Modern experimental methods and analysis at LHC

Pawel Brückman de Renstrom
Institute of Nuclear Physics P.A.N.

LHCPhenoNet Summer School,
Cracow , 7-12 September 2013

(Shamelessly using material from many other people presentations)

Preamble

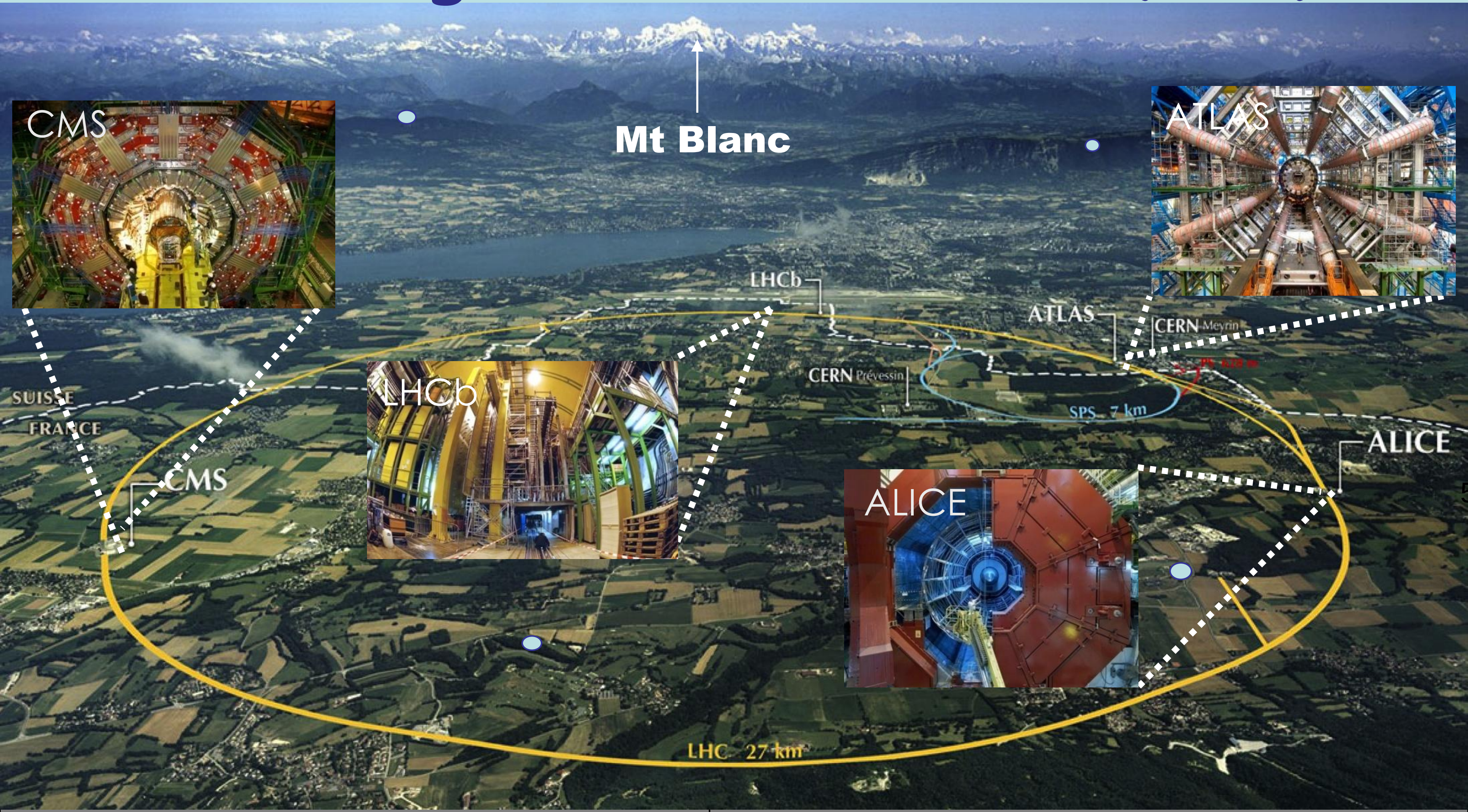
- ❖ The goal is to give you a crash course of experimental aspects of physics at the LHC.
- ❖ Unfortunately, it has to be superficial and selective.
- ❖ I take the liberty to use majority of examples from ATLAS ☺
- ❖ I will NOT speak about neither generators nor simulation.
- ❖ I will NOT talk about any particular analyses nor physics results unless for example purpose.

Outline of the lecture

- LHC: the high energy pp collisions
- The experimental setup: Spectrometers
- Luminosity determination
- Smart data taking & handling: Trigger etc.
- Reconstructing physics objects
- High-end analysis:
  - Kinematic variable reconstruction
  - Background estimation
  - Classification
  - Statistical interpretation

High energy pp collisions

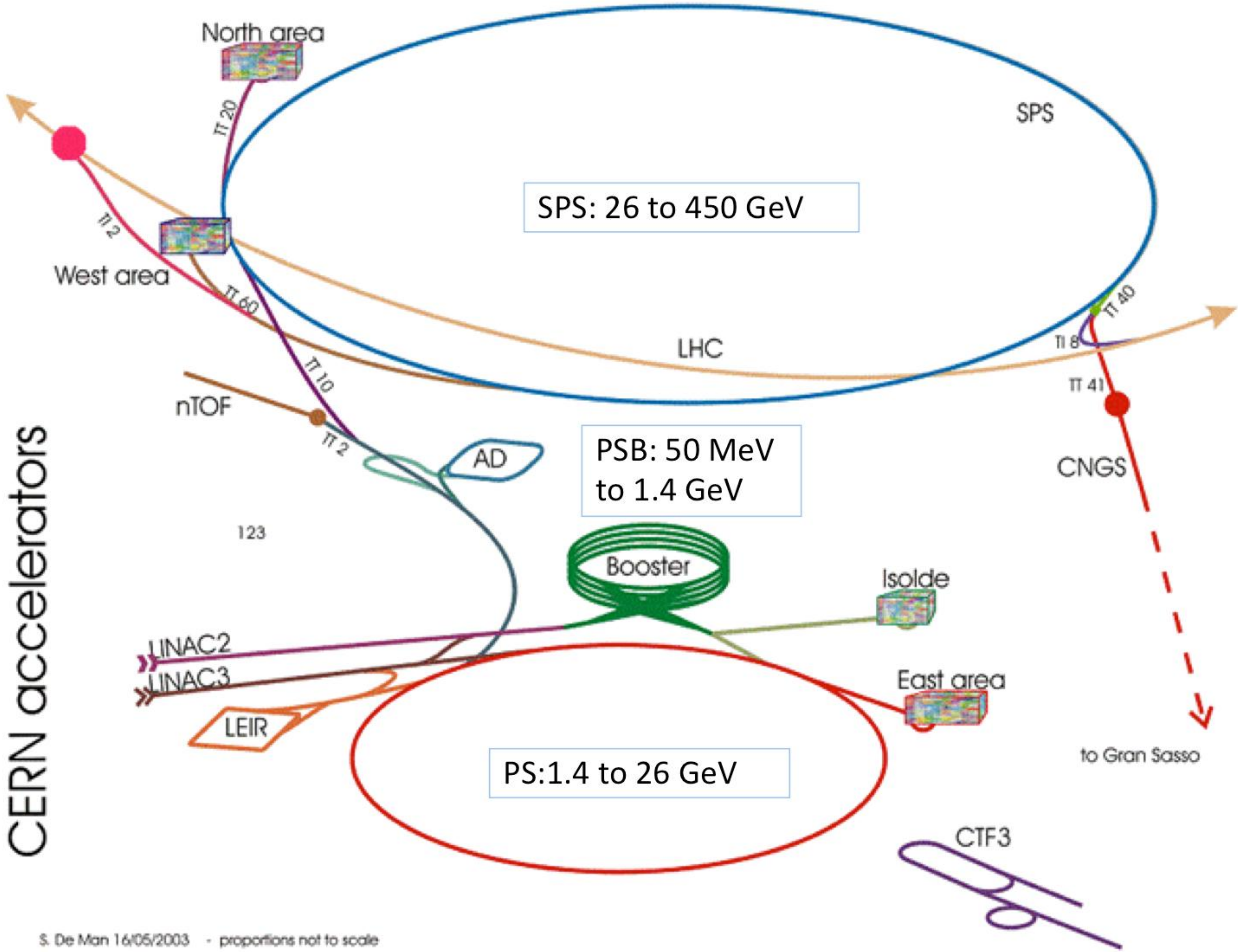
The Large Hadron Collider (LHC)



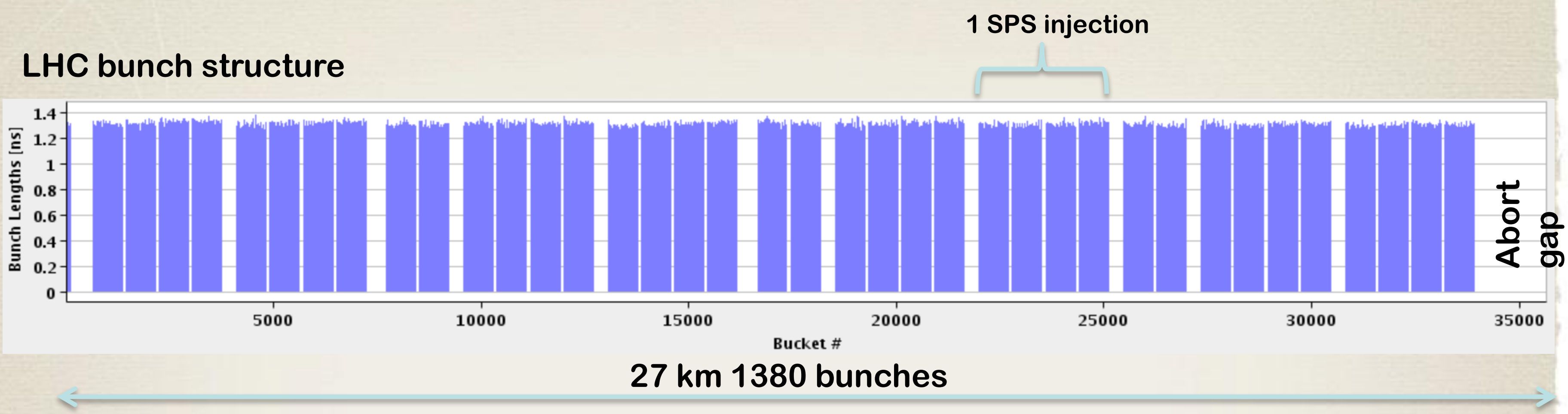
1720 Power converters
> 9000 magnetic elements
7568 Quench detection systems
1088 Beam position monitors
~4000 Beam loss monitors

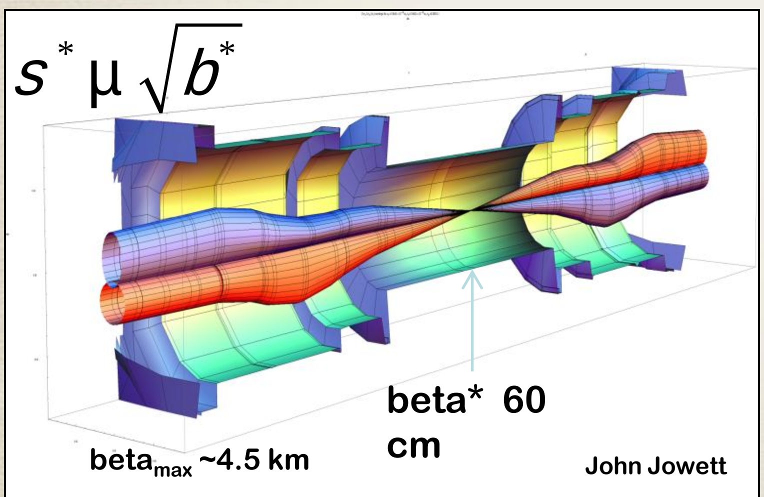
150 tonnes Helium, ~90 tonnes at 1.9 K
140 MJ stored beam energy in 2012
450 MJ magnetic energy per sector at 4 TeV

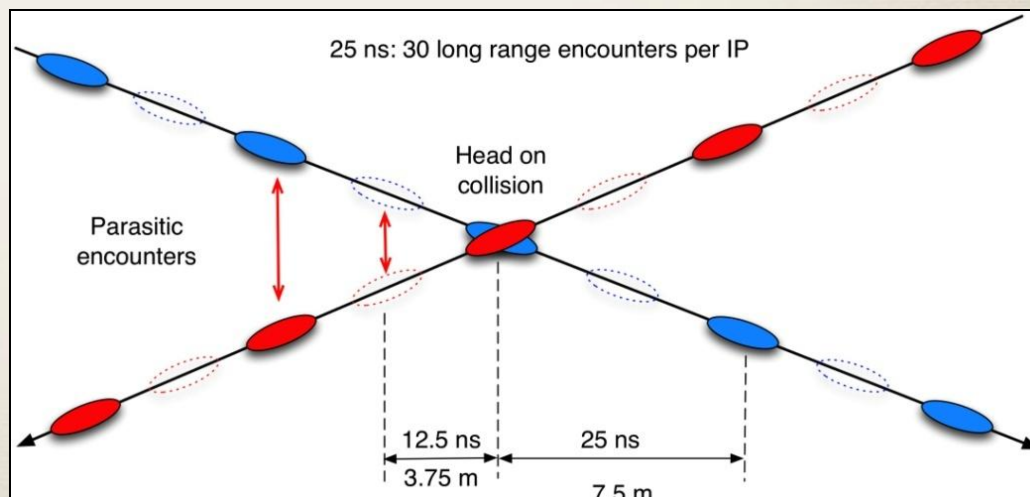
M. Lamont


CERN accelerators
North area
SPS
SPS: 26 to 450 GeV
TT 20
TI 2
West area
TT 60
TT 40
TI 8
LHC
TT 10
TT 41
nTOF
TT 2
AD
PSB: 50 MeV
to 1.4 GeV
CNGS
123
Booster
Isolde
LINAC2
LINAC3
East area
LEIR
to Gran Sasso
PS:1.4 to 26 GeV
CTF3
S. De Man 16/05/2003 - proportions not to scale

LHC bunch structure





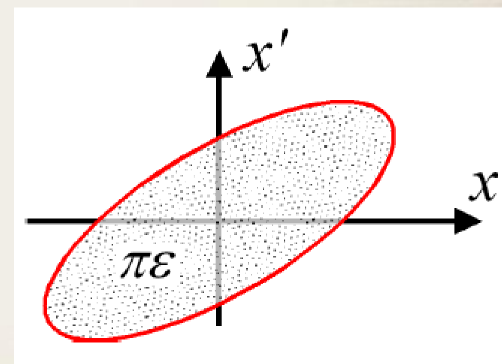


Luminosity

$$L = \frac{N^2 k_b f}{4\pi \sigma^*_x \sigma^*_y} F = \frac{N^2 k_b f \gamma}{4\pi \varepsilon_n \beta^*} F$$

N	**Number of particles per bunch**
k_b	**Number of bunches**
f	Revolution frequency
σ^*	**Beam size at interaction point**
F	Reduction factor due to crossing angle
ε	Emittance
ε_n	**Normalized emittance**
β^*	**Beta function at IP**

Round beams, beam 1 = beam 2



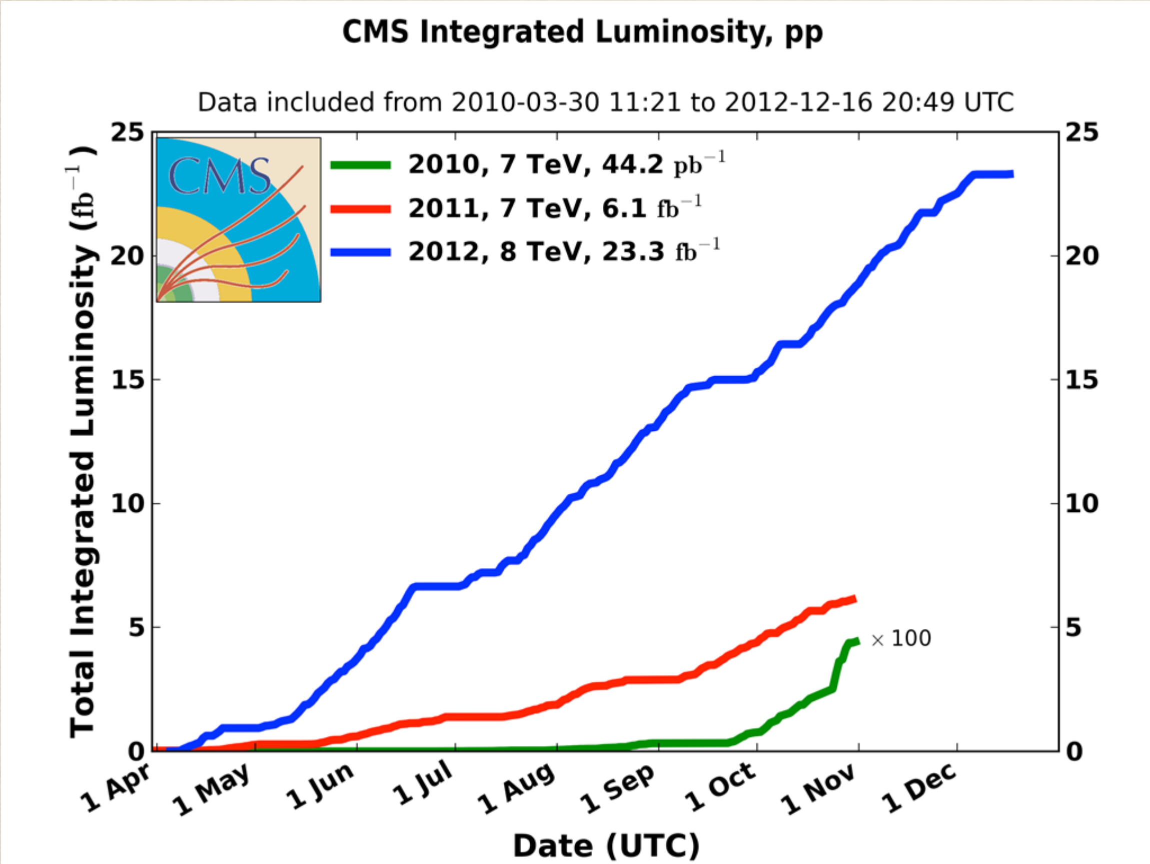
$$\varepsilon_n = \beta\gamma\varepsilon$$

$$\sigma^* = \sqrt{\beta^* \varepsilon}$$

Peak performance through the years

	2010	2011	2012	Nominal
Energy [TeV]	3.5	3.5	4	7
Bunch spacing [ns]	150	50	50	25
No. of bunches	368	1380	1380	2808
beta* [m] ATLAS and CMS	3.5	1.0	0.6	0.55
Max **bunch intensity** [protons/bunch]	1.2×10^{11}	1.45×10^{11}	1.7×10^{11}	1.15×10^{11}
Normalized **emittance** [mm.mrad]	~2.0	~2.4	~2.5	3.75
Peak luminosity [$cm^{-2}s^{-1}$]	2.1×10^{32}	3.7×10^{33}	7.7×10^{33}	1.0×10^{34}

Integrated luminosity 2010-2012



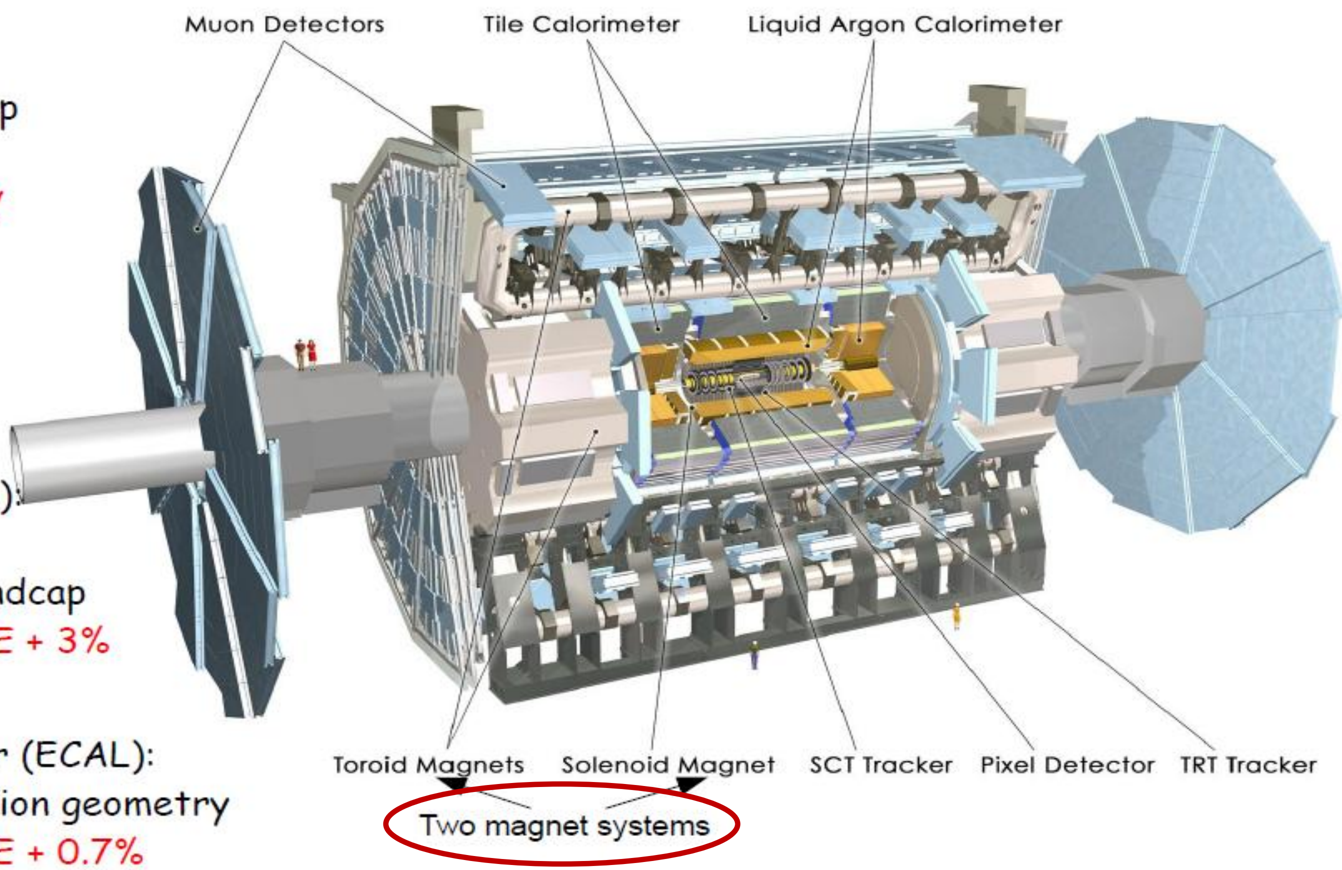
- 2010: **0.04 fb^{-1}**
  - 7 TeV CoM
  - Commissioning
- 2011: **6.1 fb^{-1}**
  - 7 TeV CoM
  - Exploring the limits
- 2012: **23.3 fb^{-1}**
  - 8 TeV CoM
  - Production

	2010	2011	2012
Max. luminosity delivered in 7 days [fb^{-1}]	0.025	0.58	1.35

Spectrometer

ATLAS detector

Muon spectrometer:
- air-core toroid magnets: 0.5 T in barrel, 1 T in endcap
- momentum resolution: 2% @ 50 GeV, 10% @ 1 TeV (combined Tracker+Muon spectrometer)



Hadronic calorimeter (HCAL):
- Fe+scint in barrel, Cu+Liquid Argon (LAr) in endcap
- resolution $\sigma(E)/E \approx 50\%/\sqrt{E} + 3\%$ (ECAL+HCAL, barrel part)

Electromagnetic calorimeter (ECAL):
- Pb+LAr technology, accordion geometry
- resolution $\sigma(E)/E \approx 10\%/\sqrt{E} + 0.7\%$

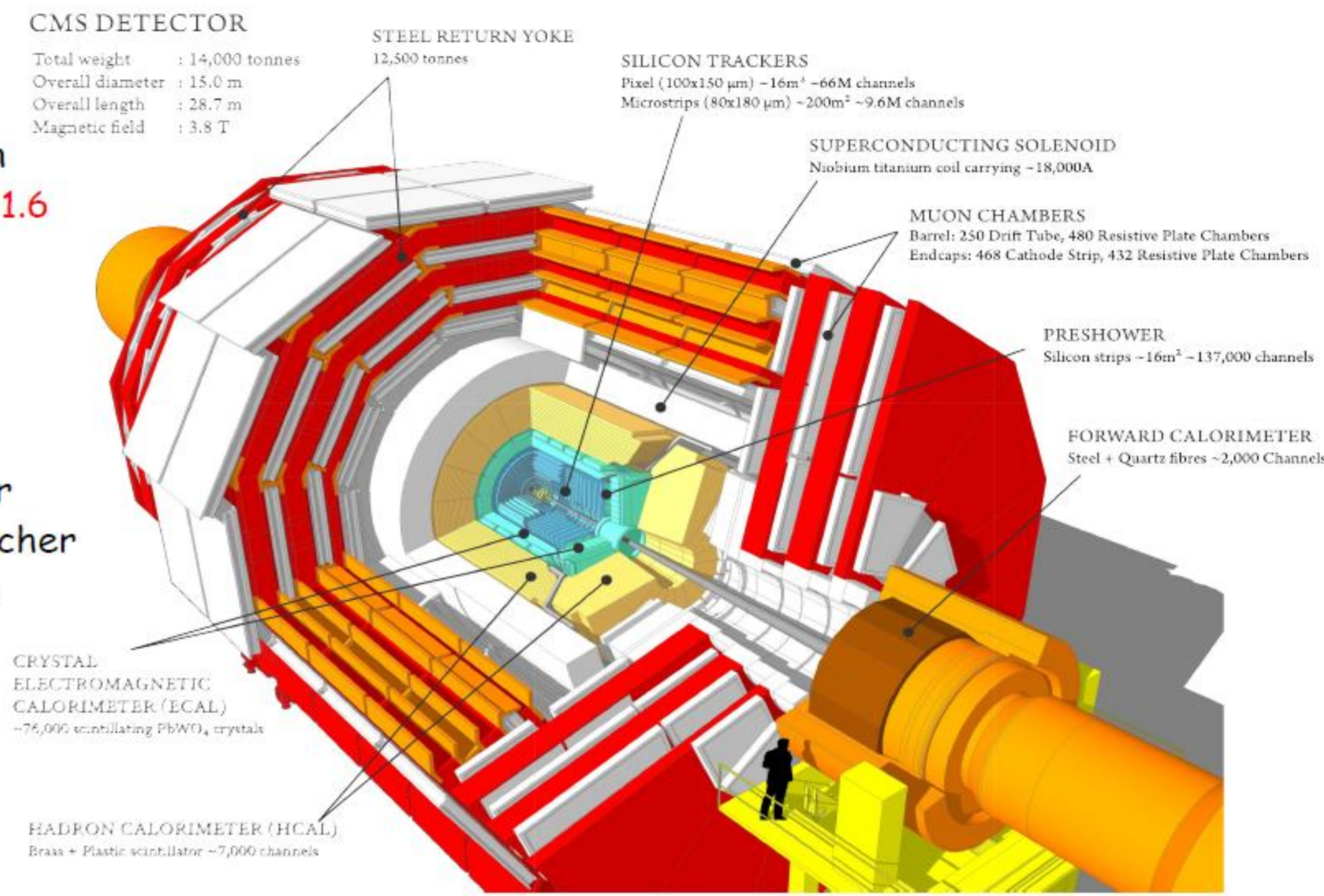
Tracker:
- Si pixels, Si strips, Transition Radiation Tracker (TRT) inside 2 T solenoid
- resolution: $\sigma(p_T^{-1}) \approx 0.36 + 13/(p_T \cdot \sqrt{\sin\theta})$ [TeV^{-1}], (θ being the polar angle wrt beam axis)

Further details in Ref:
G. Aad et al., JINST 3 (2008) S08003

CMS detector

Tracking system:

- silicon pixels and strips
- expected muon resolution
  $\sigma(p_T)/p_T$ = 1.5-2% for $|\eta|<1.6$
  at p_T=100 GeV



HCAL:

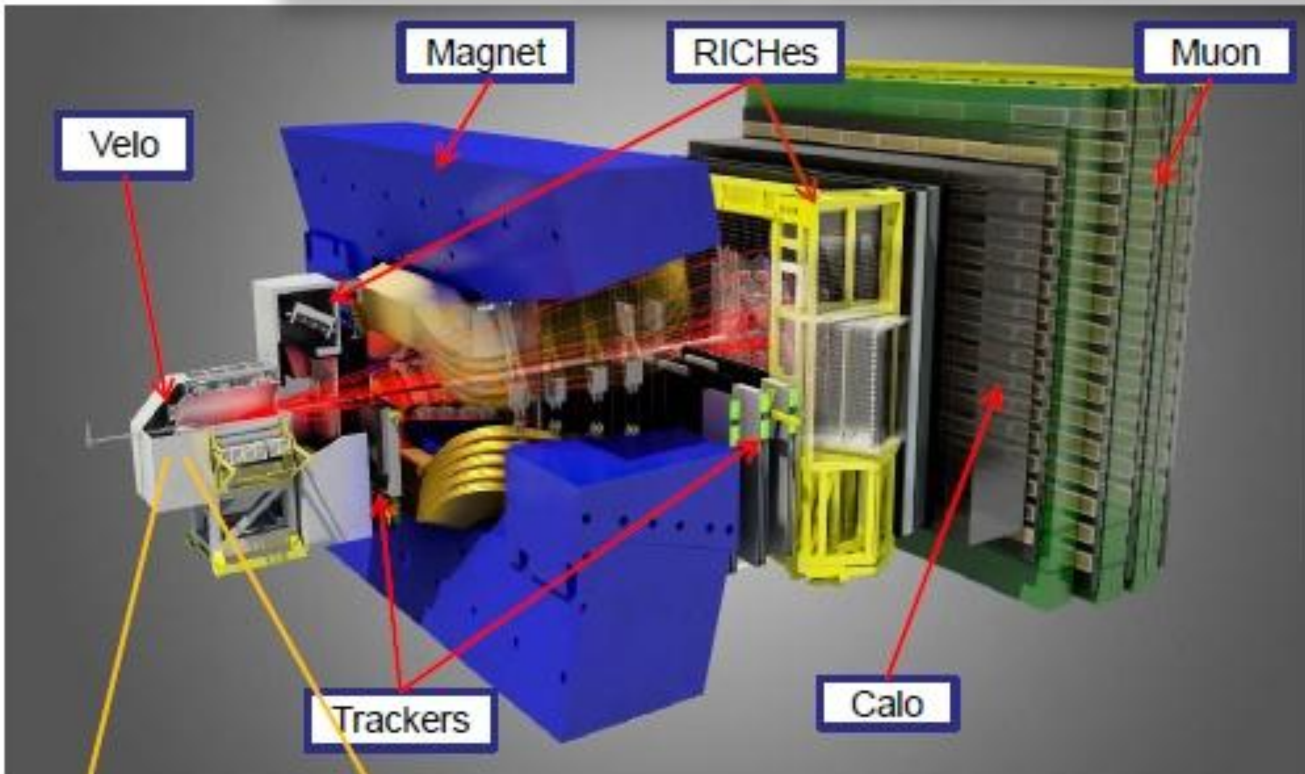
- brass + plastic scintillator
- complemented by tail catcher outside the solenoid in the central part

ECAL:

- $PbWO_4$ crystals read by APD and VPT
- testbeam resolution $\sigma(E)/E \approx 2.8\%/\sqrt{E} + 0.3\% + 0.12/E$

Details in Ref:
S. Chatrchyan et al., JINST 3 (2008) S08004

LHCb detector

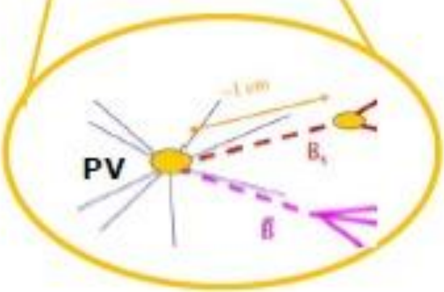


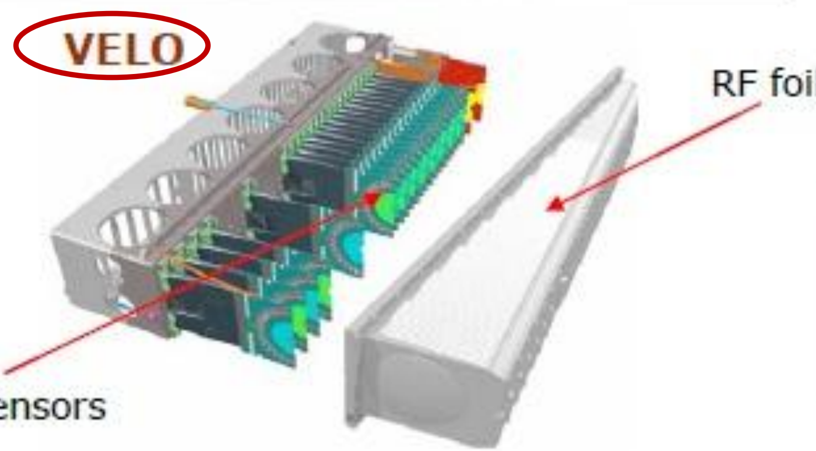
Unique acceptance
(10-300) mrad

able to access low p_T

test models with enhanced forward production

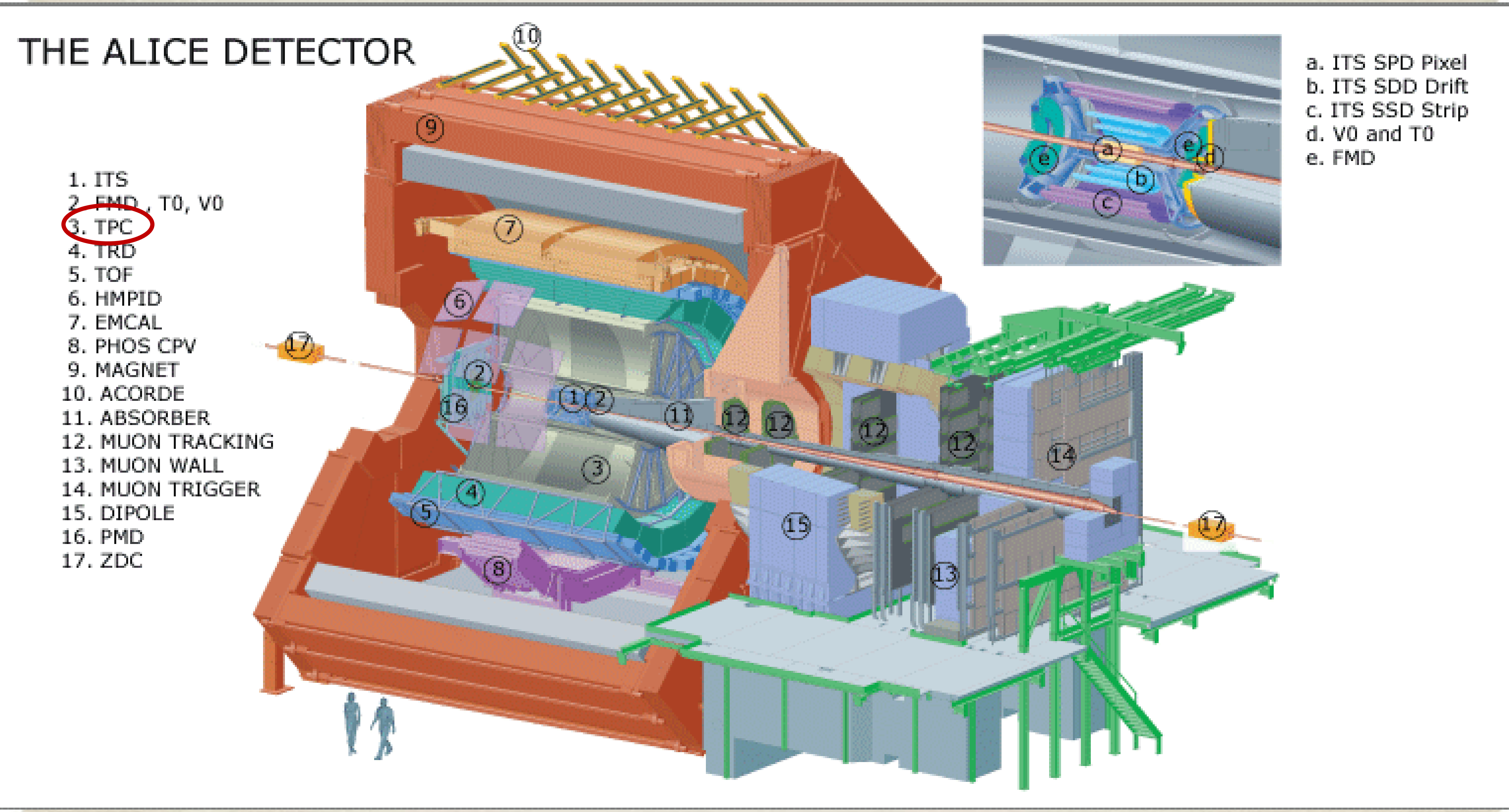
precise tracking / vertexing in $\eta \in (2\text{-}5)$





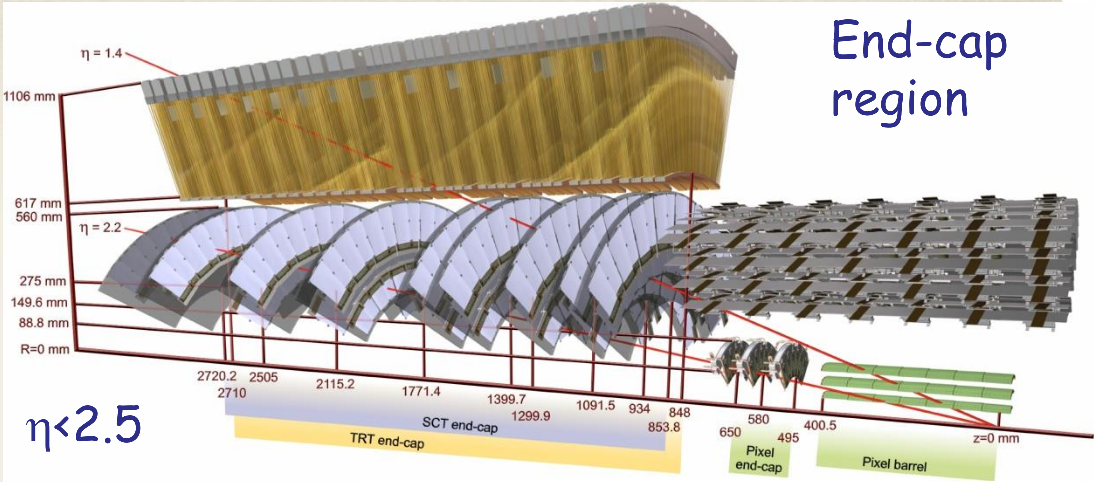
$\int L$:	~**37** pb^{-1}	*(2010)*
	~**1** fb^{-1}	*(2011)*
	~**2** fb^{-1}	*(2012)*

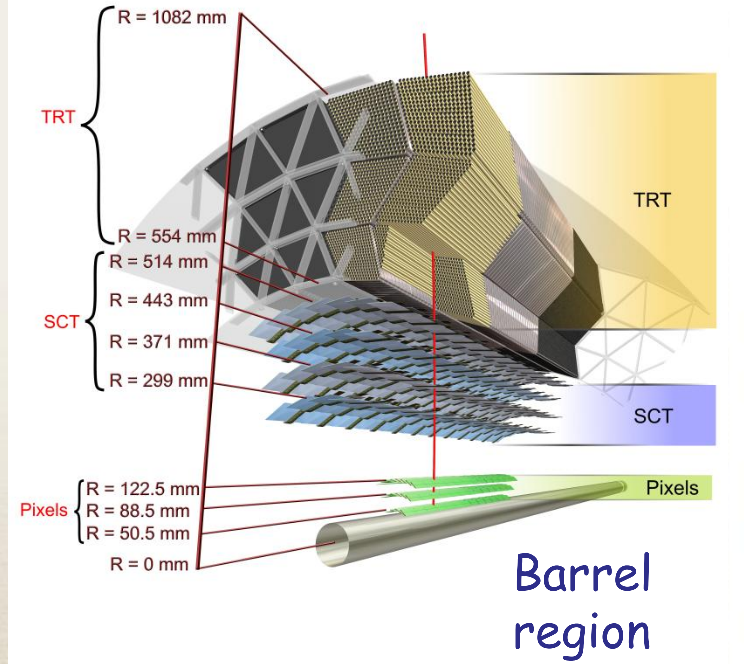
ALICE detector



ATLAS tracking

- **Inner Detector provides:**
  - Precision measurement of ch. particle trajectories
  - particle identification
  - vertex reconstruction
  - b-tagging in jets
- **Inside a 2T solenoid magnetic field are:**
  - Pixel detector:
    - 80M pixels in 3 barrels (R=5,9,12 cm) and 3 disks in each end-cap
    - 10 μm x 115 μm (RΦ x Z resolution)
  - SemiConductor Tracker (SCT):
    - 6.3M strips in 4 barrel layers and 9 disks/end-cap
    - 17 μm x 580 μm (RΦ x Z resolution)
  - Transition Radiation Tracker (TRT):
    - 350k straws with 2 mm radius
    - 130 μm (RΦ resolution)





ATLAS calorimeters

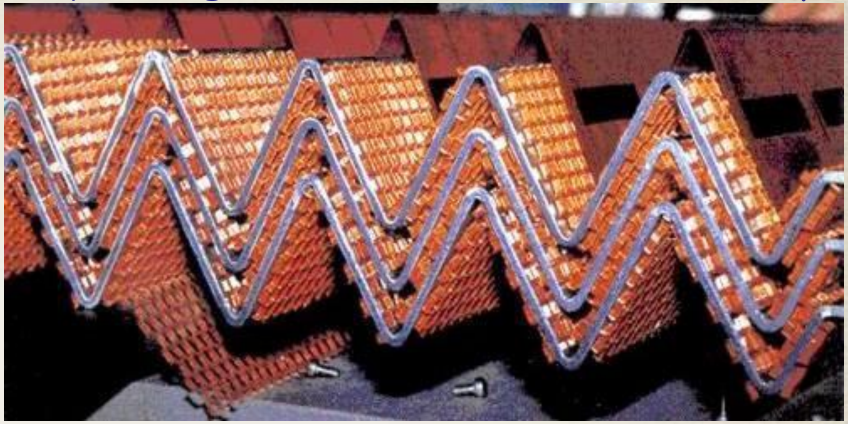
- **Electromagnetic calorimeter ($|\eta|<3.2$)**

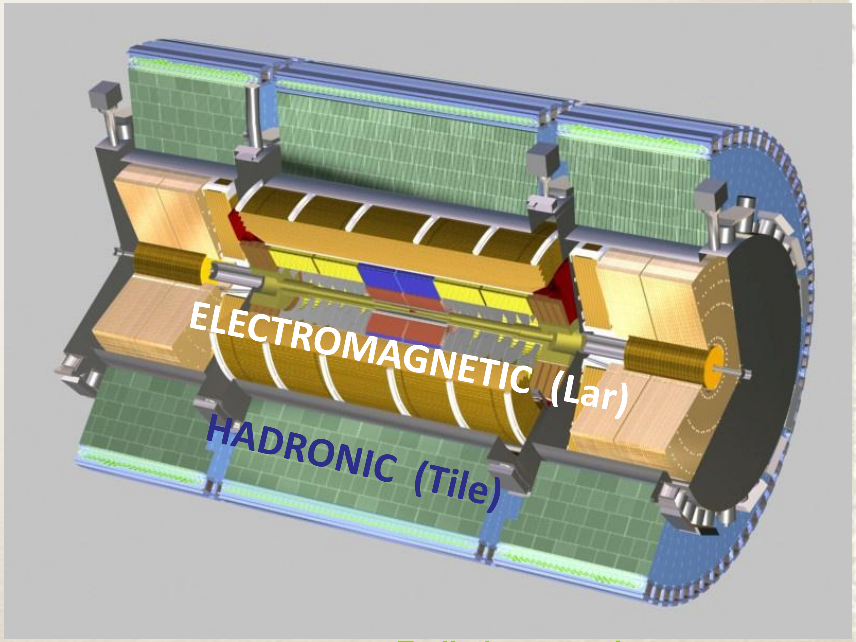
– Liquid Argon- Pb sampling calorimeter with accordion geometry
– ~180k channels
– 3 longitudinal layers + pre-sampler ($|\eta|<1.8$)
– $\sigma/E \sim 10\%/\sqrt{E} \oplus 0.7\%$
– Provides e/γ trigger, energy measurement and particle identification

- **Hadron Calorimeter:**

– Used in trigger; provides jet energy, position & E_T^{miss} measurements, helps identify muons
– Sub-divided in three regions:
• $|\eta| < 1.7$: Fe/scintillating tiles
– 3 or 4 longitudinal layers, ~20k channels
• $3.2 < |\eta| < 1.5$: Cu –LAr (HEC)
• $3.1 < |\eta| < 4.9$: FCAL Cu/W- LAr

Liquid Argon calorimeter accordion shape

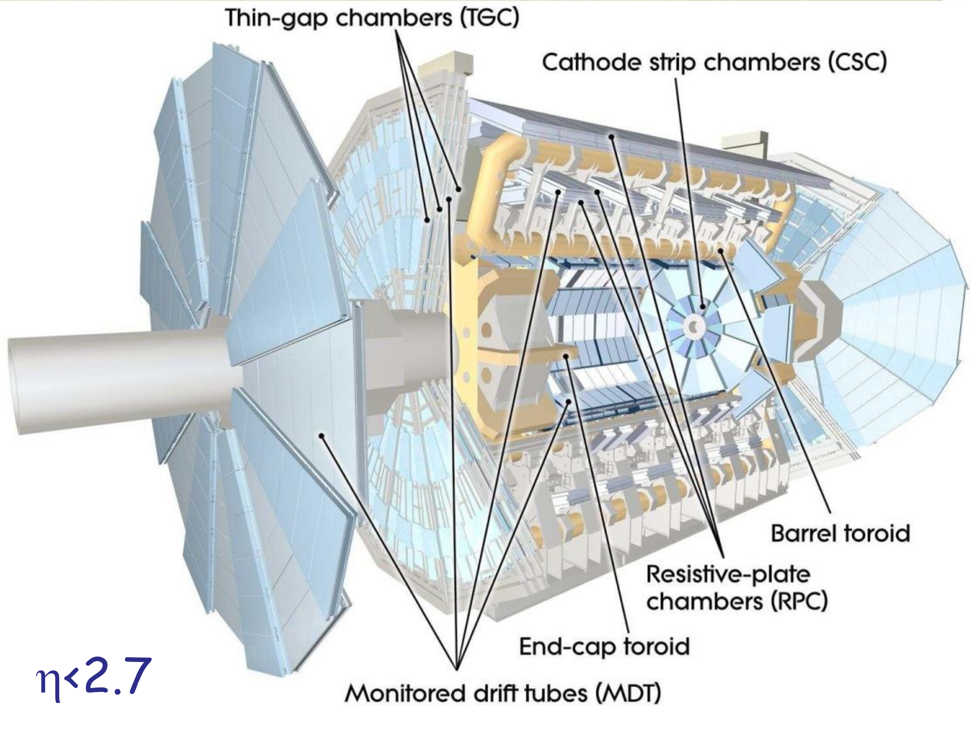




ATLAS Muon System

- **One barrel and two end-cap sections within a field of ~0.5 T provided by an air-core toroid**

- **Muon spectrometer sub-divided into different sub-detectors for precision measurement and trigger**
- Monitored Drift Tubes (**MDT**)
  - 3 layers for $|\eta|<2$ and 2 layers for $2<|\eta|<2.5$
  - hit resolution ~80μm per tube
- Cathode Strip Chambers (**CSC**)
  - 1 layer for $2<|\eta|<2.7$
  - 60μm X 5 mm resolution in $\eta \times \phi$
- Resistive Plate Chambers (**RPC**)
  - Trigger chambers: 3 layers for $|\eta|<1.05$
  - η, ϕ measurements with 1 cm resolution
- Thin Gap Chambers (**TGC**):
  - Trigger chambers: 3 layers for $2.0<|\eta|<2.5$, (η, ϕ) with 1 cm resolution



Luminosity determination

Luminosity determination

$$\frac{dN}{dt} = \sigma \cdot L \longrightarrow L = \frac{R_{inel}}{\sigma_{inel}} \longrightarrow L = \frac{\mu_{inel} n_b f_r}{\sigma_{inel}}$$

- Instantaneous luminosity at the LHC cannot be measured directly from first principles, as e.g. at LEP using Bhabha scattering.
- Elastic X-section too large to resolve single events at LHC luminosity. Besides, special optics is needed in order to count such events confidently (ALFA project)
- Extrapolation of the total $\sigma_{inel,TOT}$ from Tevatron has large uncertainty.
- **Two steps approach**: normalization of the σ_{vis} and rate counting.

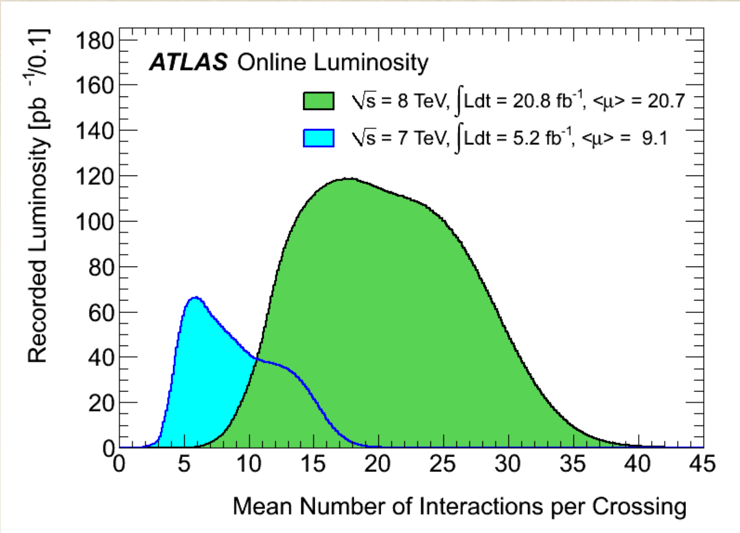
$$L = \frac{\mu_{vis} n_b f_r}{\sigma_{vis}} \quad \text{where} \quad \sigma_{vis} = \varepsilon \sigma_{inel} \quad \text{and} \quad \mu_{vis} = \varepsilon \mu_{inel}$$

Luminosity determination

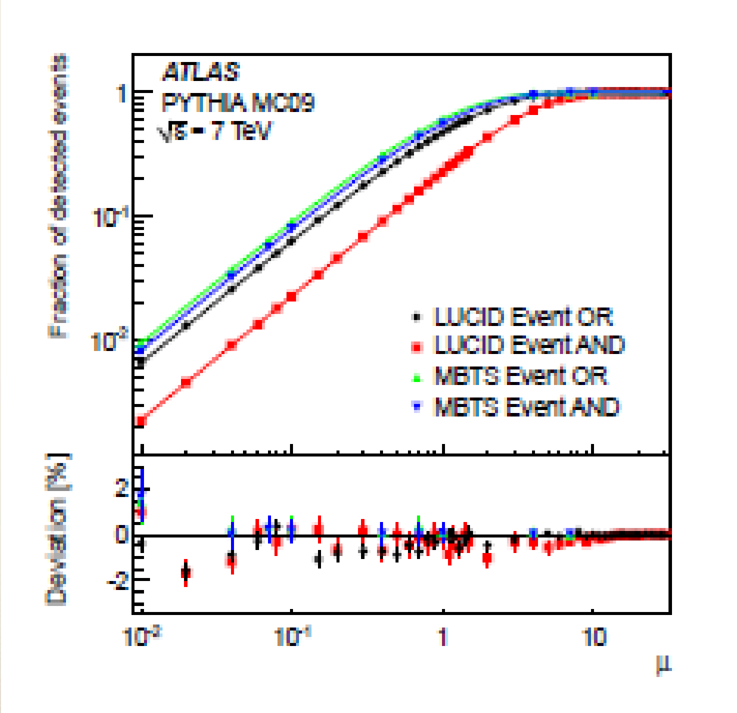
$$\mu_{vis} \approx \frac{N}{N_{BC}}$$

N ← Tagged vents

N_{BC} ← # of bunch crossings (BC)



- This method is adequate only if μ_{vis}<<1 (probability of having two interactions per BC is negligible).
- The threshold depends on the ε of the chosen observable.
- For higher μ_{vis} one needs observables proportional to the number of interactions.



Luminosity determination

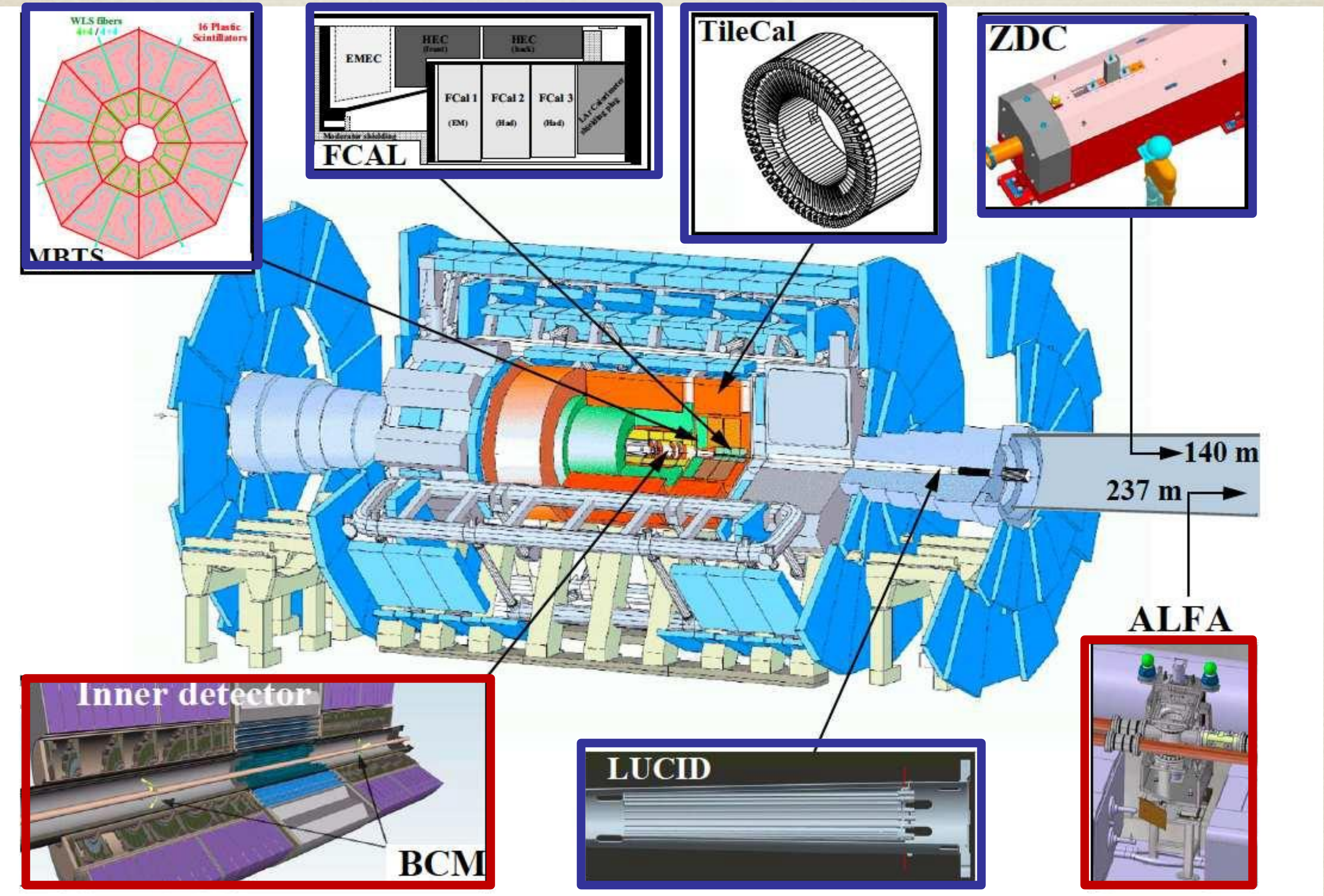
- Ansatz 1: μ follows the Poisson distribution
- Ansatz 2: number of (hits, tracks, etc.) follows Binomial distribution

$$P_{\mathrm{HIT}}(\mu_{\mathrm{Vis}}^{\mathrm{HIT}}) = \frac{N_{\mathrm{HIT}}}{N_{\mathrm{BC}} N_{\mathrm{CH}}} = 1 - e^{\mu_{\mathrm{Vis}}^{\mathrm{HIT}}}$$

- From where:

$$\mu_{\mathrm{Vis}}^{\mathrm{HIT}} = -\ln\left(1 - \frac{N_{\mathrm{HIT}}}{N_{\mathrm{BC}} N_{\mathrm{CH}}}\right)$$

ATLAS Luminosity



pp Luminosity - vdM scan

Simon van der Meer

$$L = n_b f_r n_1 n_2 \int \rho_1(x,y)\rho_2(x,y)dxdy$$

If no x-y correlation: ρ(x,y) = ρ(x) ρ(y) :

$$L = n_b f_r n_1 n_2 \Omega_x(\rho_1(x),\rho_2(x))\Omega_y(\rho_1(y),\rho_2(y))$$

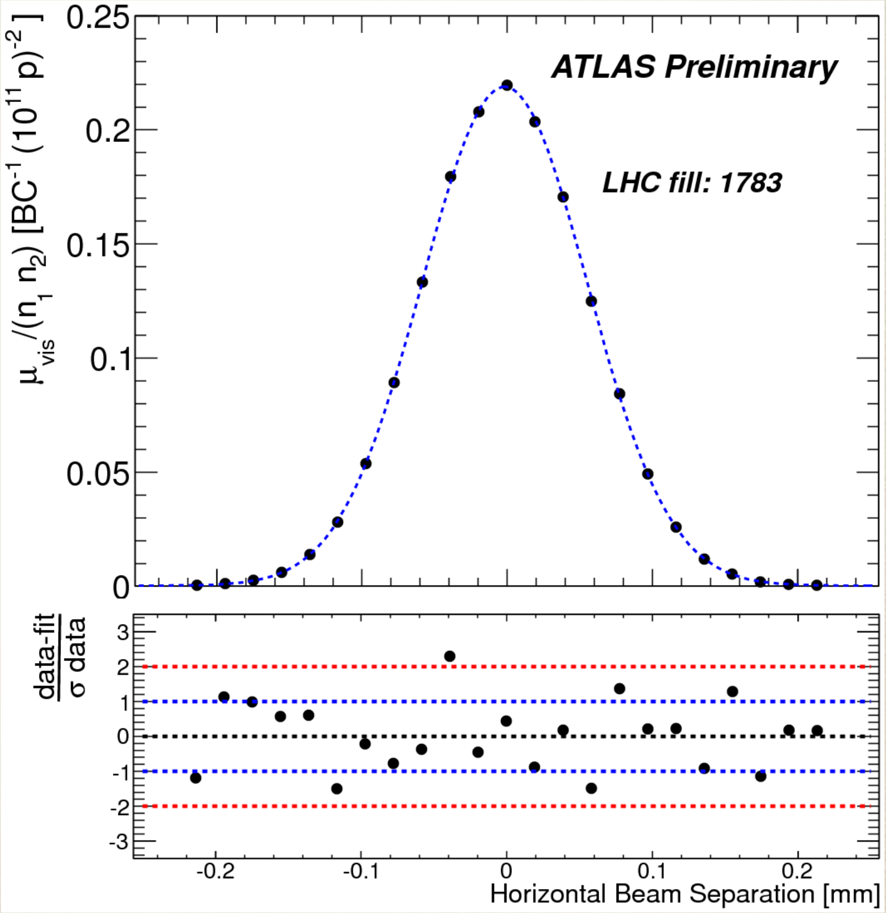
$$\Omega_x(\rho_1(x),\rho_2(x)) = \frac{R_x(0)}{\int R_x(\delta)d\delta}$$

$$\Sigma_x = \frac{1}{\sqrt{2\pi}} \frac{\int R_x(\delta)d\delta}{R_x(0)}$$

$$L = \frac{n_b f_r n_1 n_2}{2\pi\Sigma_x\Sigma_y}$$

bunch multipl.

bunch width

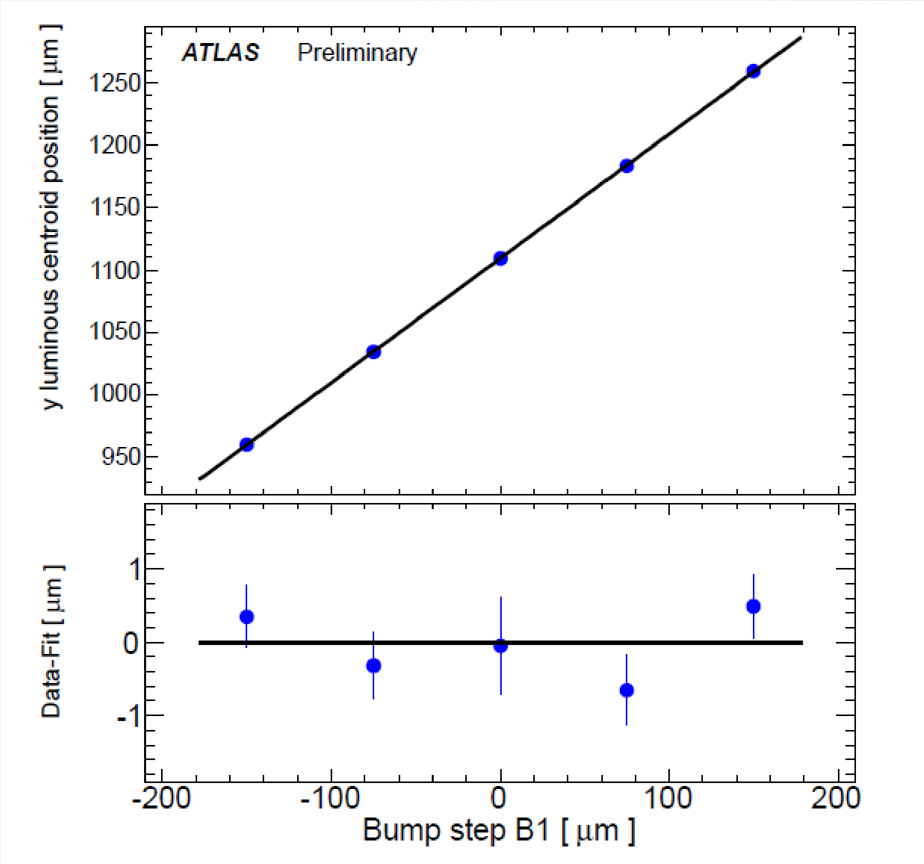


pp Luminosity – vdM scan

Simon van der Meer

- Length scale calibrated from the detector itself – beam spot reconstruction while moving meams same way.

- Ucertainty dominated by the $n_1 n_2$ determination (DCCT +FBCT)

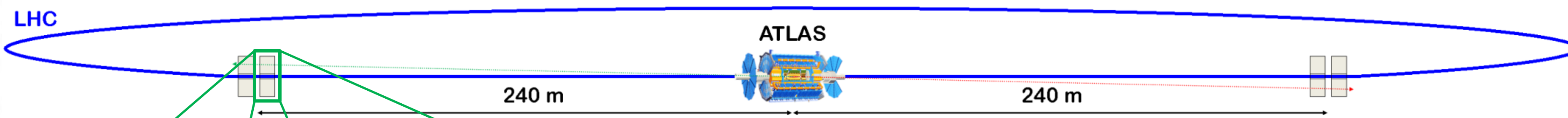


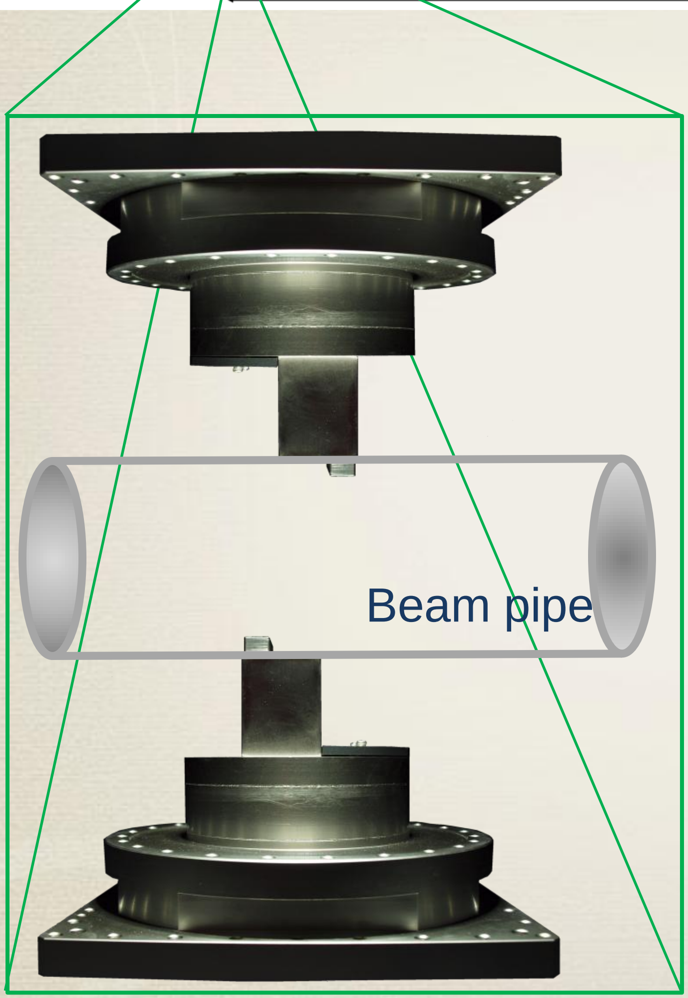
pp Luminosity – Bunch Currents

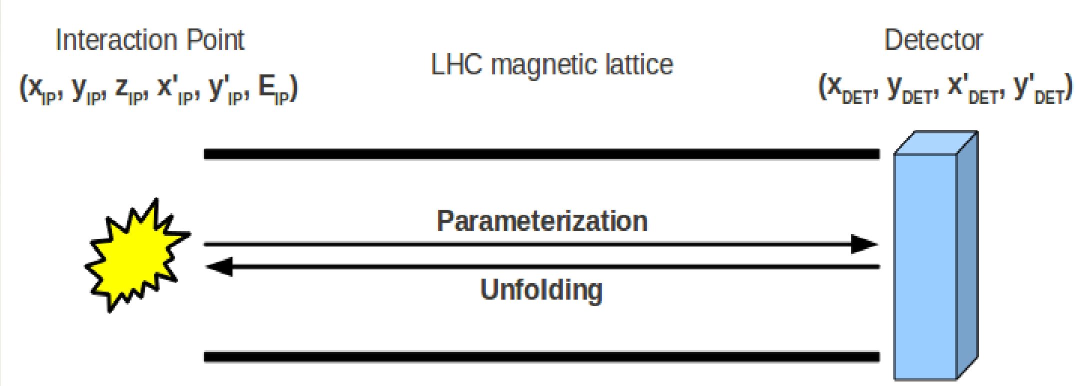
One input to the vdM scan are the bunch currents which are measured by three distinct systems:

- **DCCT** (Direct Current Current Transformer)
  - measures the total circulating bunch current
  - very high precision
- **FBCT** (Fast Beam Current Transformer)
  - measures the individual bunch population
  - only used for relative bunch charge fractions
  - normalized to the DCCT
- **BPTX** (Beam Pick-up Timing for Experiment)
  - an independent cross-check of the FBCT

ALFA - Absolute Luminosity For ATLAS





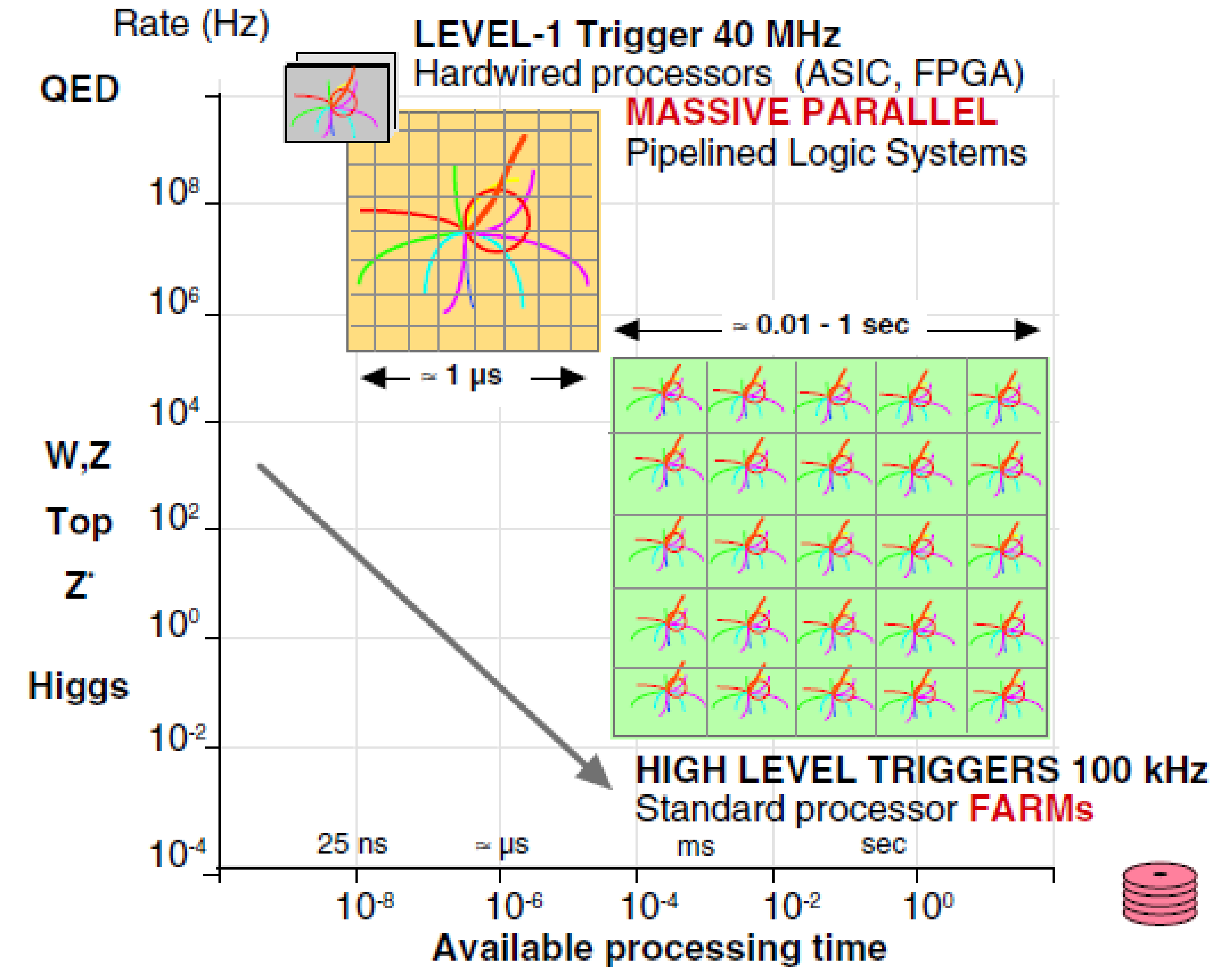


Measurement of the LHC absolute luminosity with pp elastic scattering at low intensity. Calibration of other ATLAS lumi detectors for relative lumi measurement detectors. Placed far away from the Interaction Point (240 m) inside the LHC beam pipe (Roman Pots) for precise measurement of small scattering angles enabling study of elastic and diffractive processes.

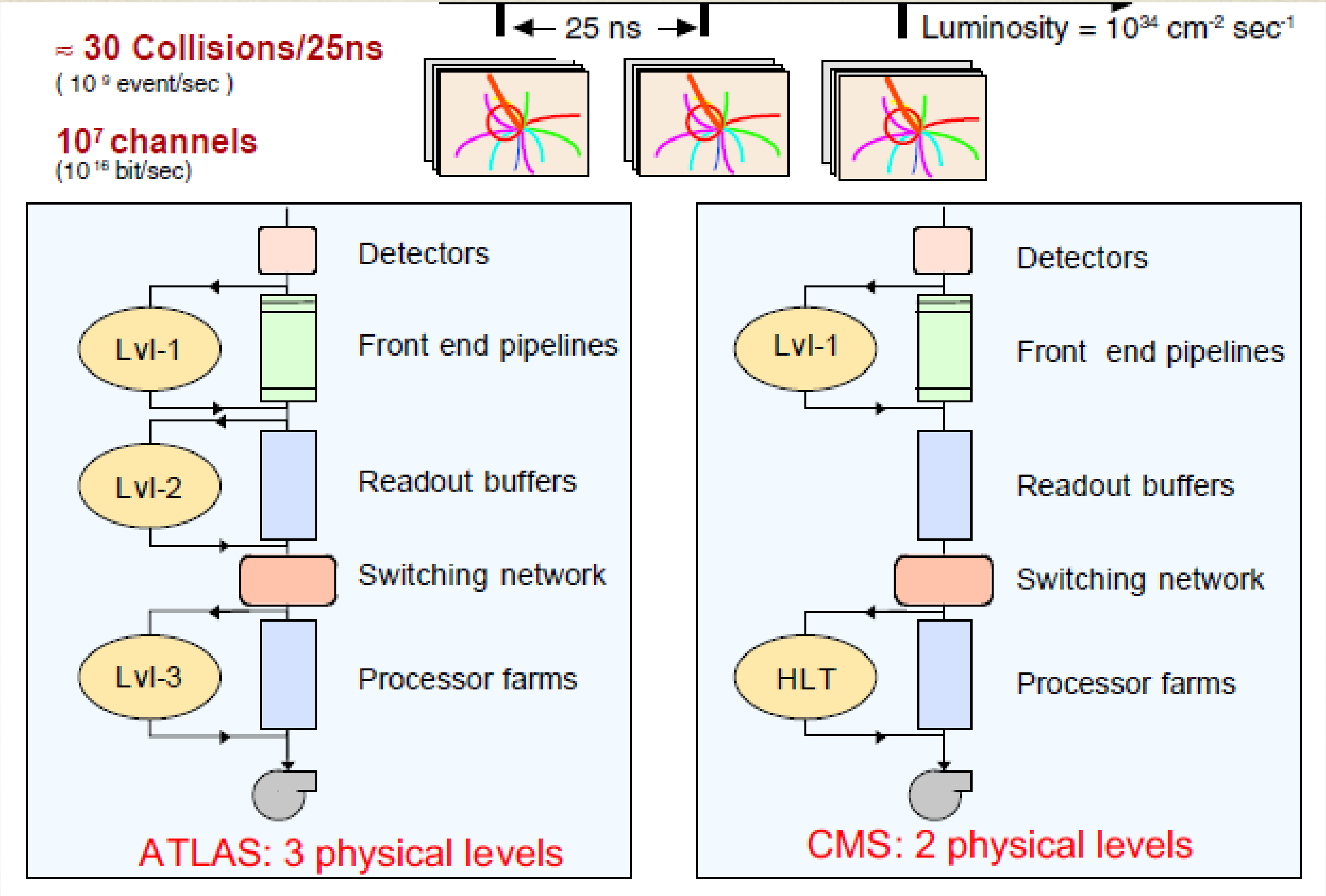
Smart data taking & handling

Trigger – the paradigm

Name of the game: TIMING



Trigger – implementations

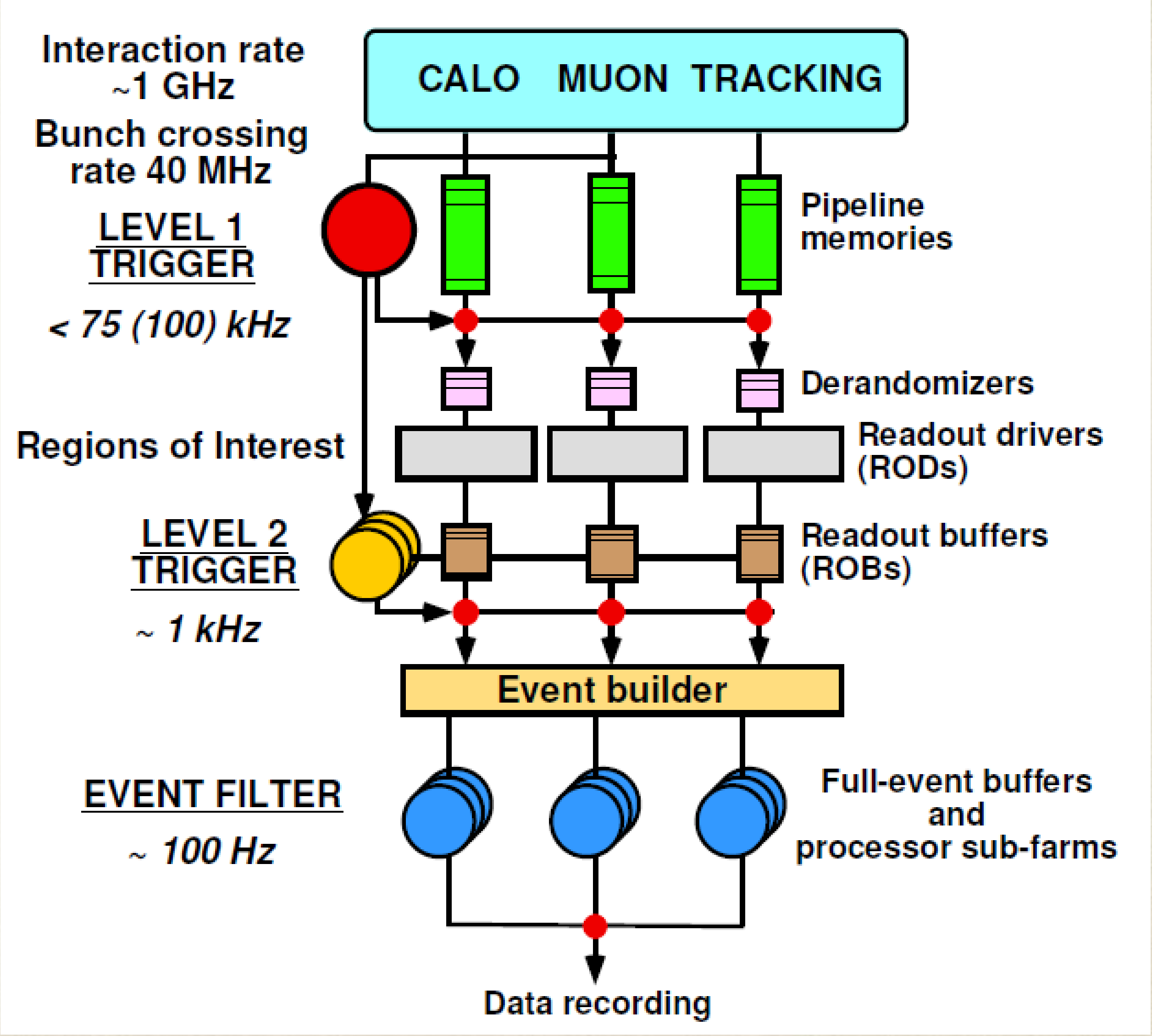


ATLAS Trigger – timing

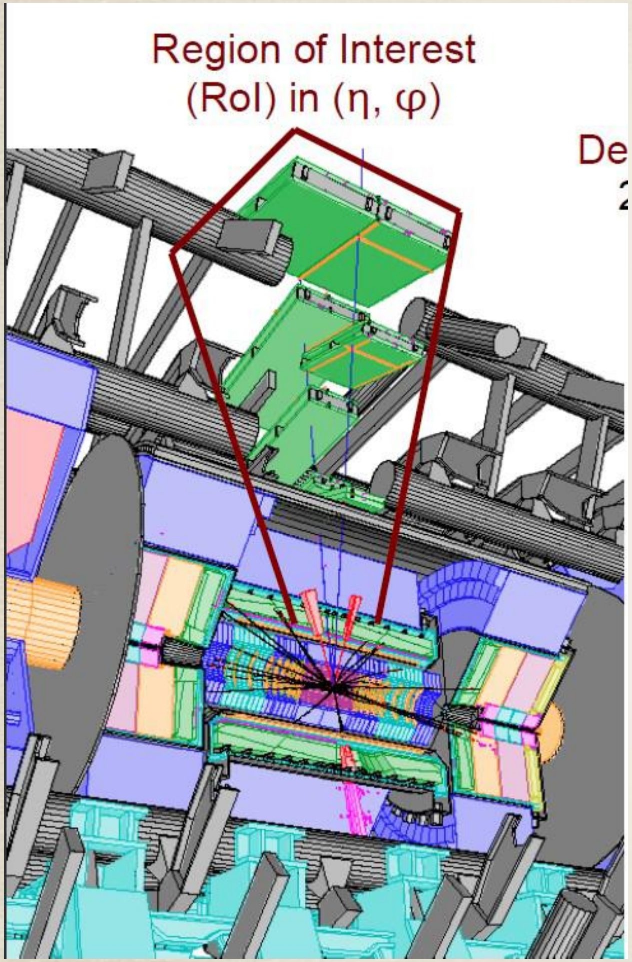
2μs latency

1-10 ms latency

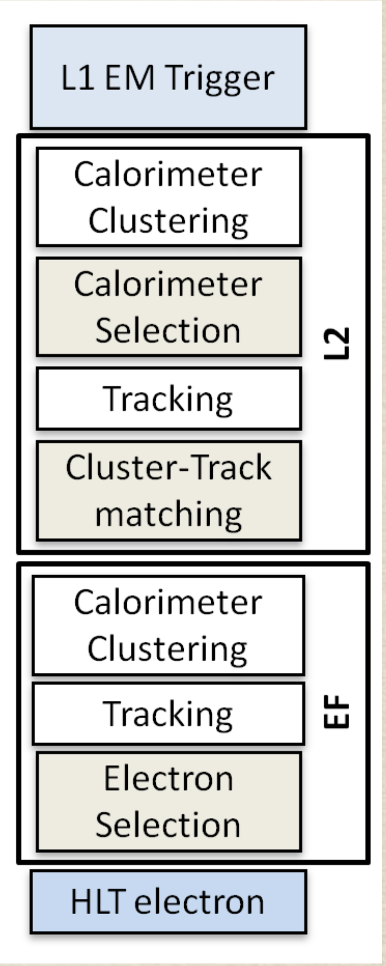
O(s) latency



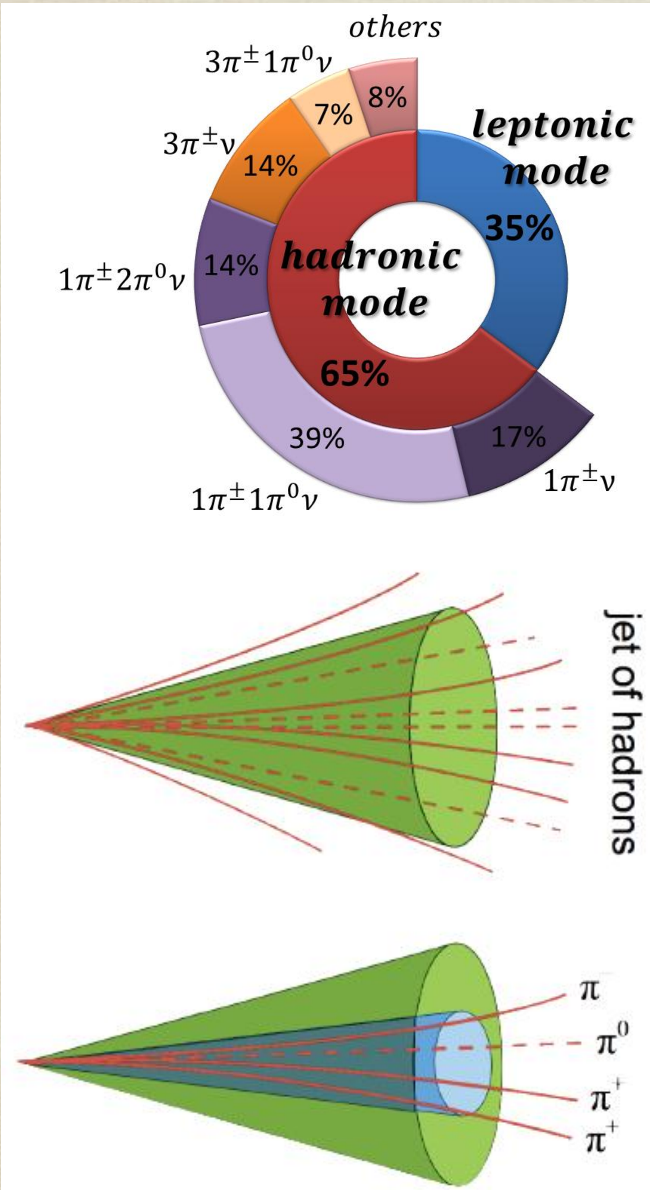
Trigger – details (ATLAS)



Level 1:
Fast, custom-build electronics finds and defines RoIs
Muon and Calorimeters only
Coarse resolution
Level 2:
Dedicated, fast software algorithms
Works on full-granularity RoI data
Level 3 (Event Filter):
Software reused from offline
Full event information available, but partly still RoI based



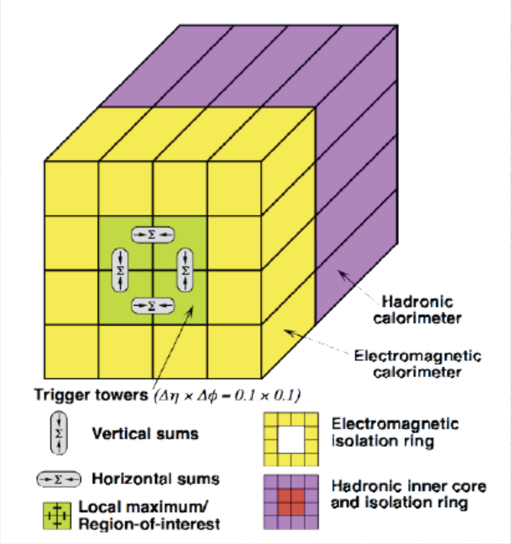
Hadronic τ trigger

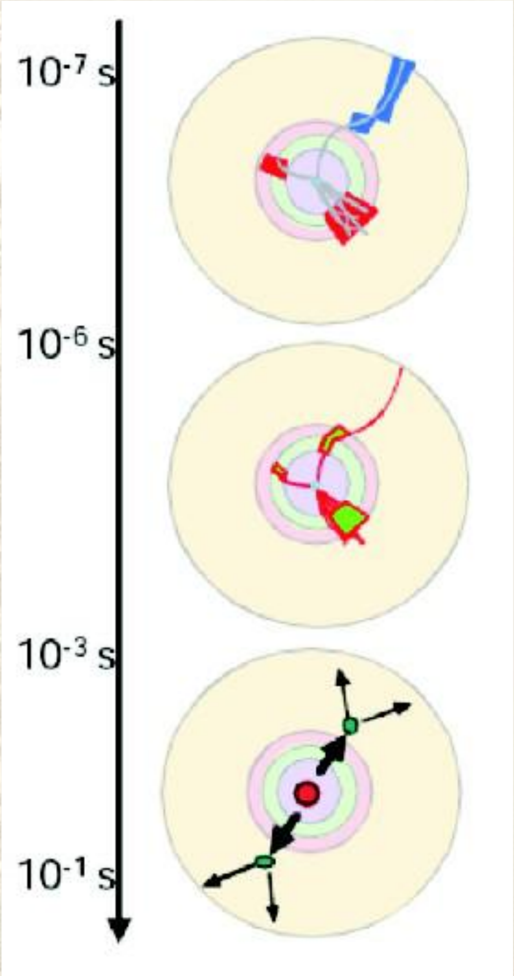


Quark or gluon initiated (QCD) jets can mimic tau jets, and are far more common in proton-proton collisions at the LHC. This important challenge is overcome by designing selection criteria that exploits the unique signature of hadronic tau decays:

- Low track multiplicity (mainly 1 or 3)
- Particles from the decay form a narrow collimated jet
- No particles around the cone containing the tau decay products (isolation)

Hadronic τ trigger





Level 1 (L1) - finds regions of activity in the detector (RoI).

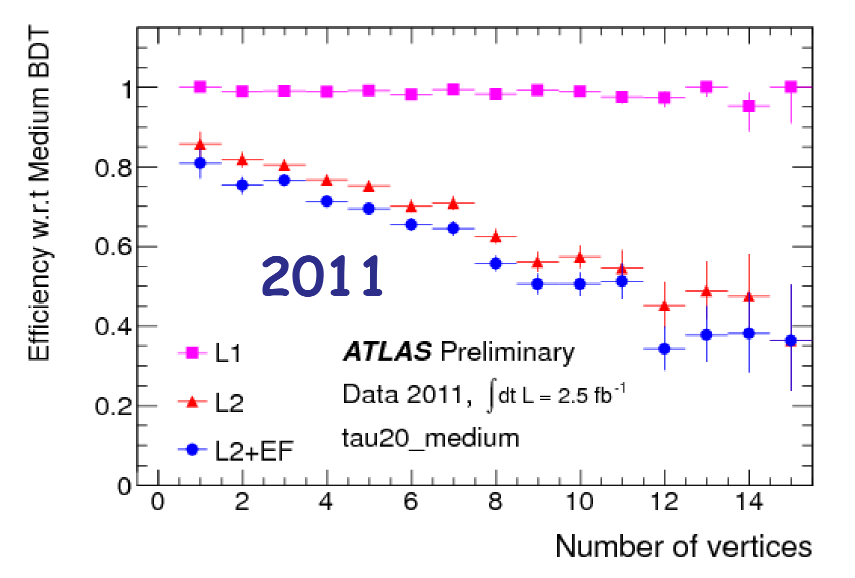
- Trigger Tower = 0.1 x 0.1(η x φ)
- Sum of several calorimeter cells
- Local maximum (0.2 x 0.2) core region should be above threshold.
- Simple, fast selection applied.

Level 2 (L2) – within the RoI, track and calorimeter cell information is combined (Δz0 criterion against pileup) tau dedicated selection on number of tracks, isolation and lateral shape in calorimeter is applied.

Event filter (EF)

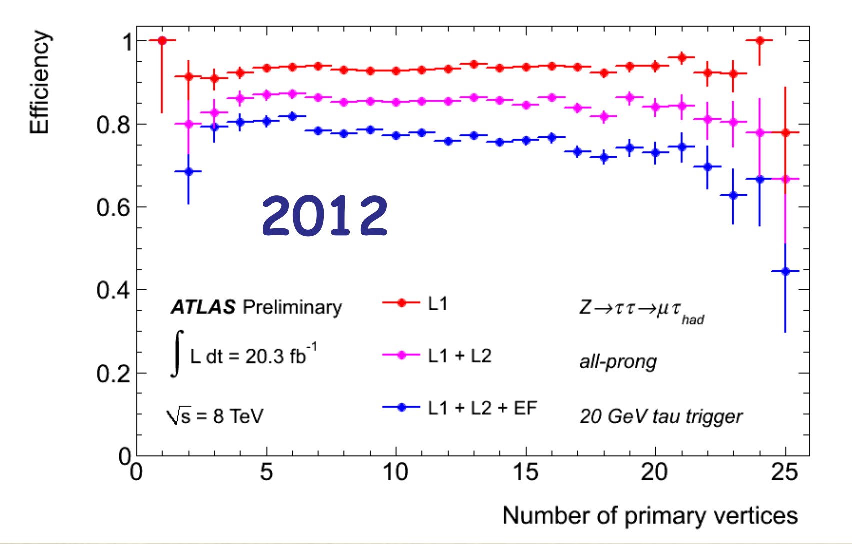
- Algorithm similar to offline reconstruction
- Calorimeter clusters with proper calibration and noise suppression
- MVA-based selection

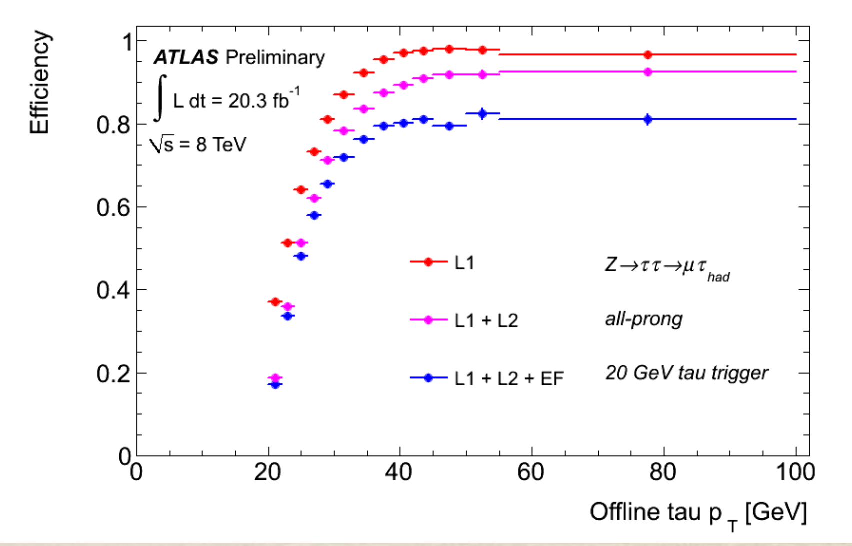
Hadronic τ trigger – performance



Significant improvements for 2012:

- ❖Much improved pileup robustness
- ❖Smaller cone sizes, Δz track cuts
- ❖EF now uses multi-variate analysis (MVA) selection to increase rejection power significantly



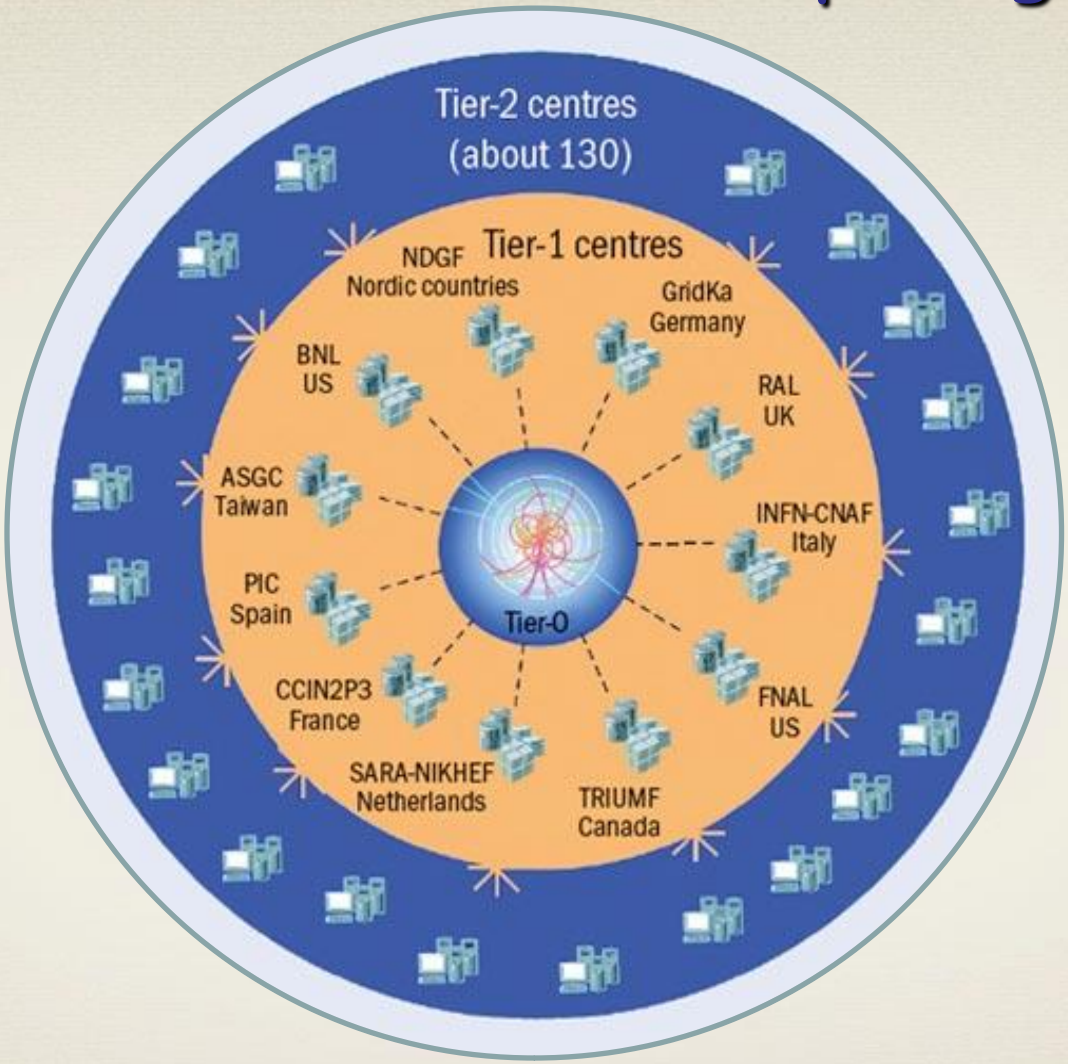


The Worldwide LHC Computing Grid

Tier-0 (CERN and Hungary): data recording, reconstruction and distribution

Tier-1: permanent storage, re-processing, analysis

Tier-2: Simulation, end-user analysis



nearly 160 sites, 35 countries

~250'000 cores

173 PB of storage

> 2 million jobs/day

10 Gb links

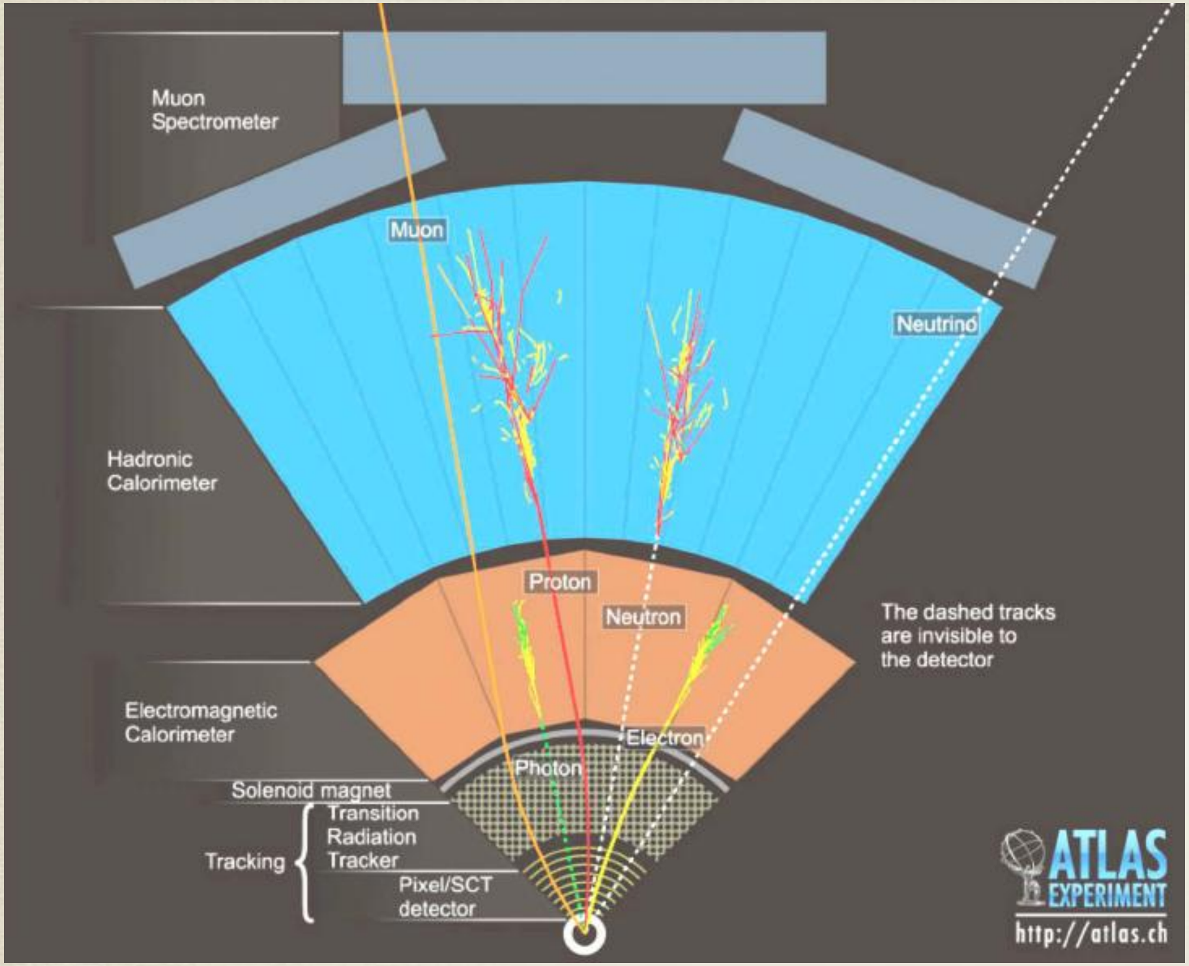
R. Heuer

WLCG:
An International collaboration to distribute and analyse LHC data

Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists

Reconstructing physics objects

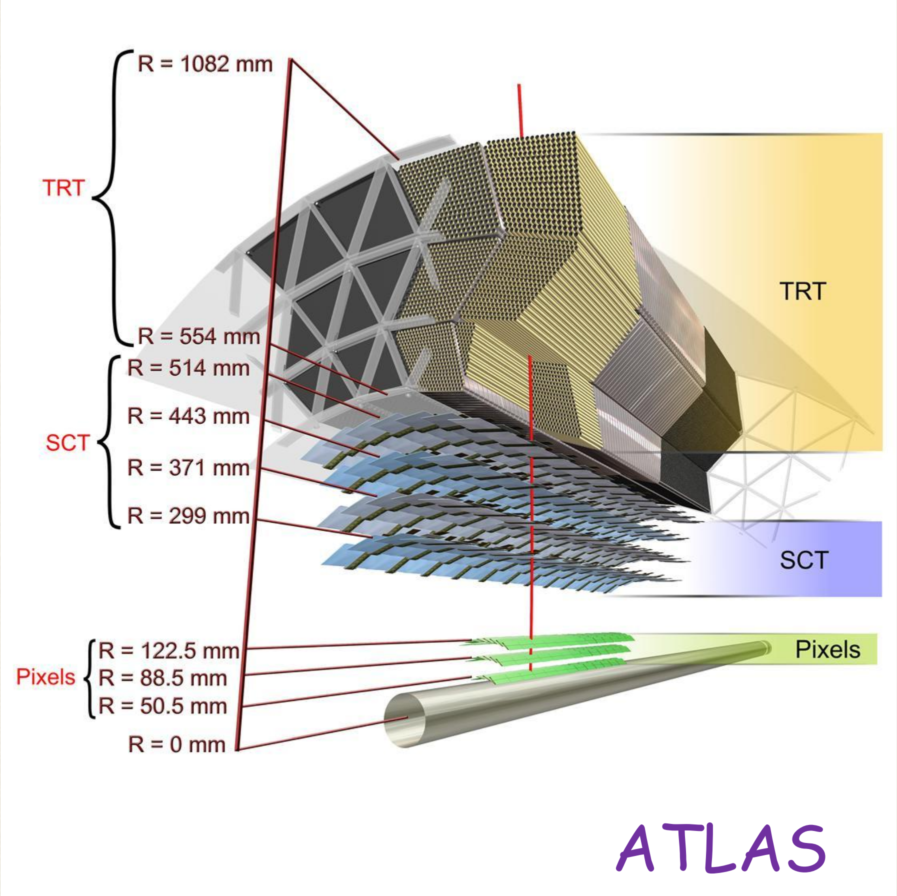
Measurements from a generic LHC spectrometer



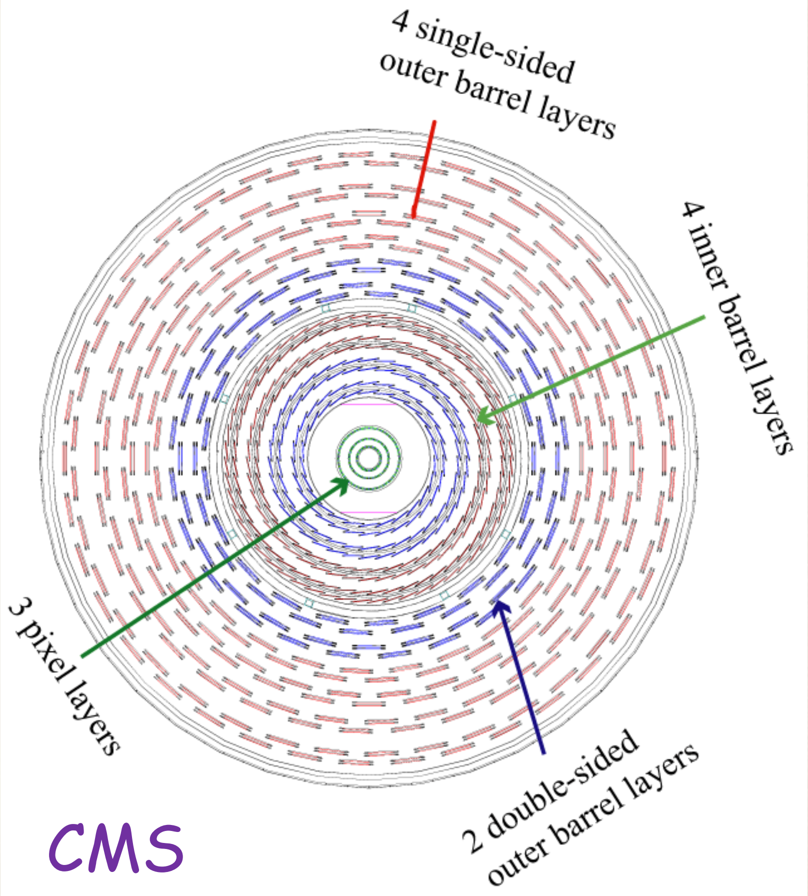
- ❖ Charged particles
- ❖ Primary and secondary vertices
- ❖ Lepton identification
- ❖ Muon spectrometer
- ❖ Electromagnetic calorimeter
- ❖ Photon identification
- ❖ Hadronic energy
- ❖ Jets
- ❖ Missing E_T ($\not{E}_T$)
- ❖ ATLAS/CMS are 4Π hermetic detectors
- ❖ Longitudinal boost unknown
- ❖ Only momentum balance in the transverse plane possible

Contemporary tracking devices

Silicon + gaseous detectors



All-silicon design



O(10k) sensing devices; intrinsic resolution O(10-100 μm)

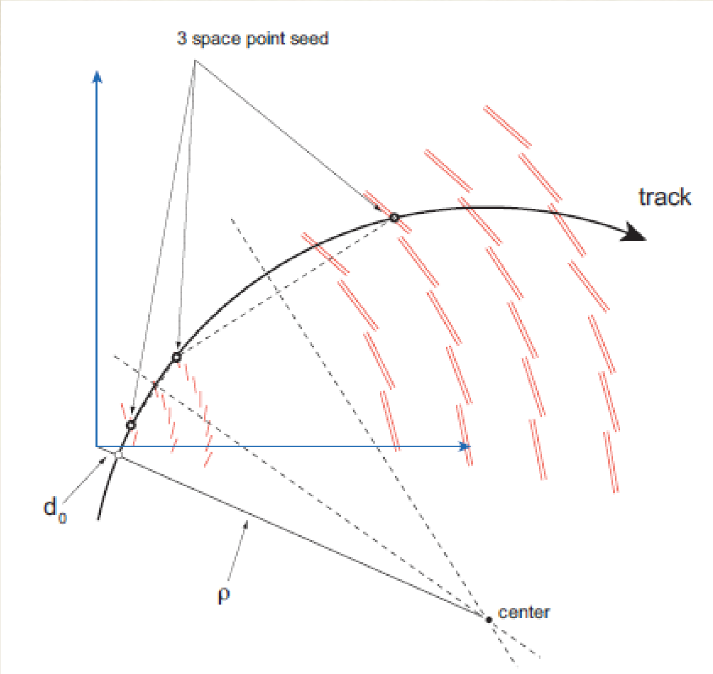
Tracking

- Seeding and candidate building (numerical propagation + Kalman filter):

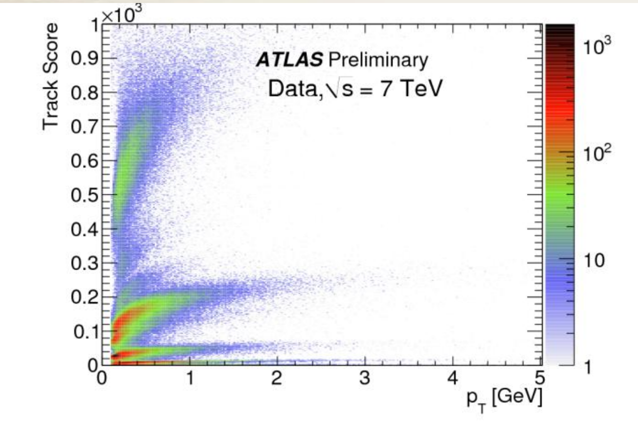
$$\boldsymbol{\pi}_{l+1} = \boldsymbol{\pi}_l + \mathrm{Cov}(\boldsymbol{\pi}_l)\mathbf{E}^{\mathrm{T}}\mathbf{W}\mathbf{r}_l \quad \text{where} \quad \mathbf{E} = \partial\mathbf{r}/\partial\boldsymbol{\pi}$$

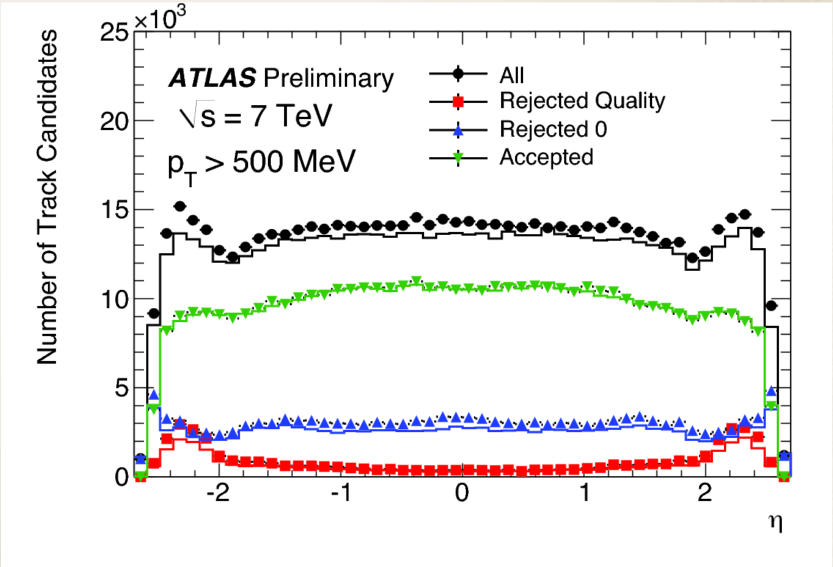
$$\mathrm{Cov}(\boldsymbol{\pi}_{l+1}) = \mathrm{Cov}(\boldsymbol{\pi}_l) - \mathrm{Cov}(\boldsymbol{\pi}_l)\mathbf{E}^{\mathrm{T}}\mathbf{W}\mathbf{E}\,\mathrm{Cov}(\boldsymbol{\pi}_l)$$

$$\mathbf{W} = [\mathbf{V} + \mathbf{E}\mathrm{Cov}(\boldsymbol{\pi})\mathbf{E}^{\mathrm{T}}]^{-1}$$



- Scorring based on number of hits, holes, momentum, fit quality, etc:





- Ambiguity resolving
- Fitting (either Kalman or Global χ^2)

Tracking & Vertexing

TRACKING

❖ Algorithms:
- Inside-out algorithm (pT>400 MeV)
- Conversions or long-lived: backtracking

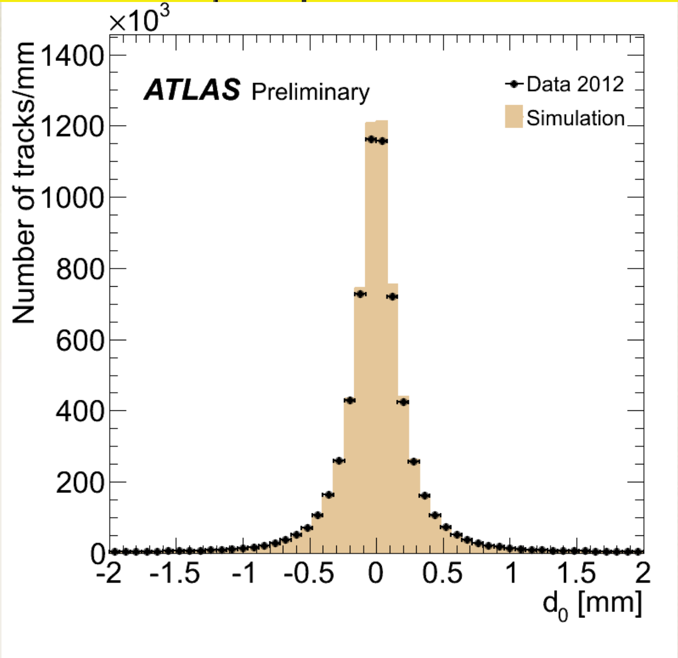
❖ Performance:
- Track reconstruction eff. computed from MC and compared to low luminosity runs
- Good resolution and MC comparison

VERTEXING

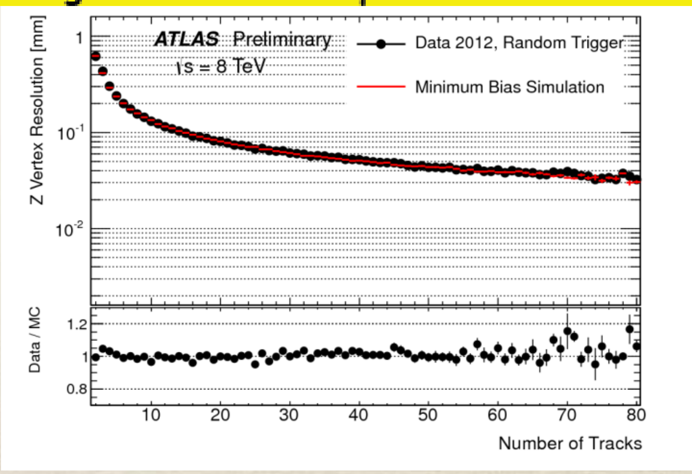
❖ Reconstruction method:
- Seed: position of z0 (2nd seed if >7σ)
- Iterative χ^2 fit of nearby tracks
- Primary vertex: with highest Σp_T^2

❖ Resolution:
- Method:
- Tracks from the same vertex are split into two sets; new vertices formed; resolution obtained from Δ
- Resolution decreases with #tracks
- Best resolution: ~20/30 μm in X/Z
- Good agreement between data and MC

Transverse impact parameter resolution

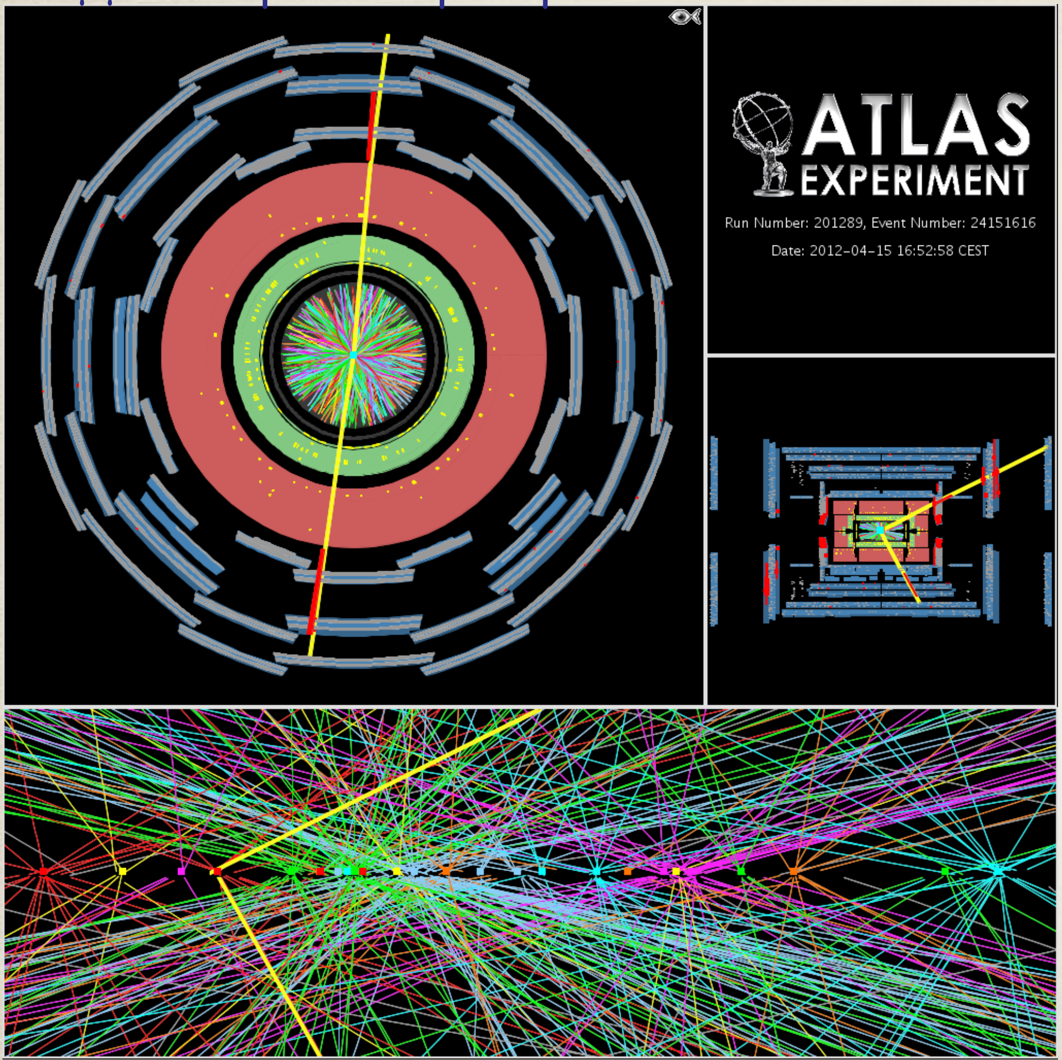


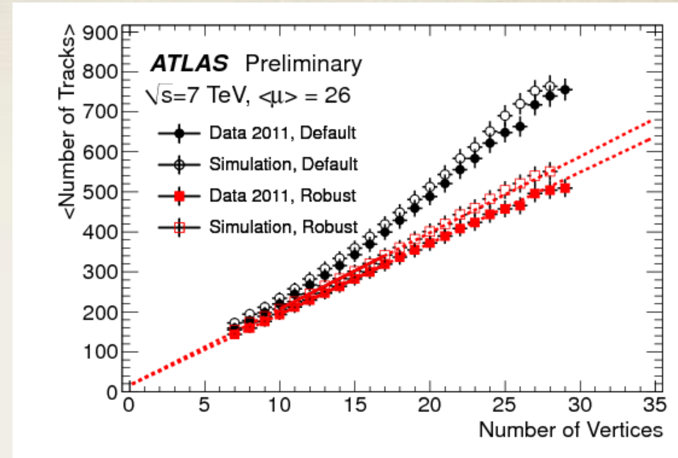
Longitudinal vertex position resolution



Tracking

Z->μμ on top of 25 pileup interactions

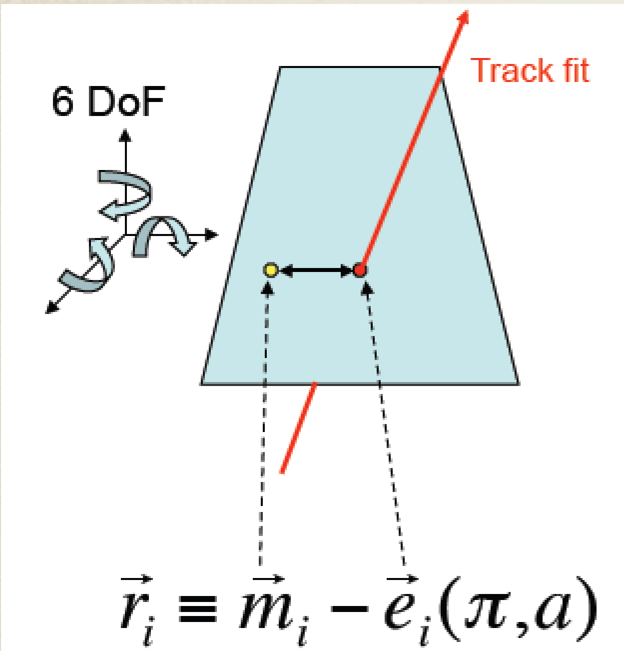




- Multiple interactions increase fake rate (combinatorial) and CPU
- Dealing with it:
  - More robust reconstruction cuts: ≥9 hits (vs 7) and 0 missing in pixel
- Result:
  - Decrease rate of fakes
  - Small effect on track efficiency

Basic track fit (linearization)

$$\pi = (d, z, \varphi, \vartheta, Q/p_T), \qquad \vec{e} \equiv \vec{e}(\pi)$$



$$\chi^2 = \mathbf{r}^T \mathbf{V}^{-1} \mathbf{r}, \qquad r_i \equiv (\vec{e} - \vec{m}) \bullet \hat{k}$$

$$\mathbf{r}(\pi) = \mathbf{r_0} + \frac{\partial \mathbf{r}}{\partial \pi}(\pi - \pi_0)$$

linear exp. aro. seed

$$\frac{d\chi^2}{d\pi} = 0$$

minimization condition

$$\pi - \pi_0 = \left(\frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \frac{\partial \mathbf{r}}{\partial \pi} \right)^{-1} \frac{\partial \mathbf{r}}{\partial \pi}^T \mathbf{V}^{-1} \mathbf{r}_0$$

Idea of the Global χ^2 alignmet approach

$$\mathbf{r}(\pi) = \mathbf{r_0} + \frac{\partial \mathbf{r}}{\partial \pi}(\pi - \pi_0) + \frac{\partial \mathbf{r}}{\partial a}(a - a_0) \qquad \frac{d\chi^2}{d\pi} = \frac{d\chi^2}{da} 0$$

simultaneous fit of all tracks and alignment parameters N+n*k pars!

Impossible to solve !!! ☹

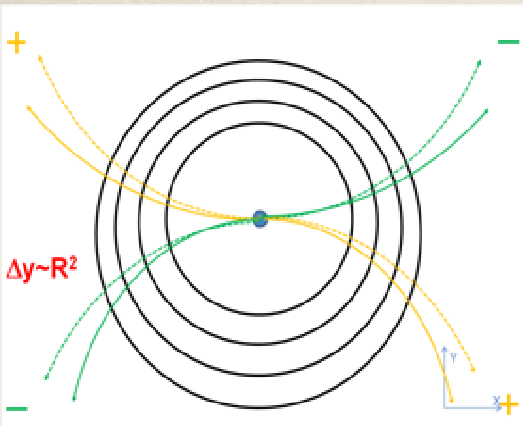
Way out!:

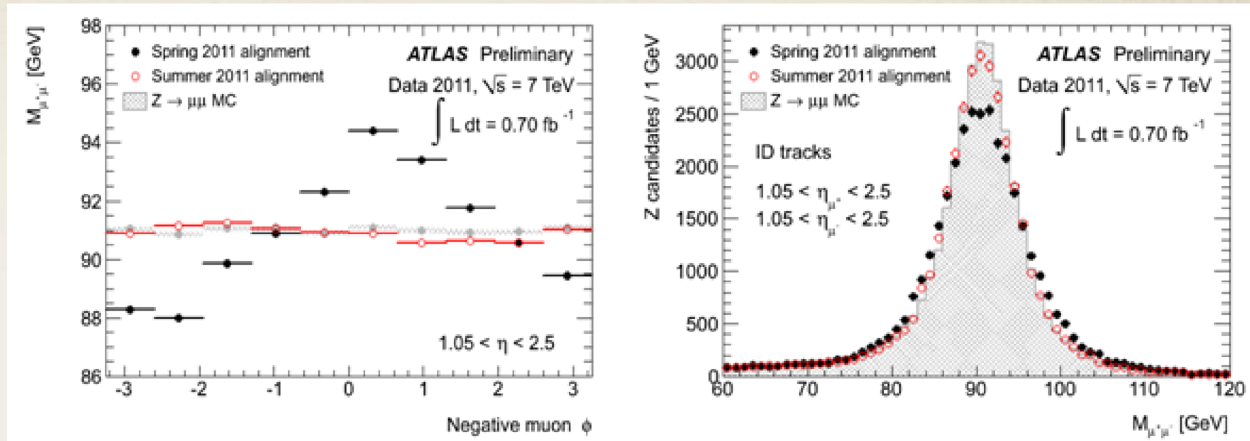
Fold the track fit in. Solve for the alignment only:

$$\mathbf{r}(\pi) = \mathbf{r_0} + \underbrace{\left(\frac{\partial \mathbf{r}}{\partial \pi}\frac{d\pi}{da} + \frac{\partial \mathbf{r}}{\partial a} \right)}_{\frac{d\mathbf{r}}{da}}(a - a_0)$$

$$a - a_0 = \left(\sum_{\text{tracks}} \frac{d\mathbf{r}}{da}^T \mathbf{V}^{-1} \frac{d\mathbf{r}}{da} \right)^{-1} \sum_{\text{tracks}} \frac{d\mathbf{r}}{da}^T \mathbf{V}^{-1}\mathbf{r}_0$$

The main challenge:
Systematic deformations due to weak modes



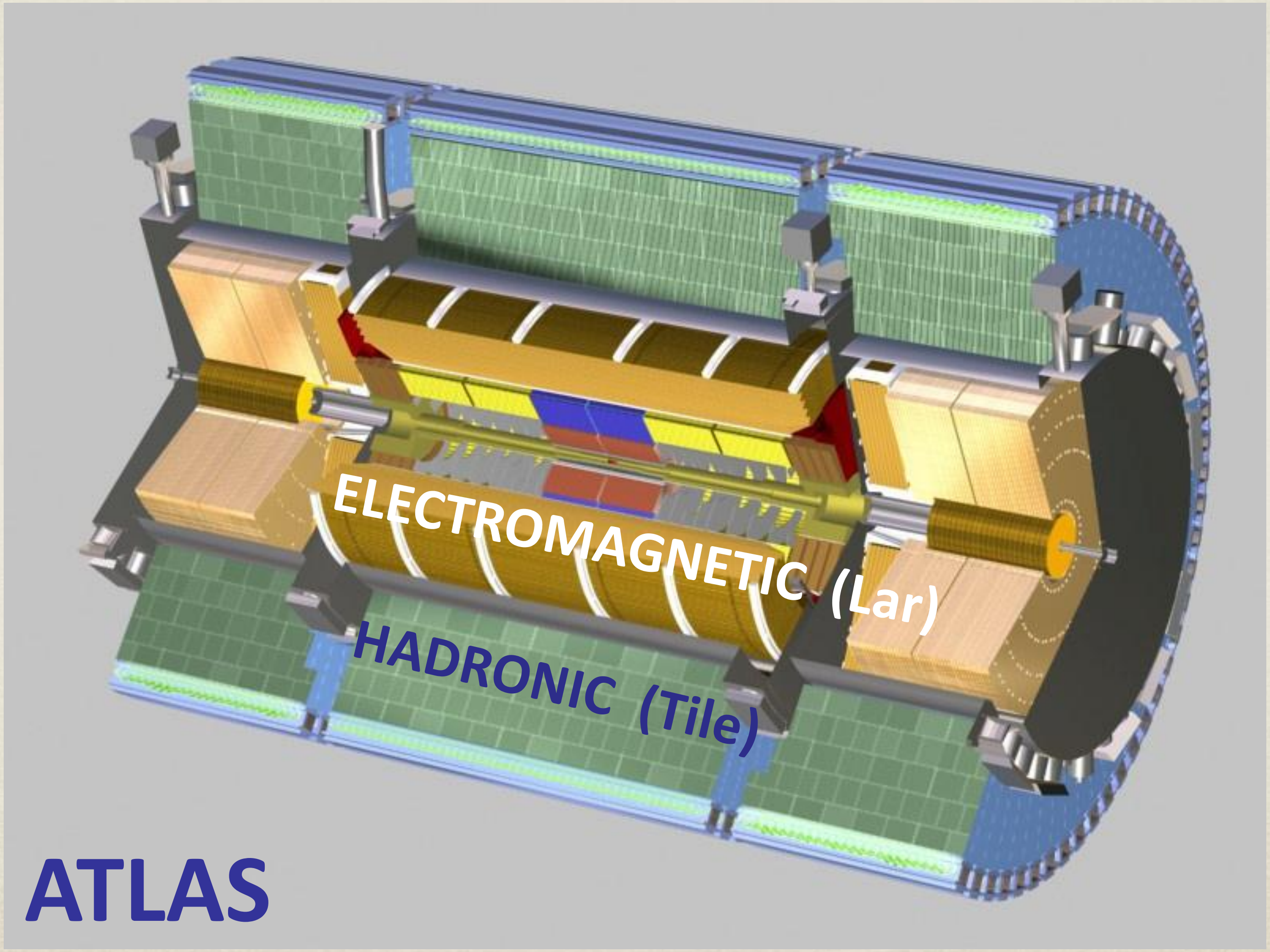


Determination of over 700,000 DoF's of the ATLAS tracking system is realized by combination of Global and Local χ^2 methods.

An involved alignment procedure using the Global χ^2 algorithm with external constraints from E/p for electrons & positrons allows to eliminate the goross effect ☺

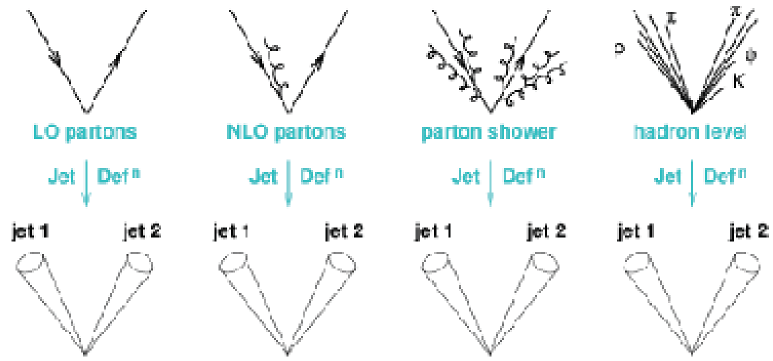
CPU intensive!

Calorimeter System
jet, electron, photon, tau, etc.



Jet reconstruction

Goal: kinematics of jet $\leftrightarrow$ kinematics of underlying (parton) physics



From G. Salam, MCNet School 2008

Apply same jet definition to objects on different levels:

1. Partons
2. **Particles** $\rightarrow$ Truth Jets
3. **Calorimeter objects (Towers, Topoclusters)** $\rightarrow$ Reconstructed Jets

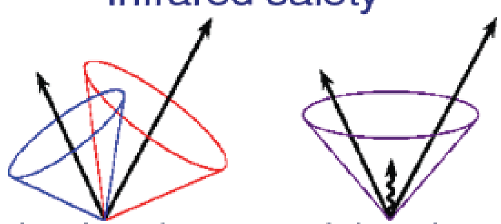
No *right* jet algorithm

Different processes $\leftrightarrow$ different algorithms / parameters.

Requirements:

1. Theoretically well behaved $\longrightarrow$ no α_s dependence of jet configuration:

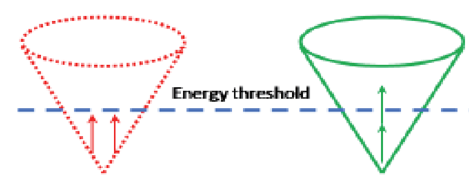
Infrared safety



Without gluon radiation: 2 different jets

With gluon radiation: only 1 jet

Collinear safety



With collinear splitting: no jet (under threshold)

Without collinear splitting: 1 jet (over threshold)

Jet reconstruction

Sequential recombination algorithms (k_t-like)

Algorithm specification: Anti-k_t

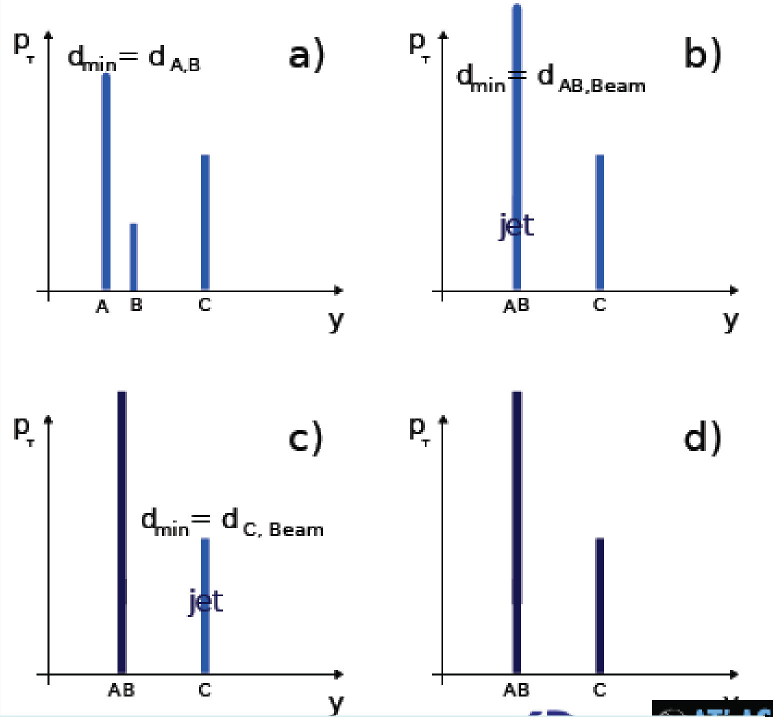
- $d_{i,j} = min(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,i}^2})\frac{\Delta R^2}{D^2}$;
  $d_{i,Beam} = \frac{1}{p_{T,i}^2}$
- **D** : algorithm parameter
- Iterate:

1. For every pair of objects i, j calculate $d_{min} = min(d_{i,j}, d_{i,beam})$
2. If $d_{min} = d_{i,j}$ recombine objects
   Else i is a jet, remove it from list [a]

- Recombination starts from **hard** objects

C. Doglioni

Idea:



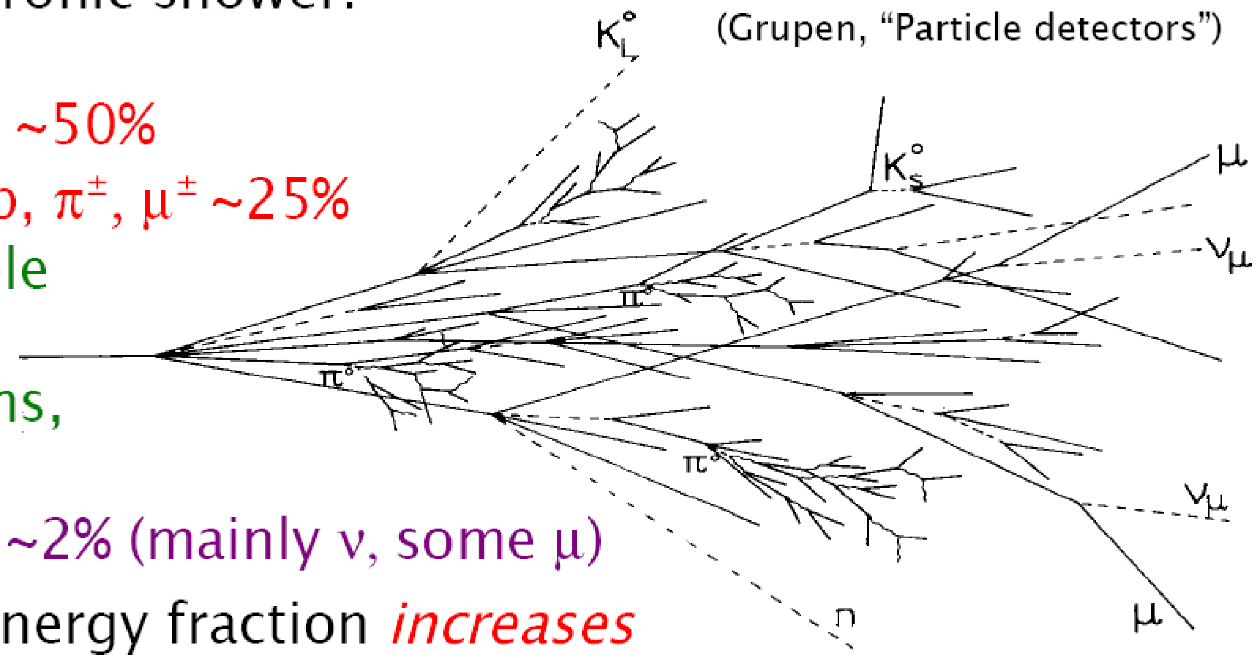
Colinear and infrared safe, soft particles recombined

Jet energy calibration

Hadronic calibration

- Contents of a hadronic shower:
  - Visible energy
    from $e^{\pm}$; $\pi^0 \rightarrow \gamma\gamma$; ~50%
    ionisation from p, $\pi^{\pm}$, $\mu^{\pm}$ ~25%
  - Hadronic invisible energy ~25% (nuclear excitations, break-ups)
  - Escaped energy ~2% (mainly ν, some μ)
- Electromagnetic energy fraction *increases* with increasing hadron energy due to production of π^0's in the shower
- Event-by-event fluctuations
- Aim of hadronic calibration: Compensate for invisible and escaped energy.

(Grupen, "Particle detectors")



Jet energy calibration

- Global calibration:

  first form larger physics objects (jets, ...)

  then calibrate the objects using weights derived for the particular object

- Local calibration:

  first compensate for invisible energy loss on local detector objects (clusters of calorimeter cell signals)

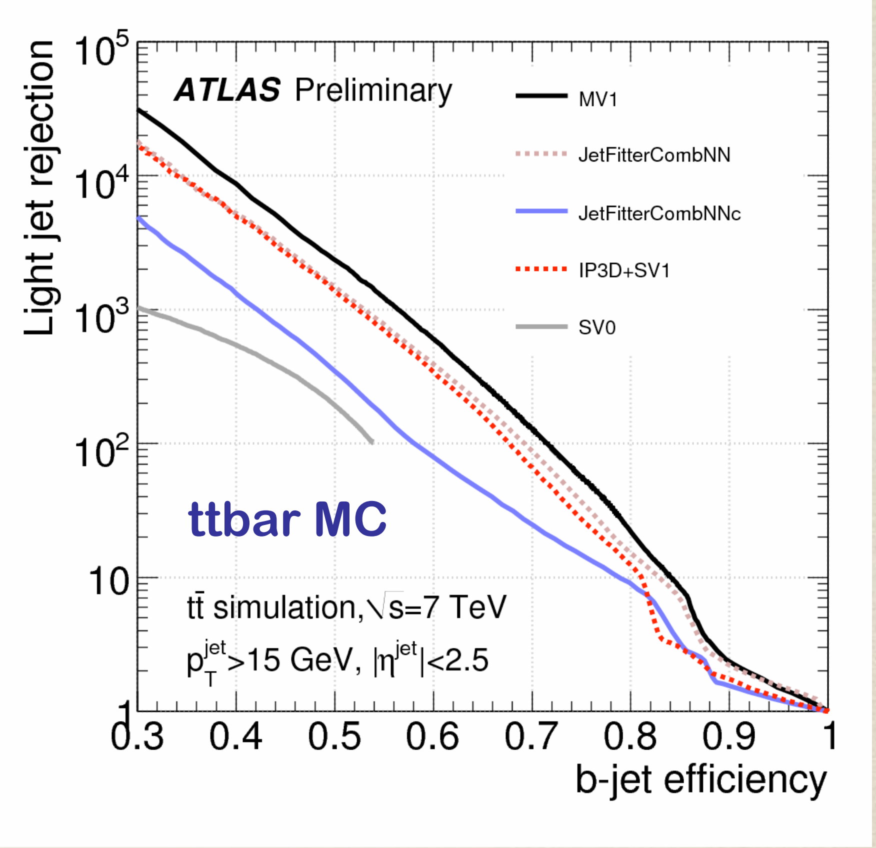
  then use calibrated clusters to form larger objects (jets etc.)

• Global calibration uses χ^2 minimization at cell level. Final scale factors are jet-algorithm dependent!
• Local calibration uses individually classified topo-clusters. More robust and versatile. This approach is gradually gaining the field.

Identification of b-jets

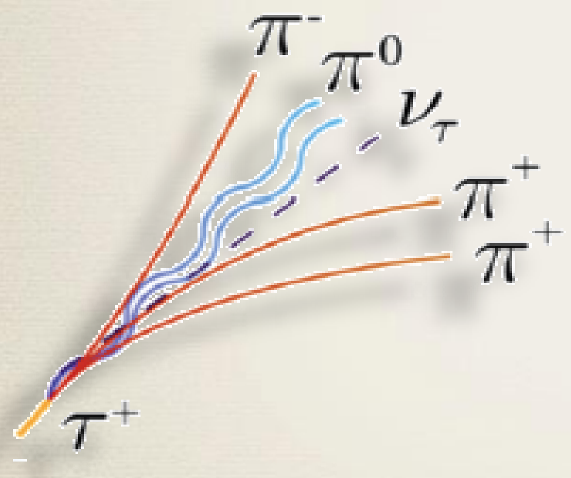
b-tagging:

- Relies on the finite lifetime of B hadrons ($\mathcal{O}$(1.5)ps).
- Different algorithms exist based on impact parameter of tracks, secondary vertex, its invariant mass, etc.
- **Neural network** combines output weights of all algorithms into a single discriminant (MV1)
- Dedicated methods measure tagging efficiency and mistag rates.



Tau lepton basics

- Mass: 1.777 GeV/c2
  the heaviest lepton
- cτ: ~87μm
  short lifetime
- decays via weak interactions



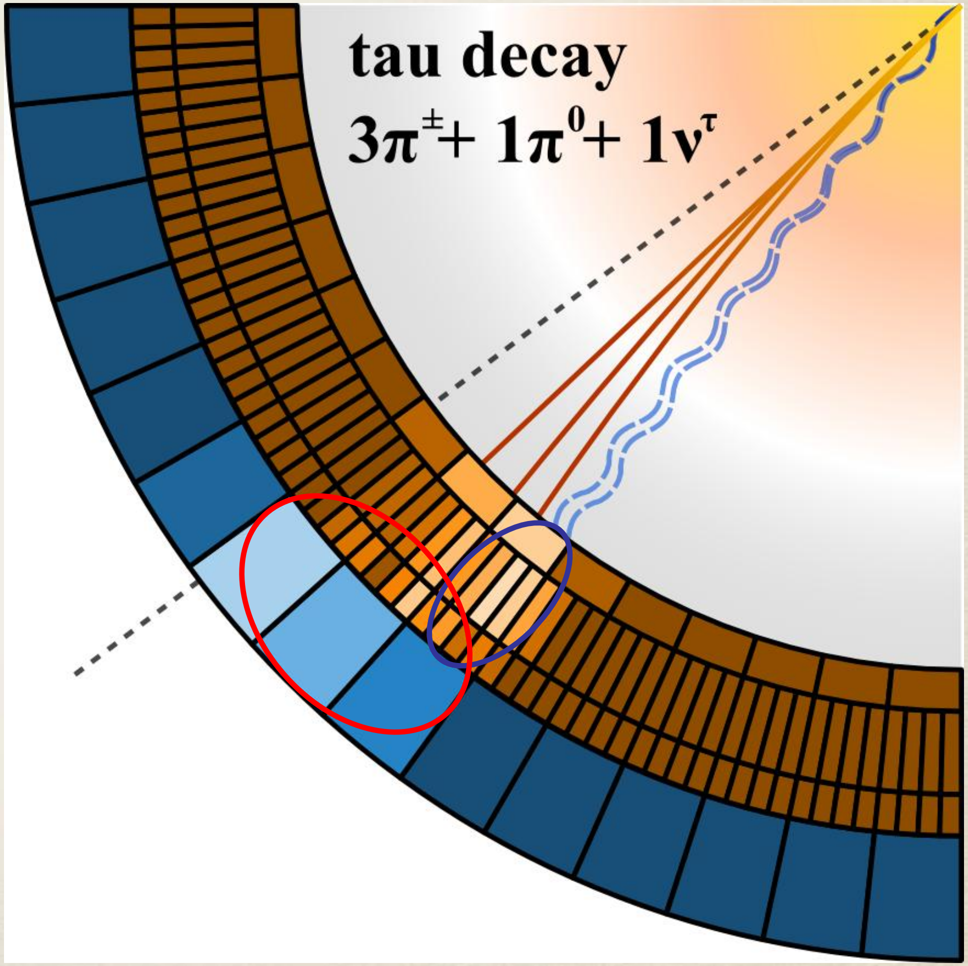
First observed in 1977 by Martin Perl et al. (SLAC-LBL)

Most important decay modes

Decay Mode	Branching Fraction
Leptonic modes	**~35%**
$\tau^\pm \to e^\pm \nu_e \nu_\tau$	18%
$\tau^\pm \to \mu^\pm \nu_\mu \nu_\tau$	17%
Hadronic modes	**~65%**
1 prong (1 charged particle)	**46%**
$\tau^\pm \to \pi^\pm \nu_\tau$	11%
$\tau^\pm \to \pi^\pm \pi^0 \nu_\tau$	26%
$\tau^\pm \to \pi^\pm \pi^0 \pi^0 \nu_\tau$	9%
3 prong (3 charged particles)	**14%**
$\tau^\pm \to \pi^\pm \pi^\pm \pi^\mp \nu_\tau$	9%
$\tau^\pm \to \pi^\pm \pi^\pm \pi^\mp \pi^0 \nu_\tau$	5%

Hadronic Tau reconstruction

❖Charged pions leave tracks (red) and energy deposits.
❖Neutral pions decay into pairs of photons (blue). No tracks but dense energy deposits.
❖Neutrinos don't leave anything.

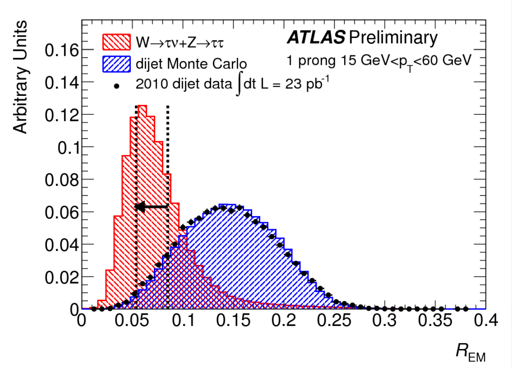


Reconstruction of hadronic τ decays

(ATLAS: **ATLAS-CONF-2011-152**) and (CMS: JINST **7 (2012) P01001**)

ATLAS (top-down)

- Start from the anti-k_T jets reconstructed from calorimeters.
- Associate charged tracks.
- Energy calibration based on MC.
- Use MVA to discriminate against QCD jets and leptons.

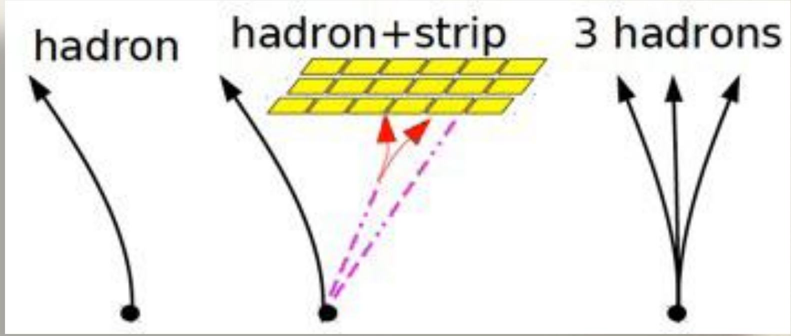


BDT-based ID:

- 60%(40%) efficiency for medium (tight)
- 2-3% (0.5%) QCD jet acceptance.

CMS (bottom-up)

- Start from particles reconstructed by the *Particle Flow* algorithm
- Construct 1-prong, 1-prong+π0's, 3-prong τ candidates.

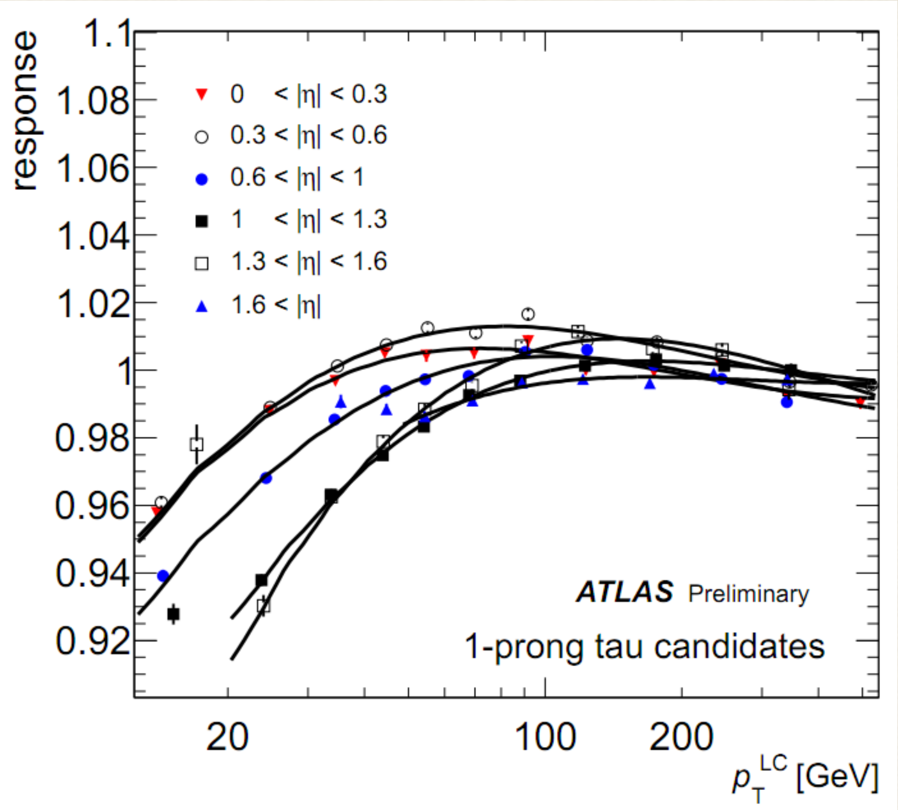


- MVA discriminant based on $\sum p_T$ of particles in rings around τ.

BDT-based ID:

- 50%(36%) efficiency for loose (medium)
- 1% (0.4%) QCD jet acceptance.

Tau energy scale



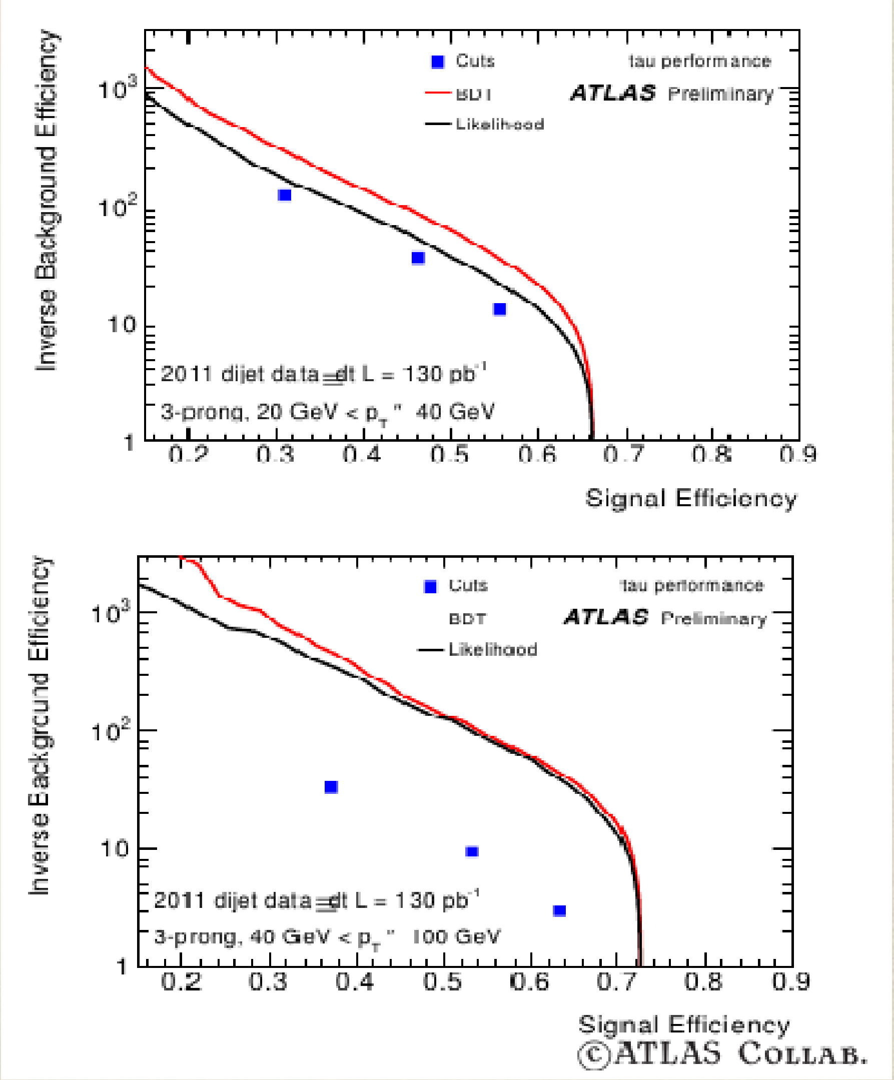
- ❖ Start from a generic calibration for topological clusters (Local Hadronic scale).
- ❖ Response functions from simulation at the Local Hadronic Scale are inverted to find the tau energy scale correction.

Tau identification

Input to **multi-variate methods**:
- Width of energy deposits
- Invariant mass of clusters
- transverse flight path signicance, etc.

Against **Jets**:
- Cut-based
- Log Likelihood Ratio
- Boosted Decision Trees (BDT)

Against **Electrons**:
- Cut-based
- BDT



High-end analysis

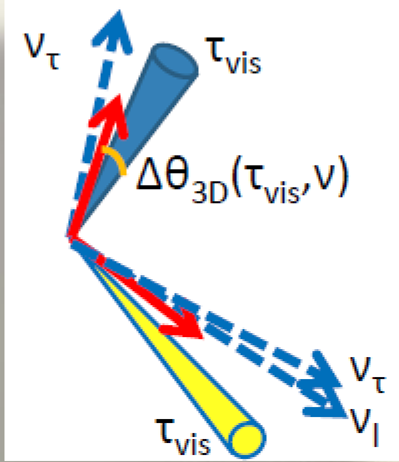
Kinematic variable reconstruction
Background estimation
Classification
Statistical interpretation

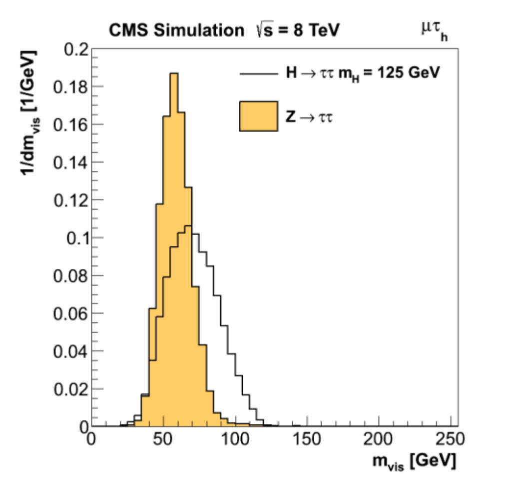
H->ττ search

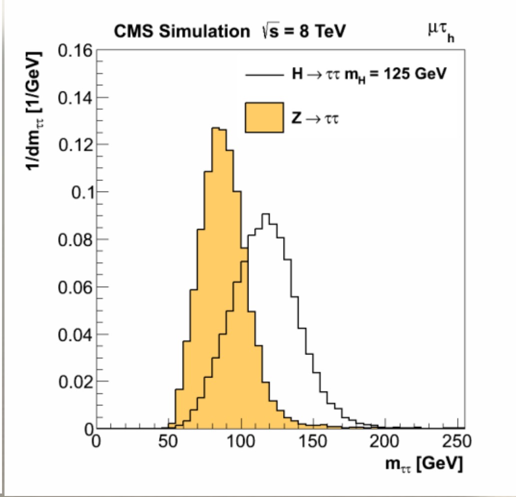
Invariant mass of the ττ system

- There are 6 to 8 parameters describing invisible neutrinos and 4 constraints ($2\times m_\tau$, $\not{E}_{Tx}$, $\not{E}_{Ty}$)
- Find max. likelihood solution accounting for the distributions of the ττ kinematics and $\not{E}_T$ resolution.

$$\mathcal{L} = -\log\left(\mathcal{P}(\Delta R_1, p_{\tau 1}) \times \mathcal{P}(\Delta R_2, p_{\tau 2}) \times \mathcal{P}(\Delta \not{E}_{T_x}) \times \mathcal{P}(\Delta \not{E}_{T_y})\right)$$







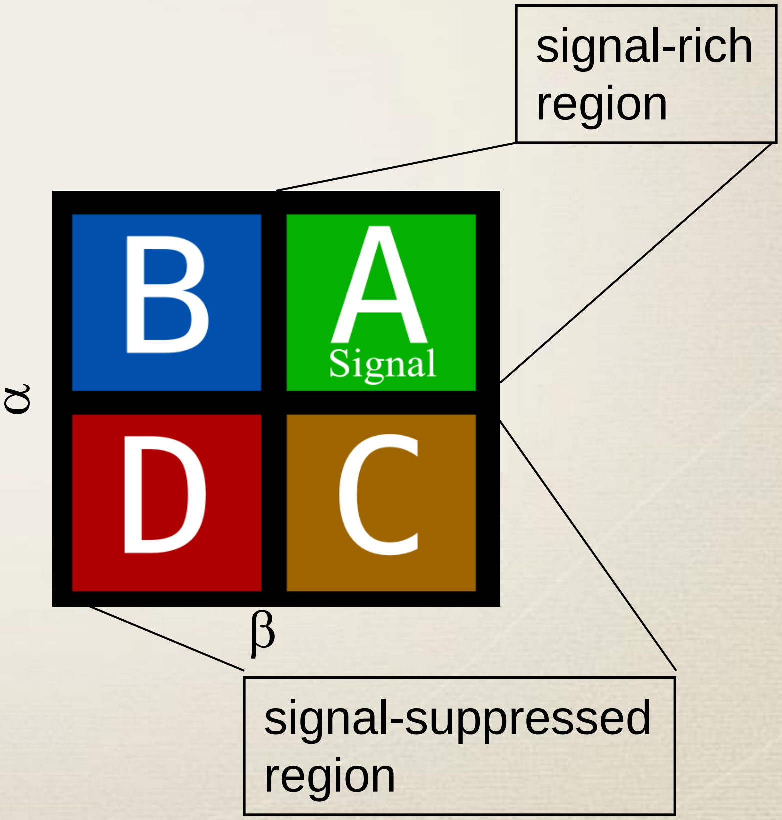
$\sigma(m_{\tau\tau}) \leq 20\%$ (depending on the channel and kinematics)

Data-driven methods – the guiding principle

- Most have a common principle which relies on identifying the control region (exclusive to the actual signal search region) which is signal suppressed but still representative for the background.
- One needs two variables (α,β) which are approximately independent:

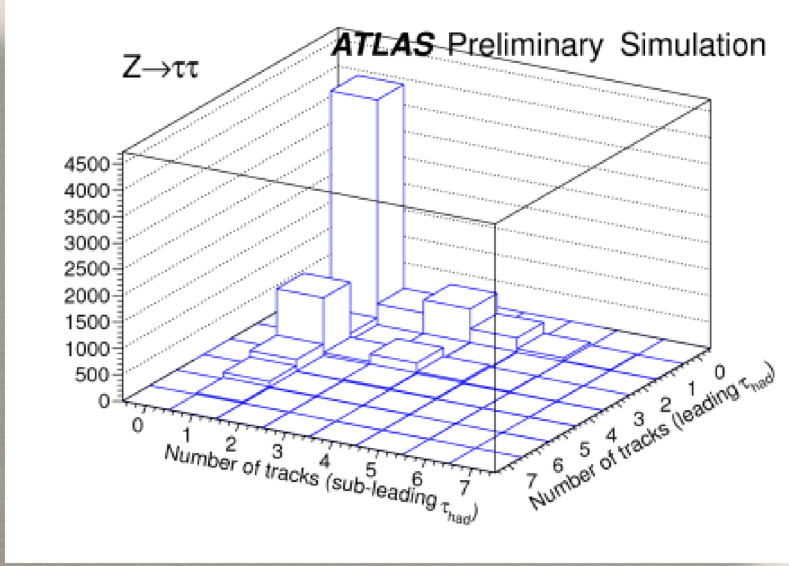
- β distribution from background events in the signal region can be estimated from data using:
  - **A=CxB/D**

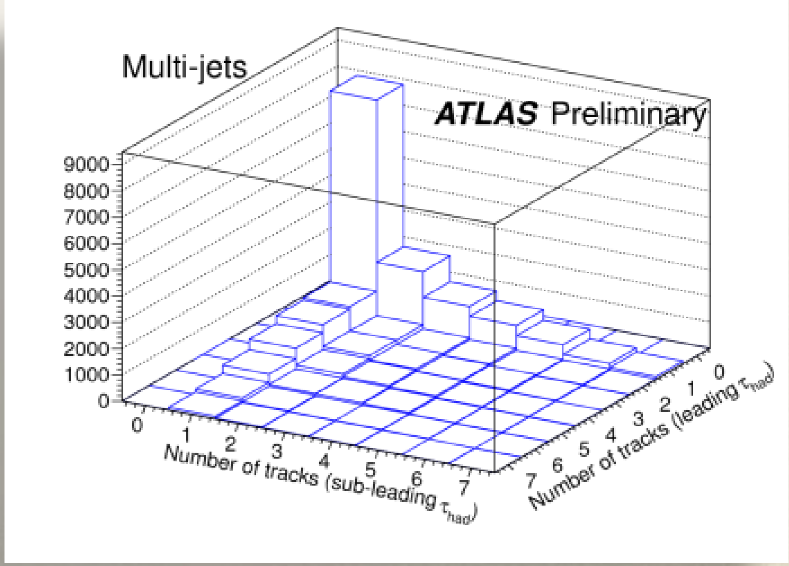
    normalisation from data-suppressed region using an independent α variable
- If there is no signal (signal region consistent with the predicted background) -> DONE
- Otherwise one can iterate subtracting the observed signal from the control sample ("new M_T method")



The template method

Contribution (of the $Z/\gamma^*\text{->}\tau\tau$ & QCD backgrounds to the $H\text{->}\tau\tau$ search in the $\tau_{had}\tau_{had}$ decay mode is estimated using 2-D track multiplicity template fit:



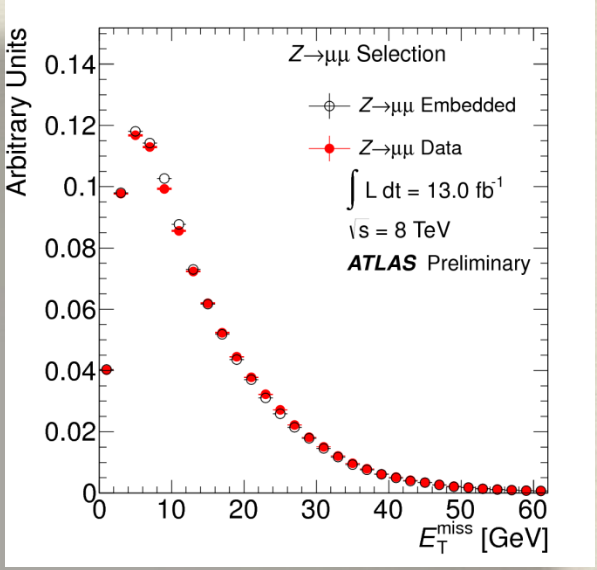


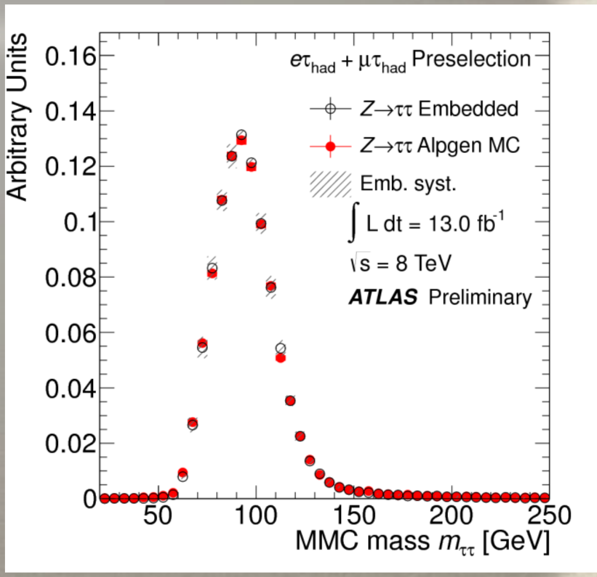
Background estimation: „embedding"

- Embedding in $Z/\gamma^*\to\mu\mu$ events: reconstructed muons are removed from data events and replaced by simulated τ decays with the same kinematics.

- Advantage: data-driven description of the entire event (except for lepton decays) leading to significantly reduced systematic uncertainties (jets, underlying event, luminosity, etc.) compared to the MC simulation.

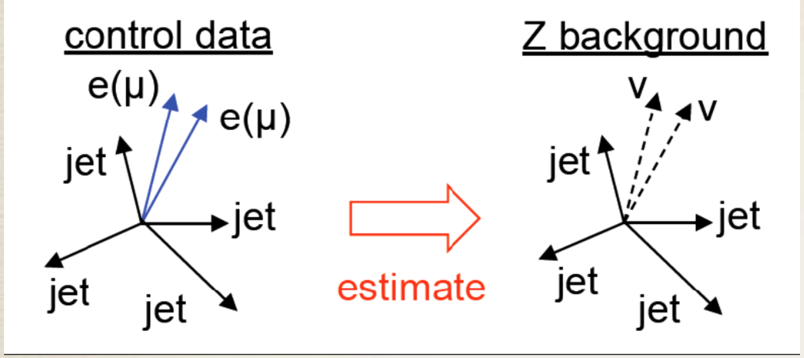
- ll, lτ_{had}, $\tau_{had}\tau_{had}$ (ATLAS & CMS)





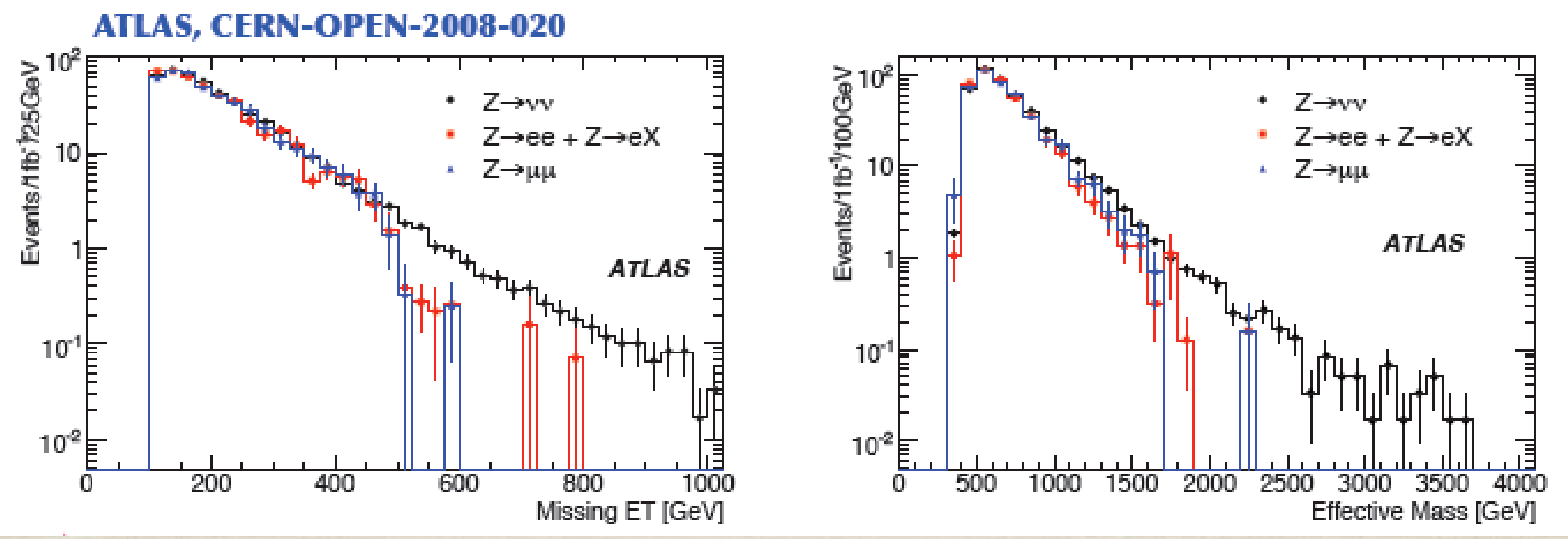
Z->νν + jets an important background to 0-lepton SUSYsearch

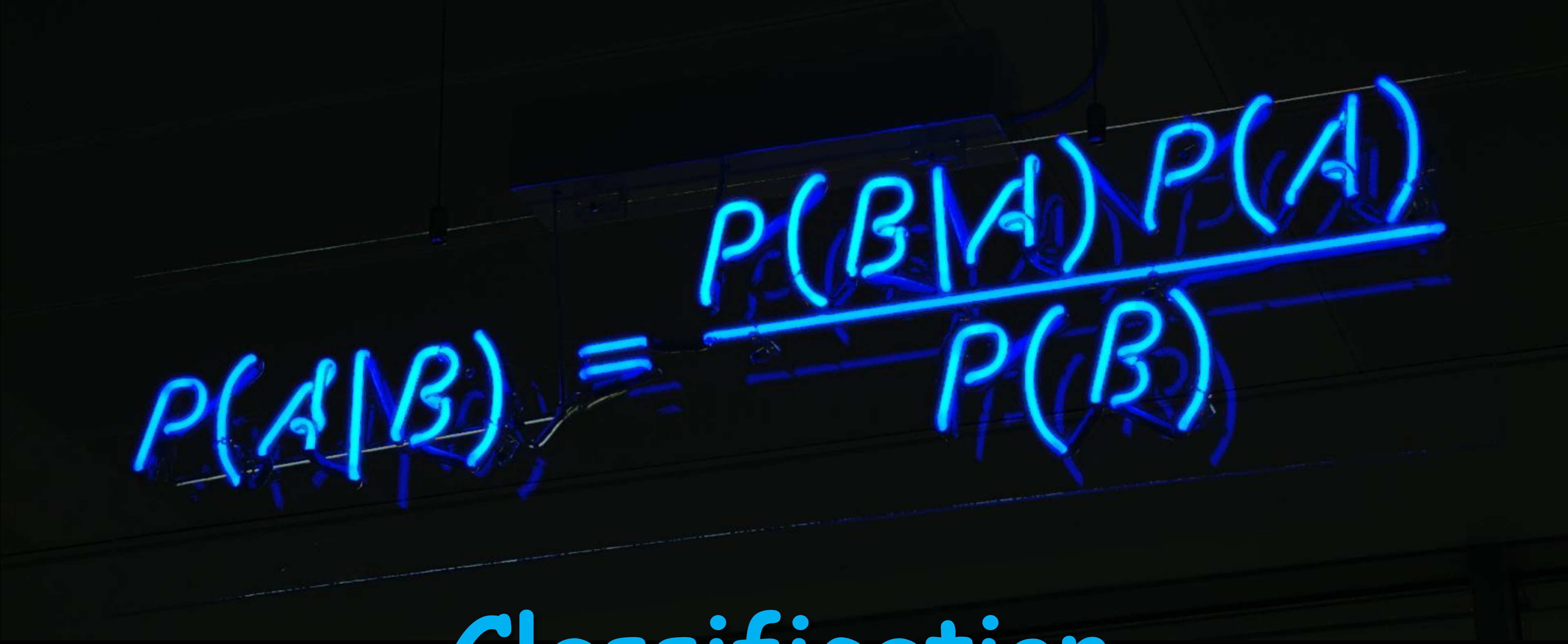
- Replace method relies on the measured Z->l+l-



Standard 0-lepton selection
+ Z->l+l-
with $p_T(l^+l^-)$ substitution for $\not{E}_T$

acceptance (η, p_T), efficiency, and Br corrections must be applied!





Classification
&
statistical interpretation

Classifiers

Identification, signal selection, etc.

Classifiers provide information in either binary format YES/NO or in terms of continuous ranking (e.g. probablility) about the identity of reconstructed objects/events, anwering: Is it our signal (S) or background (B)?. The optimal answer - Bayes classifier:

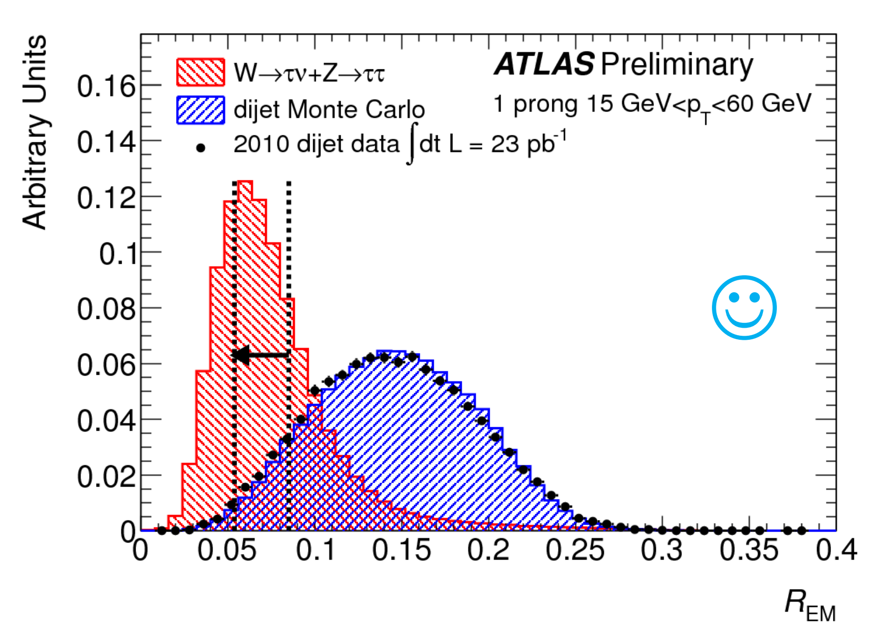
$$p(S\,|\,\mathbf{x}) = \frac{p(\mathbf{x}\,|\,S)\,p(S)}{p(\mathbf{x}\,|\,S)\,p(S) + p(\mathbf{x}\,|\,B)\,p(B)}$$

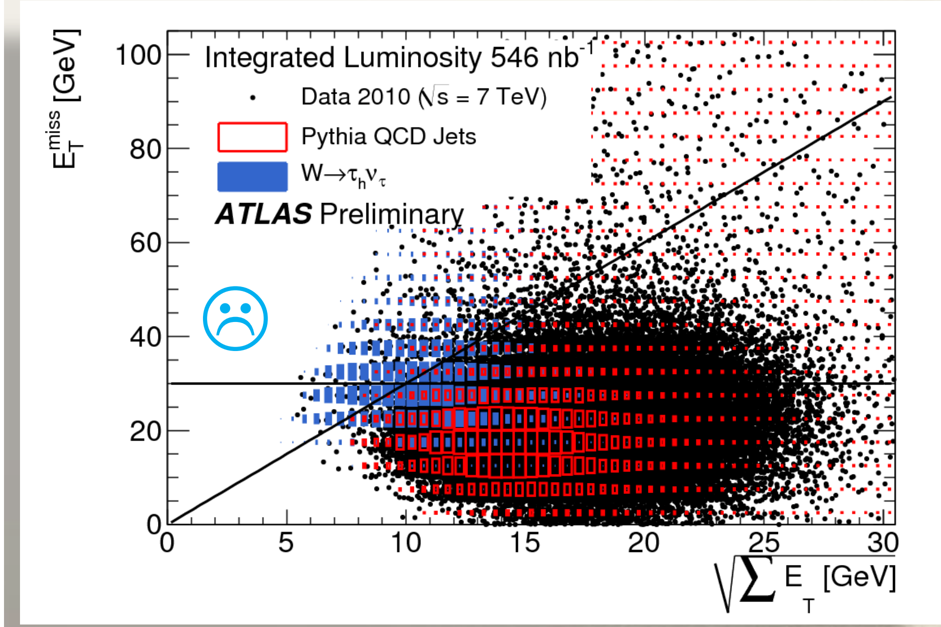
Multi-dimensional PDF impossible to construct analitically.
Numerical construction would require ∞ learning sample.
CPU intensive and practically unrealistic.
Other methods make only better or worse approximation.
Many algorithms on the market.

Classifiers

Cut-based

Cut-based classifiers are optimal in 1D problems and remain effective when more variables are involved, provided they are not correlated.

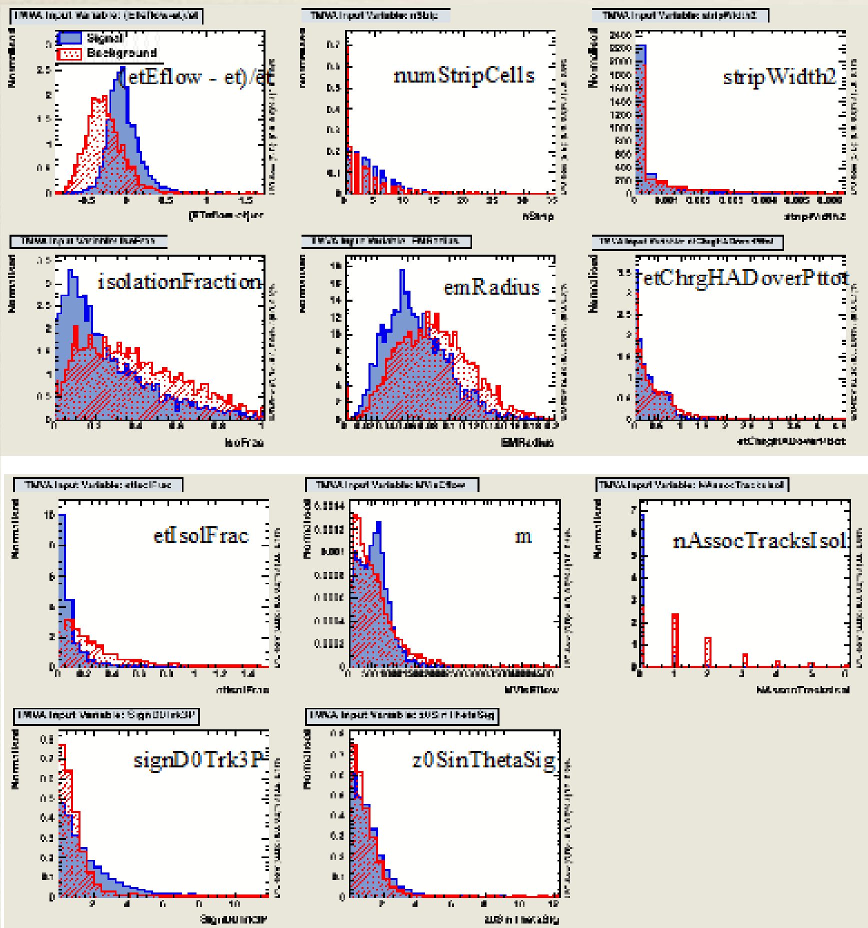




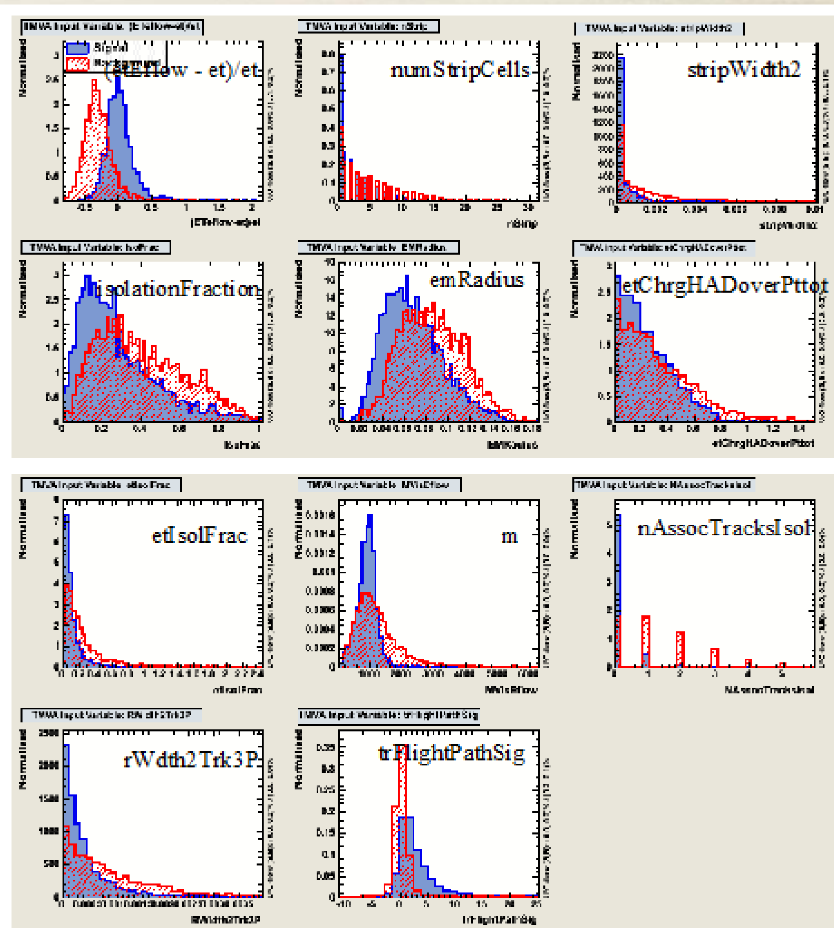
With larger dimensionality -> highly nontrivial

Classifiers

Eg: how to optimally combine all τ discriminants?



1-Prong



3-Prong

Classifiers

Projected likelihood

Likelihood ratio for event i_{event}

PDFs

discriminating variables

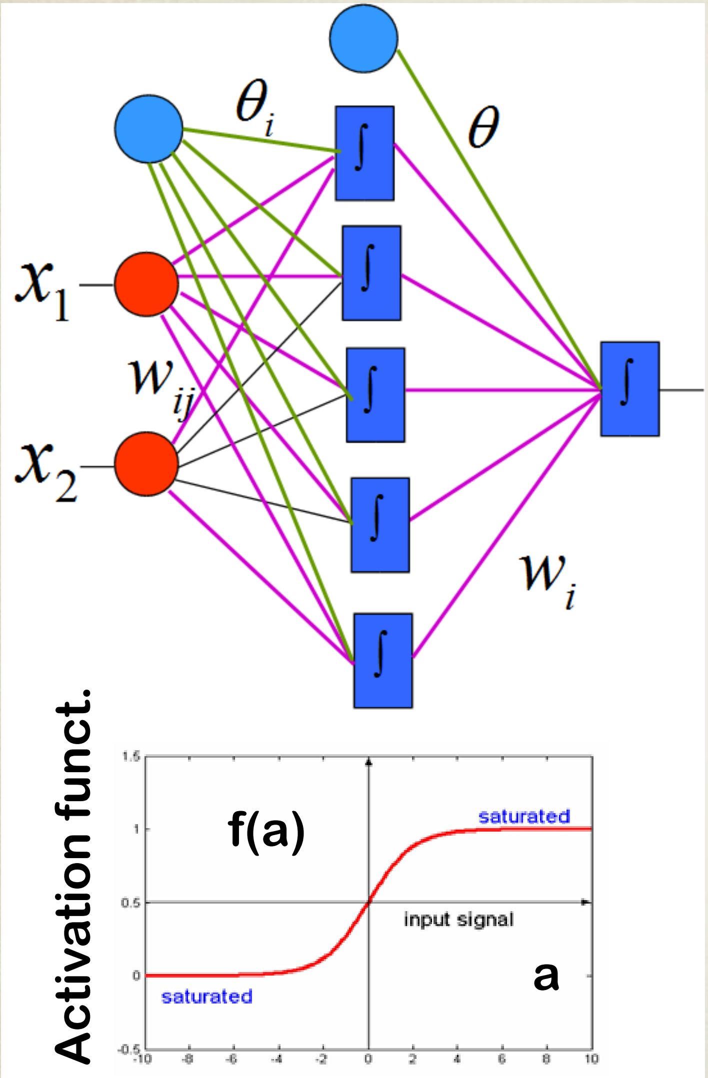
$$y_L(i_{event}) = \frac{\prod_{k \in \{variables\}} p_k^{signal}\left(x_k(i_{event})\right)}{\sum_{U \in \{species\}} \left(\prod_{k \in \{variables\}} p_k^{U}\left(x_k(i_{event})\right)\right)}$$

Species: signal, background types

- Optimal if the discriminants are independent (uncorrelated).
- This condition is rarely satisfied in practice ☹
- Widely used in signal significance analysis (1D)

Classifiers

Artificial Neural Network



Trained by (*x*, *t*) pairs

$$a_i = \sum_{j=1}^{2} w_{ij} x_j + \theta_i \quad \rightarrow f(a_i)$$

$$n(x,w) = f\left(\sum_{i=1}^{5} w_i f(a_i) + \theta\right)$$

Learning by Backpropagation:

$$\Delta w_i = -\alpha (n - t) \, \partial n / \partial w_i$$

learning speed

network output

teaching result

M. Wolter

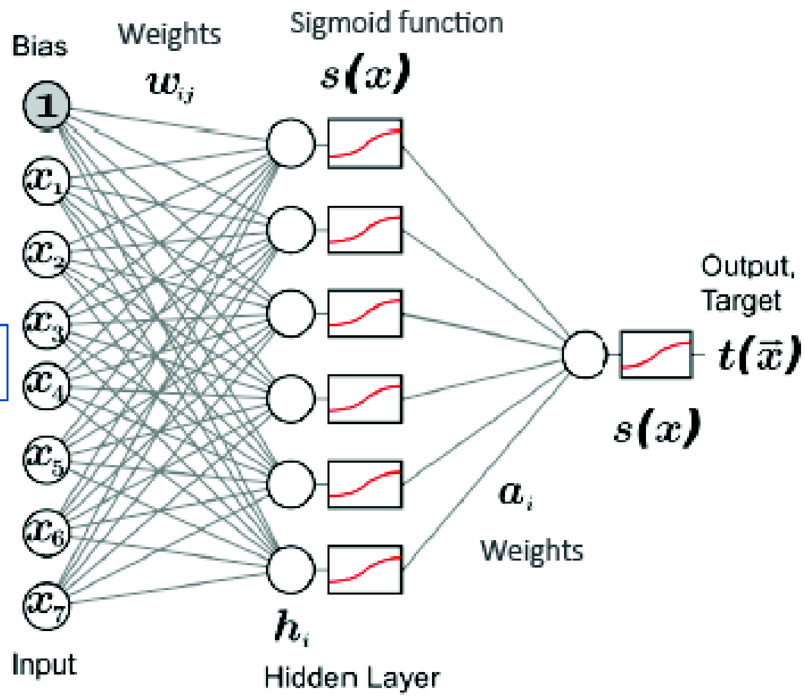
Classifiers

Artificial Neural Network – use case example

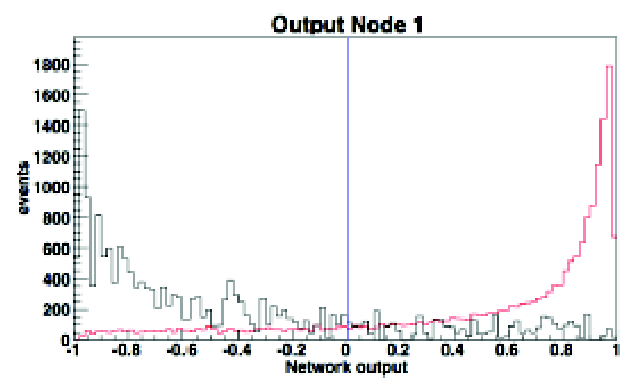
Neural Networks are trained in the 0-jet bin and 2-jet bin for m_H = 150 GeV and 180 GeV.

Training @
m_H = 150 GeV
m_H = 180 GeV

input variables



6 hidden nodes



Implementation with NeuroBayes

Search for Higgs bosons predicted in Two-Higgs-Doublet Models in the WW→lνlν decay channel

Classifiers Boosted Decision Tree (BDT)

Building a Tree:

Divide the training sample using the best separating variable.

Continue until you reach the minimal number of events or maximal purity

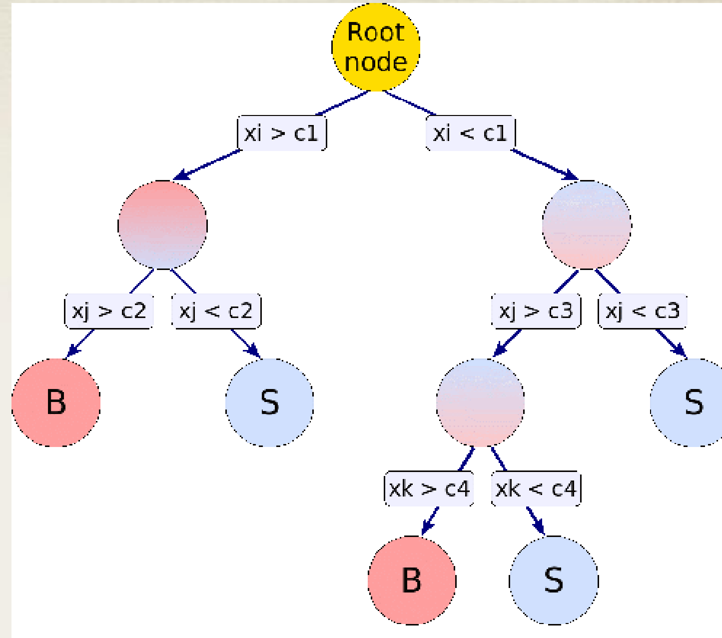
Boosting (AdaBoost):

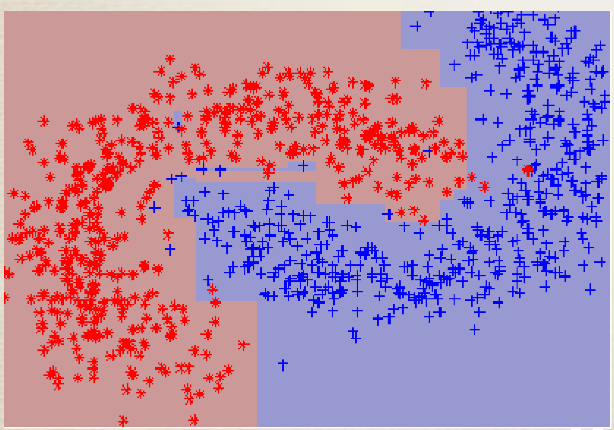
Increase the weights of the wrongly classified events.

Train a new classifier on the updated training sample.

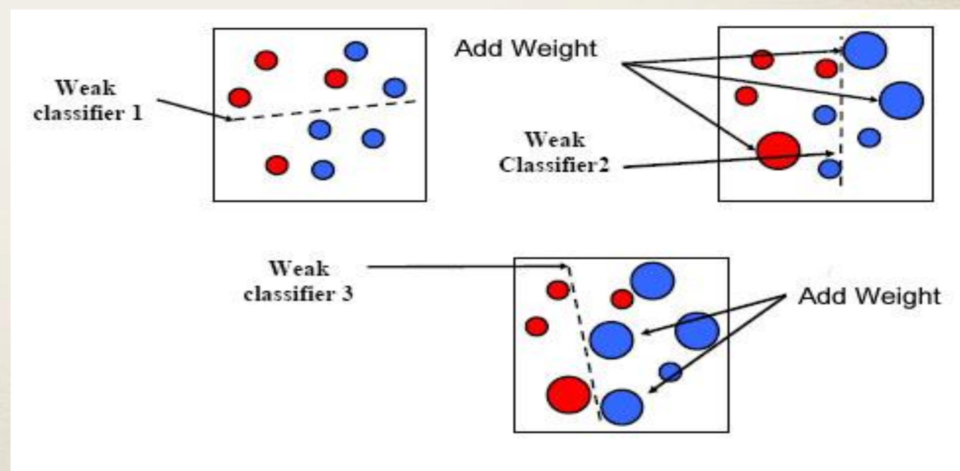
Repeat procedure => create multiple trees.

Classify data = voting of all trees (with appropriate weights)





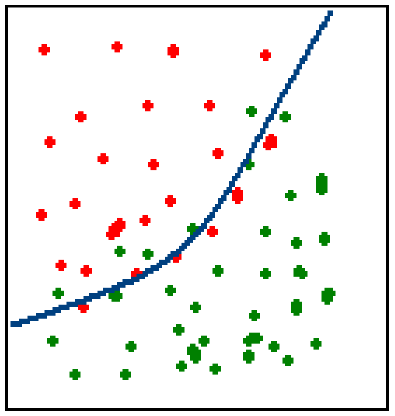
M. Wolter



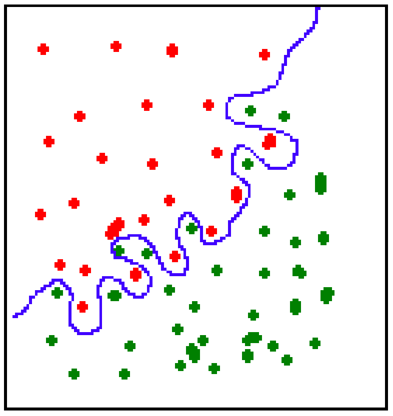
Classifiers

Beware of overtraining

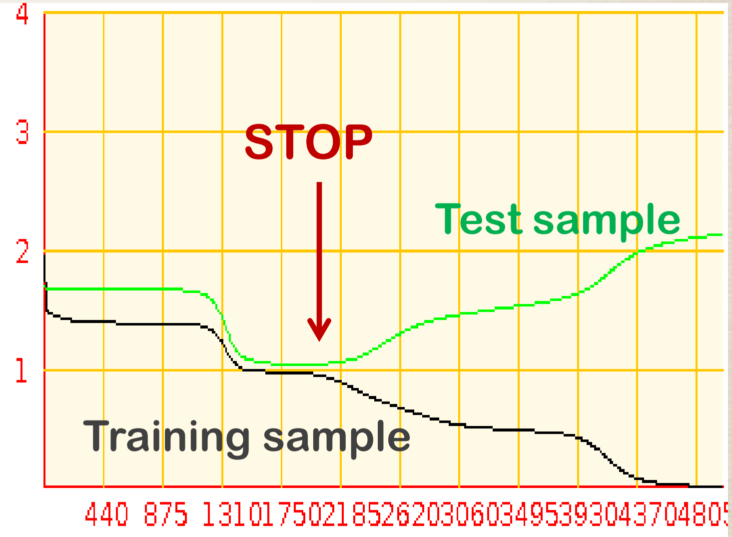
Overtraining is due to limited statistics of the training sample and e.g. inadequate structure of the NN or overgrown tree.
Generic feature of classifiers, although not all equally prone to overtraining.
BDT or Bayesian NN, are generally more resilient.



adequate



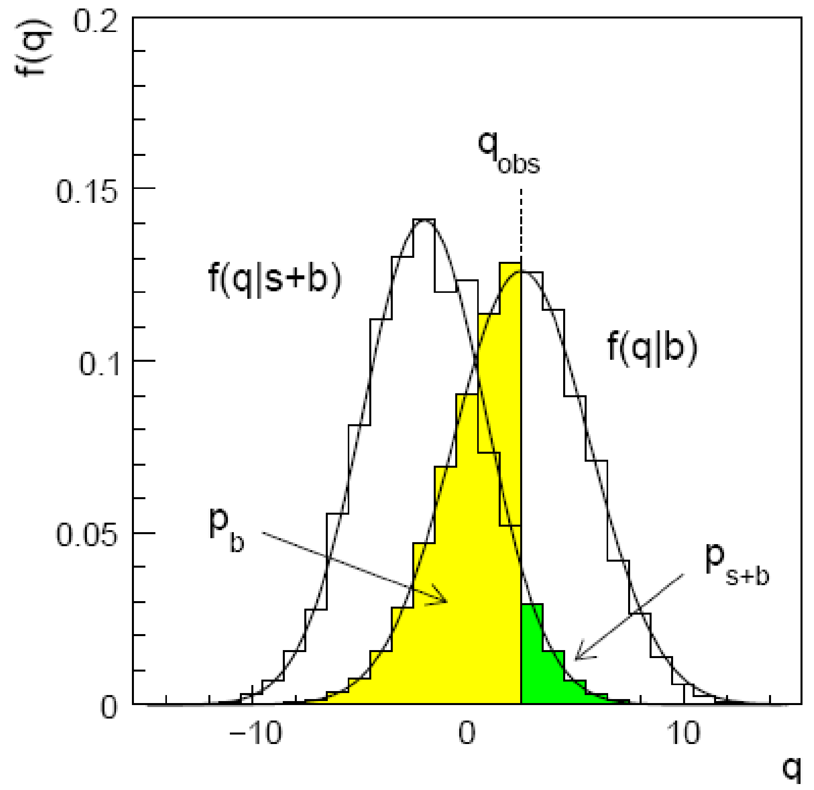
overtrained



Statistical interpretation

LHC is a discovery machine. Relevant questions:

i. Significance of observation: Are we seeing signal we search for?
ii. Confidence limit for exclusion: Can we exclude an observation with probablility e.g. ≥95%?



$$p_{s+b} = P(q \geq q_{\rm obs}|s+b) = \int_{q_{\rm obs}}^{\infty} f(q|s+b)\,dq$$

$$p_b = P(q \leq q_{\rm obs}|b) = \int_{-\infty}^{q_{\rm obs}} f(q|b)\,dq$$

CL_s method:

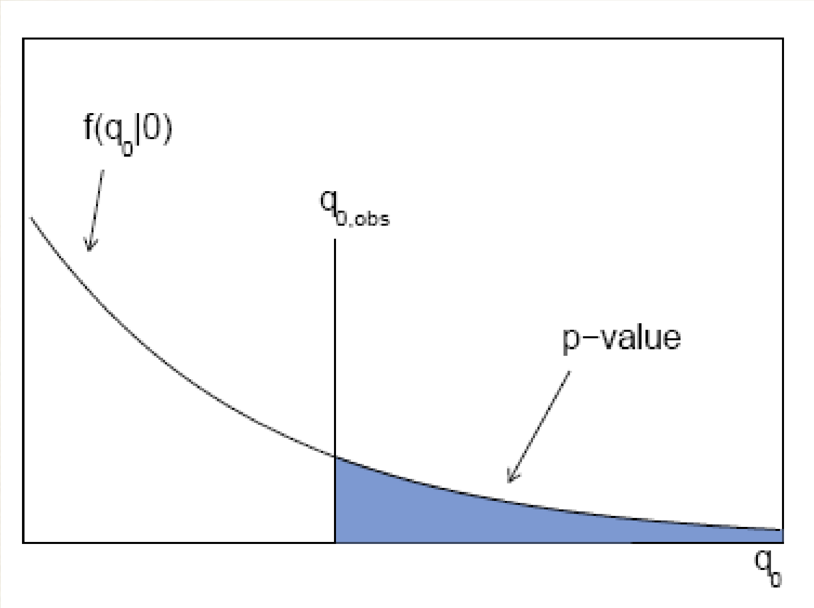
$$\frac{p_{S+B}}{1-p_B} < \alpha \left(\alpha = 0.05\right)$$

Statistical interpretation
example: Profile Likelihood

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

max L for specified μ

max L



$$q_\mu = \begin{cases} -2\ln\lambda(\mu) & \hat{\mu} \le \mu \\ 0 & \hat{\mu} > \mu \end{cases}$$

$$p_0 = \int_{q_{0,\mathrm{obs}}}^{\infty} f(q_0|0)\, dq_0$$

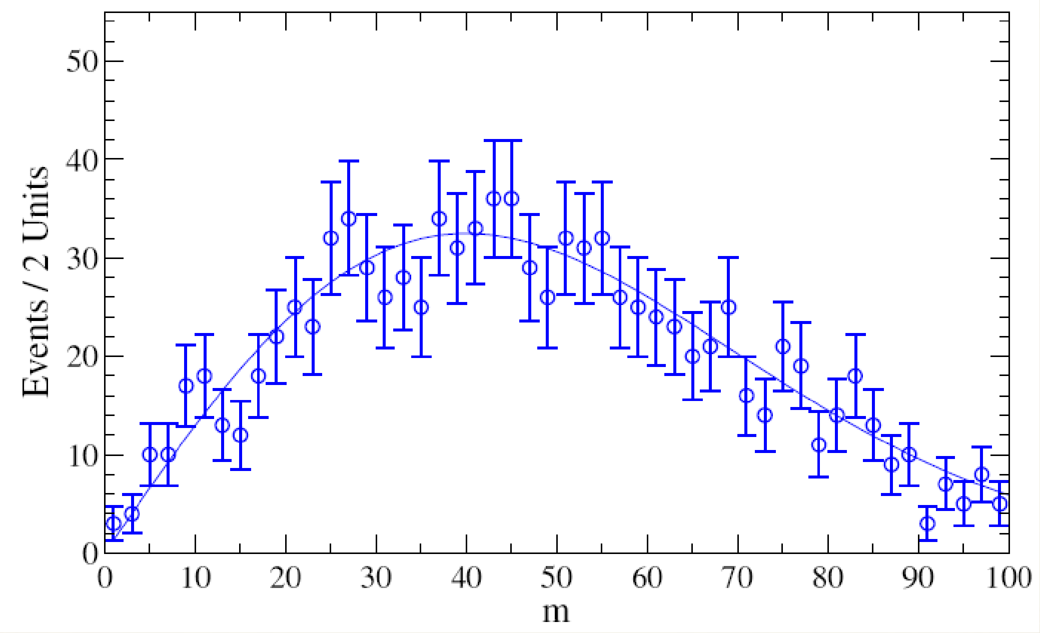
$$Z = \Phi^{-1}(1-p)$$

Significance in σ

G. Cowan

Statistical interpretation
example: Profile Likelihood

Typical example:
Search for a bump on top of the background

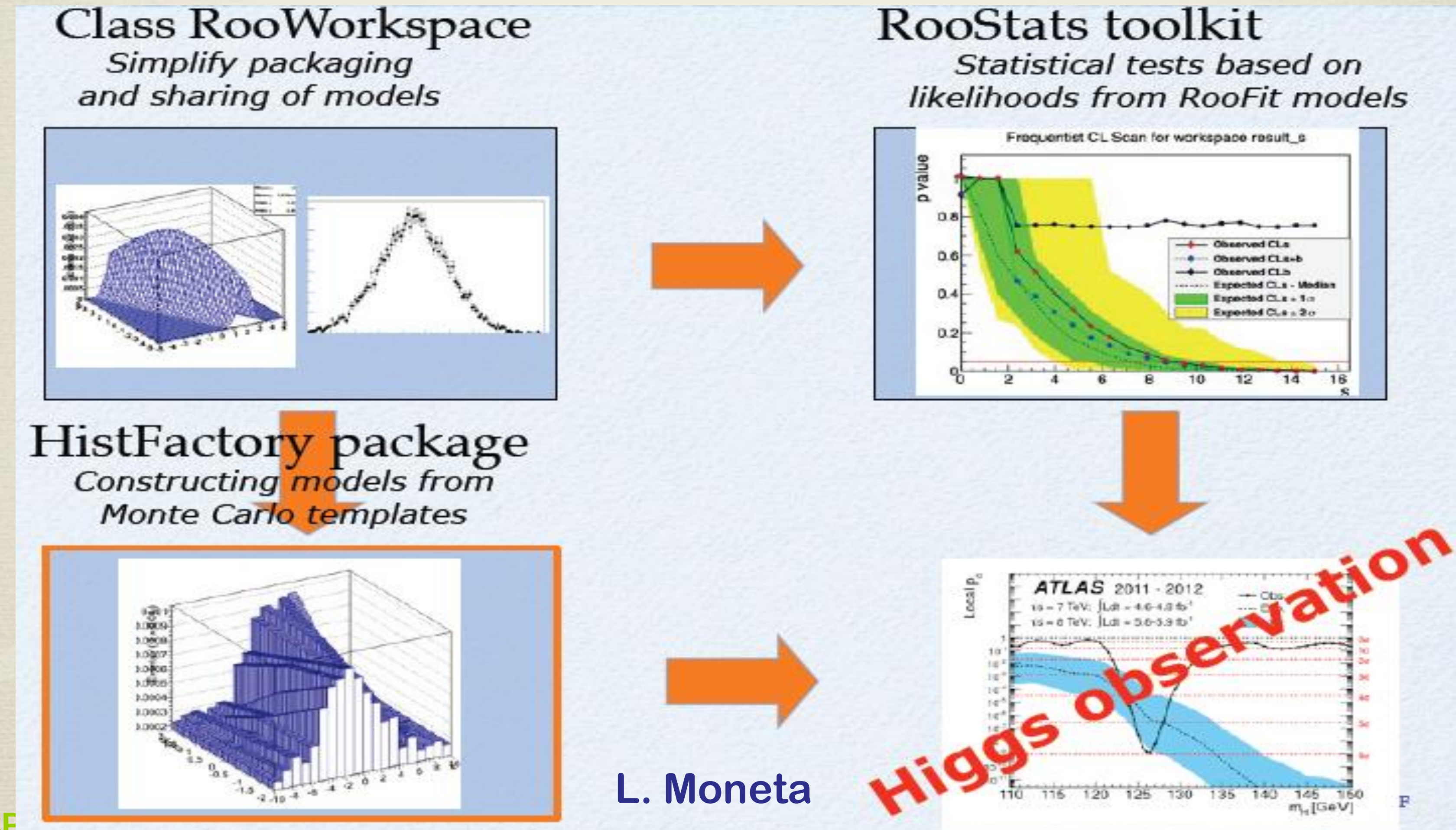


$$L(\mu,\theta)=\prod_{i=1}^{N}\frac{(\mu s_i+\theta f_{\mathrm{b},i})^{n_i}}{n_i!}e^{-(\mu s_i+\theta f_{\mathrm{b},i})}$$

G. Cowan

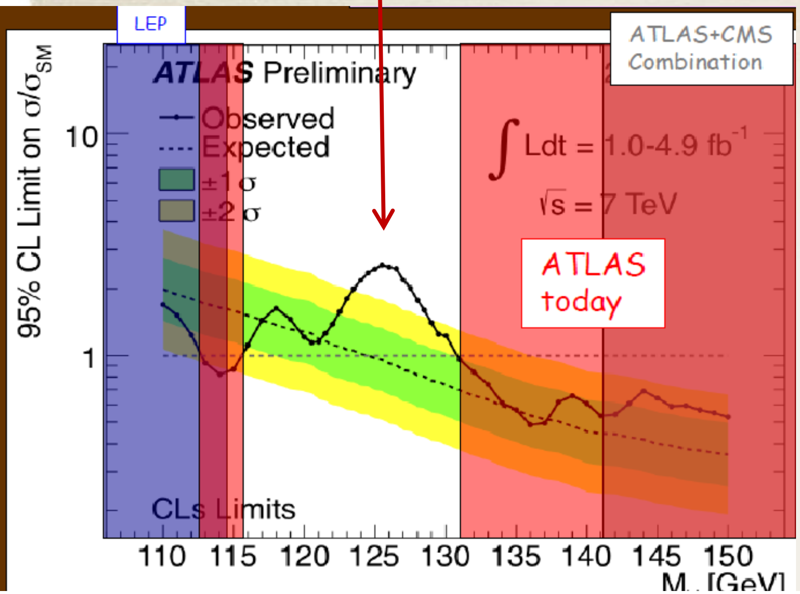
Statistical interpretation

Nothing particularly „modern" in the above.
What is really new is its recent implementation (RooStats, etc.) within the Root framework.



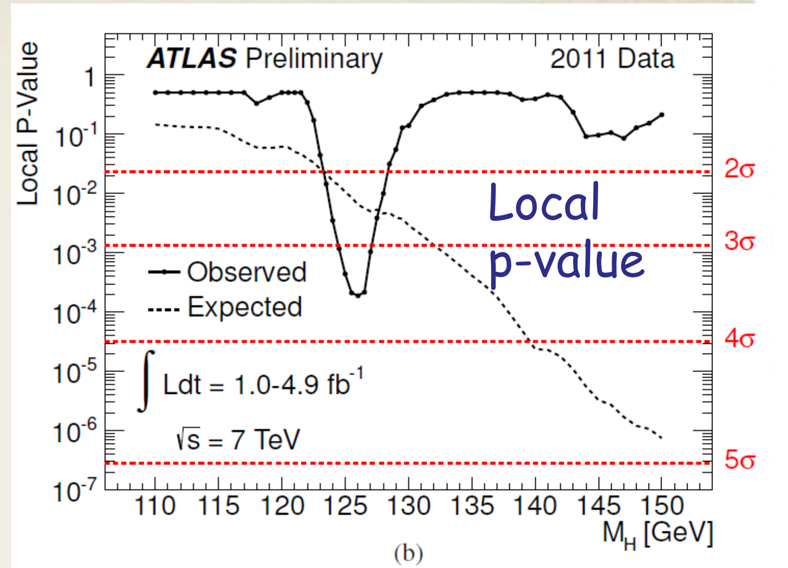
Higgs @ LHC December 2011 !!!

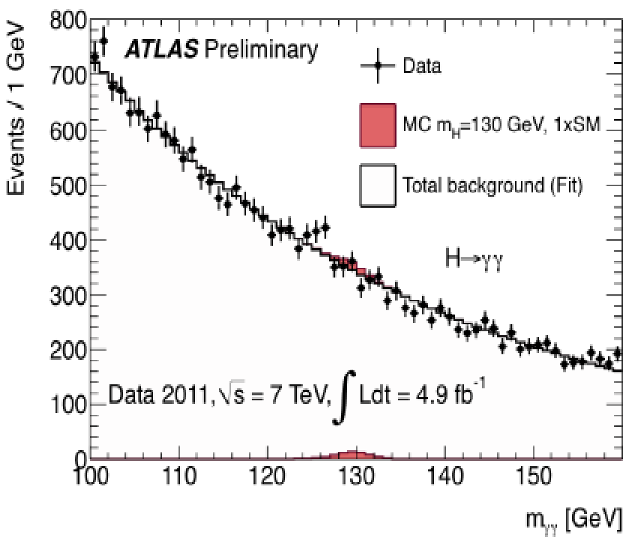
excess

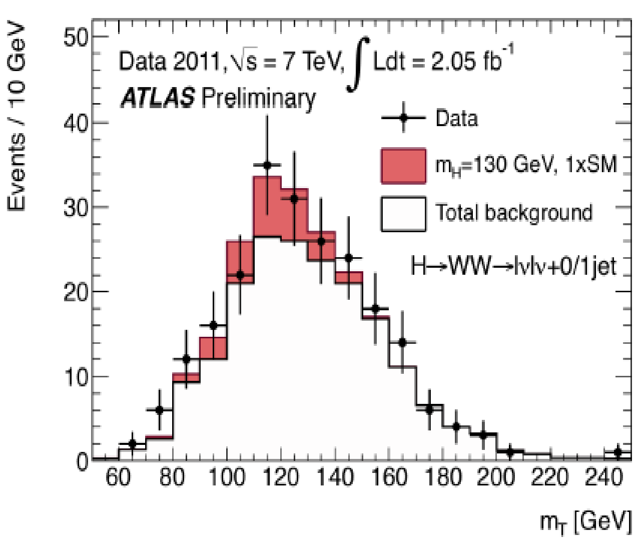


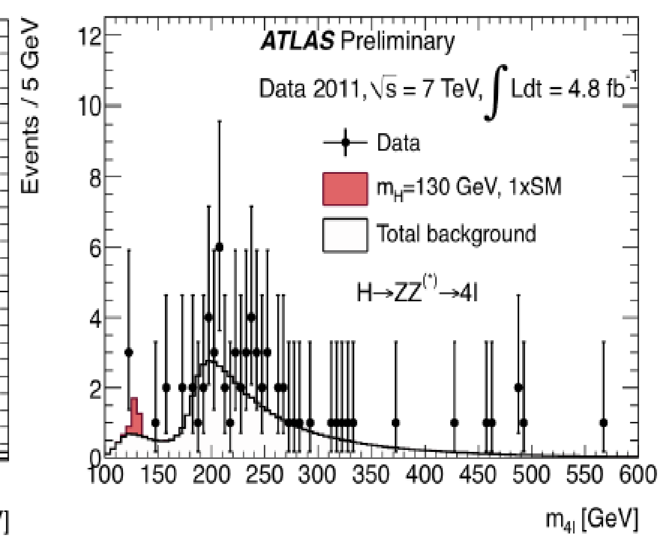
Can we claim discovery yet?

M. Wolter









Statistical interpretation

Look Elsewhere Effect

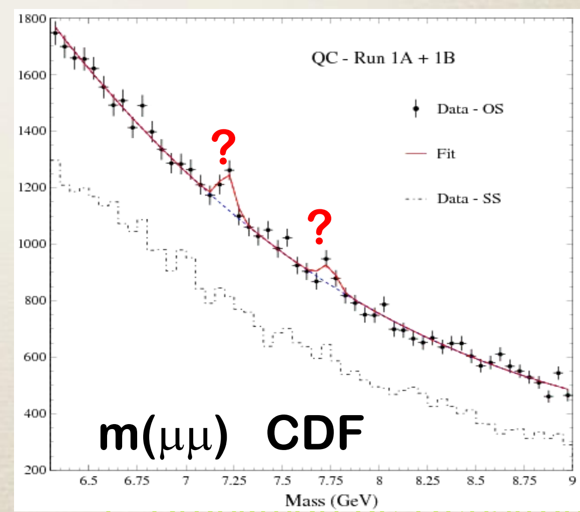
Instead of the local probability of background fluctuation we need to consider probability of a fluctuation at any point of the considered phase-space.

ATLAS (2011):

3.6 σ -> 2.3 σ

1.9×10^{-4} -> 0.01

M. Wolter



Trial factors for the look elsewhere effect in high energy physics
Eilam Gross, Ofer Vitells
Eur.Phys.J.C70:525-530,2010

Summary

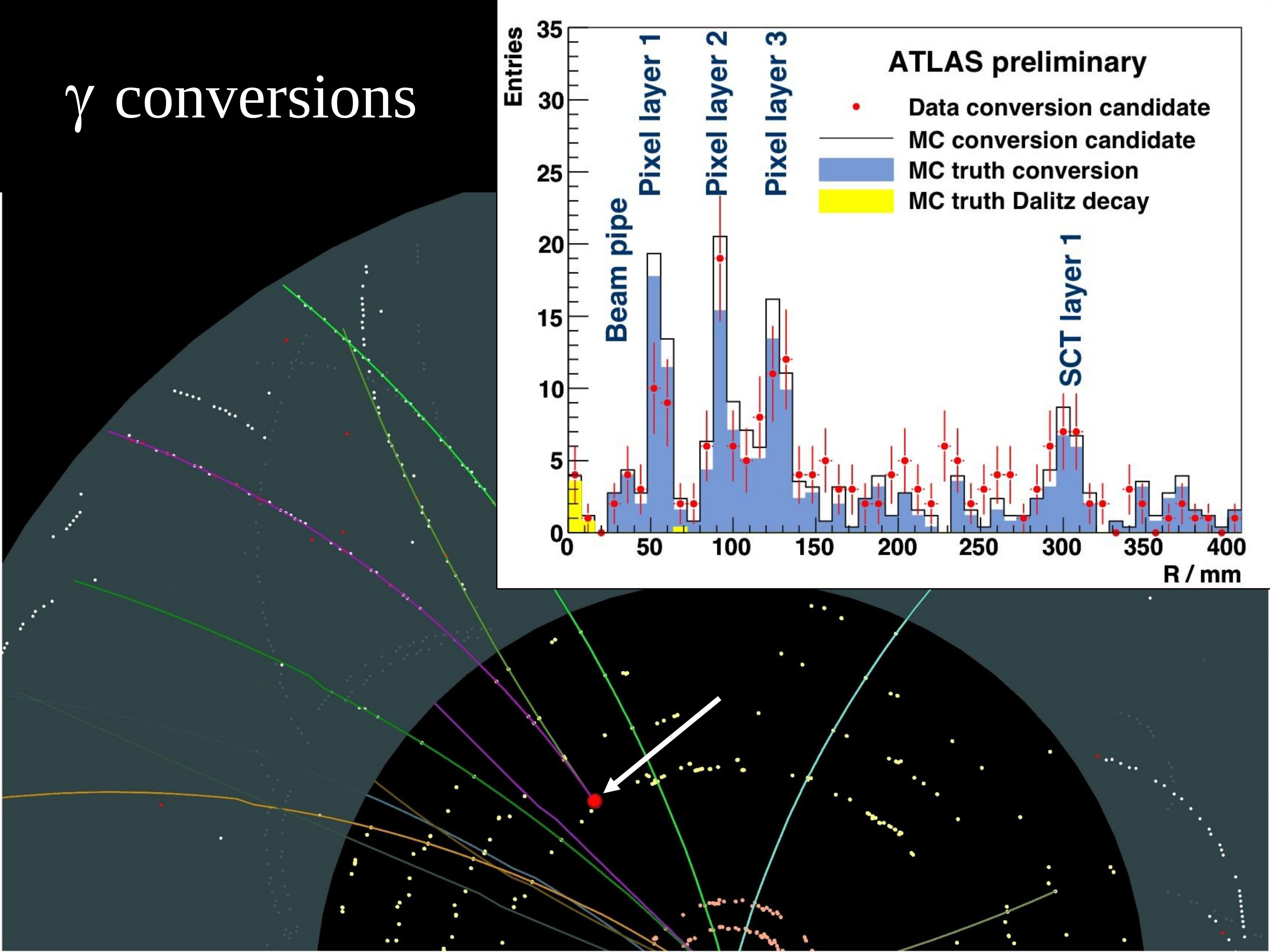
- ❖ LHC experimental setup as complex as it gets.
- ❖ Name of the game: get highest sensitivity to signal while keeping systematic uncertainties low.
- ❖ Data taking and handling possible due to phenomenal increase of computing resources. Thank you Gordon! ☺

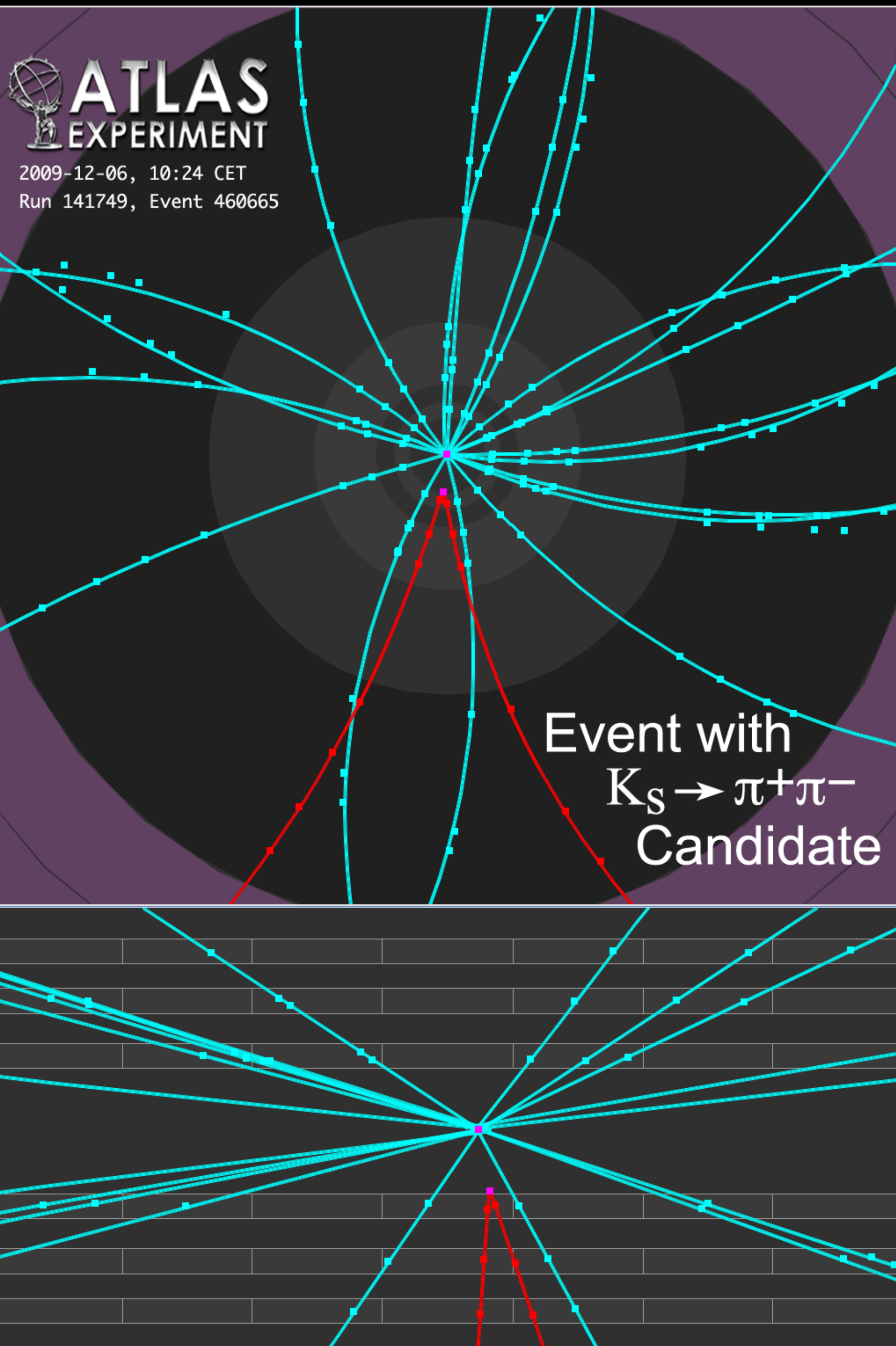
- ❖ Experimental & analysis methods not necessarily new, but promoted to unprecedented level to match up the cutting-edge research needs of the LHC.
- ❖ Nowedays, we are often limited only by our imagination!

Thank you !

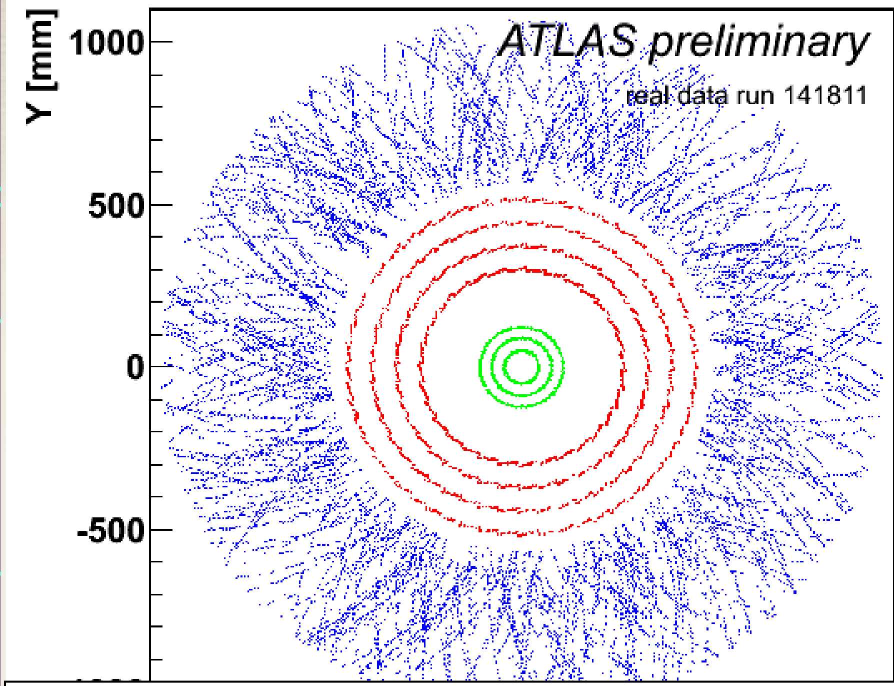
BACKUP SLIDES

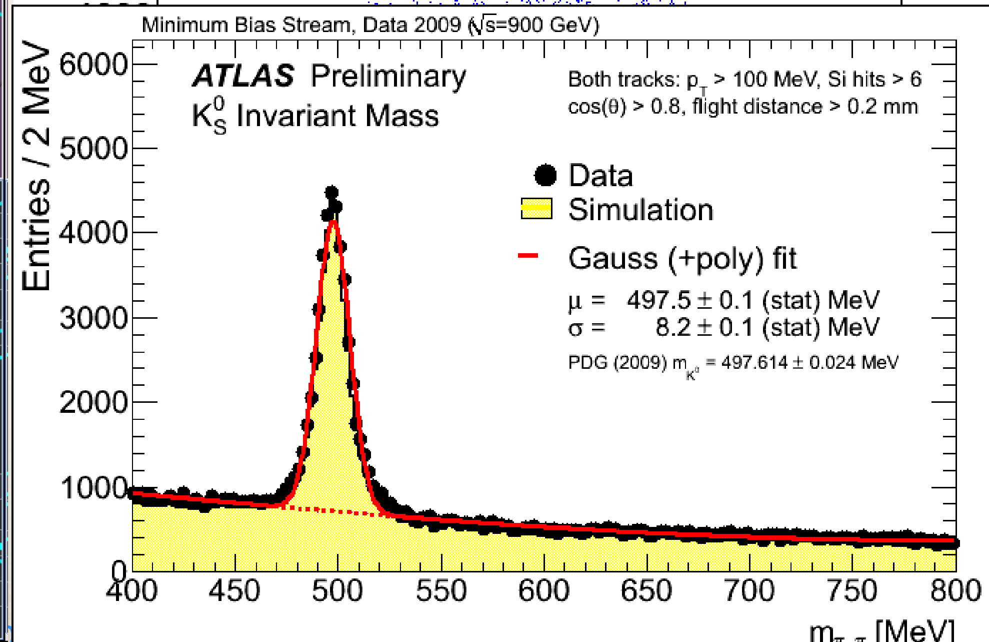
γ conversions





Scatter Plot of Hits on Tracks

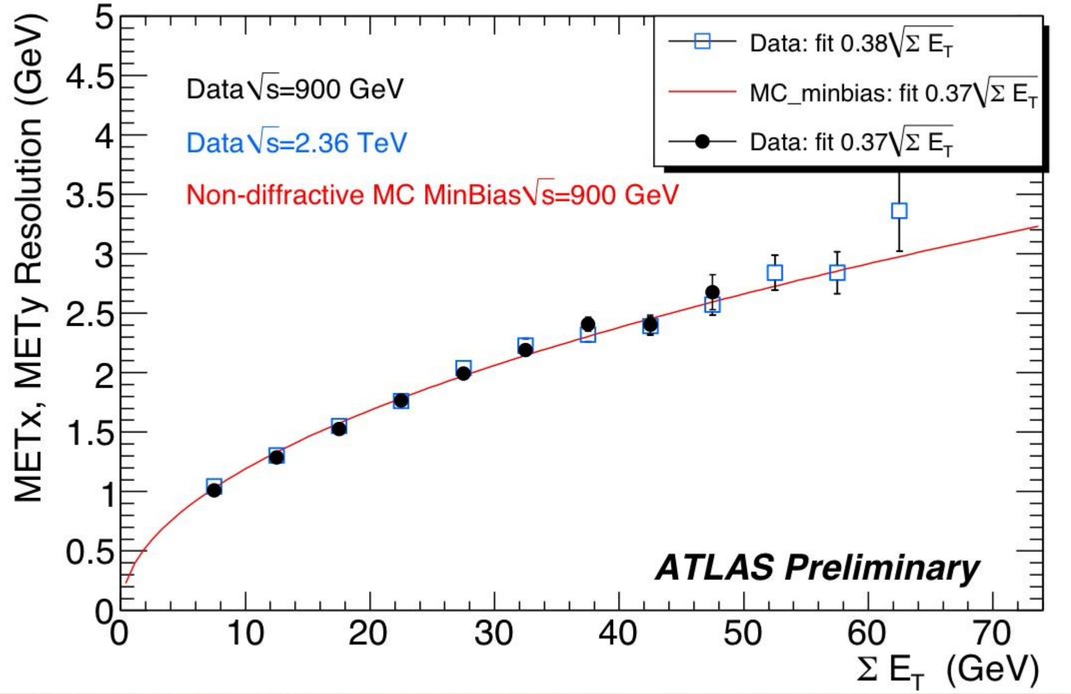


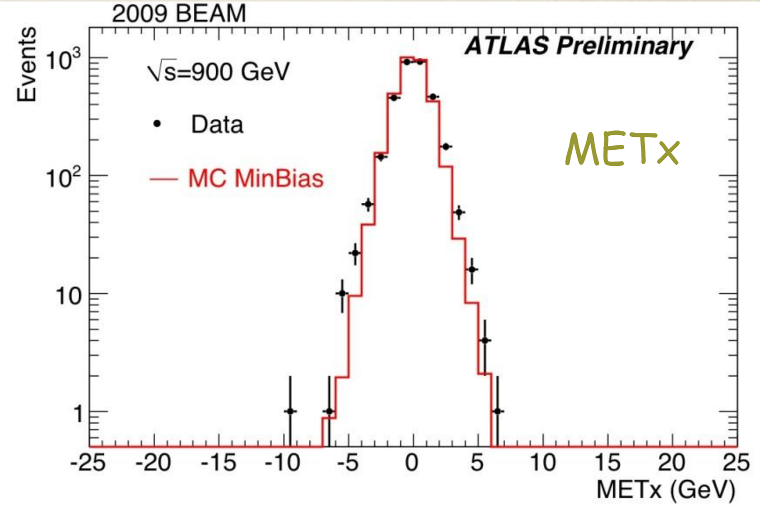


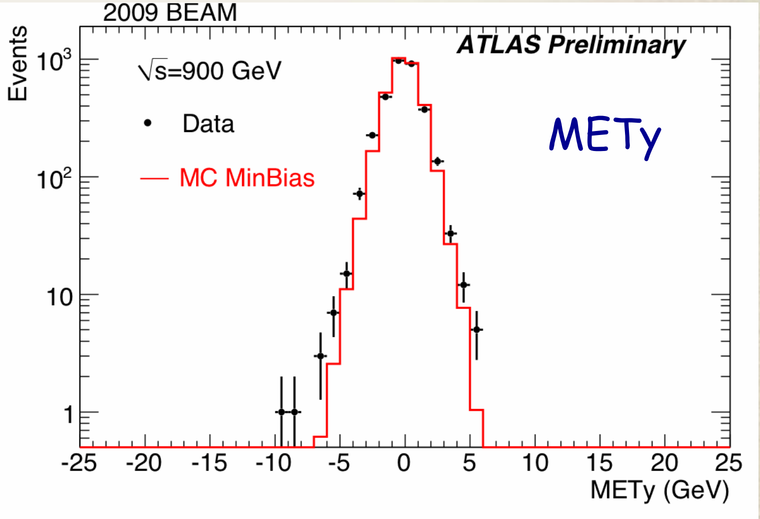
http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html

Missing transverse energy
(essential for SUSY searches!)

■ Sensitive to calorimeter performance (noise, coherent noise, dead cells, mis-calibrations, cracks, etc.) and backgrounds from cosmics, beams, …
■ Measurement over full calorimeter coverage (360^0 in φ, |η| < 5, ~ 200000 cells)

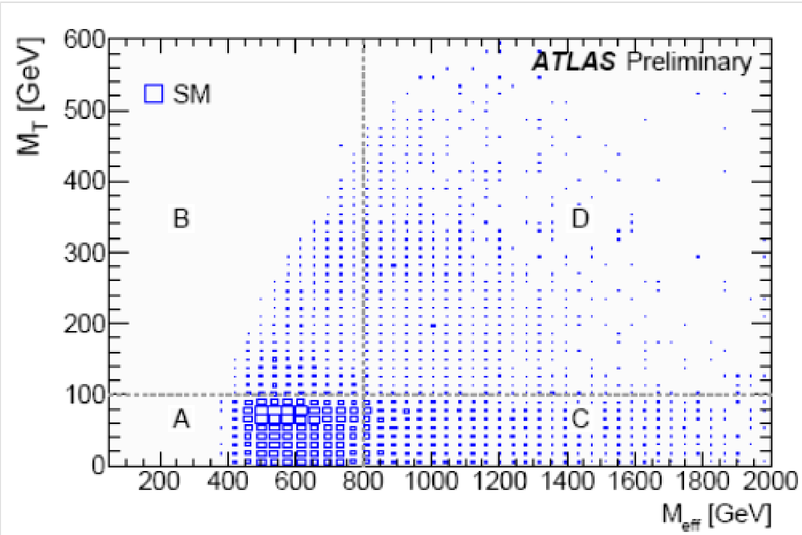


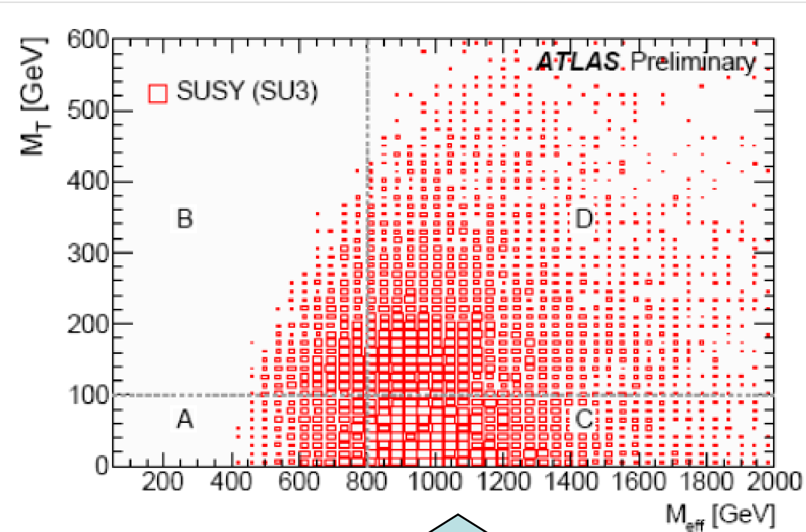




The "tiles method" - yet another way

- Variables in the SM backgrounds may exhibit correlations
- SM shapes (fractions) must be known (eg from MC)





$$\bar{N}_A = f_A^{SM}\bar{N}^{SM} + f_A^S\bar{N}^S\,, \qquad \bar{N}_B = f_B^{SM}\bar{N}^{SM} + f_B^S\bar{N}^S\,,$$

$$\bar{N}_C = f_C^{SM}\bar{N}^{SM} + f_C^S\bar{N}^S\,, \qquad \bar{N}_D = f_D^{SM}\bar{N}^{SM} + f_D^S\bar{N}^S\,,$$

$f_A^{SM}, \ldots, f_D^{SM}$ Taken from MC

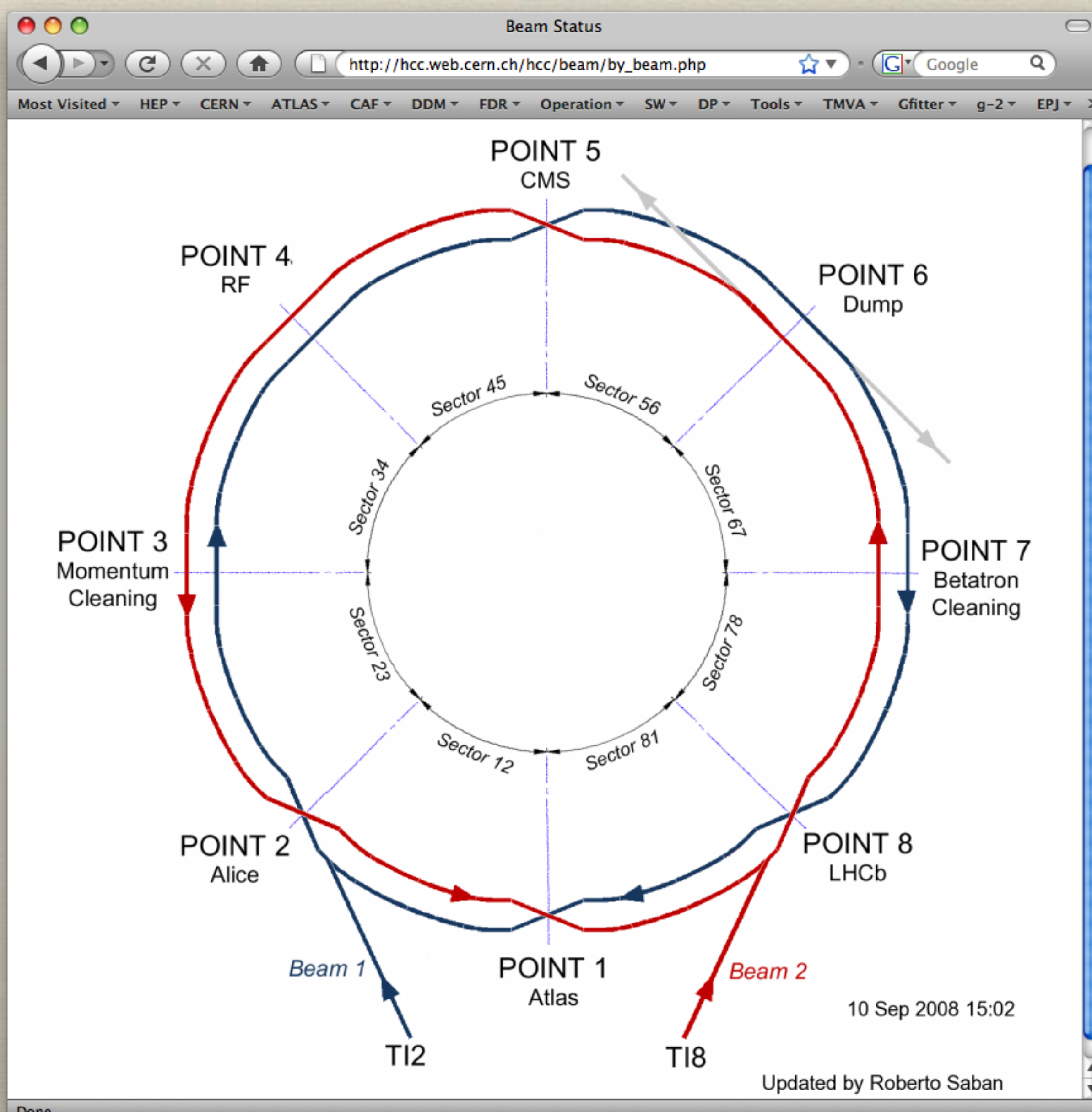
$$f_A^S = (1 - f_{M_{\rm eff}}^S)(1 - f_{M_T}^S)\,, \qquad f_B^S = (1 - f_{M_{\rm eff}}^S) f_{M_T}^S\,,$$

$$f_C^S = f_{M_{\rm eff}}^S(1 - f_{M_T}^S)\,, \qquad f_D^S = f_{M_{\rm eff}}^S f_{M_T}^S\,,$$

The two variables need to be independent in the signal!

The system is solvable without iterations!

ATL-PHYS-PUB-2009-077


Beam Status
http://hcc.web.cern.ch/hcc/beam/by_beam.php
Google
Most Visited
HEP
CERN
ATLAS
CAF
DDM
FDR
Operation
SW
DP
Tools
TMVA
Gfitter
g-2
EPJ
POINT 5
CMS
POINT 4
RF
POINT 6
Dump
Sector 45
Sector 56
Sector 34
Sector 67
POINT 3
Momentum
Cleaning
POINT 7
Betatron
Cleaning
Sector 23
Sector 78
Sector 12
Sector 81
POINT 2
Alice
POINT 8
LHCb
Beam 1
POINT 1
Atlas
Beam 2
10 Sep 2008 15:02
TI2
TI8
Updated by Roberto Saban
Done