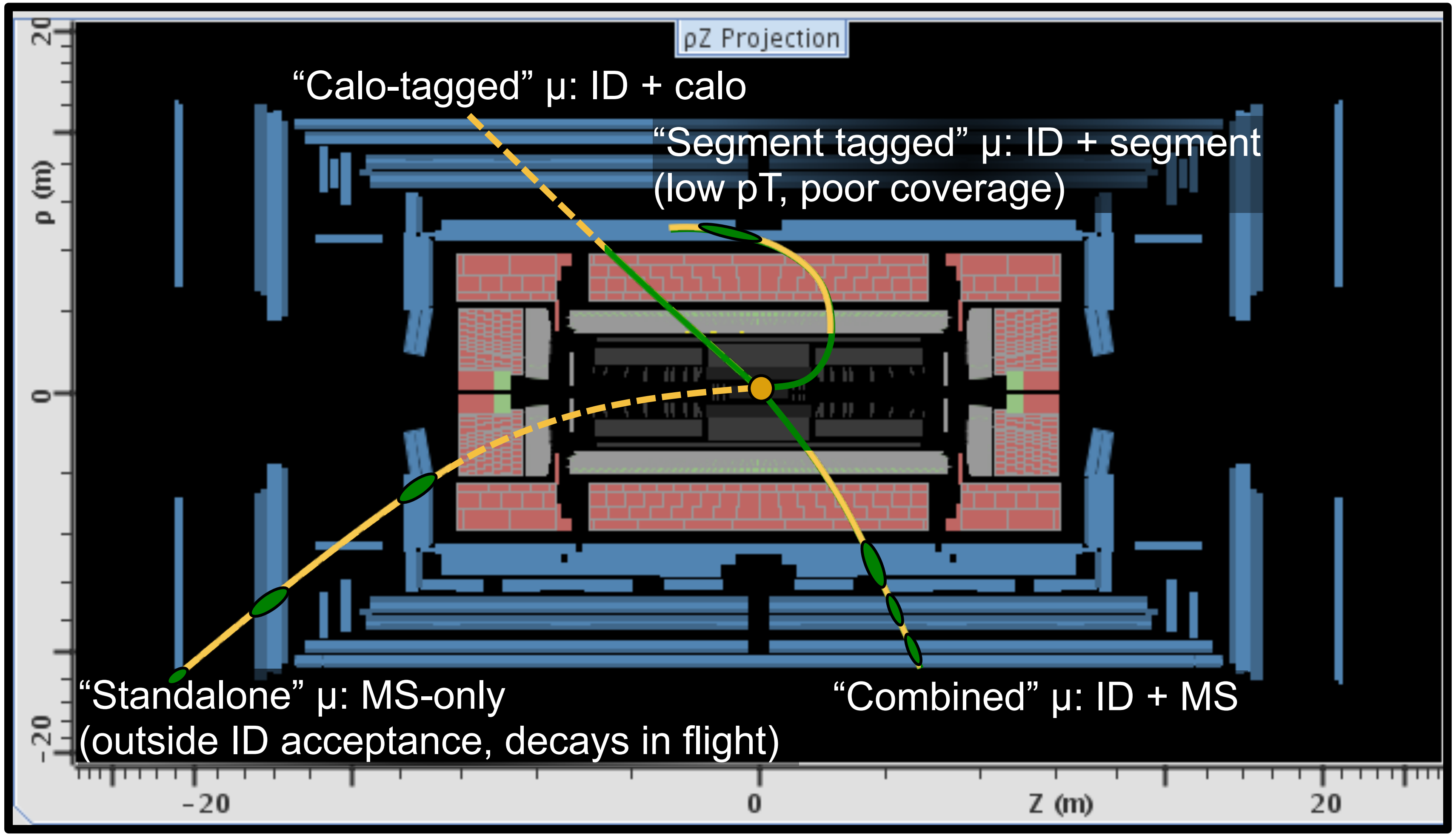
- Combine the **muon segments found in the muon detector with tracks from the tracking detector**
- Momentum of muon determined from bending due to magnetic field in tracker and in muon system
  - Combine measurements to get best resolution
  - Need an accurate map of magnetic field in the reconstruction software
  - Alignment of the muon detectors also very important to get best momentum resolution



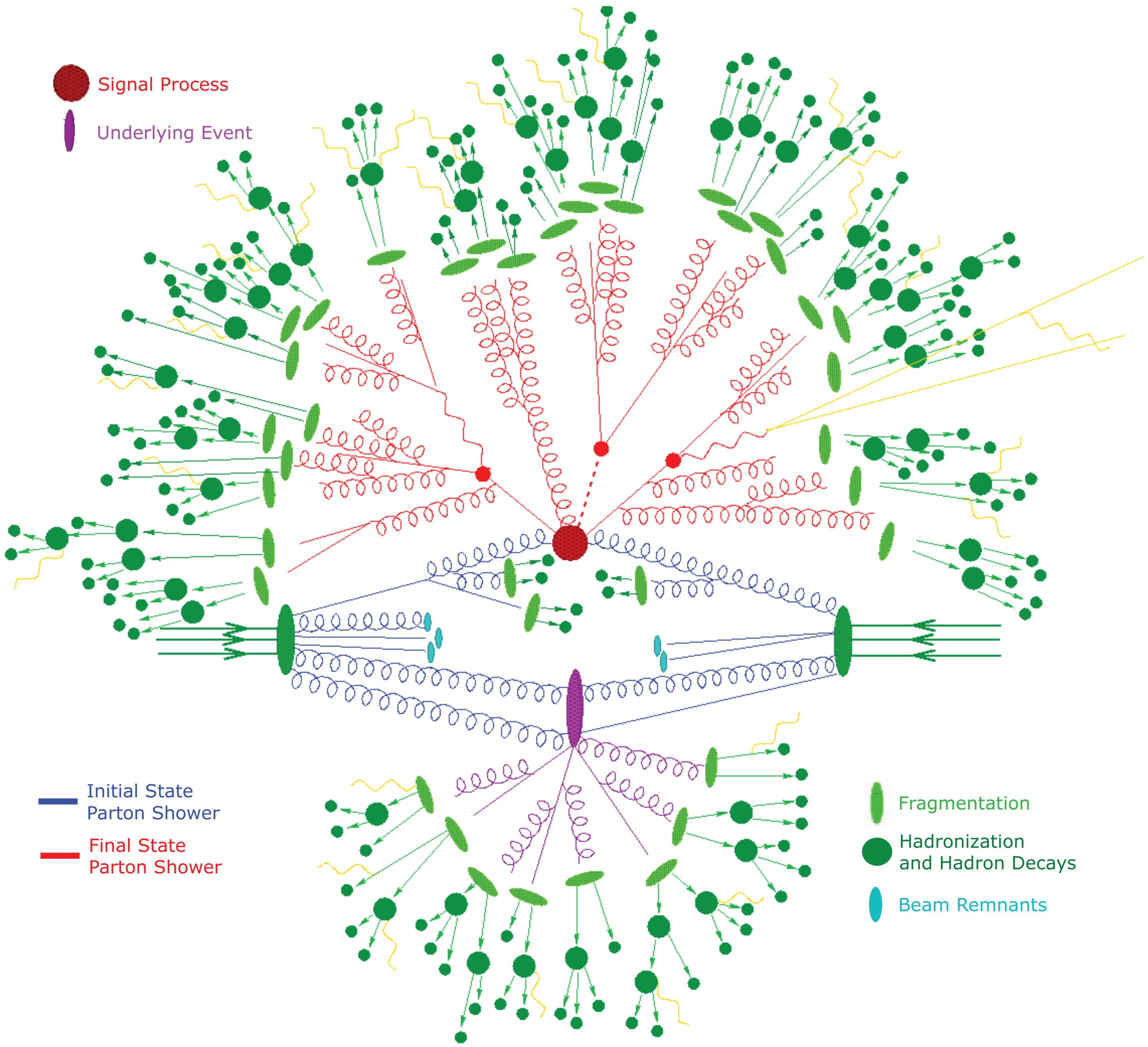
MUONS ON ATLAS

Simplified Detector Transverse View

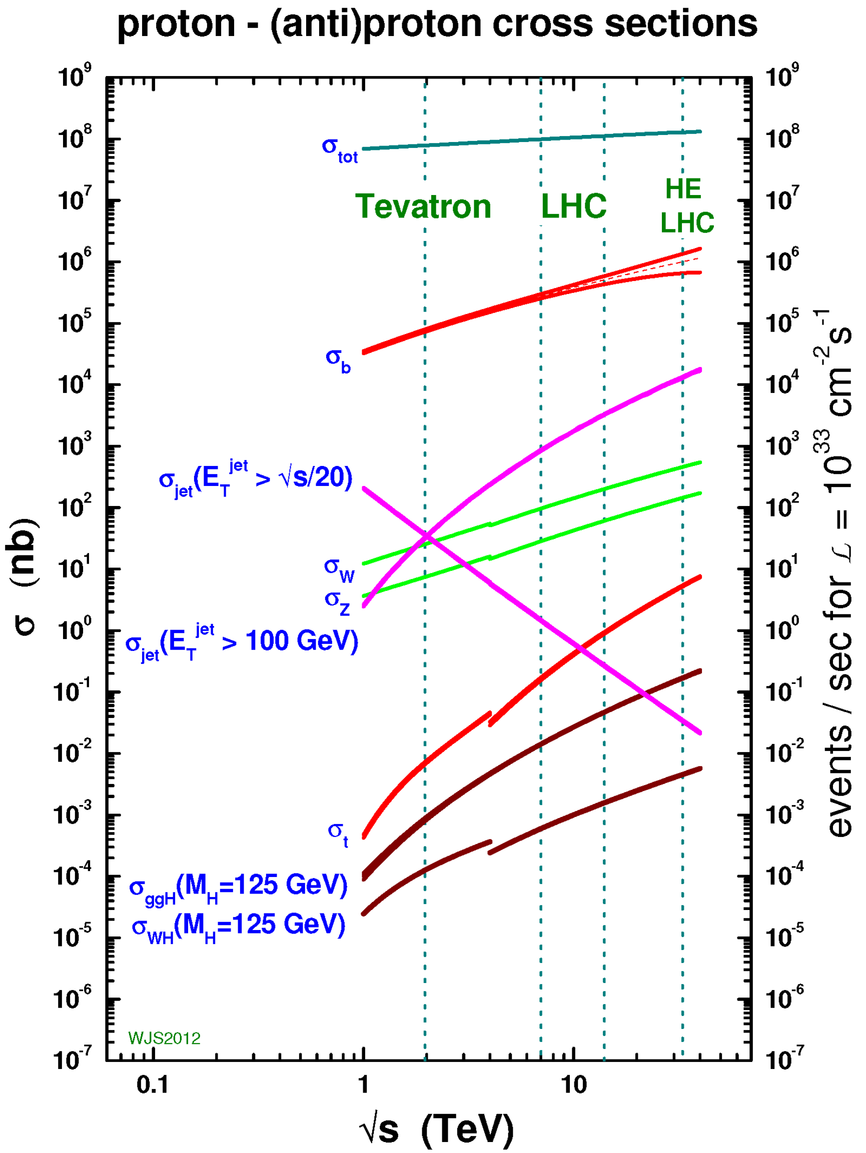
"MS"{ Muon Spectrometer
Toroids
"Calo"{ HadCAL, EMCAL
Solenoid
"ID"{ TRT, SCT, Pixels



JETS


Signal Process
Underlying Event
Initial State
Parton Shower
Final State
Parton Shower
Fragmentation
Hadronization
and Hadron Decays
Beam Remnants

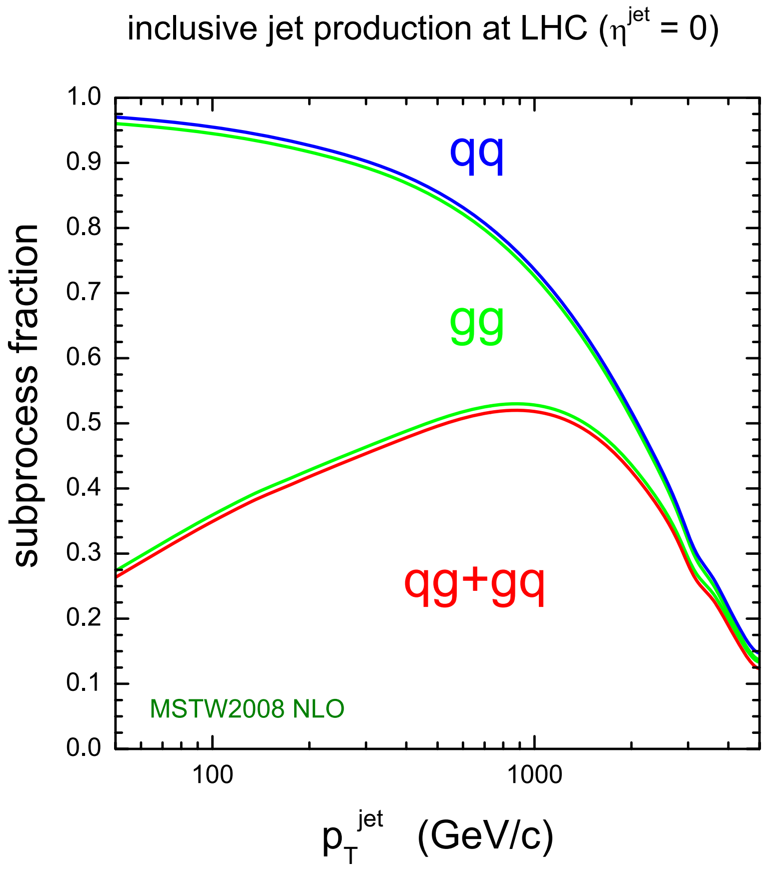
Jet production processes



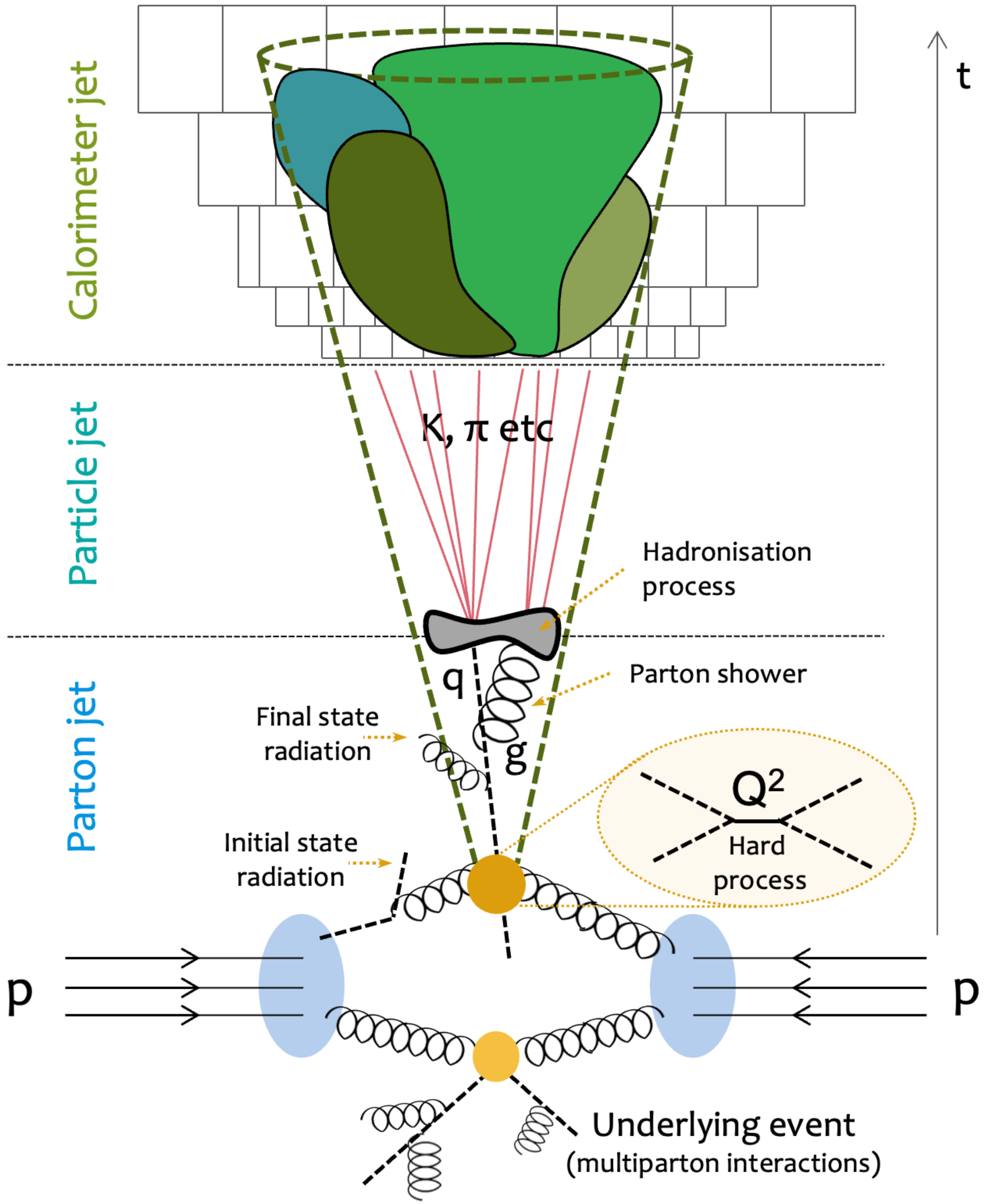
Jets are produced:

- by fragmentation of gluons and (light) quarks in QCD scattering
- by decays of heavy Standard Model particles, e.g. W & Z
- in association with particle production in Vector Boson Fusion, e.g. Higgs
- in decays of beyond the Standard Model particles, e.g. in SUSY

JETS

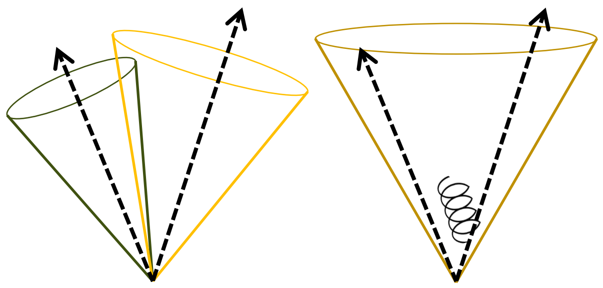


At low energy, jets are more likely produced by gluon fusion.

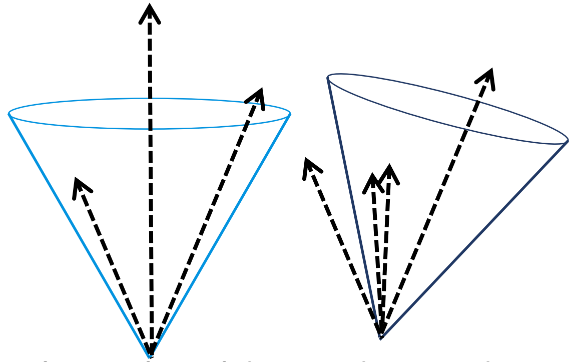


Jet Algorithms

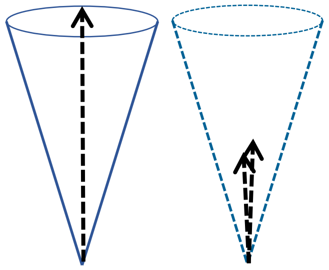
- **Theory requirements:** infrared and collinear safe



Soft gluon radiation should not merge jets



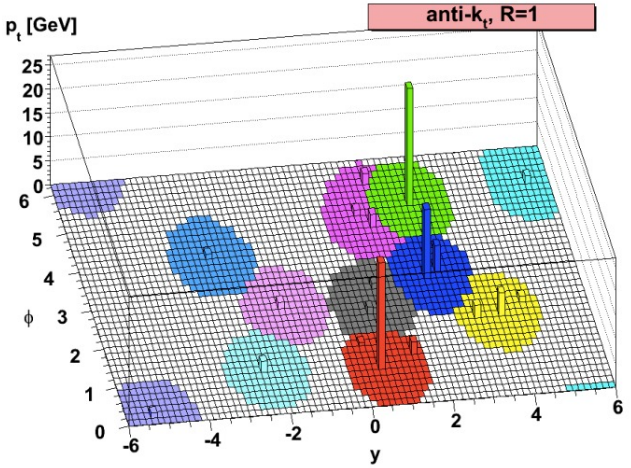
Final jet should not depend on the ordering of the seeds…



…and on signal split in two possibly below threshold

- **Experimental requirements:** Independent to detector technology and data taking conditions, easily implementable

- **Jet algorithm commonly used at the LHC: 'anti-k_t'.** A 'recursive recombination' algorithm. Starts from (topo-)clusters. Hard stuff clusters with nearest neighbor. Various cone sizes (standard R=0.4/0.5, "fat" R=1.0).



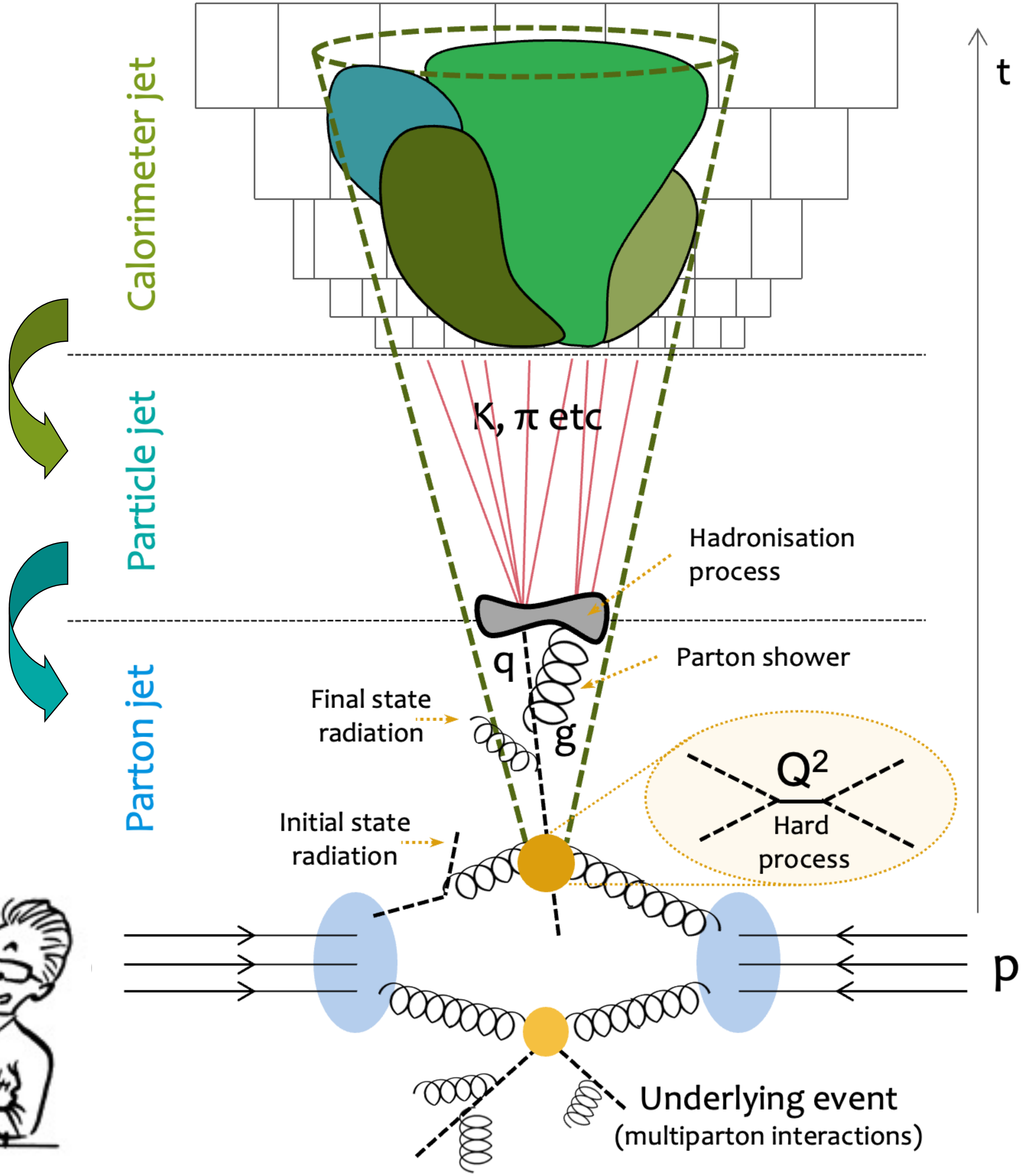
arXiv: 0802.1189

Jet Calibration

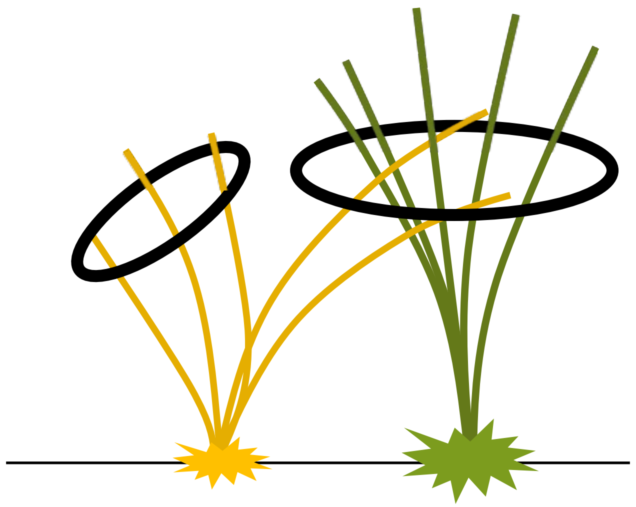
- Correct the energy and position measurement and the resolution.
- Account for:

<u>Instrumental effects</u>
Detector inefficiencies
'Pile-up'
Electronic noise
Clustering, noise suppression
Dead material losses
Detector response
Algorithm efficiency

<u>Physics effects</u>
Algorithm efficiency
'Pile-up'
'Underlying event'



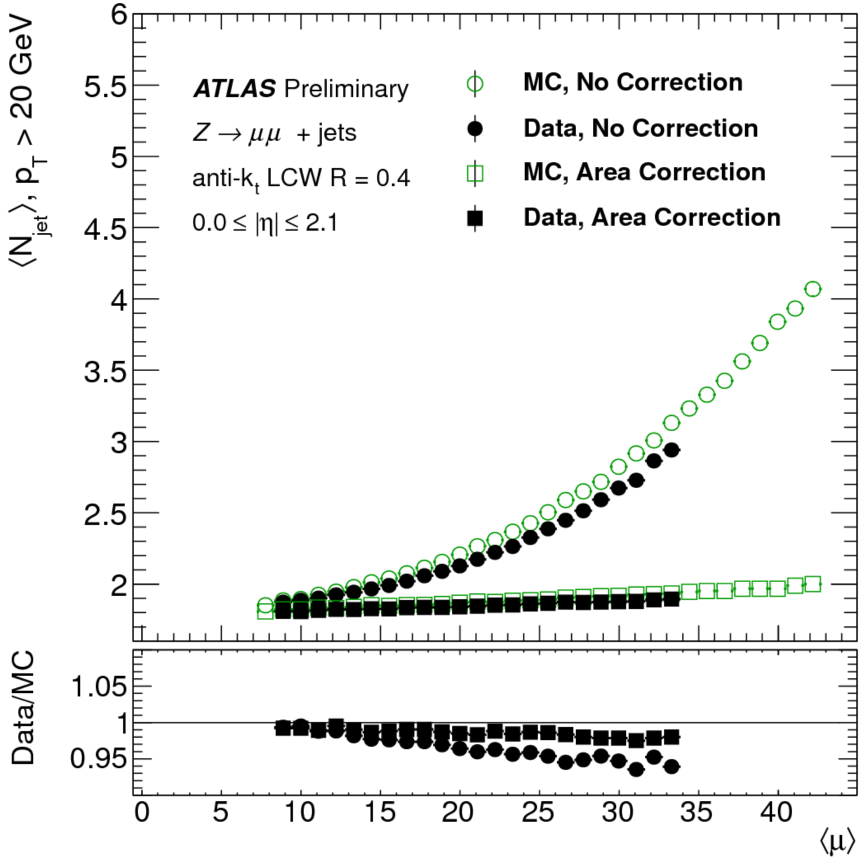
Jets and Pile-Up



Multiple interactions from pile-up

'Jet-areas' corrections

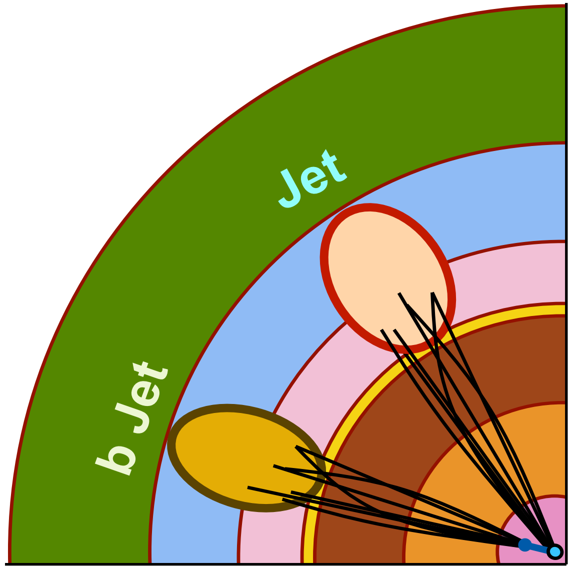
Inspired by arXiv:0707.1378

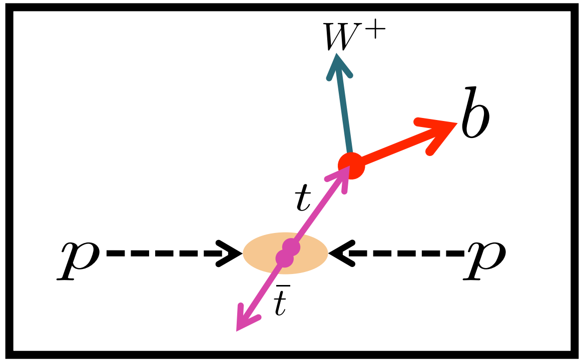


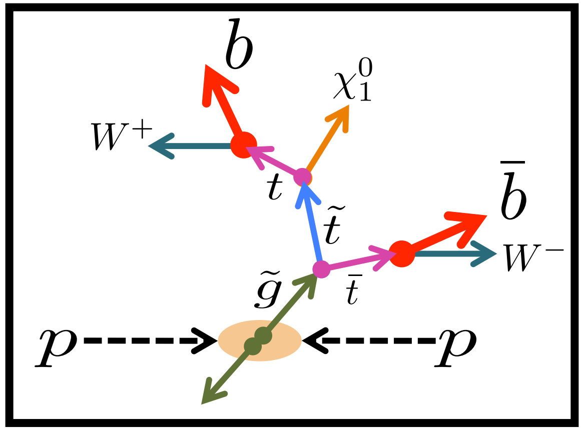
B-JETS

- b-hadrons have a lifetime of ~ 10^{-12} s.
- They travel a small distance (fraction of mm) before decaying.
- A "**displaced vertex**" creates a distinct jet, so b-jets can be tagged (**b-tagged**).
- b-tagging uses sophisticated algorithms, mostly **multi-variate (machine learning)**.

- b-jets create distinct final states, important for both **Standard Model measurements** and **searches for New Physics**.

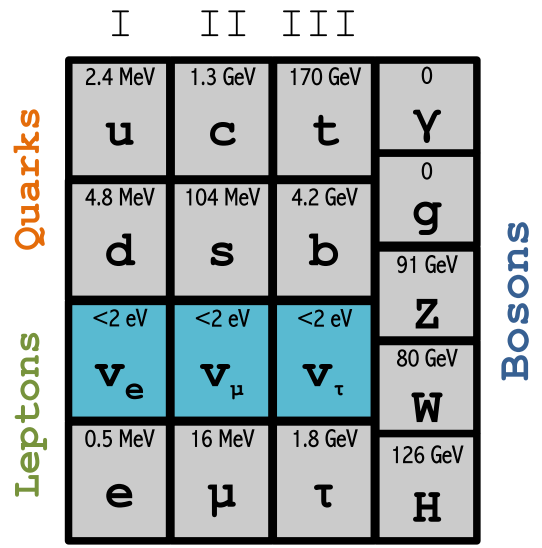


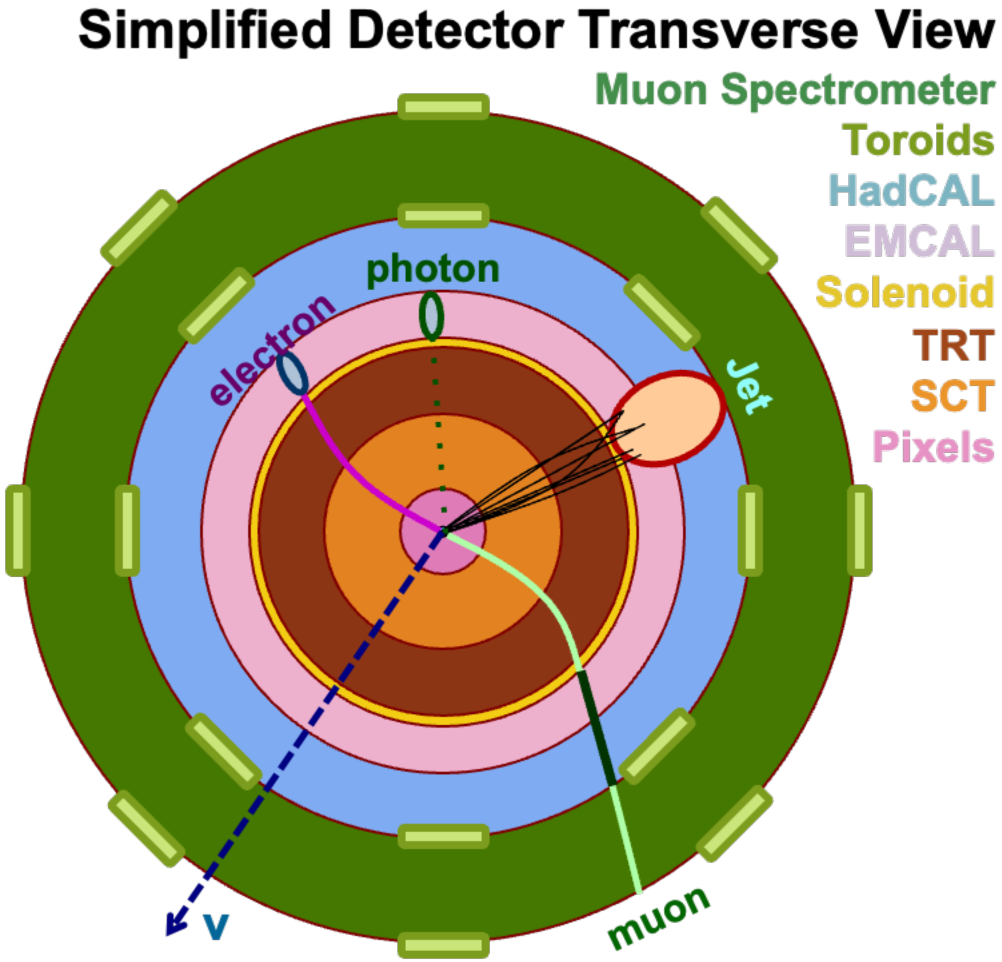





M.E.T.
THE
EXTRA-TERRESTRIAL
THE 20th ANNIVERSARY

Missing Transverse Momentum – ME_T





In the transverse plane:

$$\Sigma_i \vec{p}_{T,i} = 0$$

So for what we can't directly measure (e.g. neutrinos)

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\Sigma_i \vec{p}_{T,i}$$

Missing Transverse Momentum – ME_T

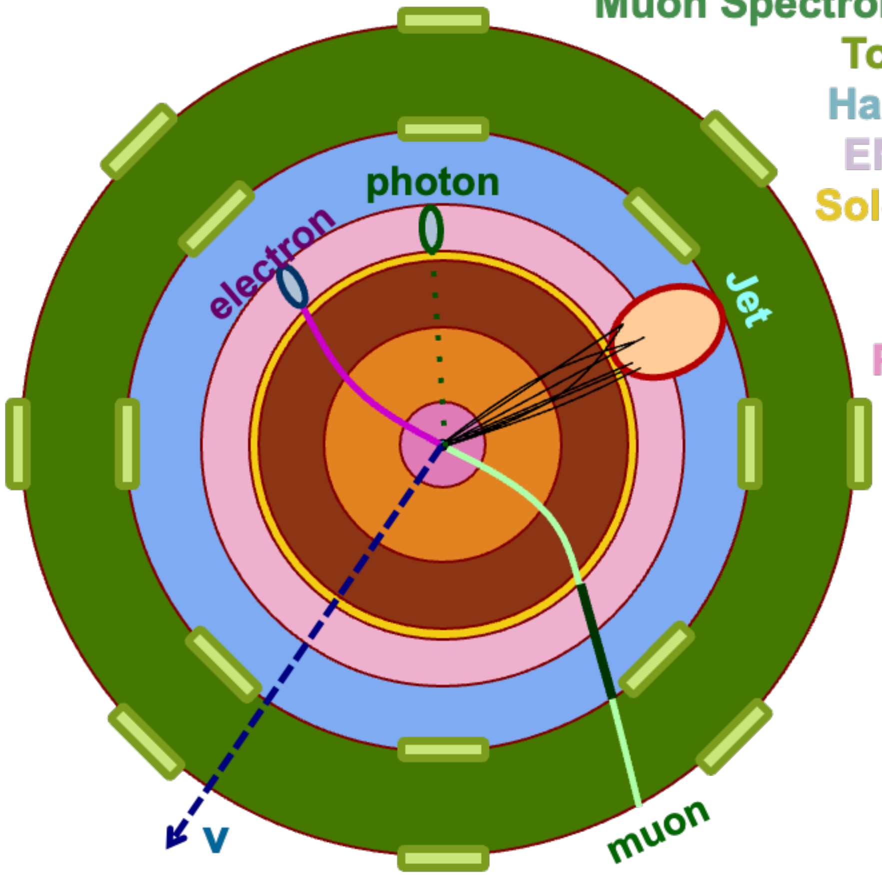
In the transverse plane:

$$\Sigma_i \vec{p}_{T,i} = 0$$

So for what we can't directly measure (e.g. neutrinos)

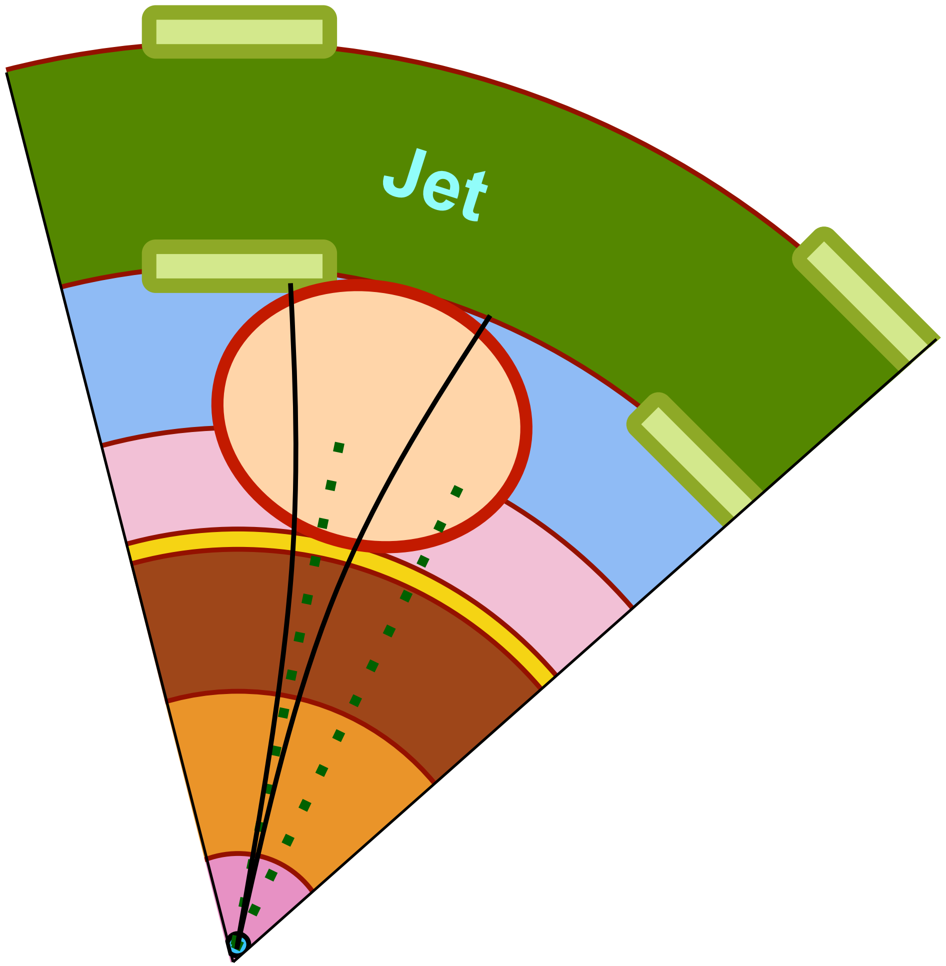
$$E_{\mathrm{T}}^{\mathrm{miss}} = -\Sigma_i \vec{p}_{T,i}$$

or Dark Matter candidates!

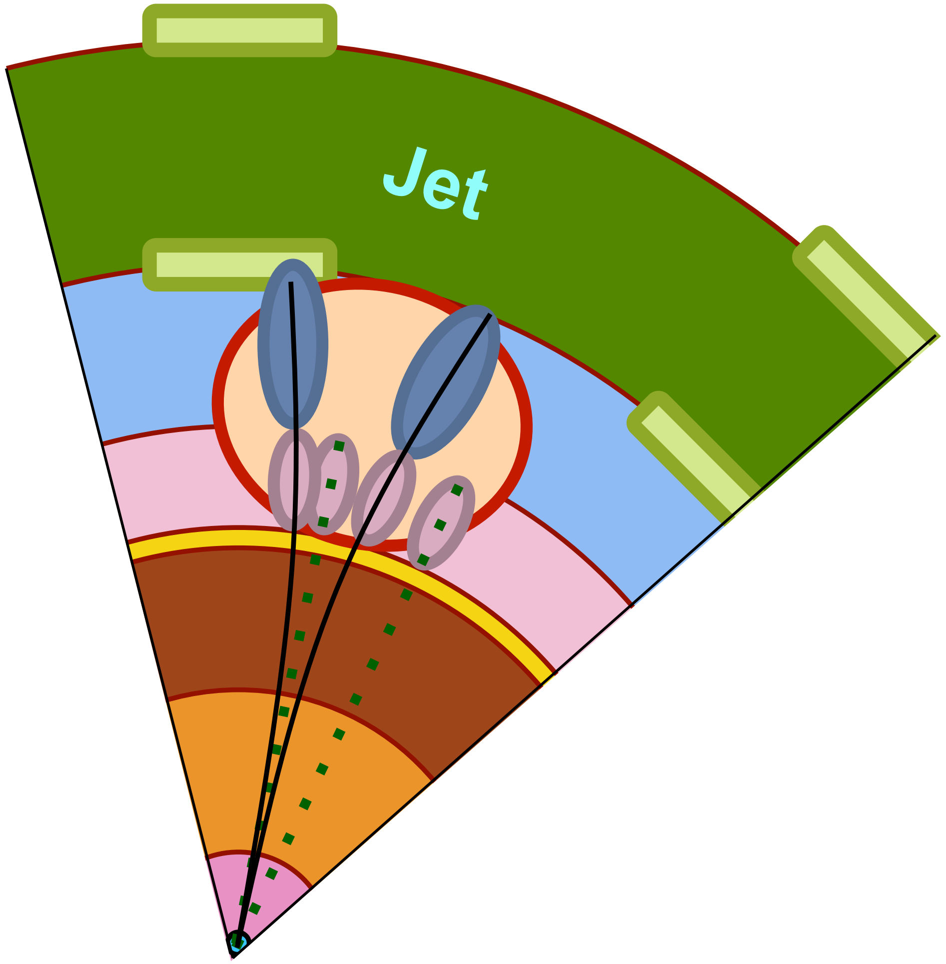


Particle Flow
for hadronic Reconstruction

Particle flow

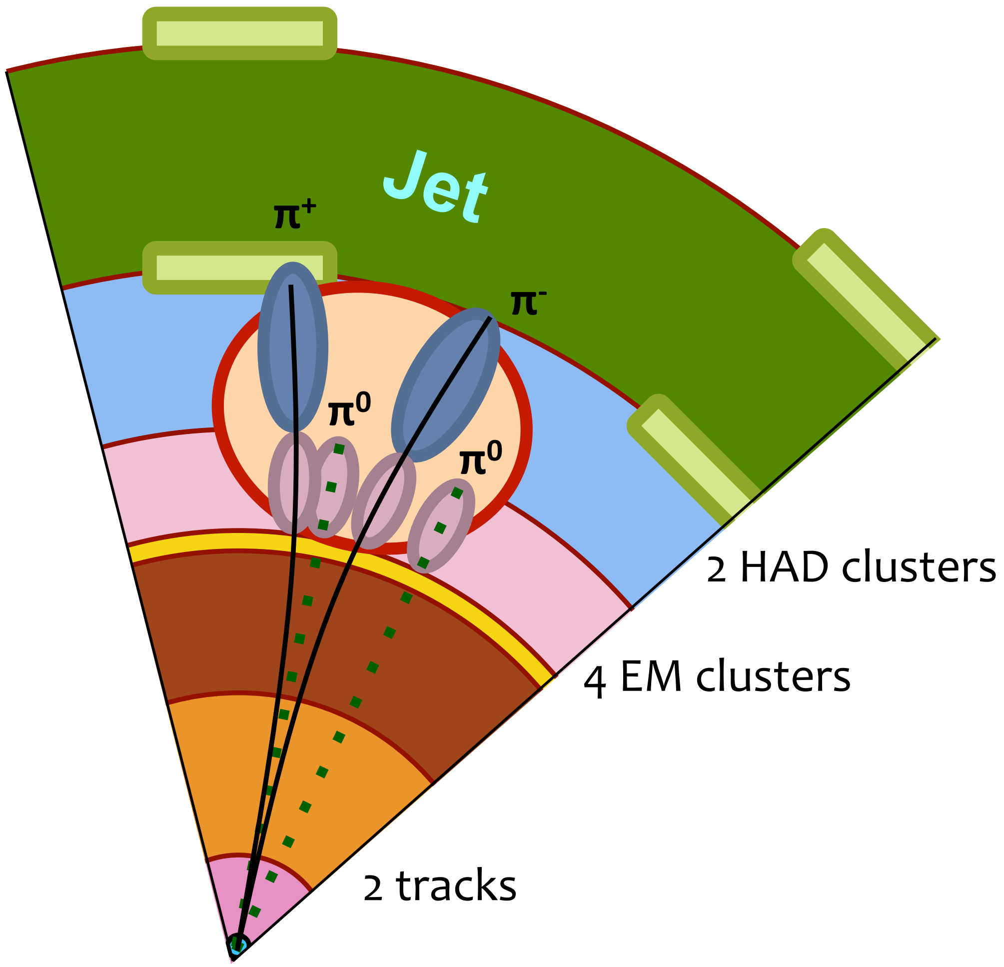


Particle flow

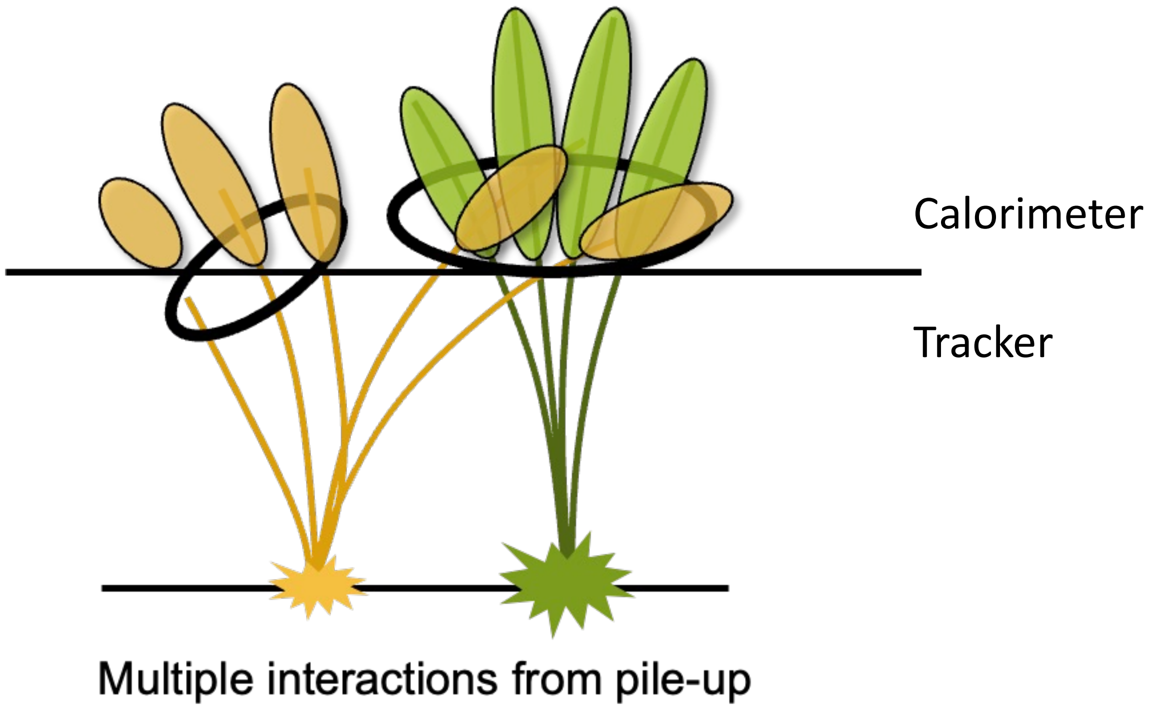


Particle Flow

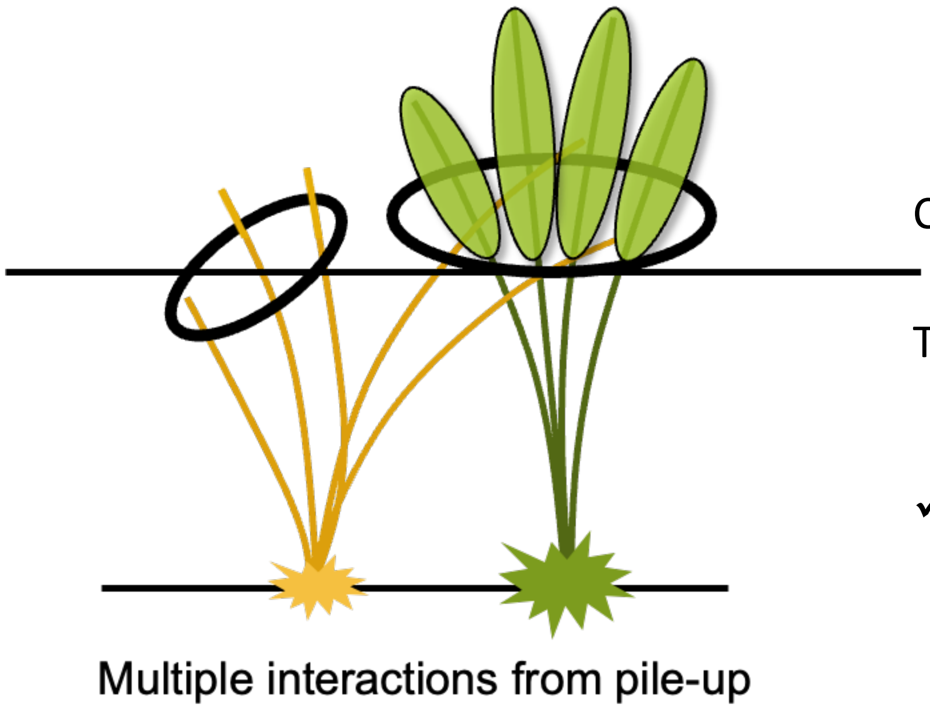
- "Flow of particles" through the detector.
- Reconstruct and identify all particles, photons, electrons, pions, …
- Use best combination of all sub-detectors for measuring the properties of the particles.
- First used at LEP (ALEPH) and then at the LHC (CMS).



Jets in Pile-up



Jets in Pile-up

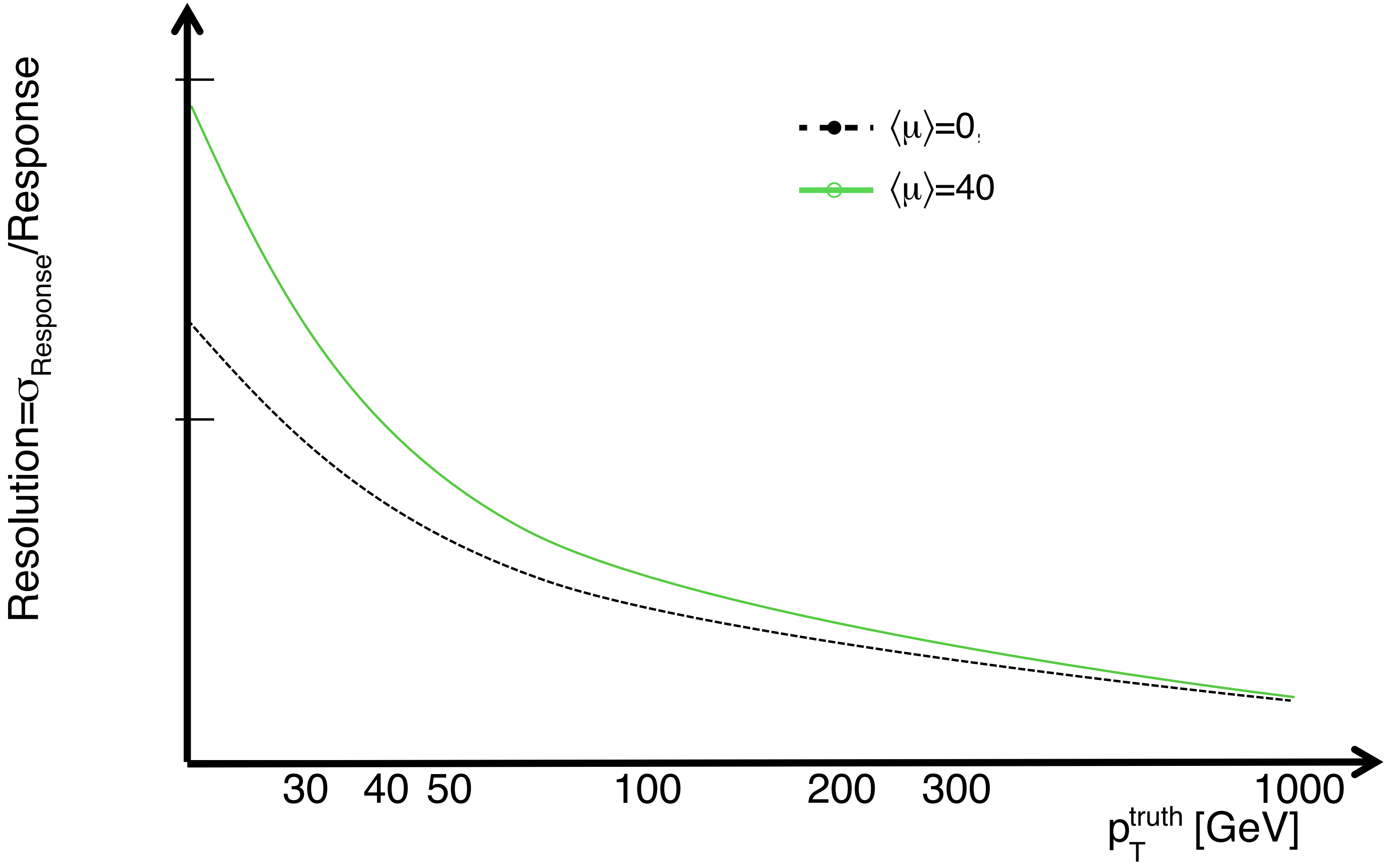


Calorimeter

Tracker

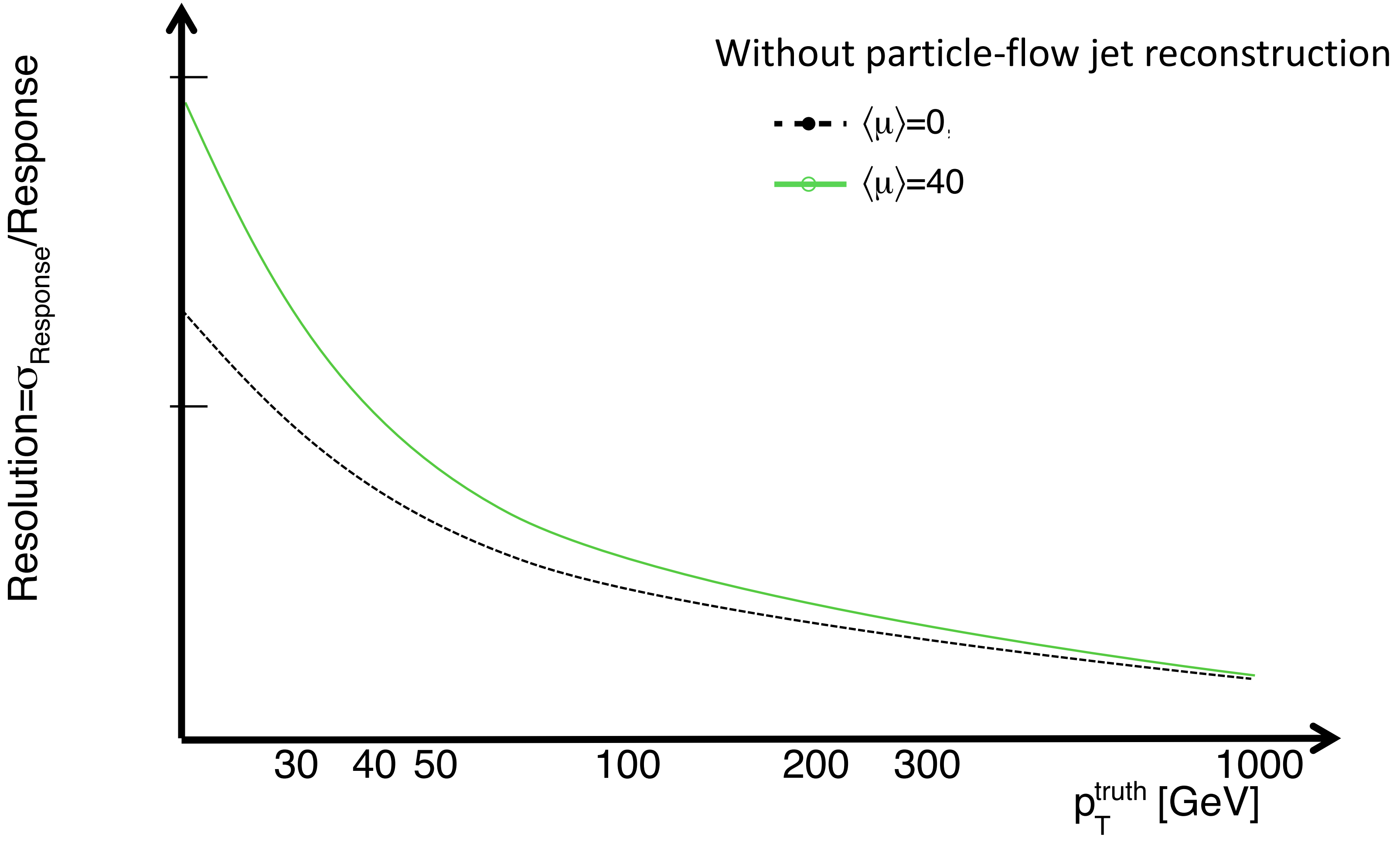
✓ Requirement that particles originate from the primary vertex.

Momentum resolution



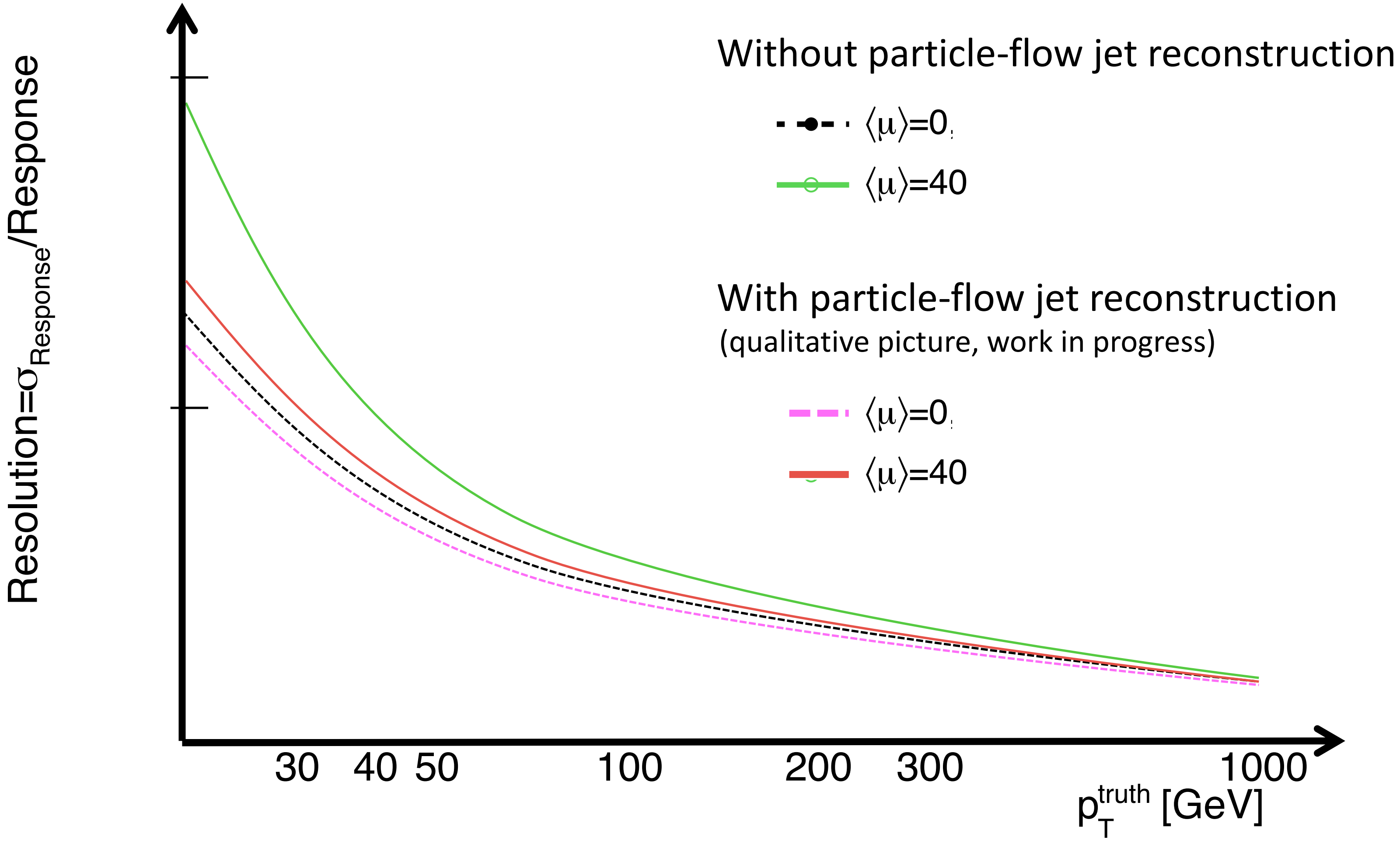
Resolution: the quality with which we measure the jet momentum.

Momentum resolution



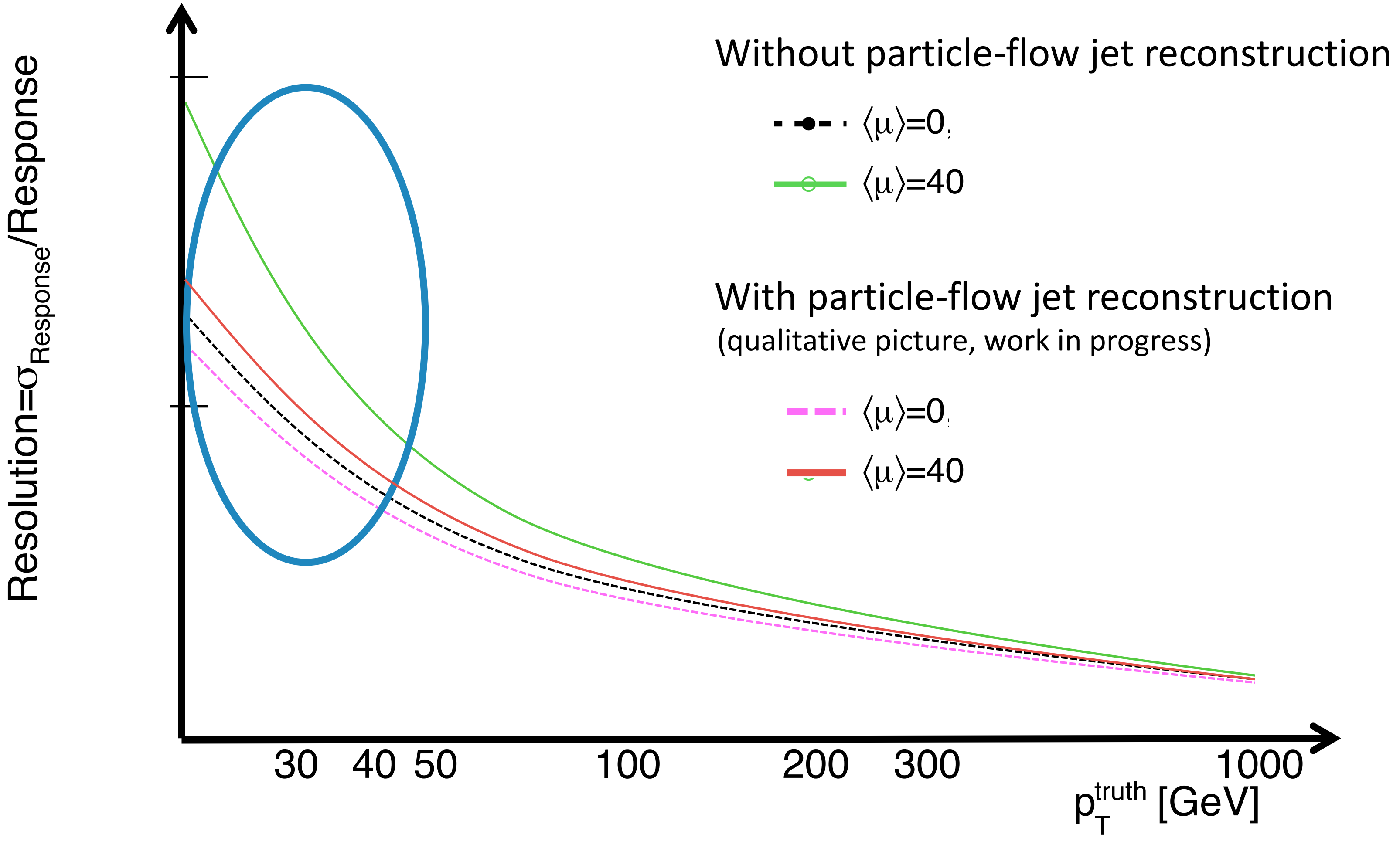
Resolution: the quality with which we measure the jet momentum.

Momentum resolution



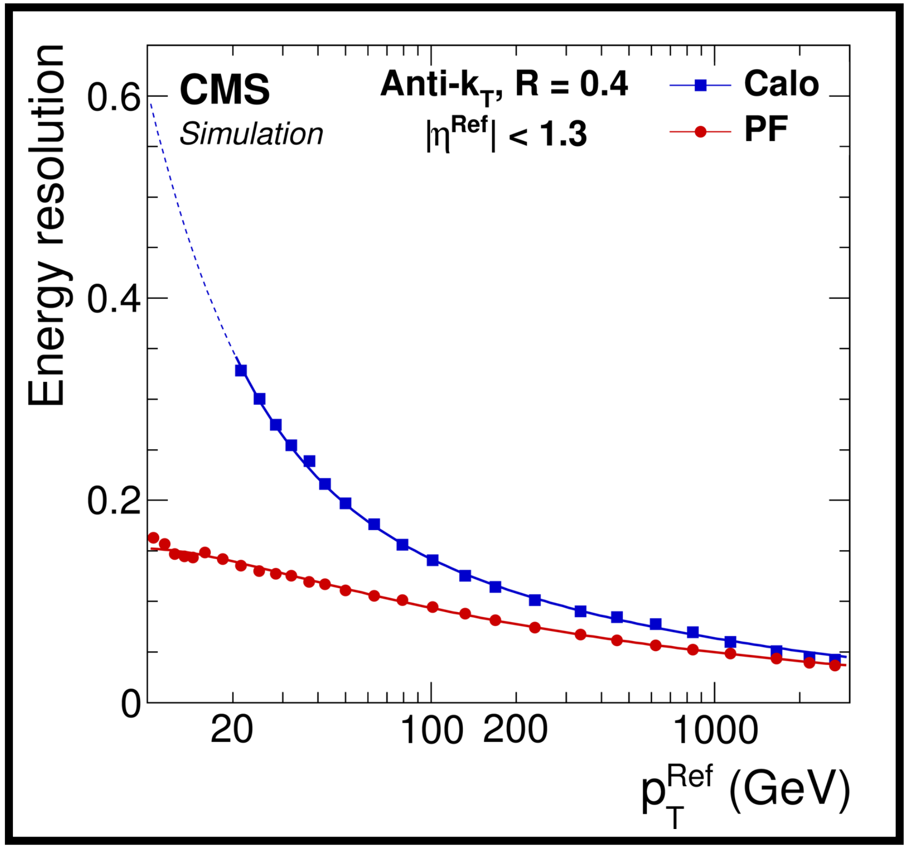
Resolution: the quality with which we measure the jet momentum.

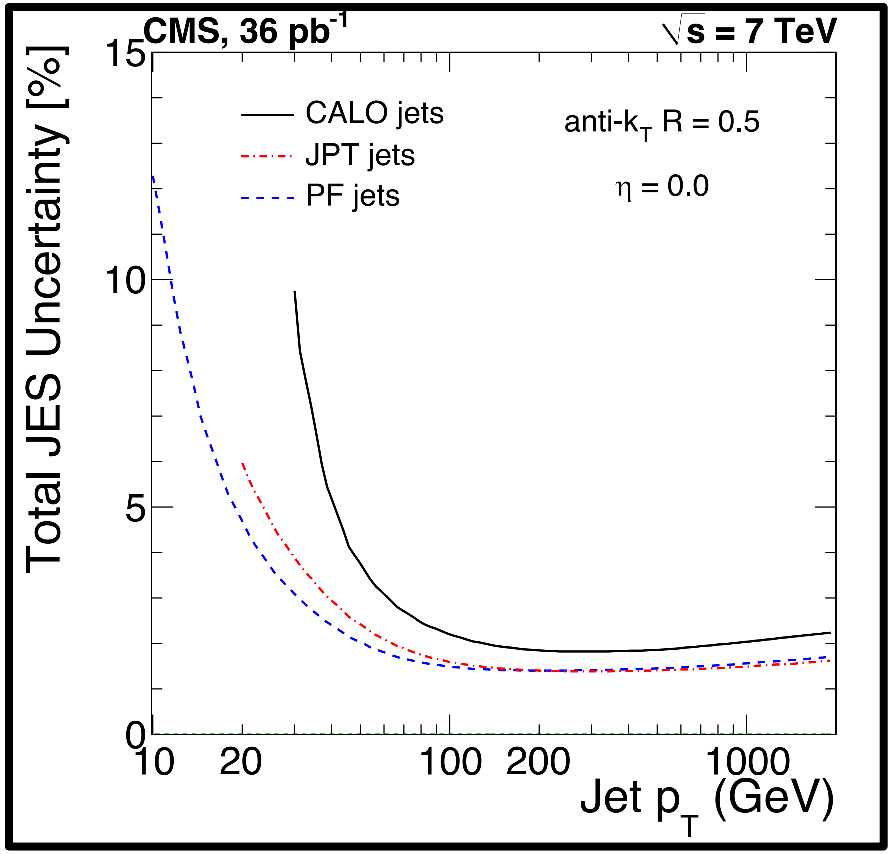
Momentum resolution



Significant improvement for low-pT jets. Similar for MET.

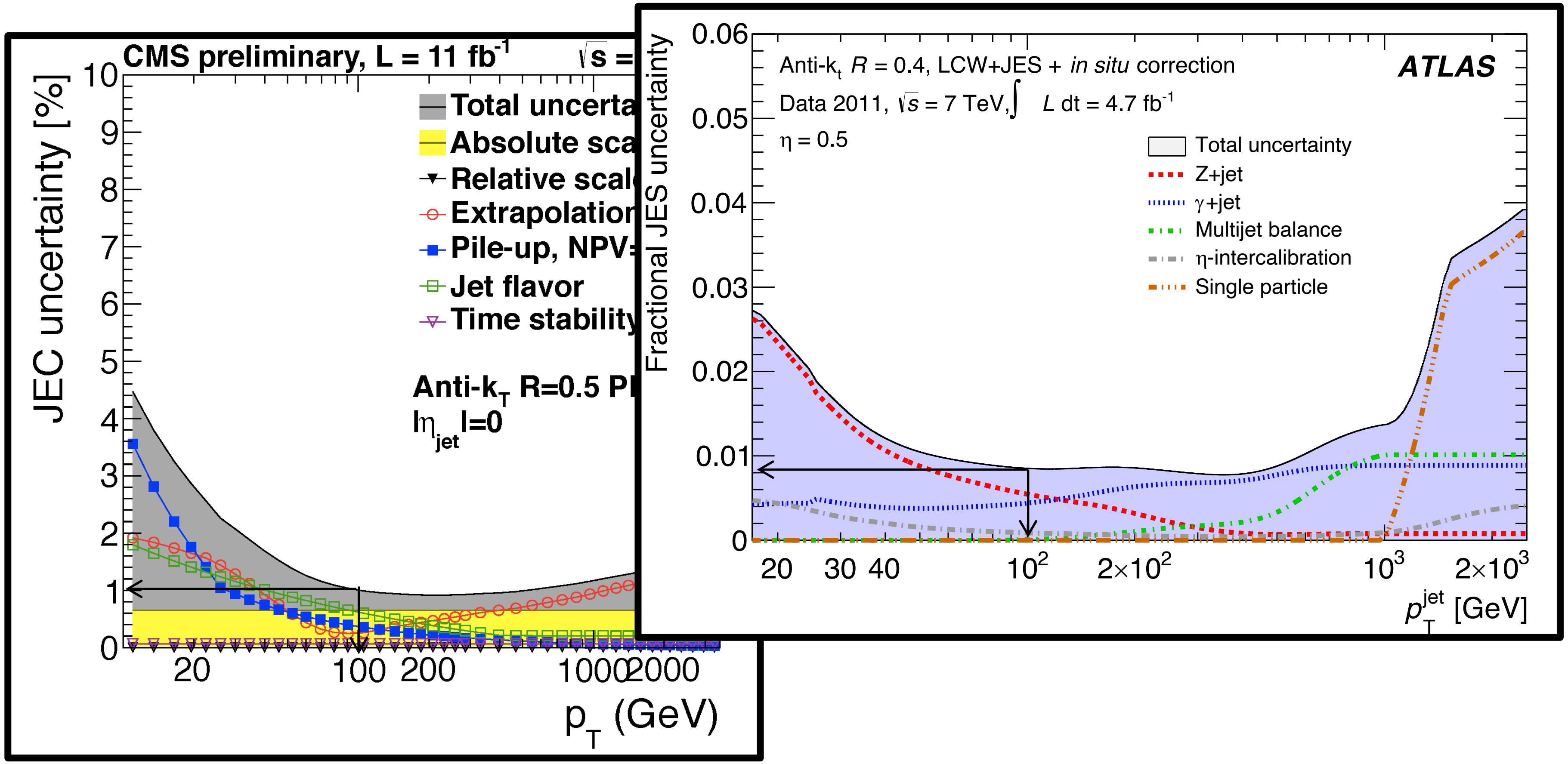
Particle Flow – performance





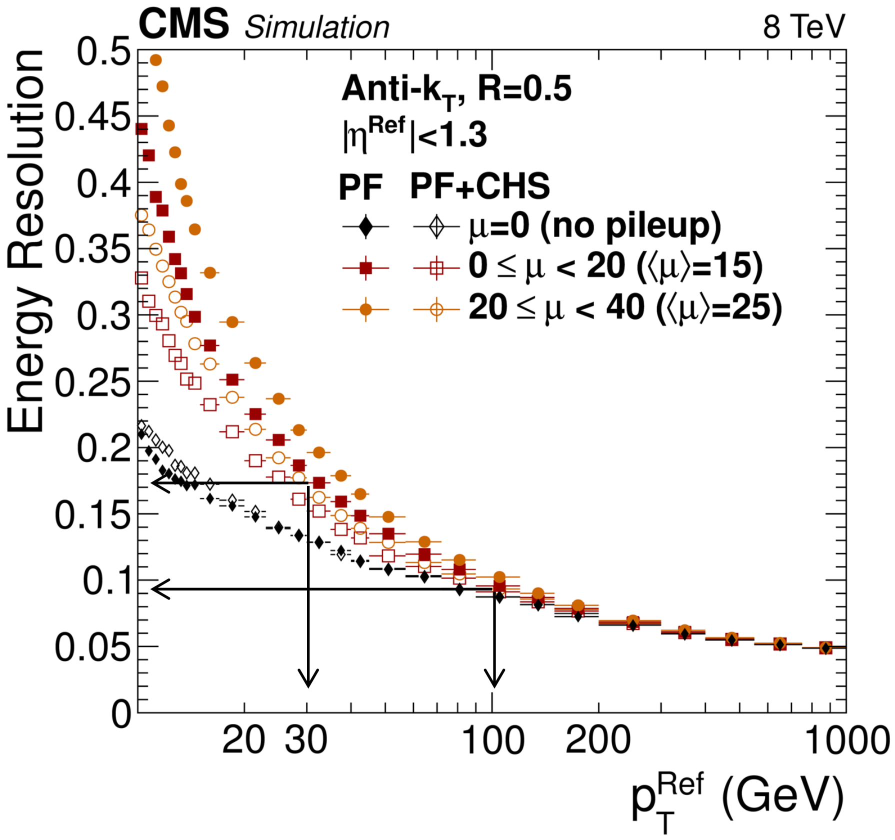
In Jet Energy resolution and uncertainty, large improvements with respect to calo jets!

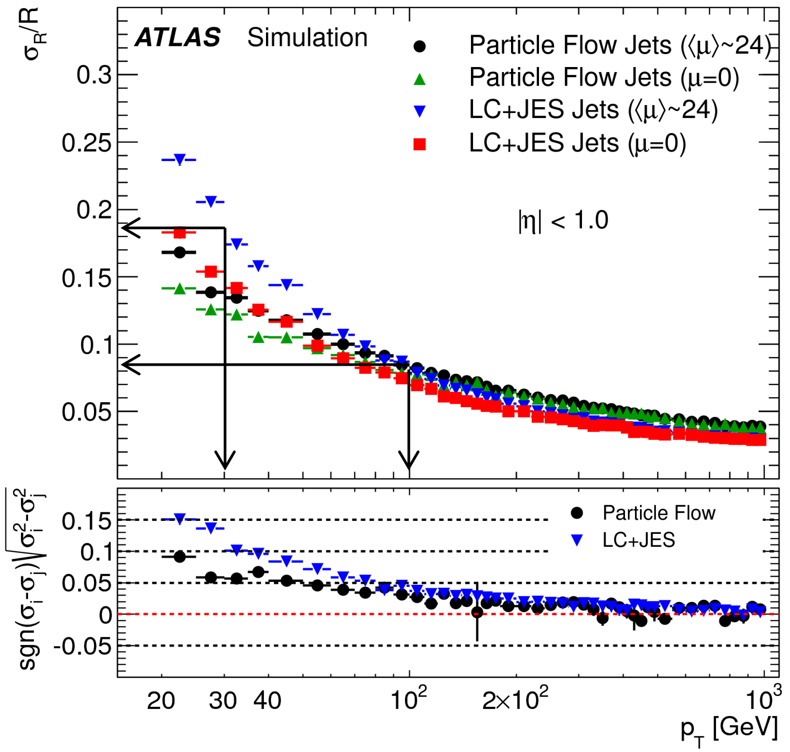
A COMPARISON



- ◎ PF jets (CMS) and calo jets (ATLAS) have similar performance.
- ◎ **Particle reconstruction always needs to be optimized depending on the detector technologies and experimental requirements.**

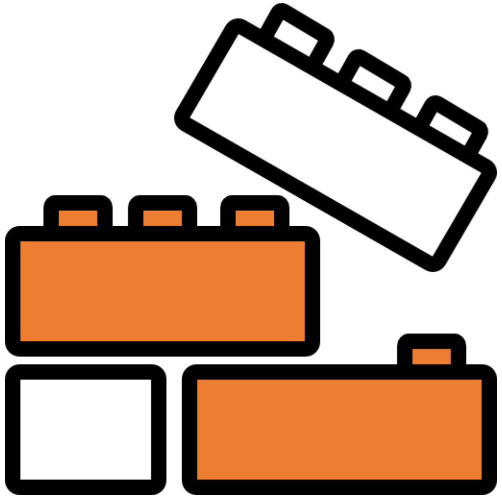
A COMPARISON





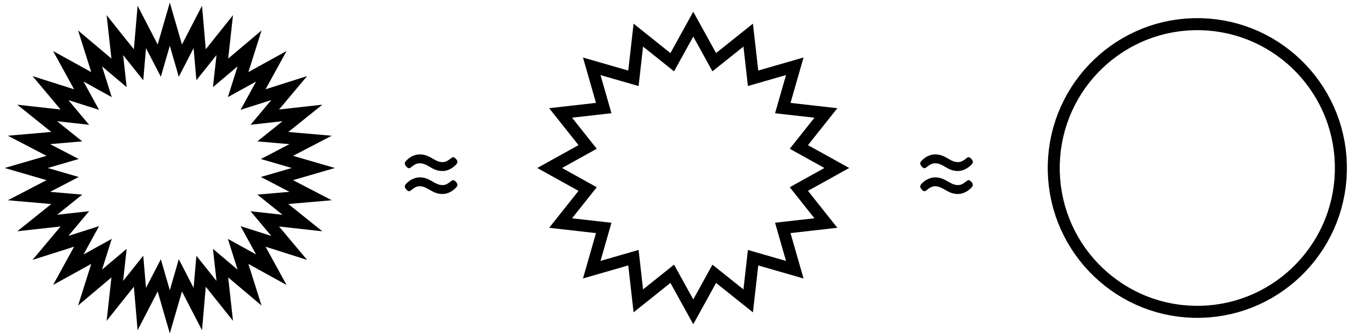
- PF jets (CMS) and calo jets (ATLAS) have similar performance.
- **Particle reconstruction always needs to be optimized depending on the detector technologies and experimental requirements.**

Online
Reconstruction



Online reconstruction

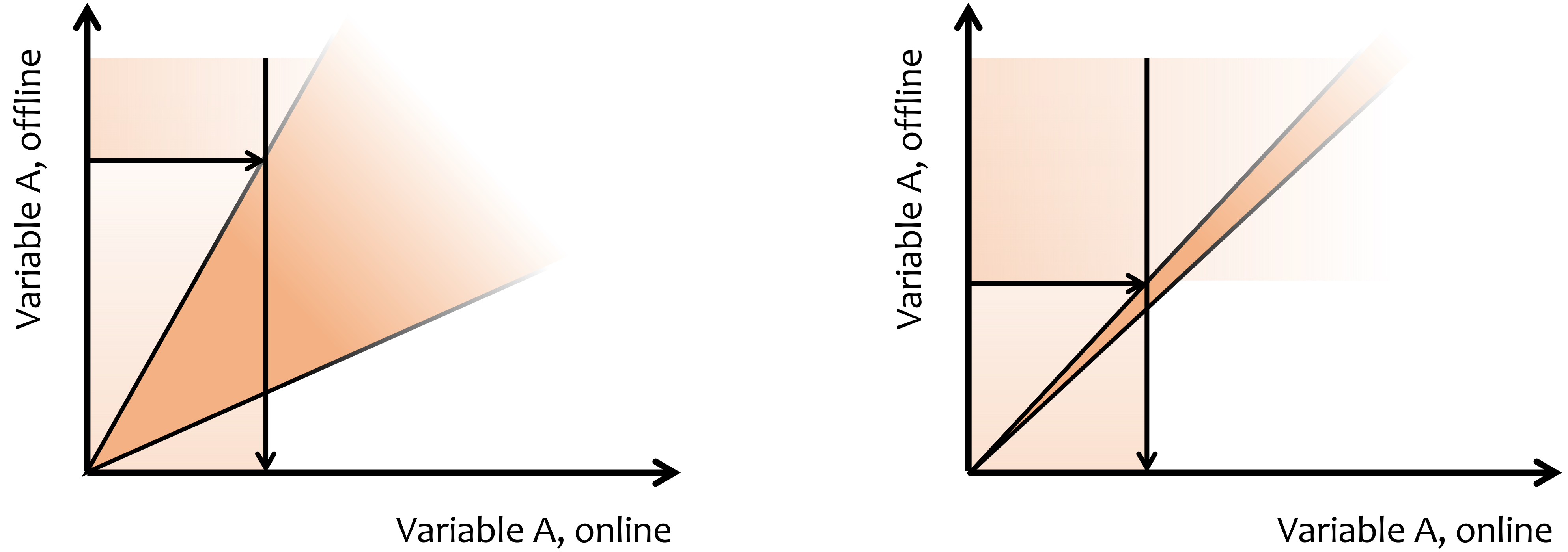
Objective: Trigger ("online") reconstruction same as "offline".
Problem: Time. Trigger decision needs to be taken fast.
Solution: Simplification.
Challenge: Clever simplification = good performance.



E.g. track reconstruction in regions of interest and simplified MET calculation.

Trigger

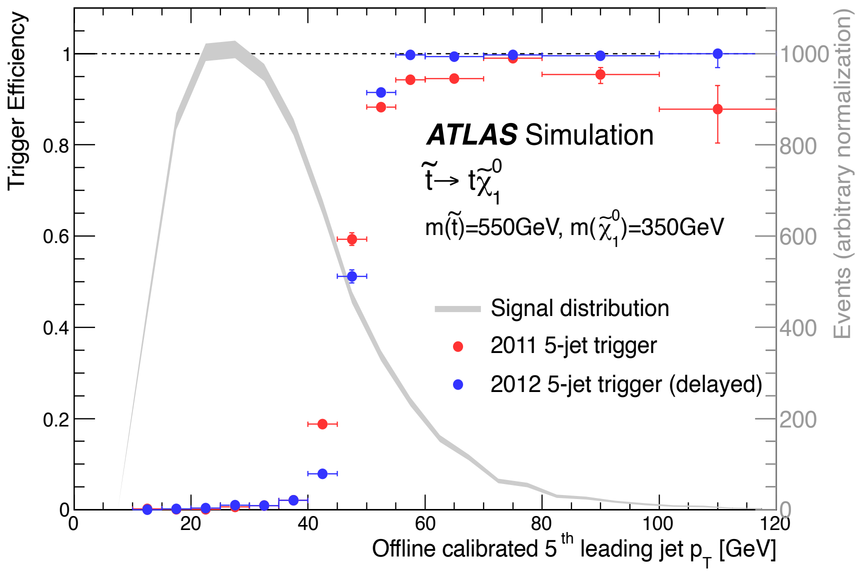
- To profit fully from an improvement in reconstruction, the relevant algorithm has to be used at the relevant trigger selections to provide **optimal online-to-offline correlation**.

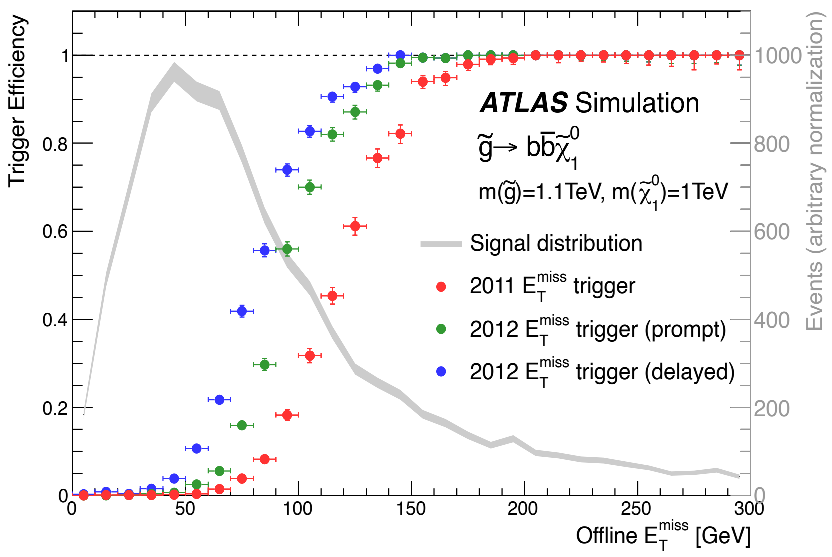


Variable A: e.g. leading jet pT

Online reconstruction

$$\text{trigger efficiency} = \frac{\#\ \text{events passing offline selection}\ \&\ \text{trigger}}{\#\ \text{events passing offline selection}}$$





Clever ideas need to be deployed to bring online closer to offline, making efficiency curves **sharper** and **plateau closer to 1.**

Efficiency measurements

Relevant beyond the trigger…

Tag and Probe

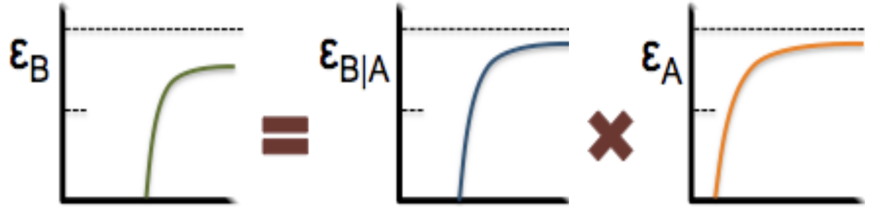
- Select events based on requirements on one object (tag) and study the response of the second object (probe), not used in the event selection, using some constraint such as the Z mass.
  - e.g. $Z \rightarrow \tau\tau$ events.
  - Typically used for measurement of the identification efficiency

Orthogonal sample

- Measure directly the efficiency on an independent, orthogonal sample.
  - e.g. jet trigger efficiency on a sample triggered by muons,

Bootstrap method

- The efficiency, ε_B, of a selection B, inclusive compared to a selection A, can be determined in a sample of events passing selection A (provided that ε_A is measurable): $\varepsilon_B = \varepsilon_{B|A} \times \varepsilon_A$.
  - e.g. trigger efficiencies, say B: tau50_loose & A: tau16_loose

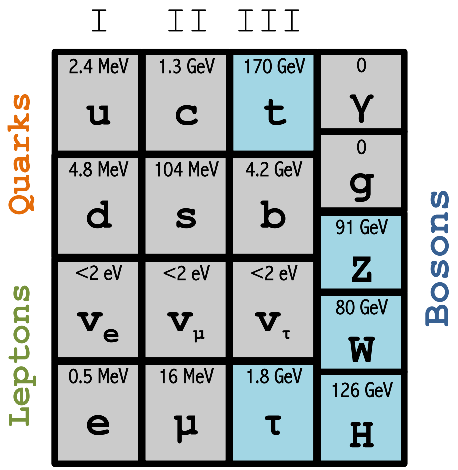


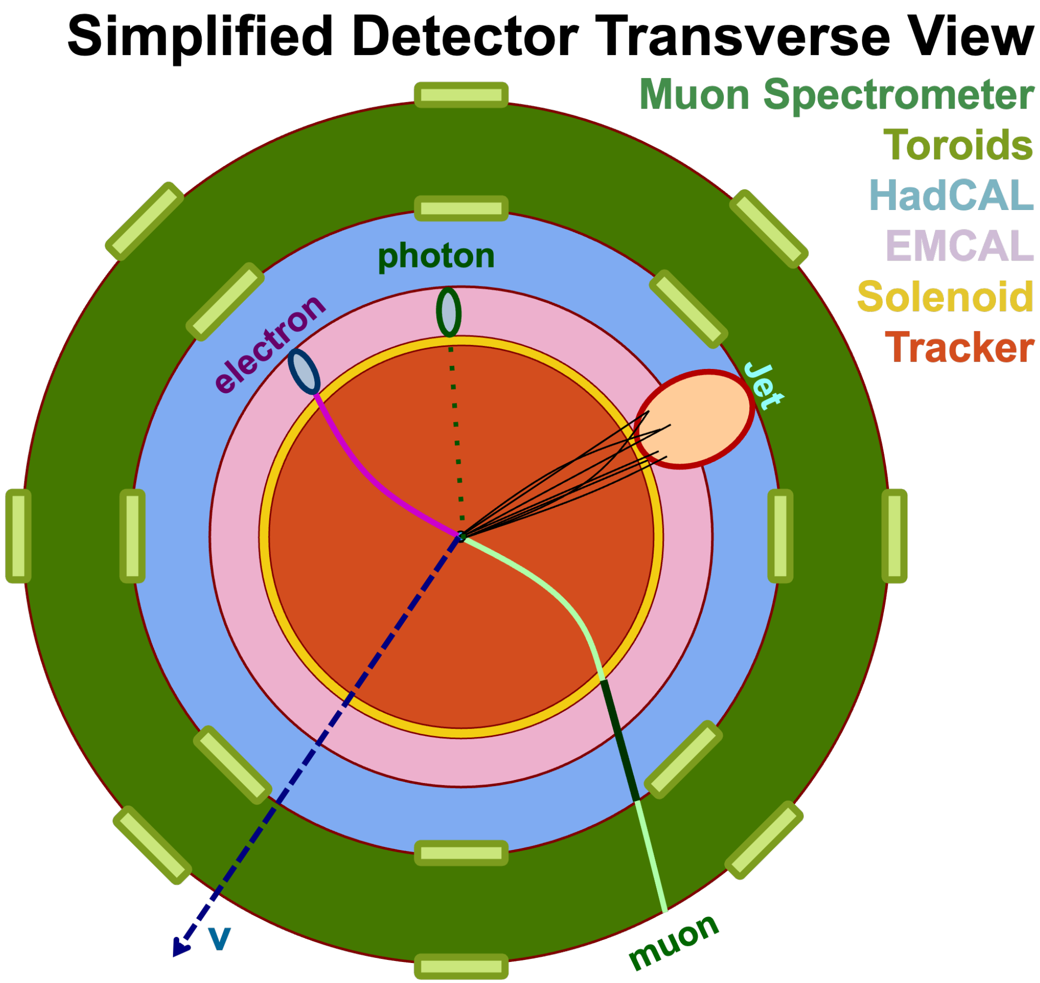
Physics Menus

Trigger selection	2015 offline threshold (GeV)	2016 offline threshold (GeV)	2017 offline threshold (GeV)	2022 offline threshold (GeV)	Representative physics case
Peak Luminosity	5×10^{33} cm^{-2} s^{-1}	1.2×10^{34} cm^{-2} s^{-1}	1.7×10^{34} cm^{-2} s^{-1}	2.0×10^{34} cm^{-2} s^{-1}	
isolated single e	25	**27**	27	27	"Main" triggers. Thrs driven by Higgs (ZH, WH), Top, SUSY.
isolated single μ	21	**27**	27	25	
di-γ	40, 30	**40, 30**	40, 30	40, 30	Higgs (H→γγ, HH→bbγγ).
di-τ (+ jet)	40, 30	**40, 30**	40, 30	40, 30	Higgs (H→ττ, HH→bbττ), SUSY.
four-jet (incl. HF)	45	**45**	45	45	SUSY, Higgs, exotics
MET	180	**200**	200	200	

Offline selections from which the triggers are "usable",
i.e. at efficiency plateau or highly efficient otherwise

Reconstructing particles



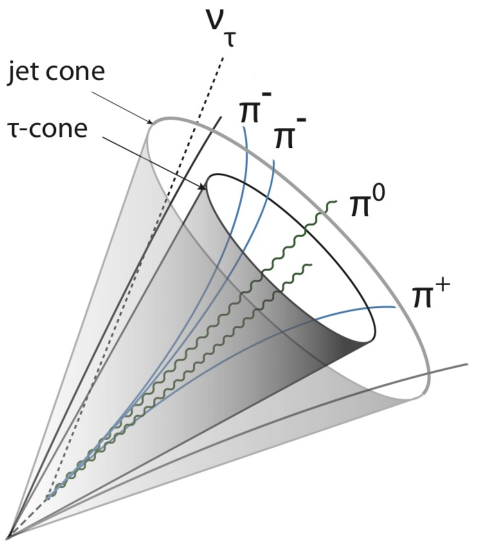


TAUS

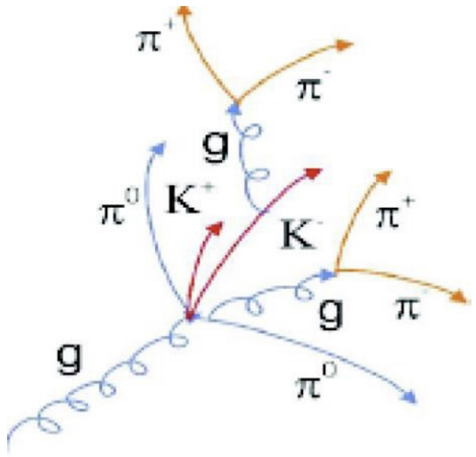
Tau Decay Mode			B.R.
Leptonic		$\tau^{\pm} \rightarrow e^{\pm} + \nu + \nu$	17.8%
		$\tau^{\pm} \rightarrow \mu^{\pm} + \nu + \nu$	17.4%
Hadronic	1-prong	$\tau^{\pm} \rightarrow \pi^{\pm} + \nu$	11%
		$\tau^{\pm} \rightarrow \pi^{\pm} + \nu + n\pi^{0}$	35%
	3-prong	$\tau^{\pm} \rightarrow 3\pi^{\pm} + \nu$	9%
		$\tau^{\pm} \rightarrow 3\pi^{\pm} + \nu + n\pi^{0}$	5%
Other			~5%

- Hadronic tau reconstruction extremely challenging
- Using **multi-variate (machine learning)** techniques based on track multiplicity and shower shapes

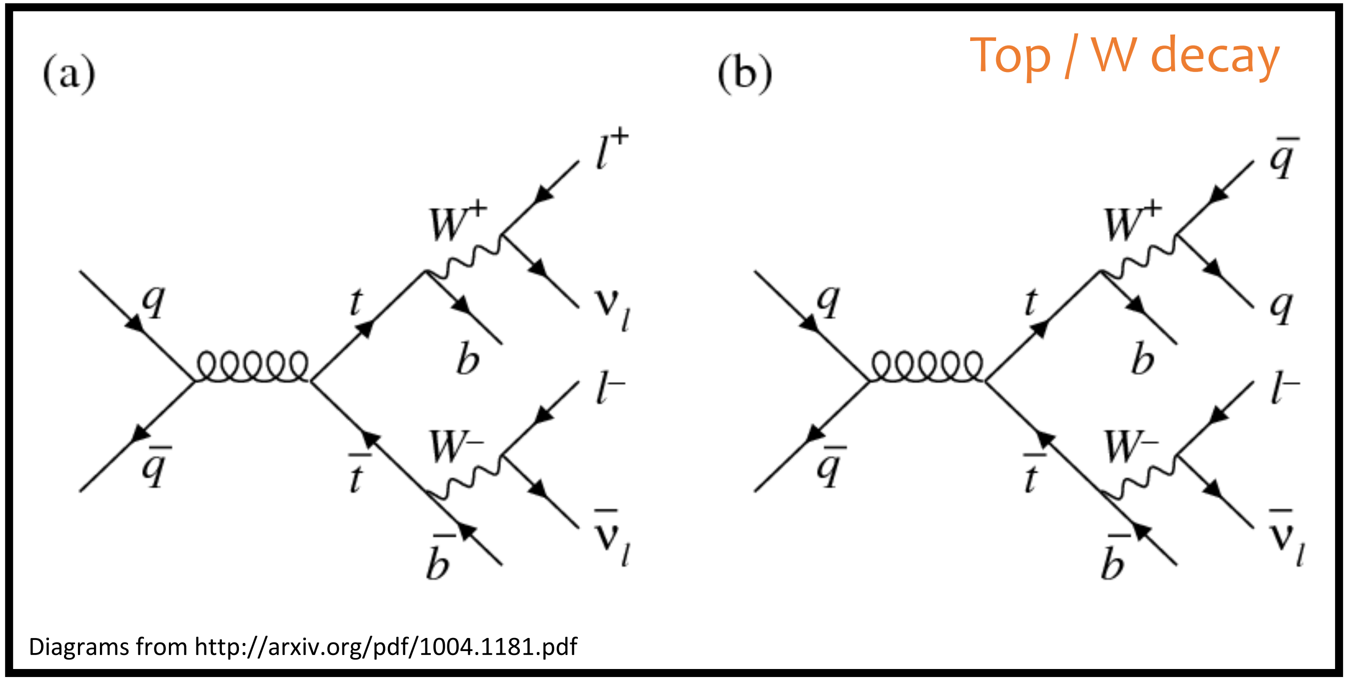
A tau jet (signal)...

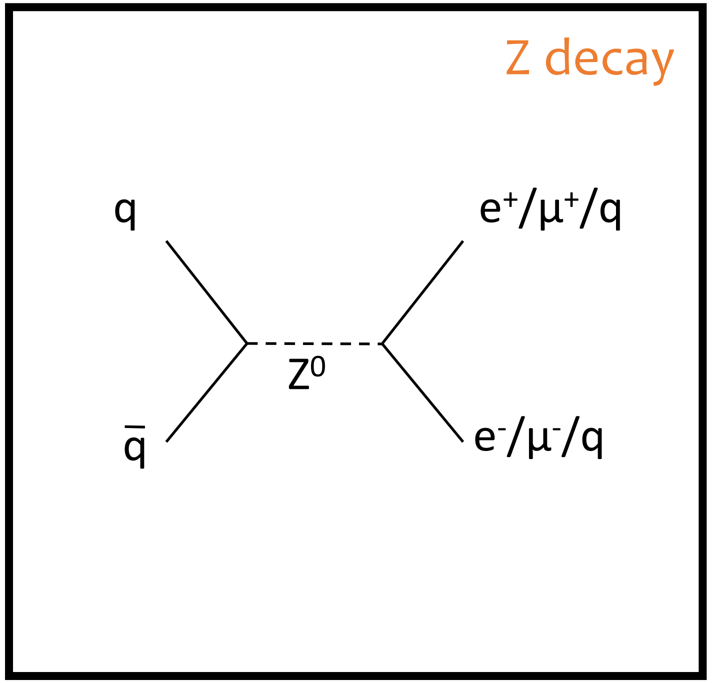


... vs. a QCD jet (background)

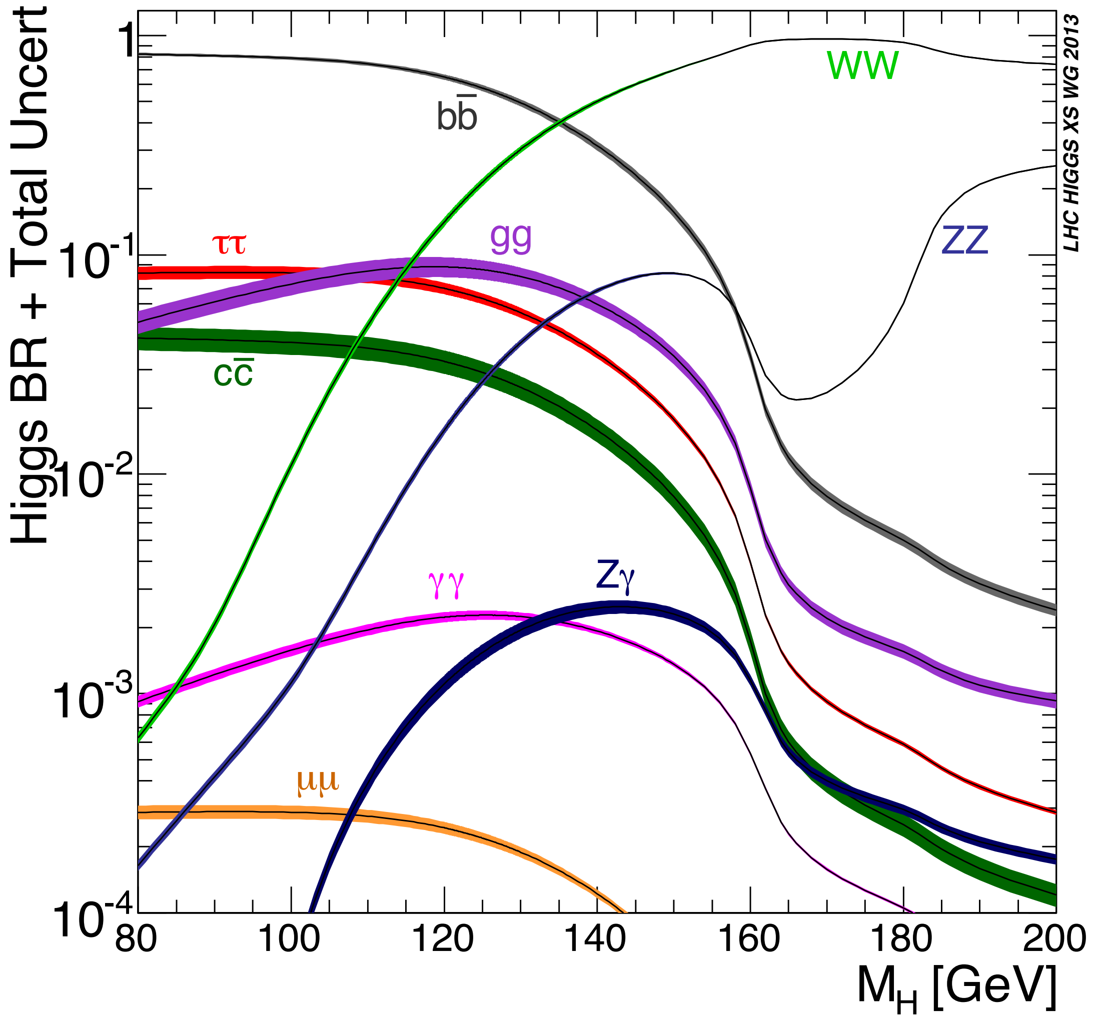


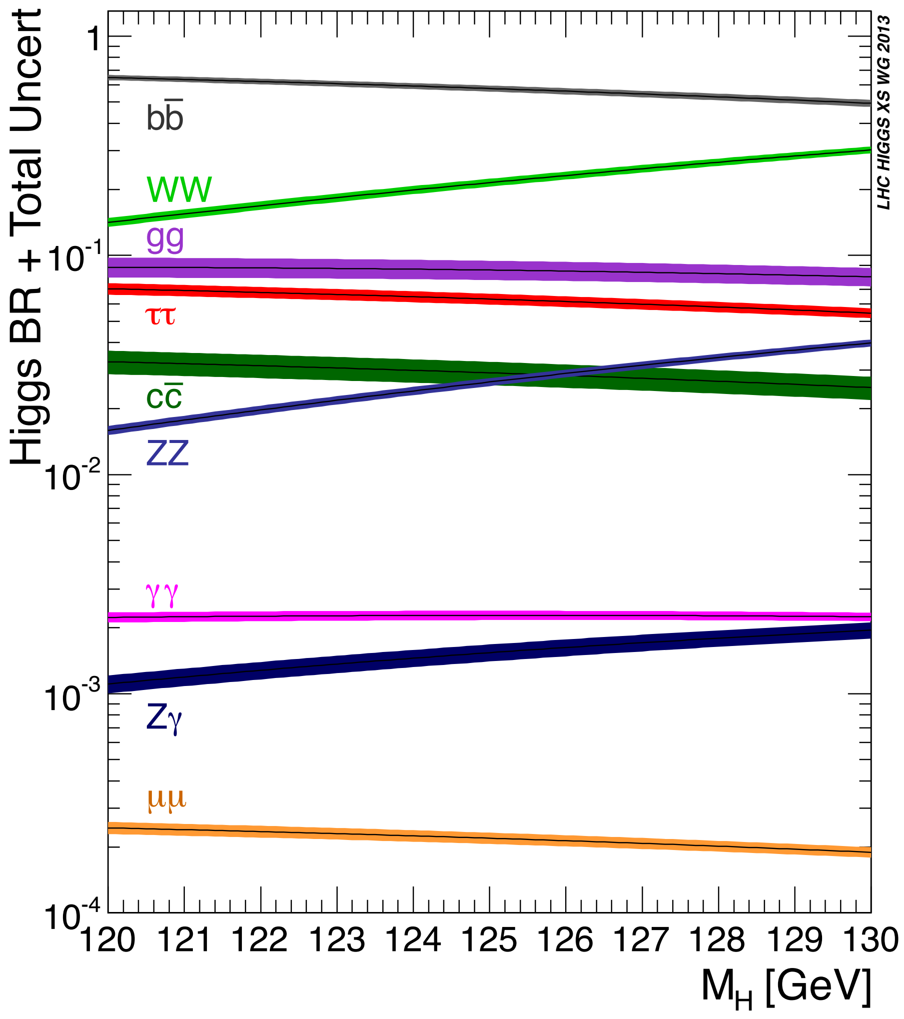
Top, W, Z


(a)
(b)
Top / W decay
q
q̄
t
t̄
W+
W−
b
b̄
l+
l−
νl
ν̄l
q̄
q
Diagrams from http://arxiv.org/pdf/1004.1181.pdf


Z decay
q
q̄
Z0
e+/μ+/q
e-/μ-/q

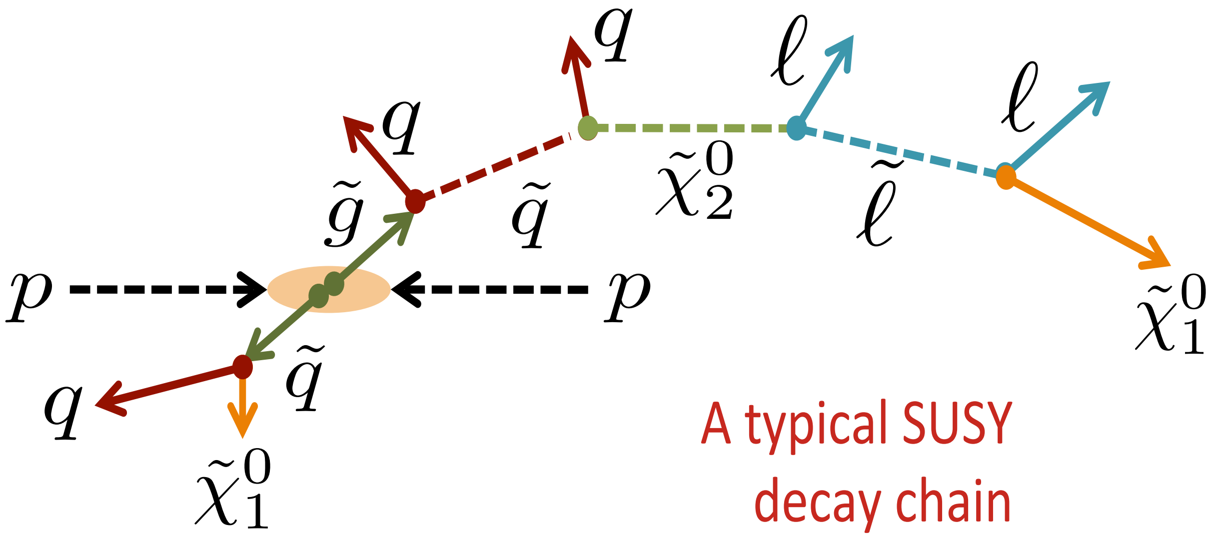
And the Higgs!





How about new particles?

- These decay to Standard Model particles or create ME_T
- E.g.



A typical SUSY decay chain

Physics analyses