

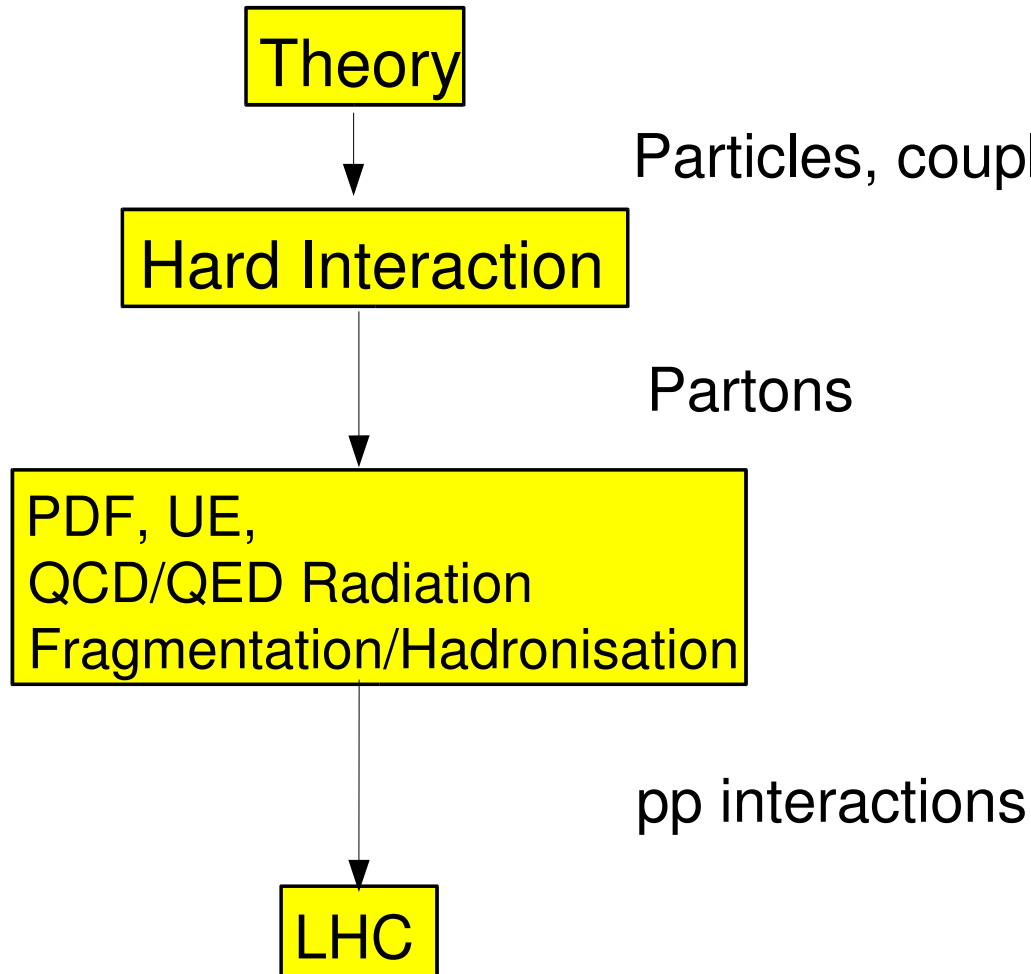
Analysis of LHC data

Giacomo Polesello

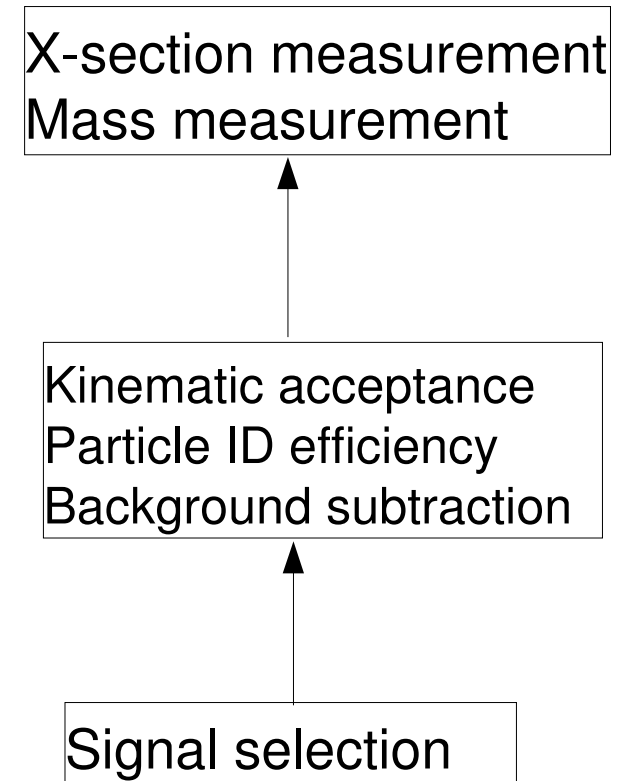
INFN, Sezione di Pavia

Data Flow in Collider experiments (1)

Data Collection

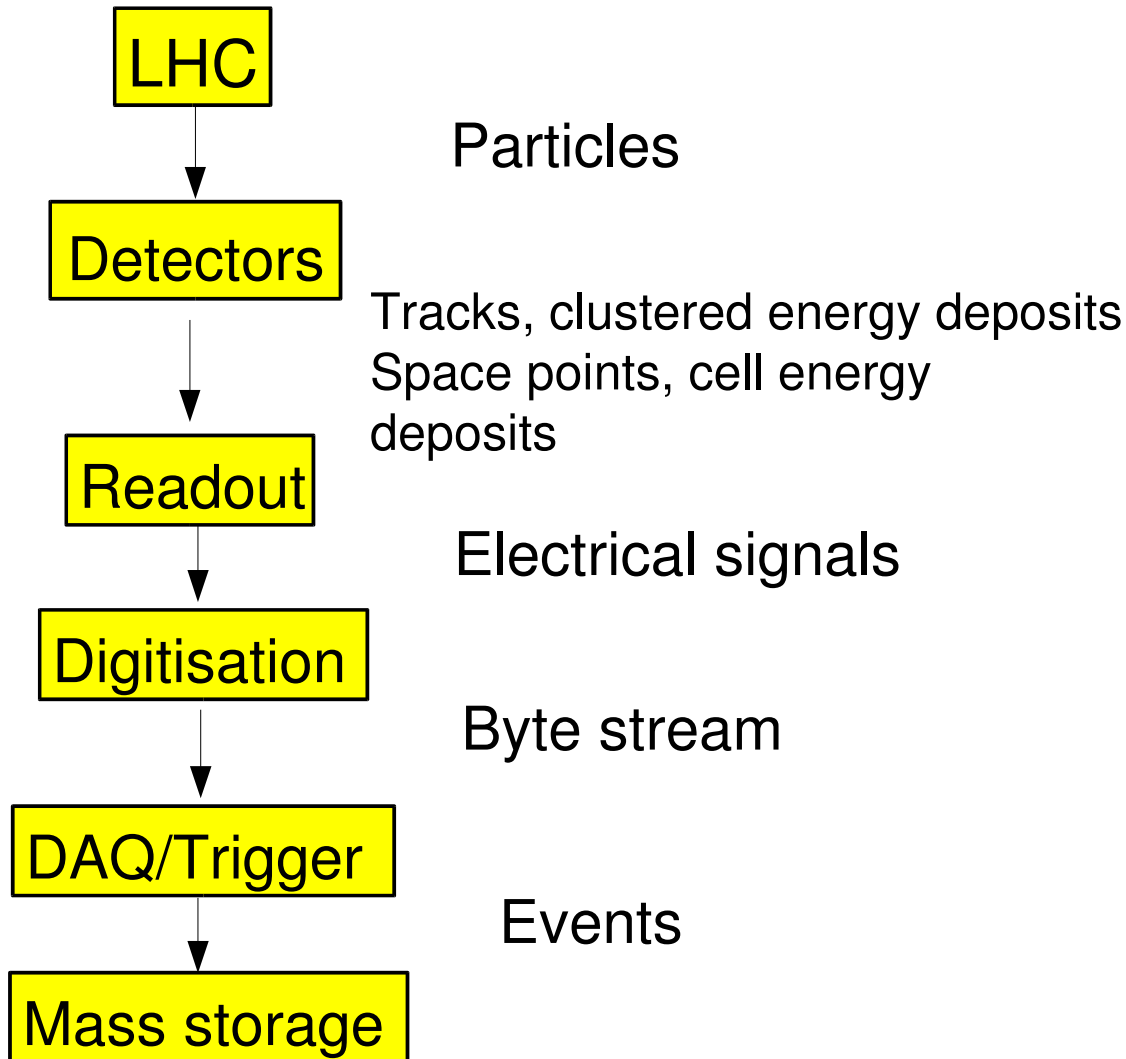


Data Analysis



Data Flow in Collider experiments (2)

Data Collection



Data Analysis

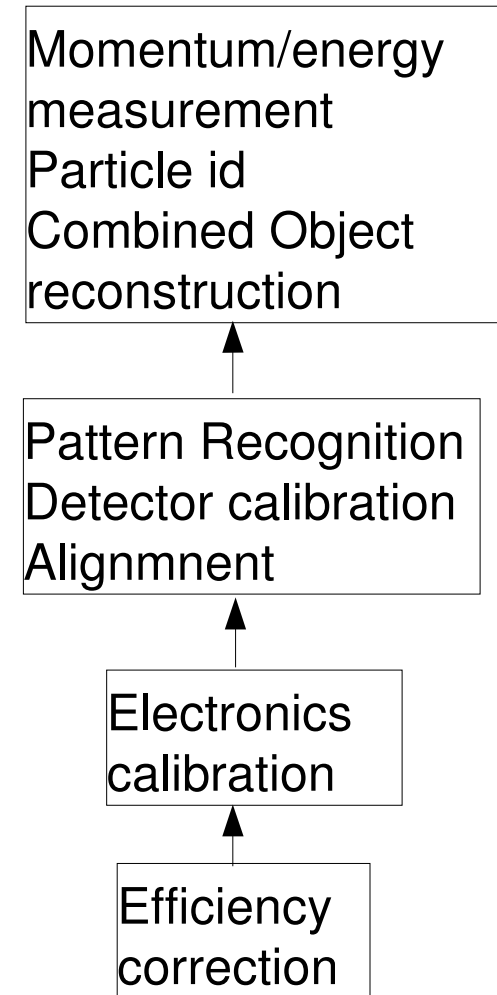


Diagram is a simplified view: in reality some of the parts proceed in parallel: e.g. some physics signals ($Z \rightarrow \ell\ell$) are also used for increase of detector understanding

Lectures will be based going up and down this data flow diagram:

- Discussion of how the physics aim influenced the requirements on the performance of the LHC and of the general purpose detectors
- Basic issues and limitation the detection and reconstruction of different objects in LHC detectors
- Description of the basic steps needed for the commissioning and understanding of a few detector subsystems
- Description of the steps for some example analyses:
 - $W \rightarrow \ell\nu$ cross-section
 - SUSY searches

LHC: pp Collider $\sqrt{s}=14$ TeV Startup: end 2009 with $\sqrt{s}=7$ TeV

Main motivations:

- Elucidate the mechanism of ElectroWeak Symmetry breaking:
 - Look for Higgs boson in allowed interval 100 GeV-1 TeV
 - In absence of low mass Higgs, study production of longitudinal gauge boson pairs.
- Find evidence for possible deviation from the Standard Model
 - Strong theoretical motivations to think that SM is only effective theory
 - In order to solve some of the theoretical difficulties with SM, deviations should be observable at \sim TeV scale

LHC Energy

$\sqrt{s} = 14$ TeV: explore the TeV scale, search for new massive particles up to 5 TeV

Maximum energy limited by the bending power needed to fit ring in 27 Km circumference LEP tunnel

$$p(\text{TeV}) = 0.3B (\text{T}) R(\text{km})$$

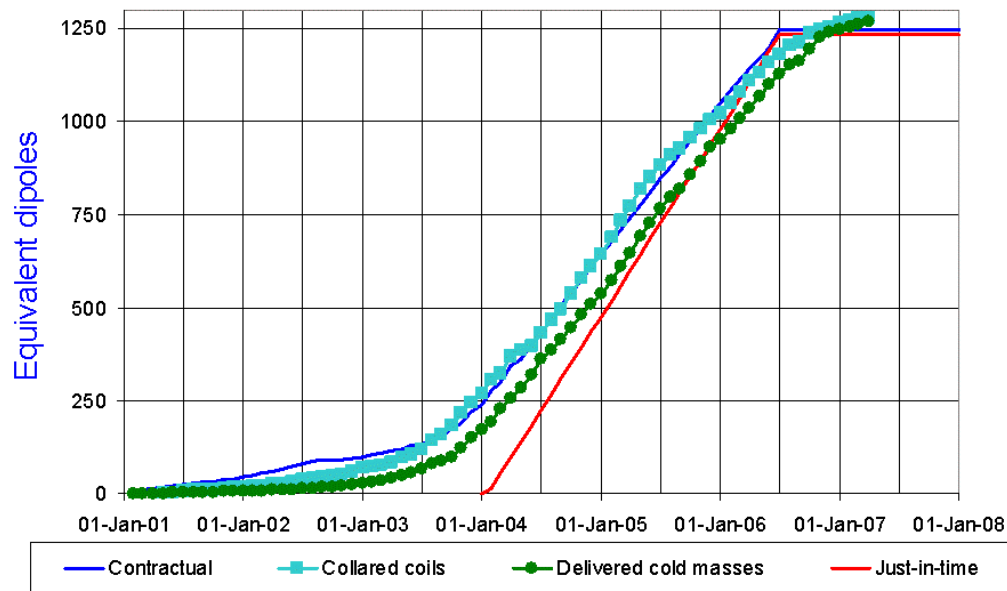


LHC Progress
Dashboard



Accelerator
Technology
Department

Dipole cold masses



LHC: $B = 8.4$ T:

~1300 superconducting dipoles
working at 1.9 K

All magnets were produced and
tested in time!

Luminosity:

$$\mathcal{L} = \frac{N}{\sigma}$$

with \mathcal{L} : Luminosity N : event frequency, σ : cross-section

Two luminosity scenarios:

- peak $\sim 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ - initial "low luminosity": $\int \mathcal{L} dt = 10 \text{ fb}^{-1}$ per year
- peak $\sim 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ - design "high luminosity": $\int \mathcal{L} dt = 100 \text{ fb}^{-1}$ per year

Benchmark: ensure detection of Higgs boson in the range 100 GeV-1 TeV

$$\begin{array}{l} m(H) \sim 100 - 150 \text{ GeV} \\ m(H) = 1 \text{ TeV} \end{array} \left| \begin{array}{l} H \rightarrow \gamma\gamma \\ H \rightarrow WW \rightarrow \ell\nu jj \end{array} \right. \left| \begin{array}{l} \sigma \times BR \times \epsilon \sim 10 - 20 \text{ fb} \\ \sigma \times BR \times \epsilon \sim 2 - 3 \text{ fb} \end{array} \right. \left| \begin{array}{l} S/B \sim 1/50 \\ S/B \sim 1/2 \end{array} \right.$$

Discovery when statistical significance for signal $S/\sqrt{B} > 5 \rightarrow$

Required integrated luminosity for discovery (no K -factors):

- $H \rightarrow \gamma\gamma$: ~ 1000 events $\sim 100 \text{ fb}^{-1}$
- $H \rightarrow WW$: ~ 50 events $\sim 20 \text{ fb}^{-1}$

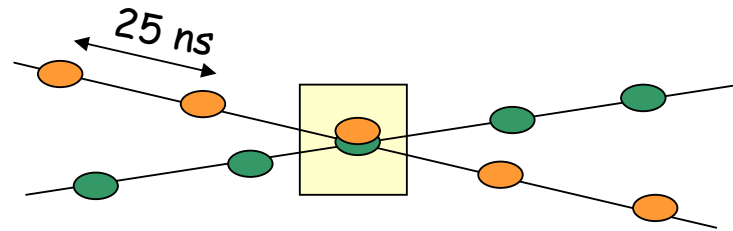
How is luminosity \mathcal{L} achieved?

If two beams containing n_1 and n_2 particles collide with a frequency f :

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi\sigma_{beam}^2}$$

with σ_{beam} gaussian transverse beam profile

LHC values: $n_1 = n_2 = 10^{11}$, and $\sigma_{beam} \sim 16 \times 10^{-6}$ m, determined by the physics of colliding beams.



To achieve $\mathcal{L} = 10^{34} \text{ cm}^{-2}\text{s}^{-1}$, LHC has to run with a bunch crossing every 25 ns

Inelastic proton-proton cross-section at $\sqrt{s} = 14$ TeV is ~ 70 mb \Rightarrow

LHC interaction rate at high luminosity: $\sim 7 \times 10^{-2} \times 10^{-24} \times 10^{34} = 7 \times 10^8$ Hz

40 MHz crossing frequency: $\Rightarrow \sim 25$ superimposed interactions per crossing
(pile-up)

Characteristics of pile-up interactions

Soft partonic interactions: describe with non-perturbative phenomenological models

Collider jargon: "**Minimum bias**": experimental definition: depends on experiment's trigger. **Usually associated to non-single diffractive events**

Measured at $S\bar{p}pS$ and Tevatron, large uncertainties in extrapolation to LHC

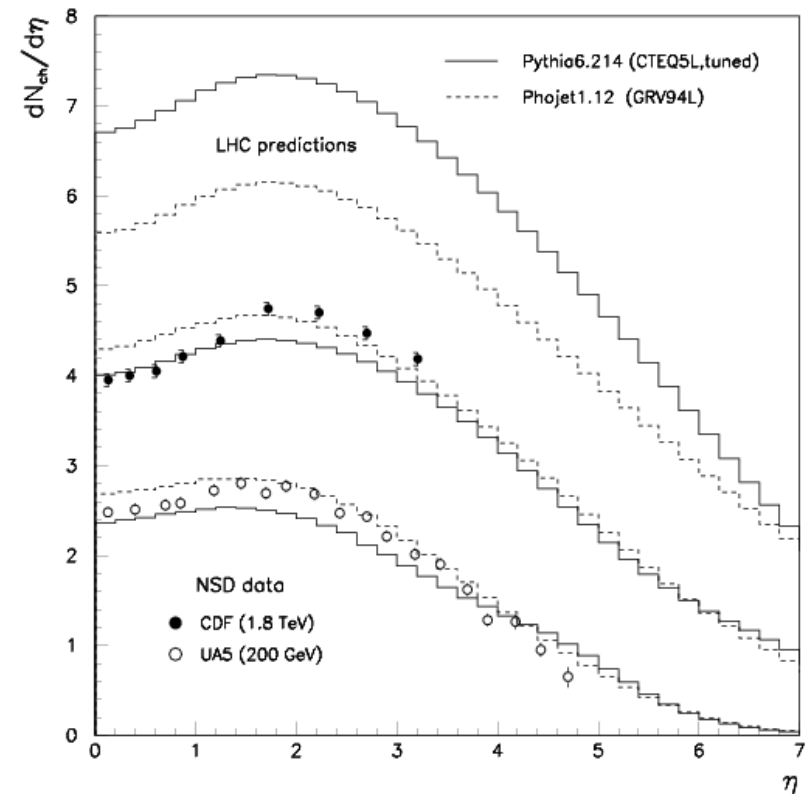
Main features:

~ 7 charged particles per unit of rapidity \Rightarrow

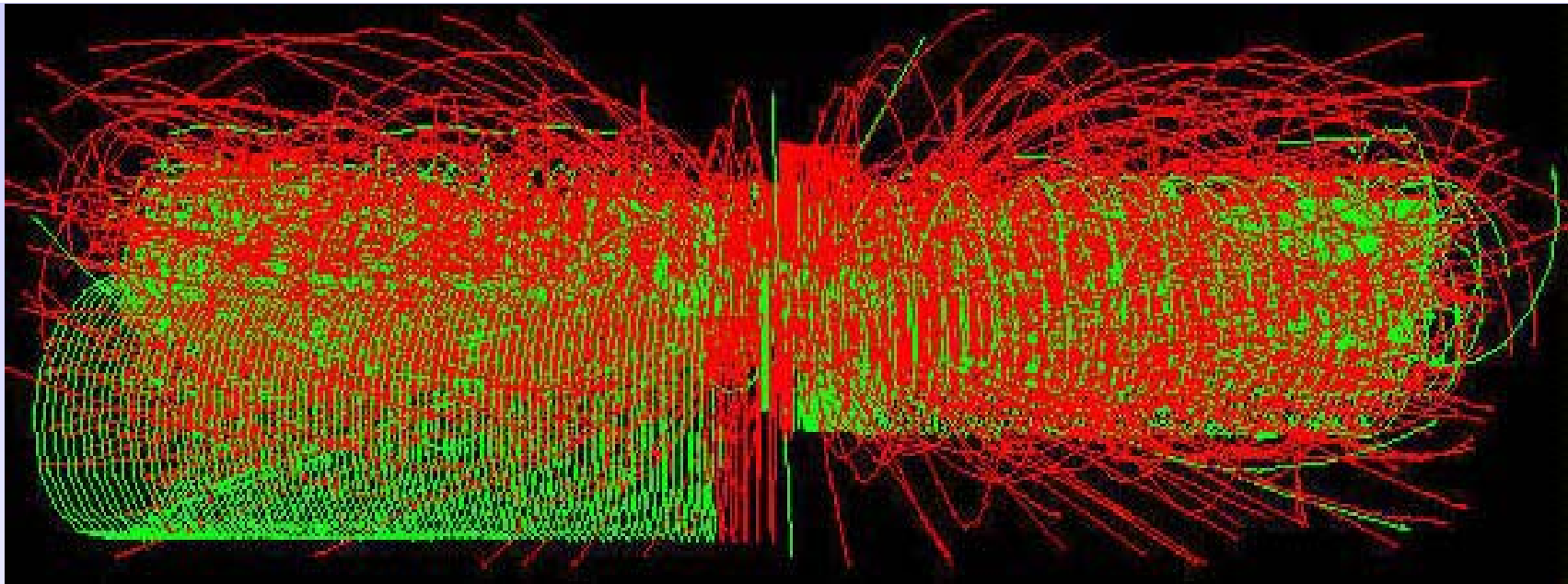
~ 100 charged particles over $|\eta| < 2.5$ per crossing at low luminosity

Significant radiation damage from interaction!

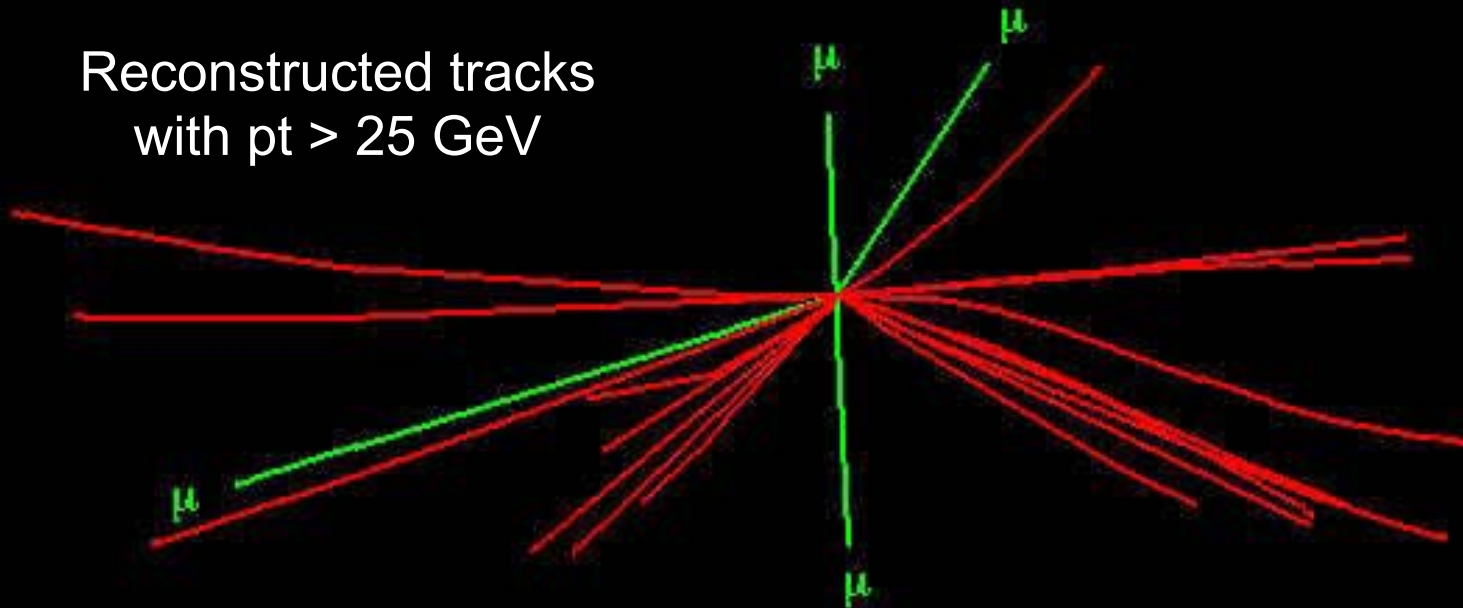
$\langle p_T \rangle \sim 500 \text{ MeV} \Rightarrow$ can select interesting particles by cut in p_T



Example: $h \rightarrow 4\mu$ event in CMS at high luminosity



Reconstructed tracks
with $p_t > 25$ GeV



Large impact on detector design:

- **Speed:**

LHC detectors must have fast response otherwise integrate over too many bunch crossings

Typical response time: 20-50 ns \rightarrow integrate over 1-2 bunch crossings

\Rightarrow very challenging readout electronics

- **Granularity:**

LHC detectors must be highly granular to minimise probability that pile-up particles in same detector element as interesting object

\Rightarrow Large number of electronics channels

- **Radiation hardness:**

High flux of particles from pp collisions \Rightarrow high radiation environment

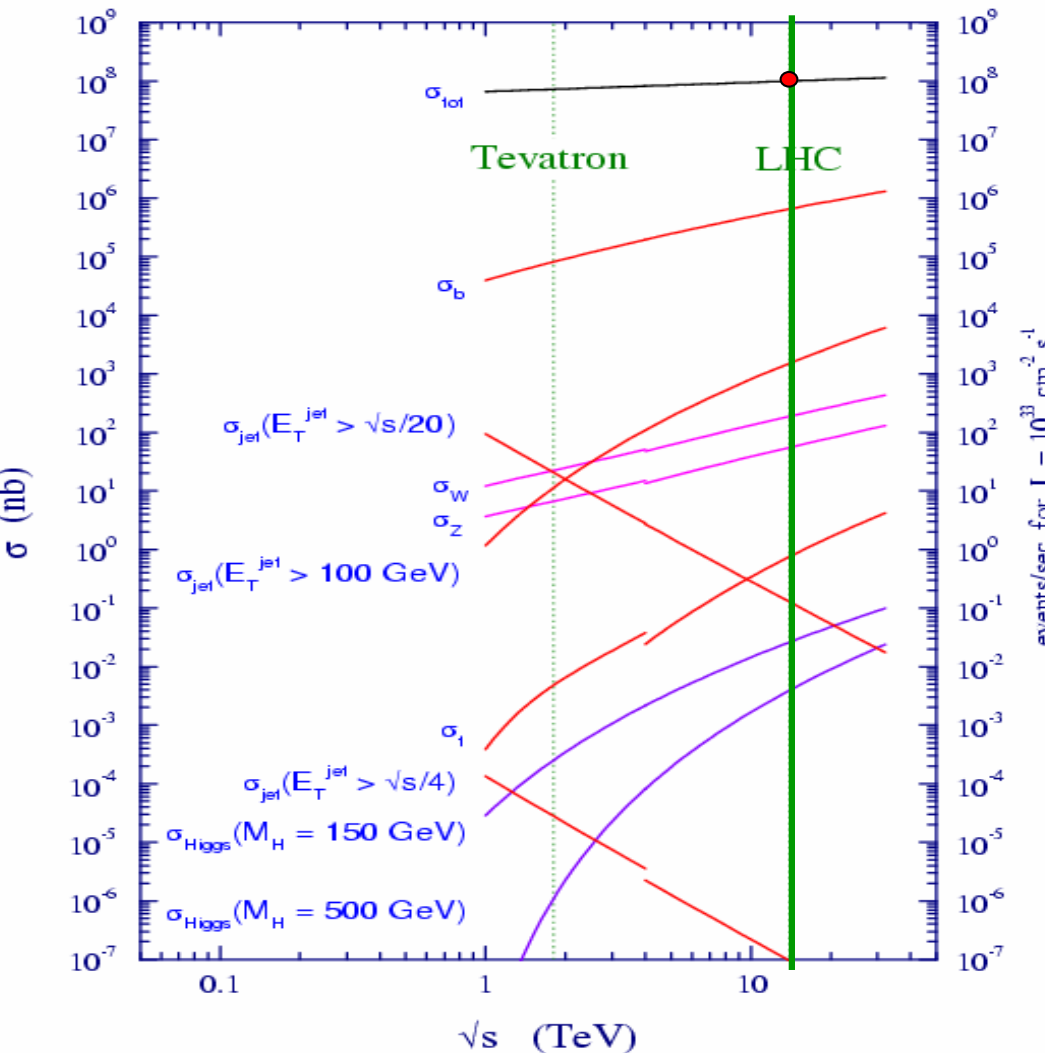
In 10 years of LHC data: up to $10^{17} n \text{ cm}^{-2}$, up to 10^7 Gy

Radiation decrease like d^2 from beam: detectors near beam pipe mostly affected

\Rightarrow Need radiation resistant detector technologies especially at high $|\eta|$

\Rightarrow Need also radiation hard electronics

Backgrounds to discovery physics



High p_T events dominated by QCD jet production:

- Strong production
- Many contributing diagrams

$$\sigma_{jet}(E_T^{jet} > 100 \text{ GeV}) \sim \mu\text{b}$$

Signal processes rare:

- Involve heavy particles:

$$\sigma_{\tilde{q}\tilde{q}}(m(\tilde{q}) \sim 1 \text{ TeV}) \sim \text{pb}$$

- Have weak cross-section

$$\sigma_{Higgs}(m(Higgs) = 100 \text{ GeV}) \sim 30 \text{ pb}$$

QCD background from 5-6 orders of magnitude larger than signals

Overwhelming QCD backgrounds in exclusively hadronic channels

\Rightarrow rely on final states involving γ , leptons, \cancel{E}_T , b-jets \Rightarrow pay additional price in BR

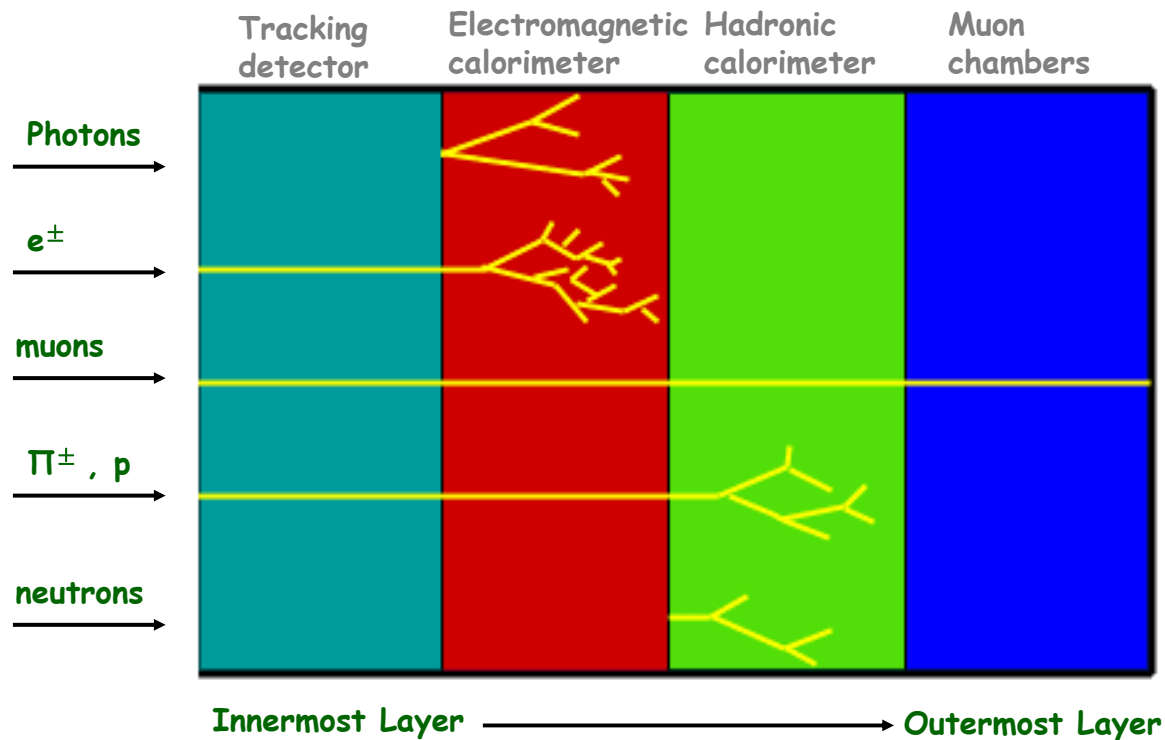
Generic Collider detector

Do not know how new physics will manifest itself:

⇒ Detectors must be sensitive to as many particles and signatures as possible:

$e, \mu, \tau, \nu, \gamma$, jets, b – quarks

Implement this through a layered structure of subdetectors round the beam pipe



ATLAS and CMS detectors

Schema of previous page implemented in both detectors:

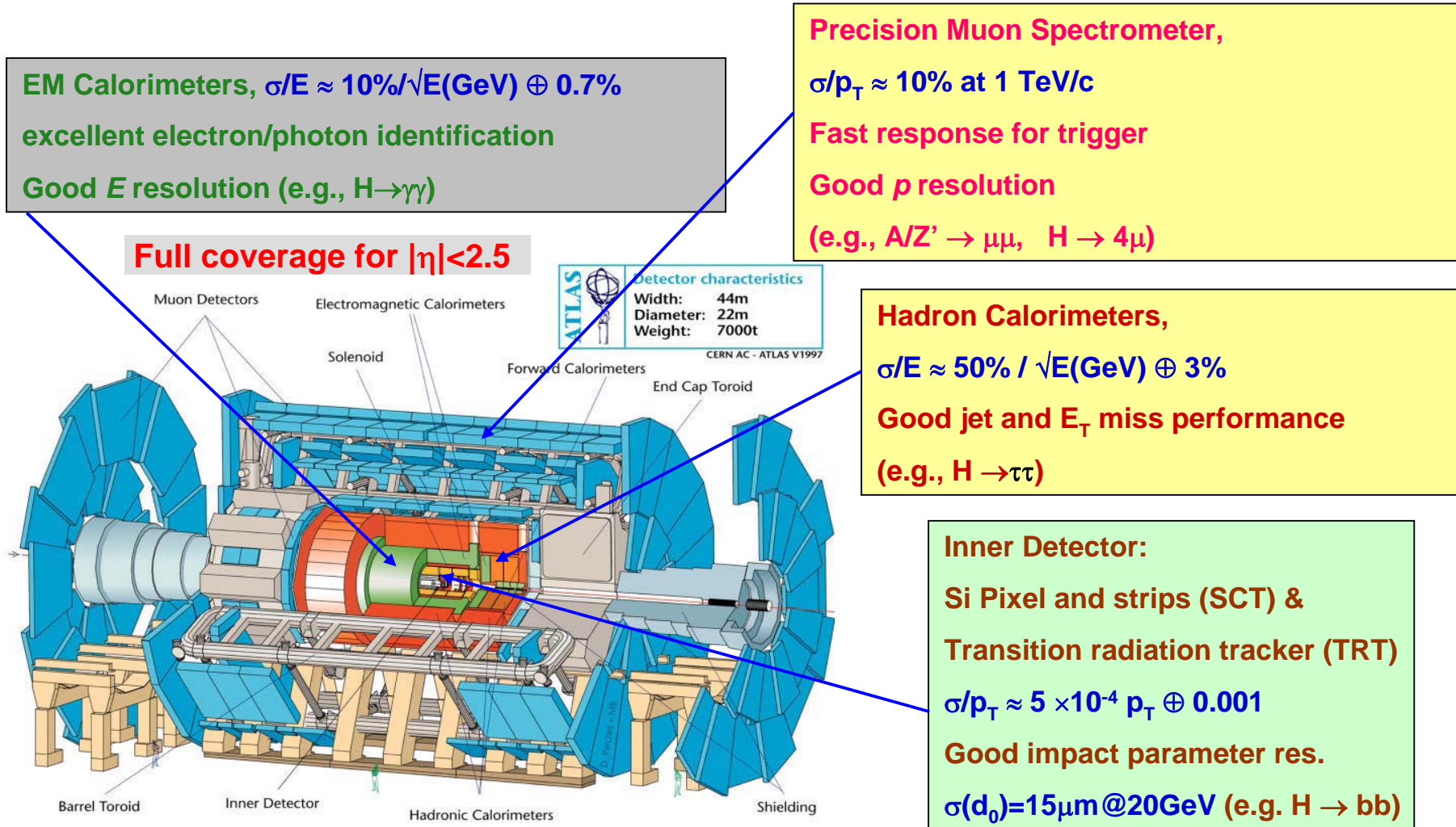
- Momentum/charge of **tracks and secondary vertices** (e.g. from b -quark decays) measured in **central tracker**. Excellent momentum and position resolution required
- Energy and position of **electrons and photons** measured in **electromagnetic calorimeters**. Excellent position and energy resolution required
- Energy and position of **hadrons and jets** measured mainly in **hadronic calorimeters**.

Good coverage and granularity required

- **Muons** identified and momentum measured in **external muon spectrometer** (+ central tracker). Excellent resolution required.
- **Neutrinos** “detected and measured” through measurement of **missing transverse energy \cancel{E}_T** . Calorimeter coverage over $|\eta| < 5$ needed

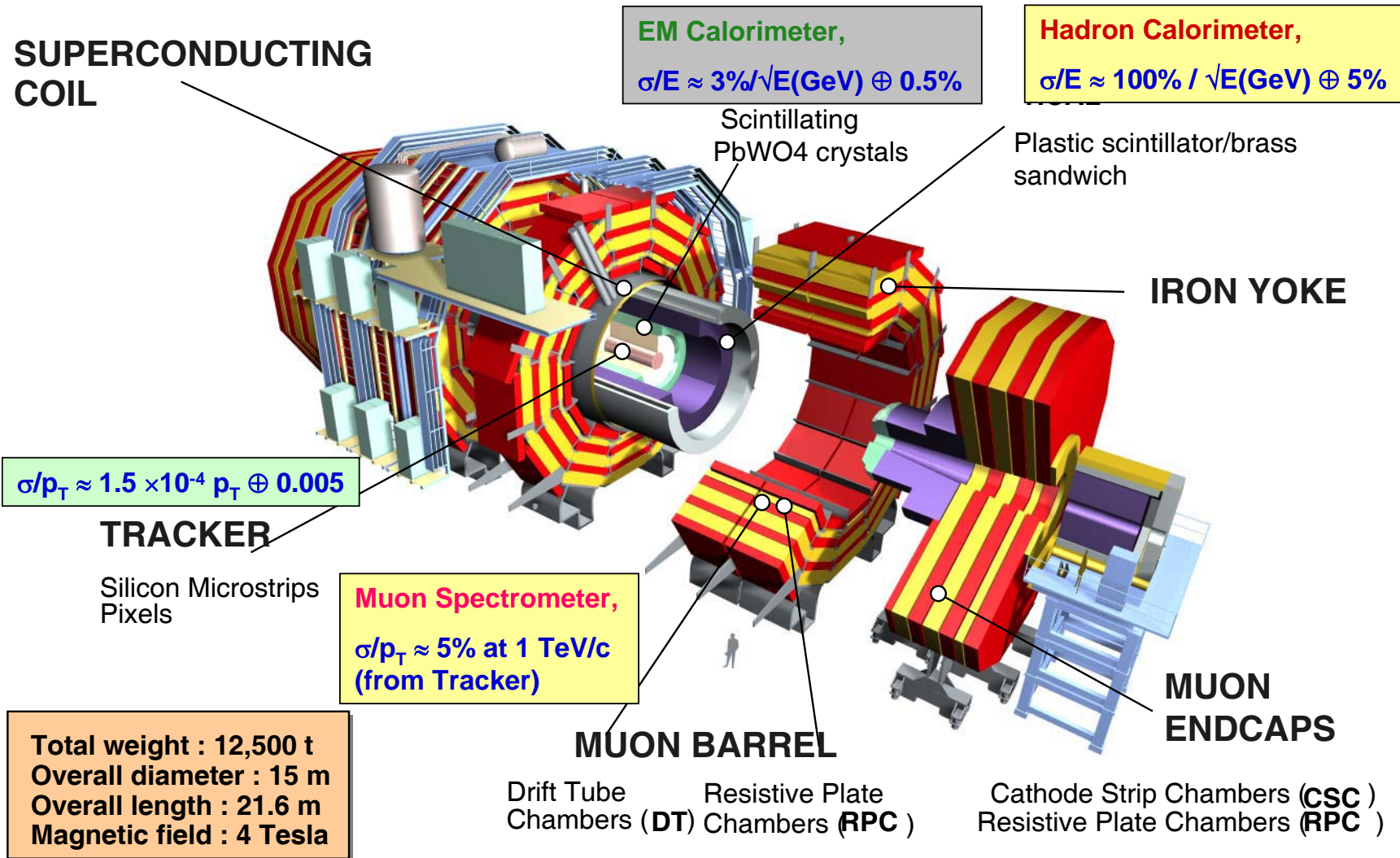
Difference between detectors is in choice of technologies for different subcomponents and in configuration of magnetic field

ATLAS detector



Magnets: solenoid (Inner Detector) 2T, air-core toroids (Muon Spectrometer) ~0.5T

CMS detector



	ATLAS	CMS
MAGNET (S)	Air-core toroids + solenoid in inner cavity Calorimeters outside field 4 magnets	Solenoid Calorimeters inside field 1 magnet
TRACKER	Si pixel + strips TRD → particle identification B=2T $\sigma/p_T \sim 5 \times 10^{-4} p_T \oplus 0.01$	Si pixel + strips No particle identification B=4T $\sigma/p_T \sim 1.5 \times 10^{-4} p_T \oplus 0.005$
EM CALO	Pb - liquid argon $\sigma/E \sim 10\%/\sqrt{E}$ uniform longitudinal segmentation	PbWO ₄ crystals $\sigma/E \sim 3-5\%/\sqrt{E}$ no longitudinal segm.
HAD CALO	Fe-scintillator + Cu-liquid argon (10 λ) $\sigma/E \sim 50\%/\sqrt{E} \oplus 0.03$	Cu-scint. (> 5.8 λ + catcher) $\sigma/E \sim 65\%/\sqrt{E} \oplus 0.05$
MUON	Air → $\sigma/p_T \sim 7\%$ at 1 TeV standalone	Fe → $\sigma/p_T \sim 5\%$ at 1 TeV combining with tracker

A few examples of required performance:

- Lepton measurement: $p_T \sim \text{GeV} \rightarrow 5\text{TeV}$ ($b \rightarrow lX, W', Z'$)

- Mass Resolution ($m \sim 100 \text{ GeV}$):

$$\sim 1\% \quad (H \rightarrow \gamma\gamma, 4l)$$

$$\sim 10\% \quad (W \rightarrow jj, H \rightarrow bb)$$

- Calorimeter coverage: $|\eta| < 5$ (E_T^{miss} , forward jet tag)

- Particle identification :

$$\epsilon_b \sim 50\% \quad R_j \sim 100 \quad (H \rightarrow bb, \text{SUSY})$$

$$\epsilon_\tau \sim 50\% \quad R_j \sim 100 \quad (A/H \rightarrow \tau\tau)$$

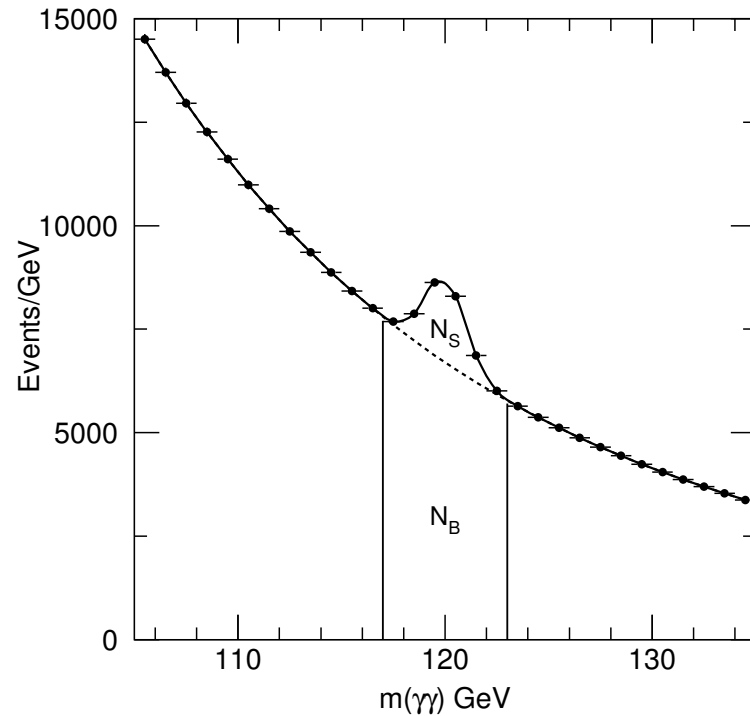
$$\epsilon_\gamma \sim 80\% \quad R_j \sim 10^3 \quad (H \rightarrow \gamma\gamma)$$

$$\epsilon_e > 50\% \quad R_j \sim 10^5$$

- Trigger: 40 MHz \rightarrow 100 Hz reduction

Example of optimisation of discovery potential

Suppose a narrow particle $X \rightarrow \gamma\gamma$ is produced:



Signal significance:

$$S = \frac{N_s}{\sqrt{N_B}}$$

N_s = number signal events

N_B = number of background events

$\sqrt{N_B}$ = error on number of background events for large numbers

Otherwise use Poisson statistics

$S > 5$: signal is larger than 5 times the error on background.

Gaussian probability that background fluctuates up by more than 5σ : 10^{-7}

LHC detector performance optimised for case of $H \rightarrow \gamma\gamma$

Parameters for maximising significance

- Detector resolution (σ_m):

If σ_m increases, need to correspondingly enlarge peak region to keep same number of signal events $\Rightarrow N_B$ increases by same factor ($B \sim flat$)

$$\Rightarrow S = N_s / \sqrt{N_B} \text{ decreases} \Rightarrow S \sim 1 / \sqrt{\sigma_m}$$

This is only valid for $\Gamma_H \ll \sigma_m$ If Higgs broad, detector resolution not relevant

- Integrated luminosity L :

$$N_s \sim L \quad N_B \sim L \quad \Rightarrow S \sim \sqrt{L}$$

- Signal selection efficiency ϵ

$$N_s \sim \epsilon. \text{ If background rejection constant: } S \sim \epsilon$$

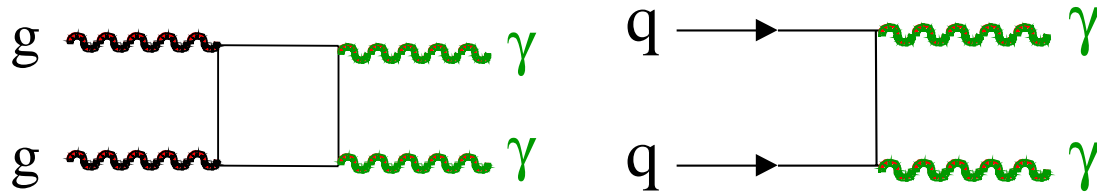
- Rejection of reducible backgrounds R .

$$\text{If they dominate and } N_s \text{ constant : } S \sim \sqrt{R}.$$

Typically try to achieve R such that reducible backgrounds \ll irreducible ones

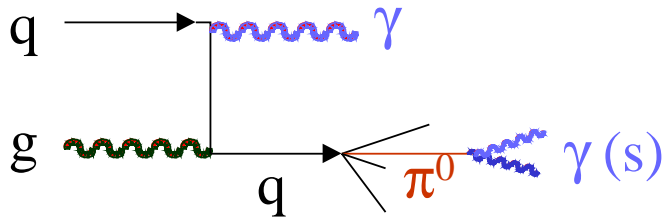
$H \rightarrow \gamma\gamma$: backgrounds

- $\gamma\gamma$ production, irreducible (same final state as signal)



$$\frac{\sigma_{\gamma\gamma}}{\sigma(H \rightarrow \gamma\gamma)} \sim 60$$

- γ +Jet, Jet+Jet where one or both jets fake a photon, reducible



$$\frac{\sigma_{\gamma j}}{\sigma_{\gamma\gamma}} \sim 10^3 \quad \frac{\sigma_{jj}}{\sigma_{\gamma\gamma}} \sim 10^6$$

Need rejection factor of $\sim 10^3$ on jets for $\epsilon_\gamma = 80\%$ in order to bring reducible background below $\gamma\gamma$ background

Fake photons dominated by π^0 , design EM calo with excellent granularity to achieve desired γ/π^0 rejection

$H \rightarrow \gamma\gamma$: irreducible background

Irreducible background: can not be suppressed with detector cuts

$$\sigma_{\gamma\gamma} = 2 \text{ fb/GeV for } m_{\gamma\gamma} = 100 \text{ GeV; } \Gamma_H < 1 \text{ MeV} \Rightarrow \text{Significance} \sim 1/\sqrt{\sigma_m}$$

$$\sigma_m^2 = 2E_1E_2(1 - \cos \theta_{\gamma\gamma}) \quad \text{Two contributions to } \sigma_m: \sigma_E \text{ and } \sigma_\theta$$

ATLAS

Lead-liquid argon sampling calorimeter:

$$\frac{\sigma(E)}{E} \sim \frac{10\%}{\sqrt{E}}$$

Longitudinal segmentation: measure γ direction

$$\sigma(\theta) \sim \frac{50\text{mrad}}{\sqrt{E}}$$

When available use reconstructed conversions or reconstructed primary vertex

ATLAS result for $m_H \sim 120 \text{ GeV}$: $\sigma_m = 1.46 \text{ GeV}$

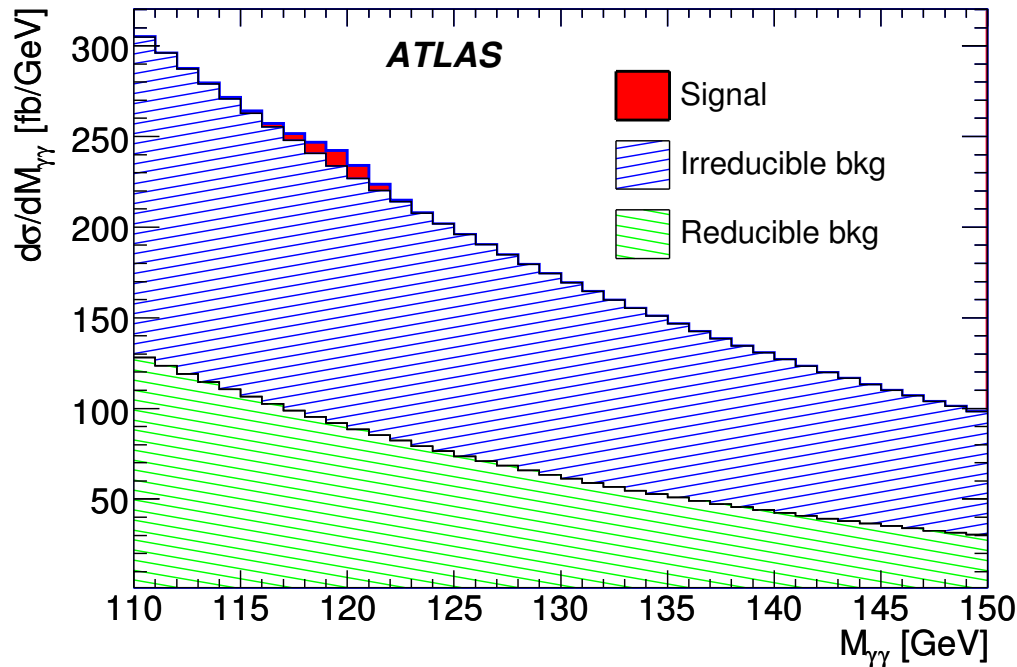
CMS

Crystal calorimeter:

$$\frac{\sigma(E)}{E} \sim \frac{2 - 5\%}{\sqrt{E}}$$

No longitudinal segmentation: vertex measured from secondary tracks from spectator partons, difficult at high L, available only for 80% of events

$H \rightarrow \gamma\gamma$: resulting ATLAS performance



Two isolated photons

$$P_t(\gamma_1) > 40 \text{ GeV}$$

$$P_t(\gamma_2) > 25 \text{ GeV}$$

$$|\eta| < 2.37, \text{ excluding crack region}$$

Signal efficiency within $\pm 1.4\sigma_m$ of the peak is 26%, contributions from:

- Kinematic cuts
- Photon Id and isolation cuts ($\sim 80\%$ per leg)
- Mass bin ($\sim 73\%$)

For $m(H) = 120 \text{ GeV}$, 1 fb^{-1} , 25.4 ev. signal, 947 bg. ($N_s/N_B \sim 2.5\%$)

Background dominated by $\gamma\gamma$ events, can be determined from sidebands

Improved sensitivity by combining with analyses on exclusive channels with 1 or more jets.

Electron-photon identification (ATLAS)

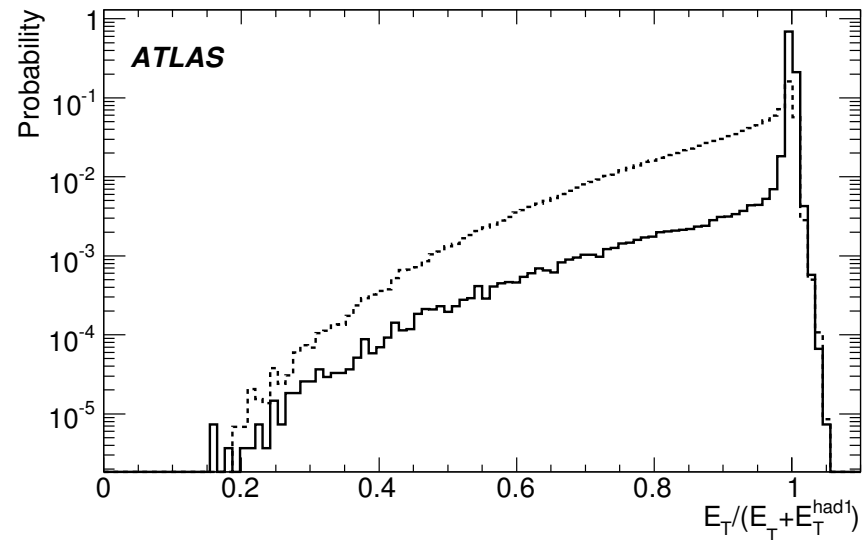
Separate electrons/photons from the overwhelming background of QCD jets

Reject charged hadrons in jets through longitudinal and lateral energy deposition pattern (lateral and longitudinal segmentation). Identify EM object

Example: show ratio of energy in EM compartment to sum of energy in EM compartment and in first hadronic compartment.

Electrons: full line

Jets: dotted line



Main remaining background : fragmentation of quarks/gluons where a π^0 carries away most of the momentum, with the decay $\pi^0 \rightarrow \gamma\gamma$

Distinguish two photons from π^0 decay from single photon through detailed study of

EM shower in Calorimeter: High EM calo granularity to separate two photons

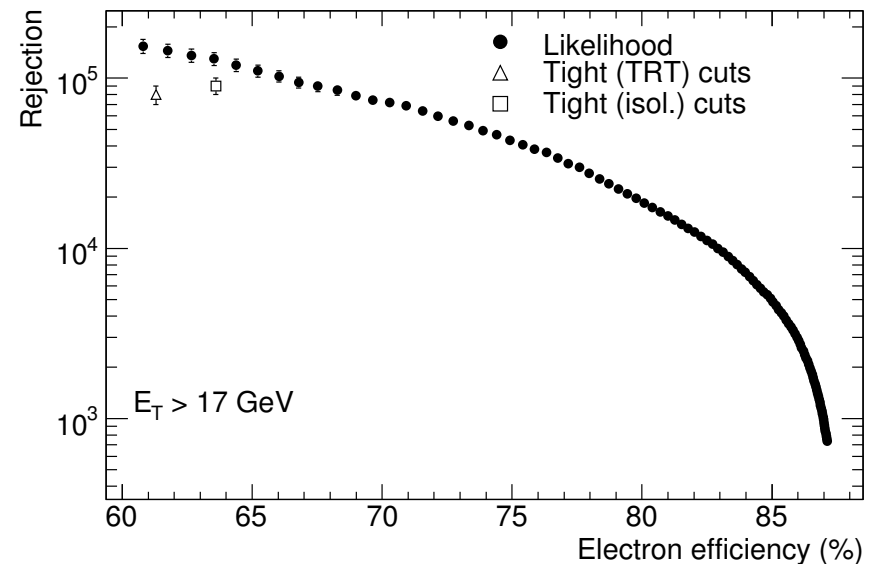
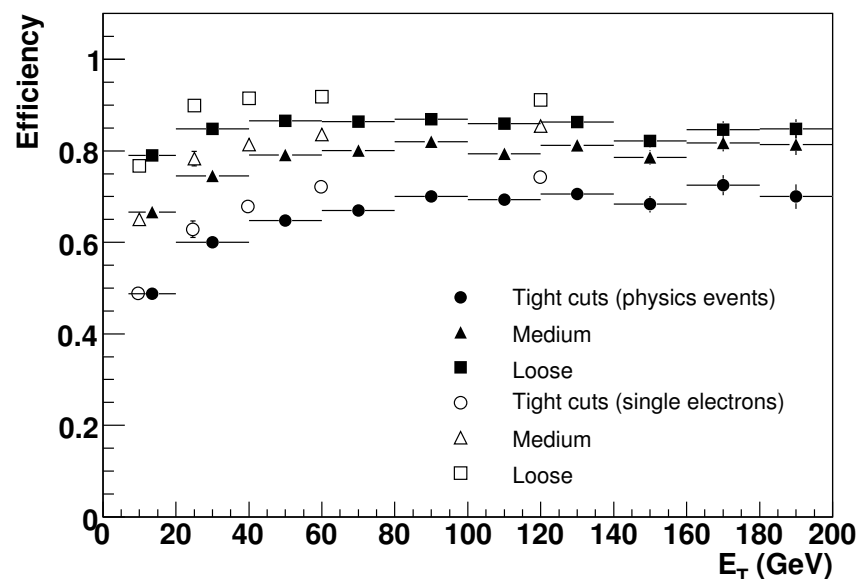
Electron identification (2)

If track from π^\pm superimposed to EM cluster can fake electron

Use matching between position/momentum of track and position/energy of EM cluster to reject fake electrons: Require excellent EM energy and position resolution

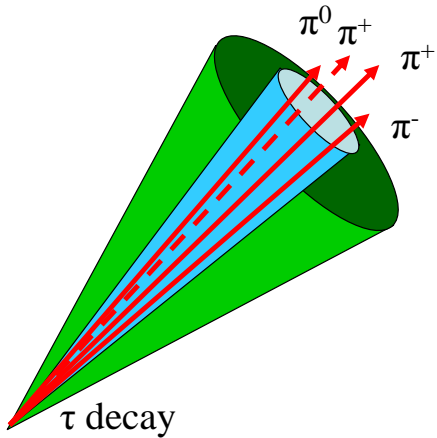
Further cuts include an isolation request, which suppresses electrons from B decays, and a Transition Radiation Signal

In ATLAS defined three default sets of requirements, based on combination of the above cuts, yielding different efficiencies



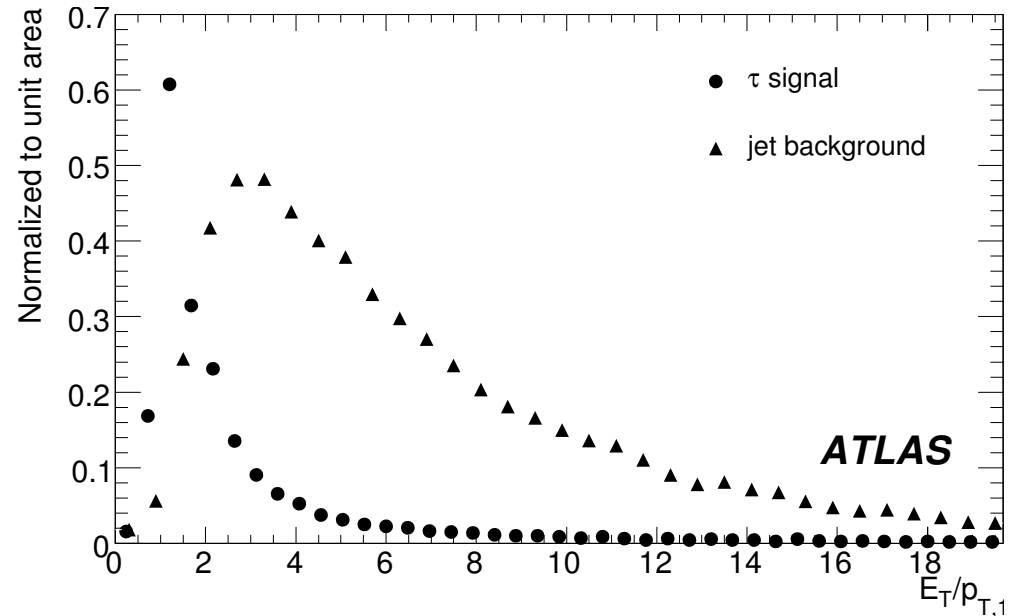
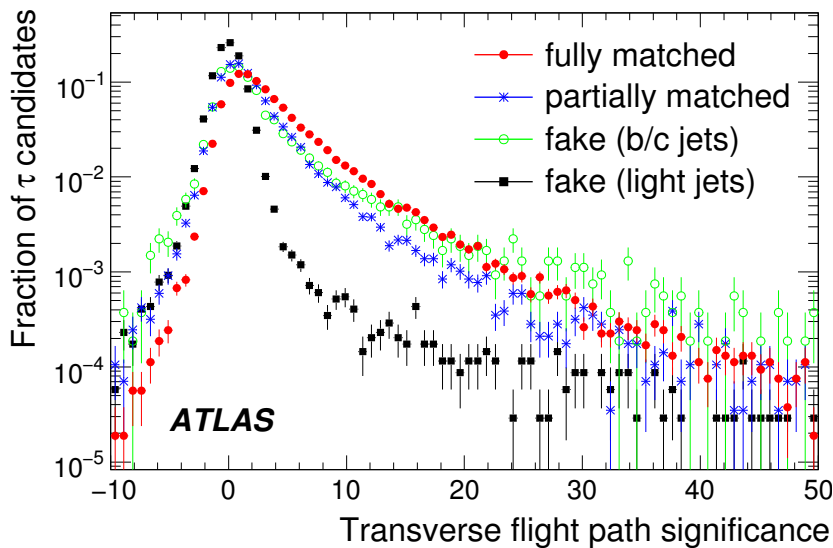
Identification of τ hadronic decays

Taus decay hadronically into one or three π^\pm plus π^0 and ν_τ



Exploit difference between hadronic τ decays and QCD jets:

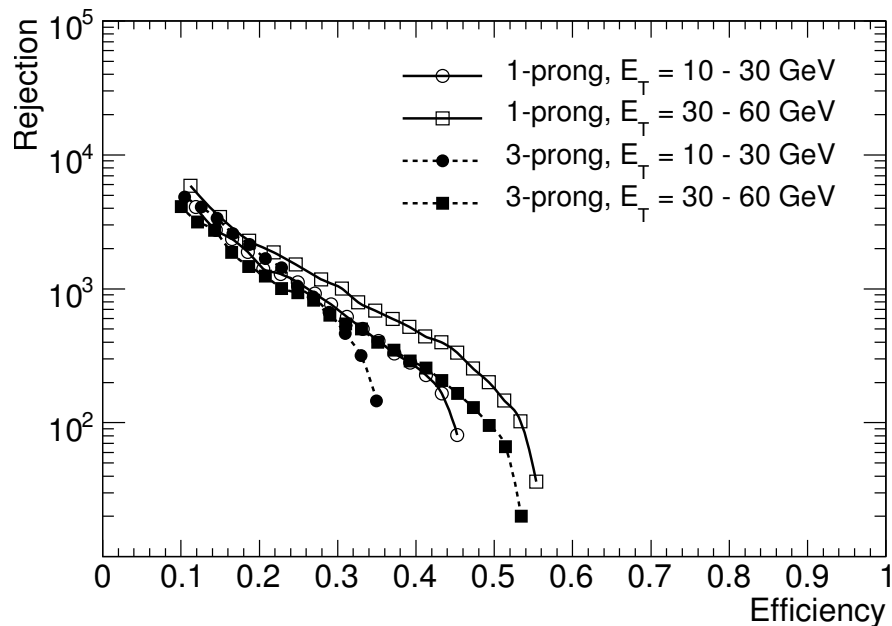
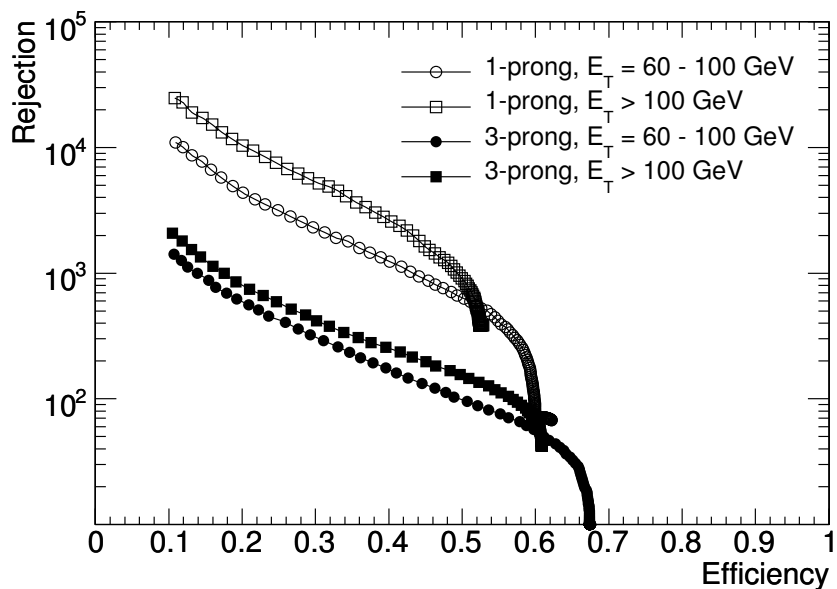
- Low track multiplicity ($1 < N_{tr} < 3$)
- Collimated track topology for 3-prong decays
- Narrow jet in calo
- Large fraction of energy carried by leading track (right plot)
- Non zero Impact parameter (left plot)



τ identification (2)

Two complementary approaches in ATLAS:

- Calo-based, seeded by calorimeter cluster, build likelihood on discriminat variables (right)
- Track-based, seeded by tracks in inner detector, energy is recostruced from tracks and EM energy deposit. Perform discrimination based on a NN (left)

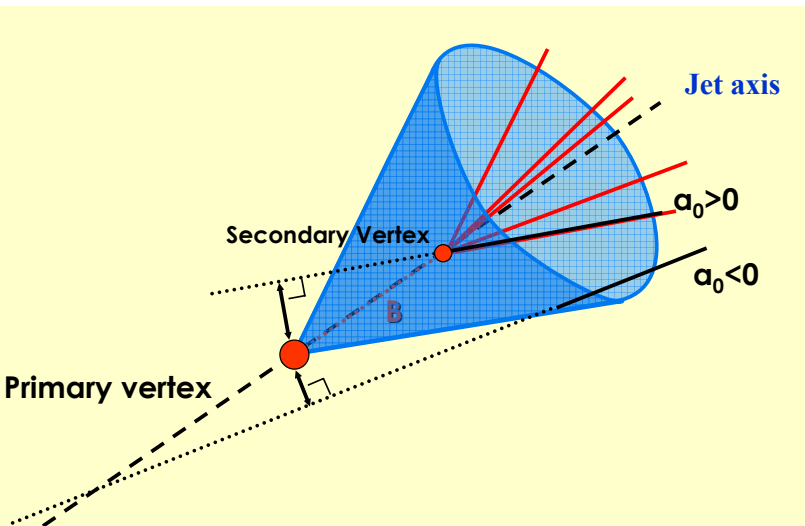


Track-based method effective for low- p_T τ , calo-based at high p_T

Strong difference between 1 and 3-prong decays dor calo method

For 50% efficiency achieve rejection between 100 and 1000

B-tagging



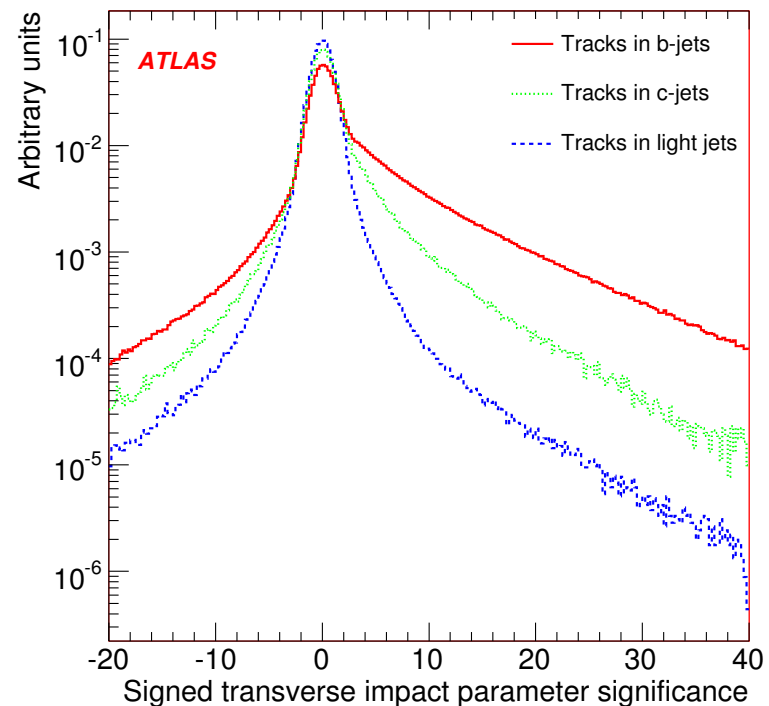
b-hadrons decay a few mm away from interaction vertex. Two main features:

- Long decay path
- Secondary vertex inside jet

Decay path measured through **impact parameter**: minimum distance from primary vertex

Distribution of significance of impact parameter: symmetric for tracks from fragmentation of **light quarks**

Enhancement on positive side for tracks from **b-hadron decays**

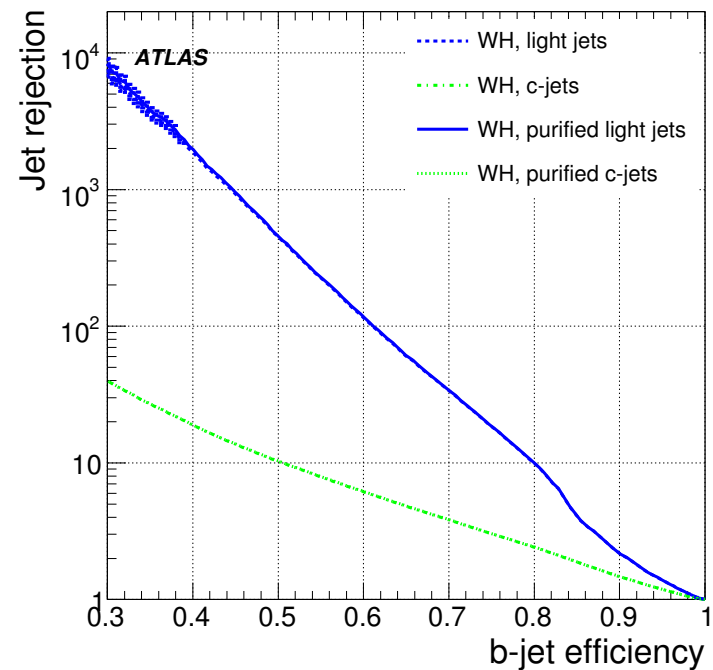
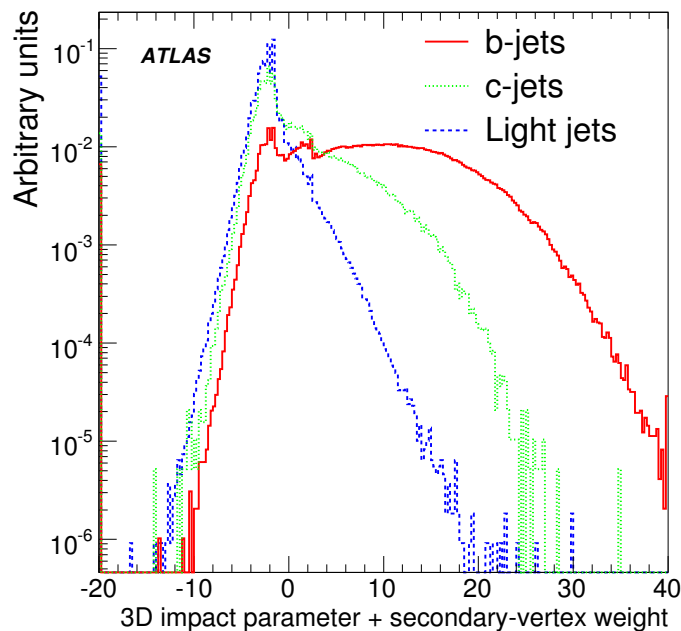


B-tagging (cont)

For a jet, build likelihood function from the impact parameter of the tracks associated to it and from variables for secondary vertex

ATLAS: Study samples of fully simulated WH , ttH , $t\bar{t}$ events

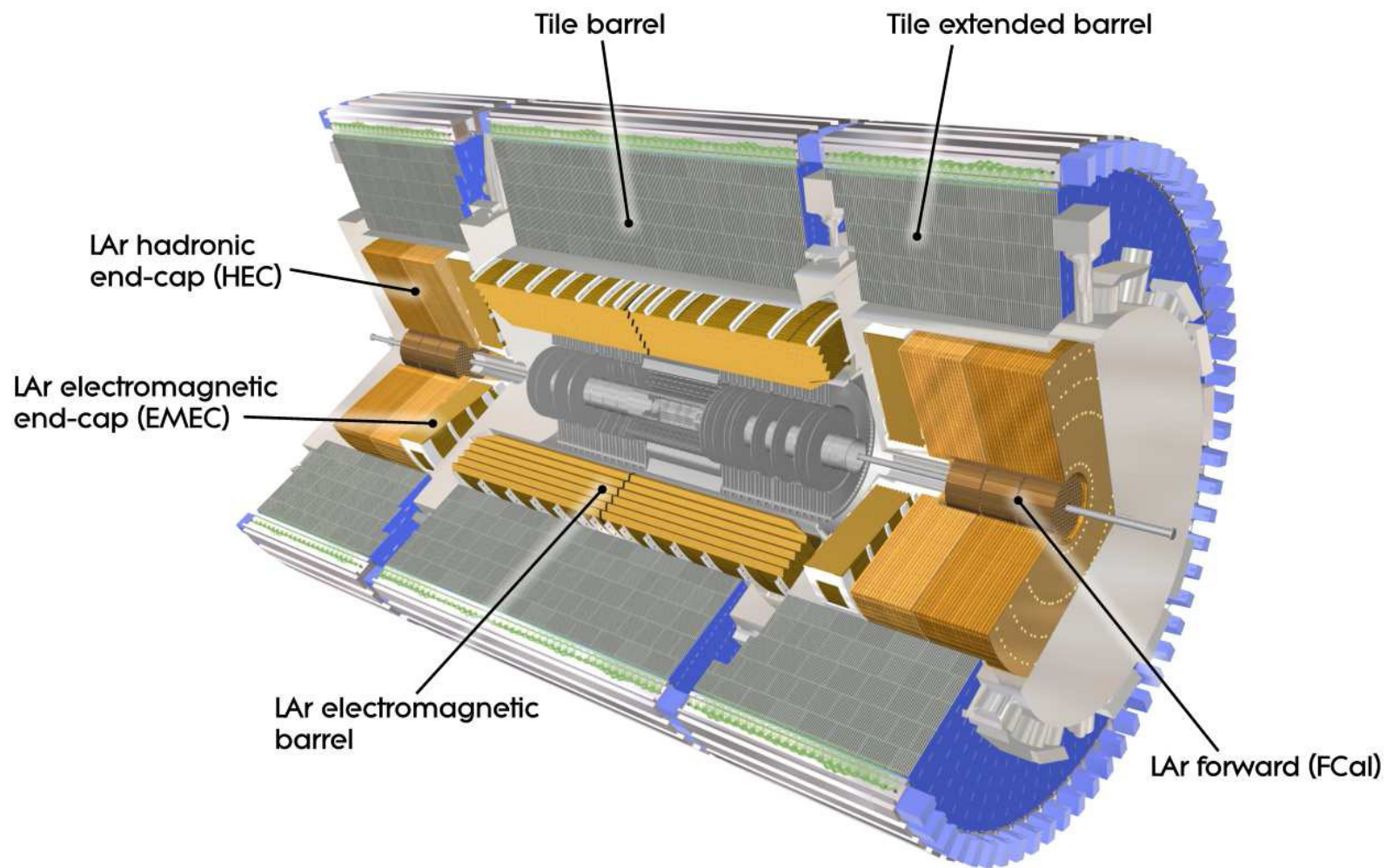
Measure rejection on QCD jets as a function of tagging efficiency



For WH sample observe rejection factor of 100 on light jets for $\epsilon_b = 60\%$

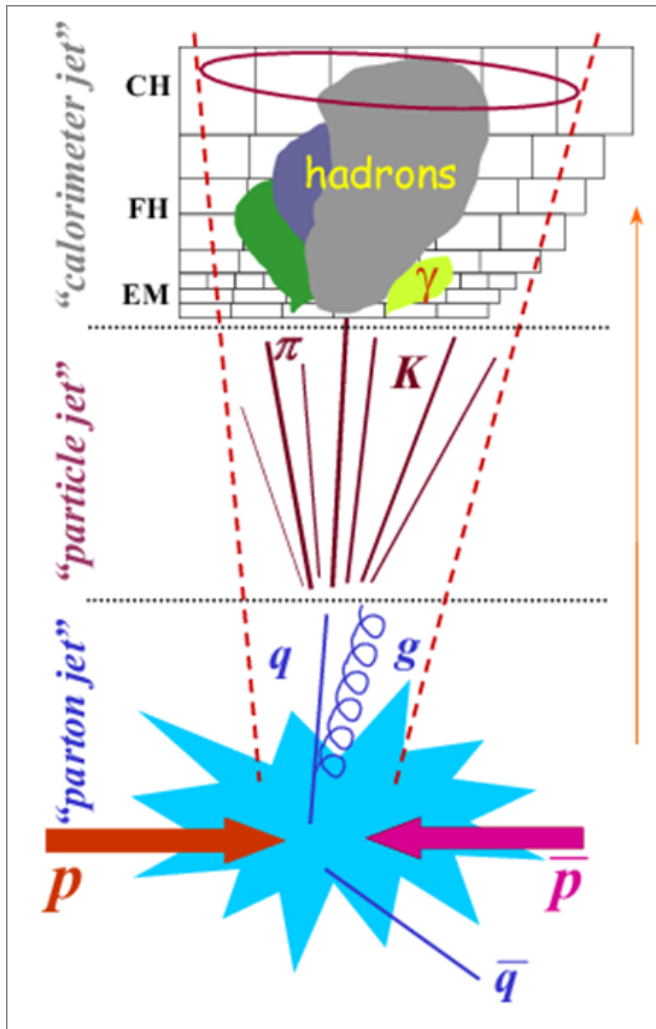
The efficiency/rejection curve is a function of the used sample.

ATLAS Calorimeter system



Jet reconstruction

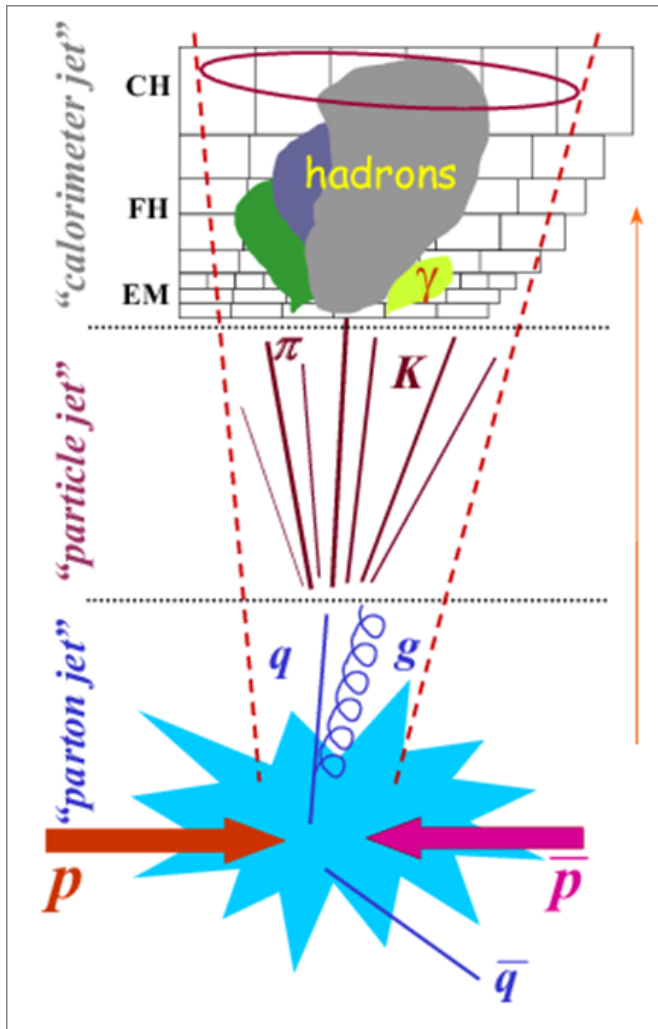
Jet reconstruction: starting from the calorimeter signals obtain the kinematics of the particle jet or of the parton jet, depending on where we want to perform Data-Theory comparison



Calorimeter jet and Particle jet obtained running the same jet algorithm on calo cell or on stable particles

Crucial element in Data-Theory comparison, needs to satisfy theoretical requirements (infrared-safe) and experimental requirements (e.g. low sensitivity to soft part of event)

Jet reconstruction phase 1: the Calorimeter jet



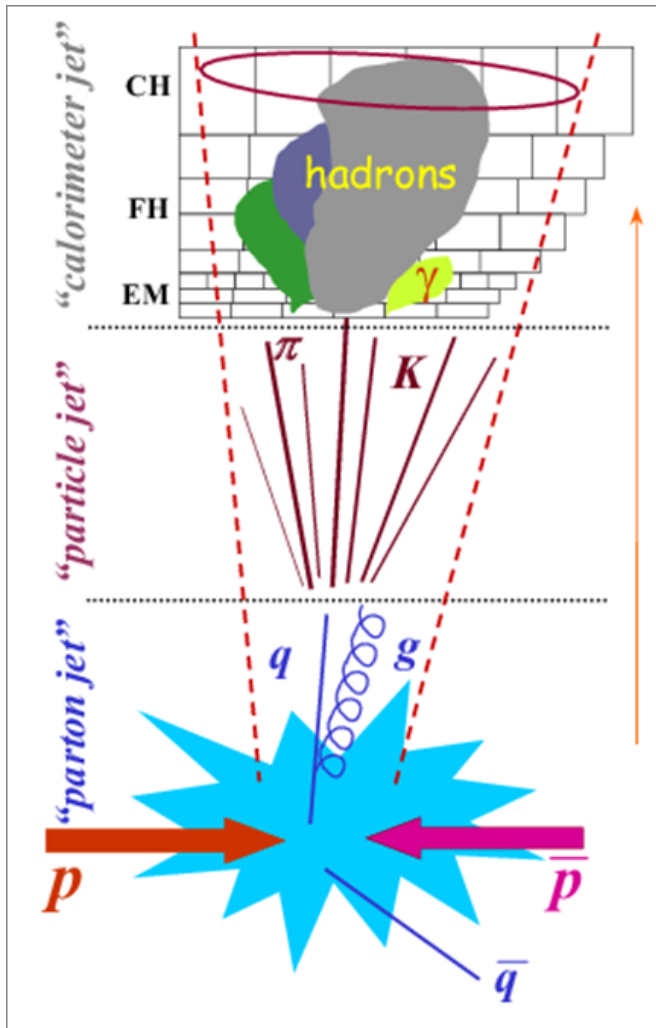
Hadronic signal definition: cell energy deposits are clustered to build the base objects for jet reconstruction

Algorithms for suppression of electronic and pile-up noise are applied at this stage

In ATLAS two types of base objects:

- Towers of dimension $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$
- 3D energy blobs

Jet reconstruction phase2: to Particle jet

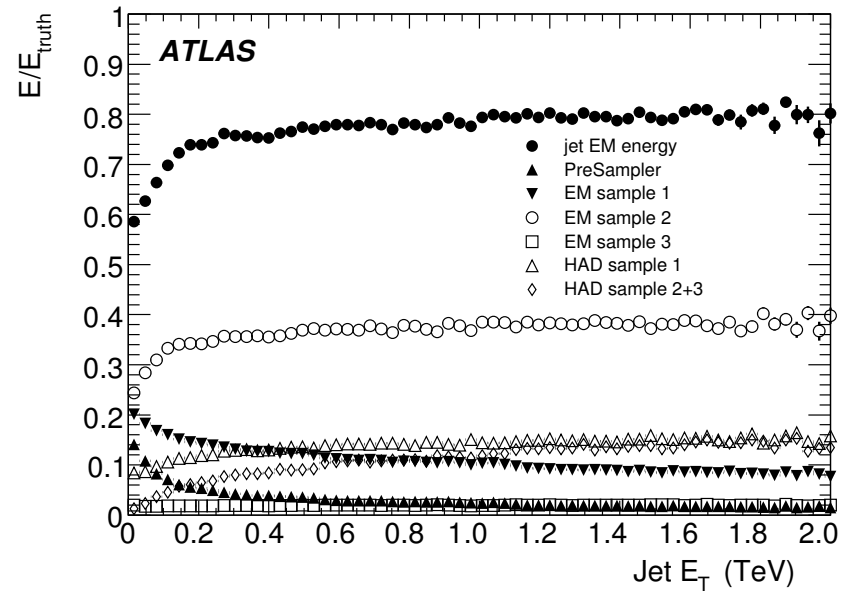
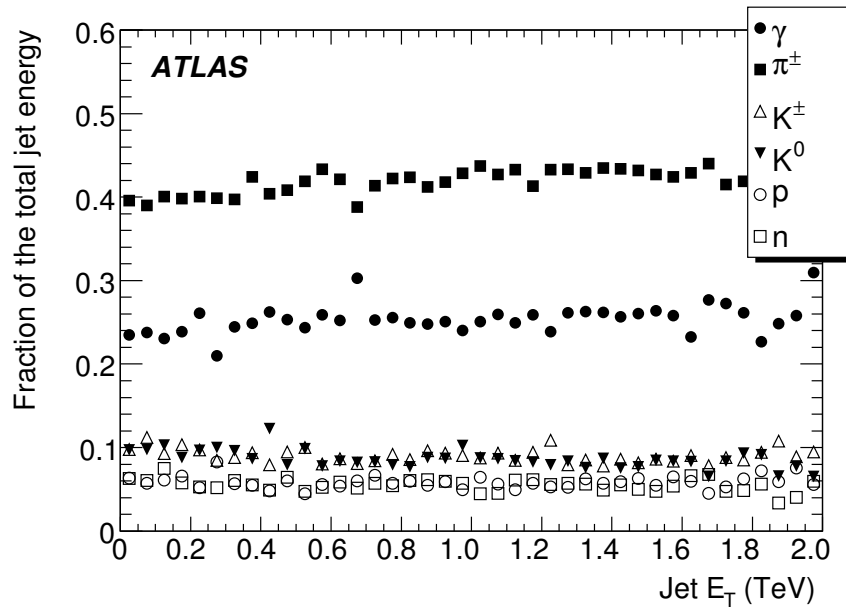


Correct for detector effects to get from calorimeter jet to particle jet:

- Non compensation of calorimeters (response to hadrons \neq response to electrons)
- Effects of cracks, dead material, losses in material in front of calorimeters
- Longitudinal leakage
- Magnetic field effect

This step relies on excellent MC simulation of calorimeter response validated on test beam

Jets: Composition and Energy deposit LAR/Tile



Jet energy carried by particle types

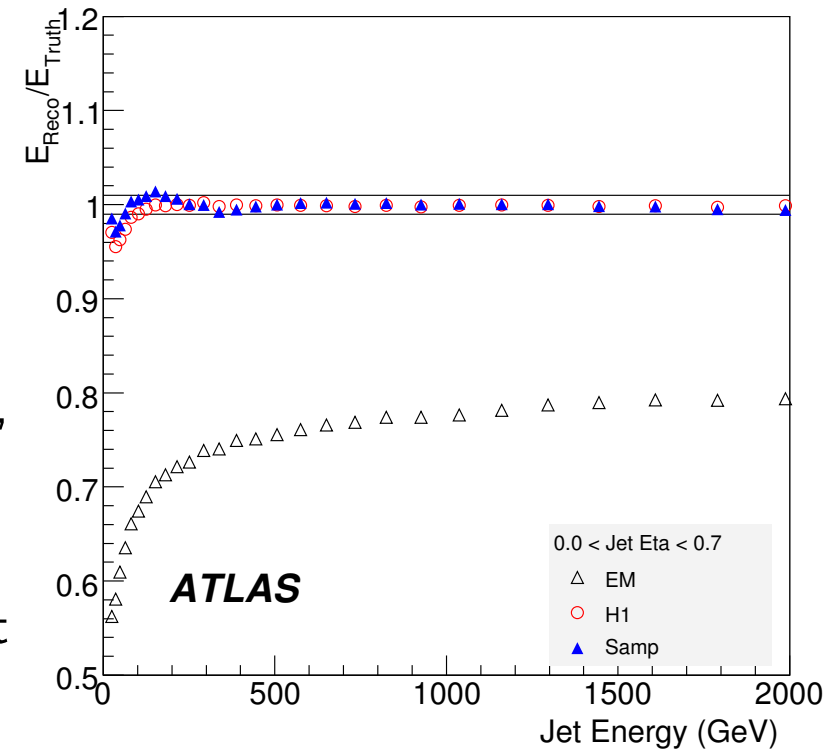
e.g. 40% π^\pm , 25% γ

60-80% of true jet energy in LArg EM

calorimeter, for $|\eta| < 0.7$ central region

Calorimeter Calibration

- Lar and TileCal non-compensating ($e/h \neq 1$)
- Must apply a calibration scheme
- Example: cell weighting
- Electromagnetic showers: high energy density, no weighting
- Hadronic showers: lower energy density, weight to scale up energy

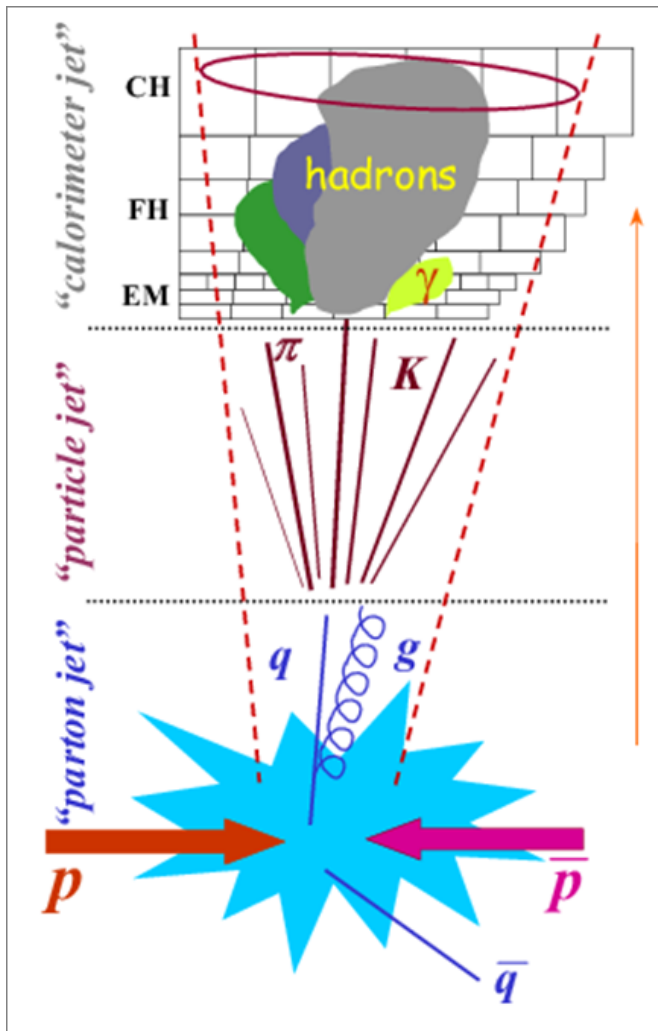


$$E_{jet} = \sum_{cells} E_i \times W_i(\rho_i)$$

$$\rho_i = E_i/V_i$$

Jet response before and after cell weighting

Jet reconstruction phase 3: go to partonic energy



Theoretically not well defined, but need this step e.g. when one wants to measure the mass of a heavy resonance decaying to jet-jet

Correct for:

- Energy loss outside of jet clustering
- Added energy from underlying event

This step will be performed with the data, using processes where a jet recoils against a well measured object: W, Z, γ , or $W \rightarrow jj$ in top decays. Goal is 1% precision on jet energy scale

The \cancel{E}_T variable

Total energy and momentum in final and initial state equal: $(E_f, \vec{p}_f) = (E_i, \vec{p}_i)$.

Initial state in collider: $E_i = \sqrt{s}$, $\vec{p}_i = 0$

For e^+e^- collider all final state particles in detector acceptance $\vec{p}_f = 0$, if non-interacting particle with momentum \vec{p}^ν produced $(\nu, \tilde{\chi}_1^0)$, then $\vec{p}_f \neq 0$, and

$$\vec{p}^\nu = \vec{p}^{miss} = -\sum_j \vec{p}_j = -\vec{p}_f$$

With j running on all the detected particles

For hadron collider: particles from spectator quarks undetected in the beam pipe, $\vec{p}_f \neq 0$

$\sum_k (\vec{p}_T)_k \sim 0$ for particles k outside acceptance $\Rightarrow \sum_j (\vec{p}_T)_j = 0$ for particles j in acceptance.

If non-interacting particle with momentum \vec{p}_T^ν produced:

$$\vec{p}_T^\nu = \vec{p}_T^{miss} = -\sum_j (\vec{p}_T)_j$$

Approximate \vec{p}_T^{miss} by \cancel{E}_T , vector sum of energy deposition in calorimeter cells:

$$E_x^{miss} = \sum_j E_j \sin \theta_j \cos \phi_j \quad E_y^{miss} = \sum_j E_j \sin \theta_j \sin \phi_j$$

j runs on cells with energy deposition and $\phi_j(\theta_j)$ respectively azimuthal and polar angle of cell j

Experimental measurement of E_T

Based on assumption that all the energy is measured in the calorimeters or seen as muons in the spectrometers

Multi-step procedure correcting for experimental effects:

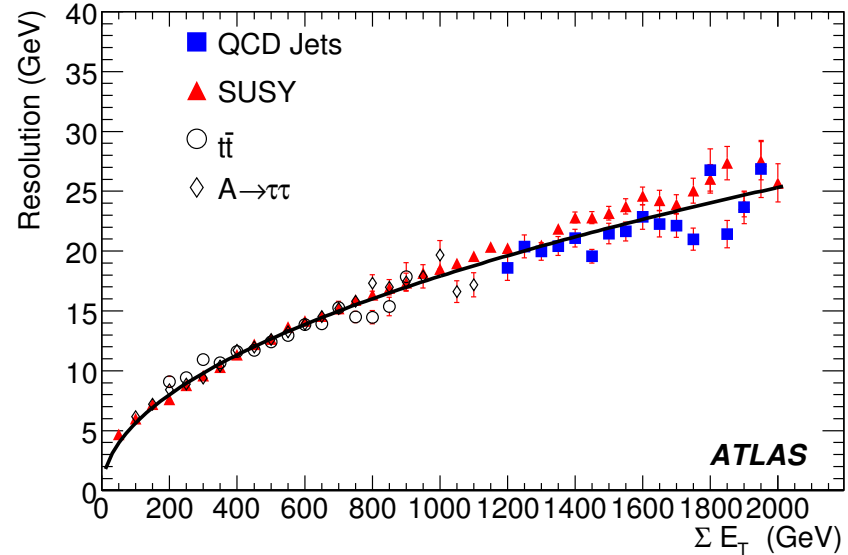
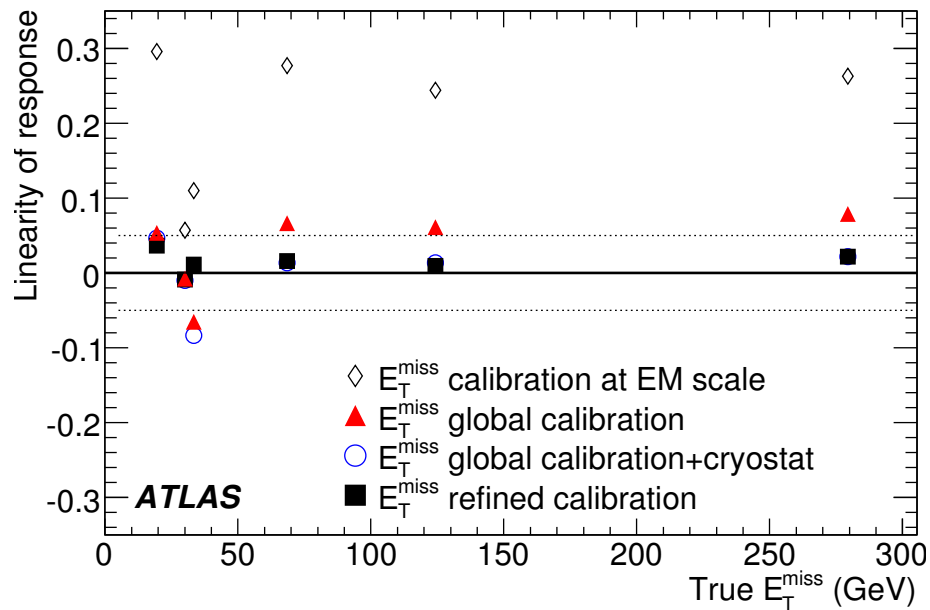
- Consider all cells calibrated at EM scale surviving an electronic noise suppression procedure
- Calibrate using global calibration weights based on energy density in cells, and perform vector sum
- Apply corrections for energy lost in cryostats
- Apply correction for detected muons
- Perform final refinement recalibrating cells depending on the reconstructed object they are assigned to

ATLAS procedure, very similar procedure for CMS. Additional step in CMS is improvement of resolution by using information from tracks

\cancel{E}_T performance

Measurement resolution estimated on MC by plotting the difference between true and estimated \cancel{E}_T separately on each of the components

Resolution can be fitted as $0.57 \cdot \sqrt{\Sigma E_T}$



.

Backup