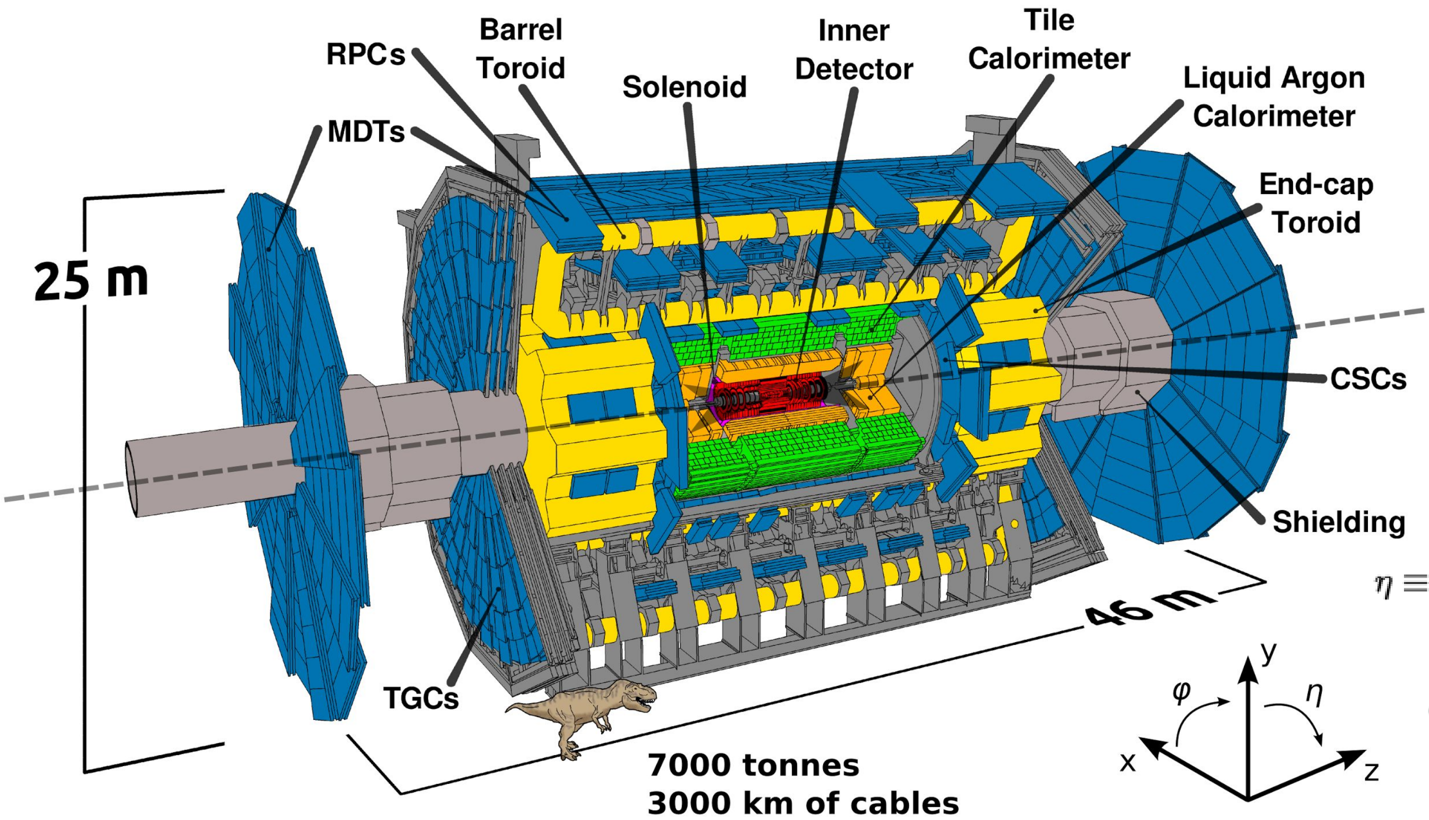
Jets, b-jets, taus and missing transverse momentum

Noel Dawe

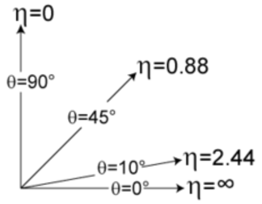
CoEPP Annual Workshop
20 Feb, 2017
Adelaide

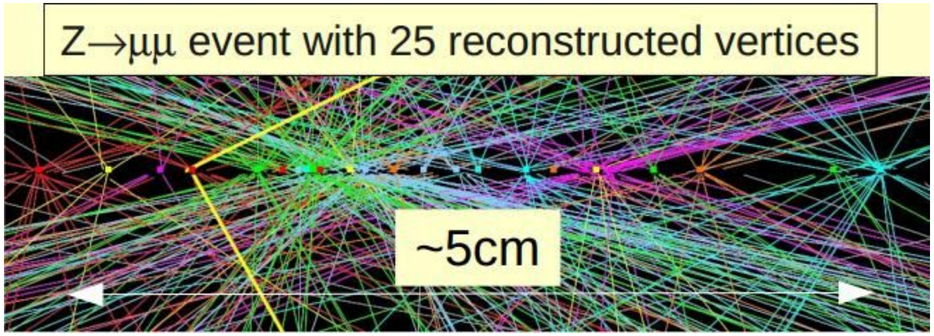
A few things first...

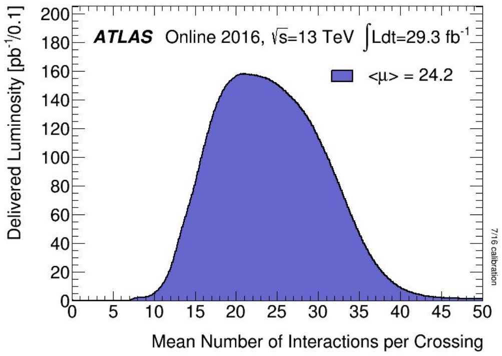


7000 tonnes
3000 km of cables

$$\eta \equiv -\ln\left[\tan\left(\frac{\theta}{2}\right)\right],$$



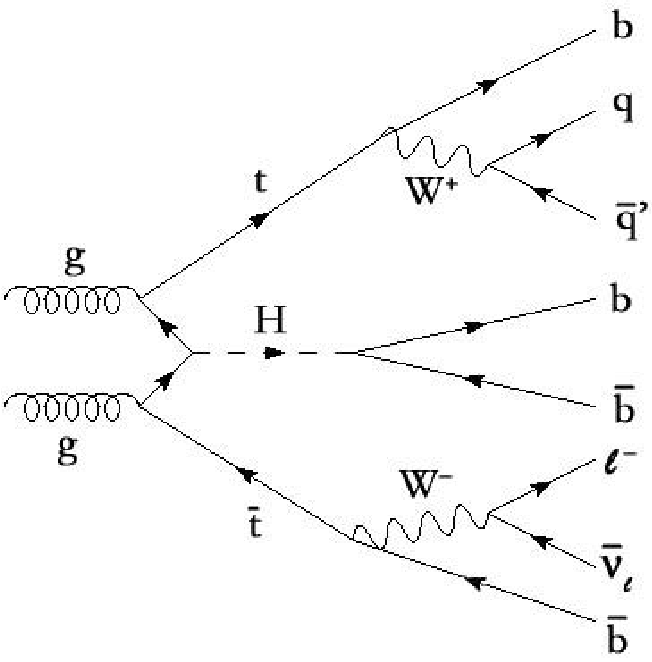


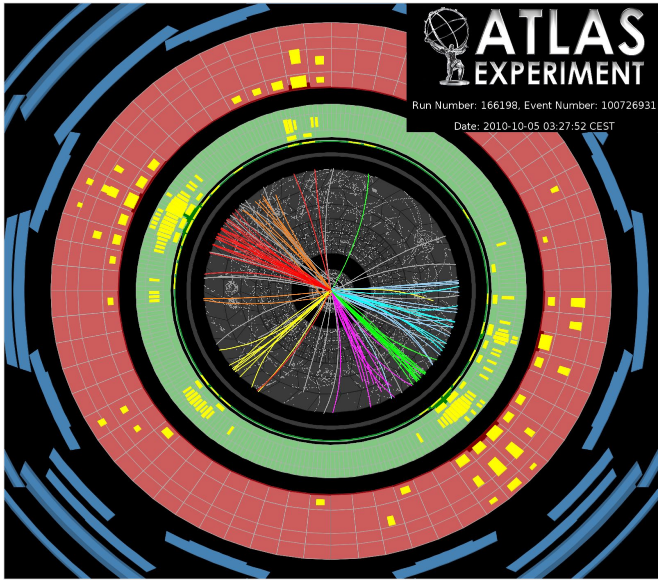


Jets…

Theorists calculate final states
with quarks and gluons

Experiments measure hadrons…





Need to understand transition from partons to hadrons (fragmentation and hadronization) and reverse-engineer it! Hadrons are colour singlets while quarks/gluons are colour triplets/octets so this process can never be perfect.

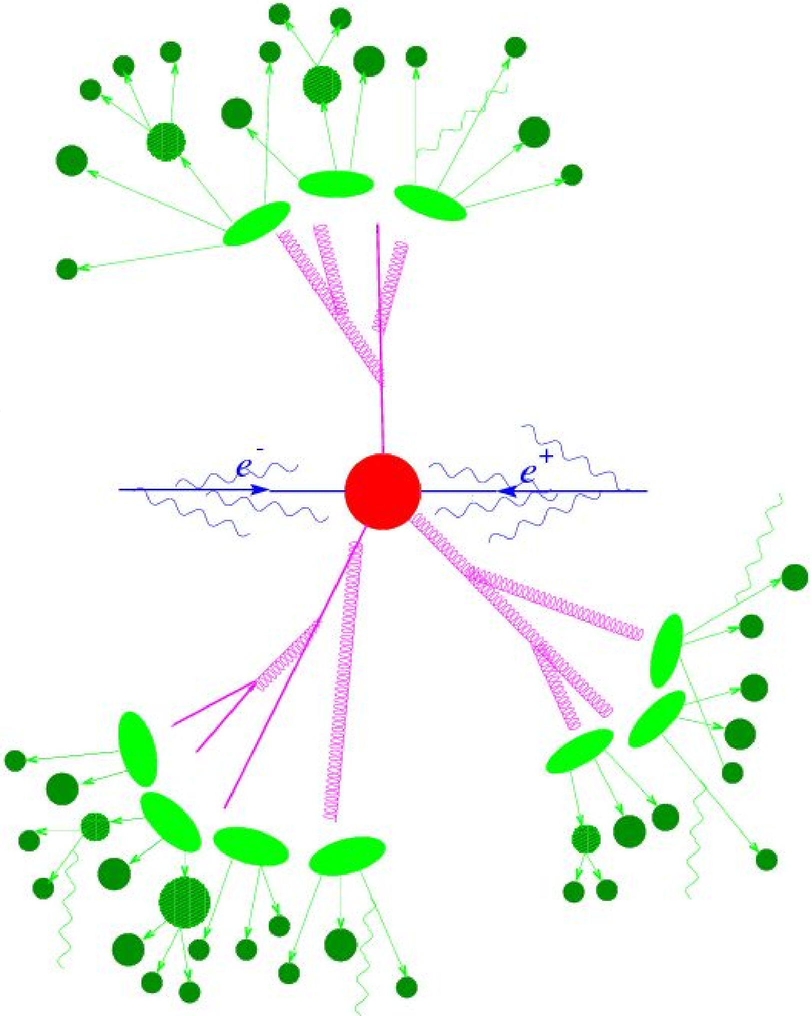
What are jets?

Quarks and gluons (partons) are essentially free particles at short distances ($<10^{-2}$ fm)

When accelerated, energy builds up in a gluon field at a rate of ~ 1 GeV / fm and fragments into new quark-antiquark pairs at ~ 1 fm.

Quarks radiate gluons and gluons split into quark-antiquark pairs. Children nearly collinear with parent partons (collimated showers).

Proceeds down to a low energy (~ 1 GeV) where **hadrons are formed.**



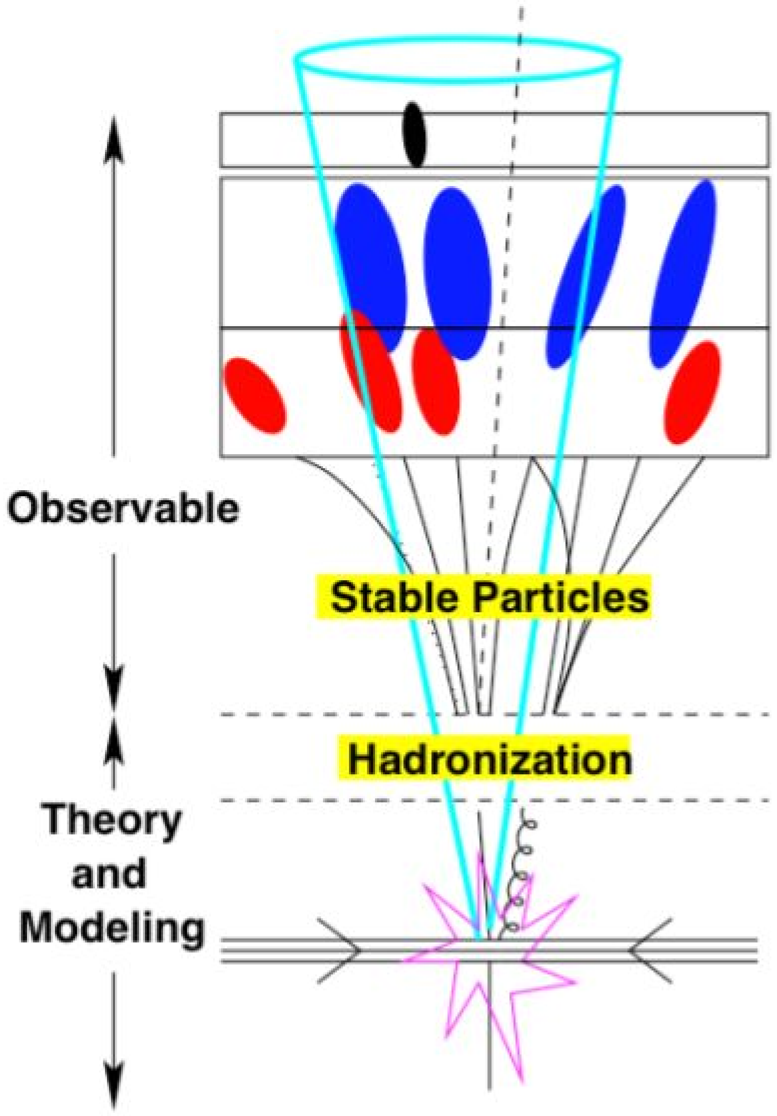
What are jets?

Jet finding is the approximate attempt to reverse-engineer the quantum mechanical processes of this fragmentation and hadronization.

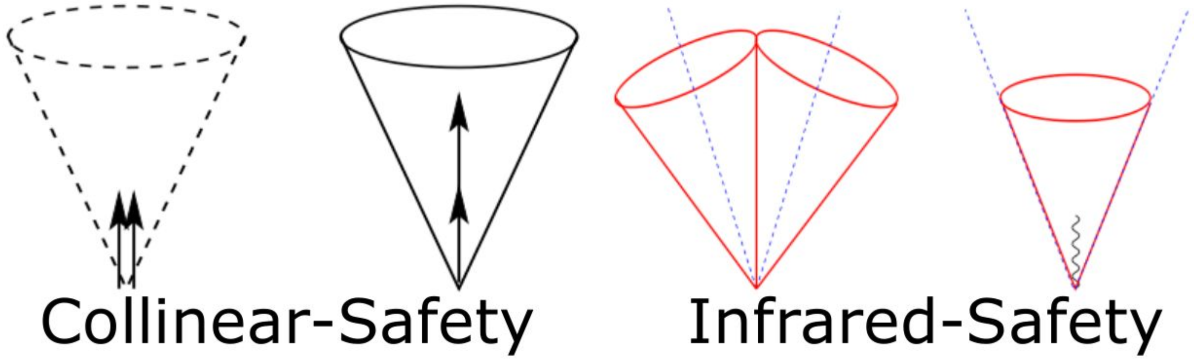
Jets are observable "objects" that can relate experimental observations to theory predictions formulated in terms of quarks and gluons.

Reconstructed at detector level from **clusters of calorimeter cells or tracks.**

Reconstructing jets is not a unique procedure!
There are several different approaches.



Jet Algorithm Requirements



Collinear splitting shouldn't change jets!

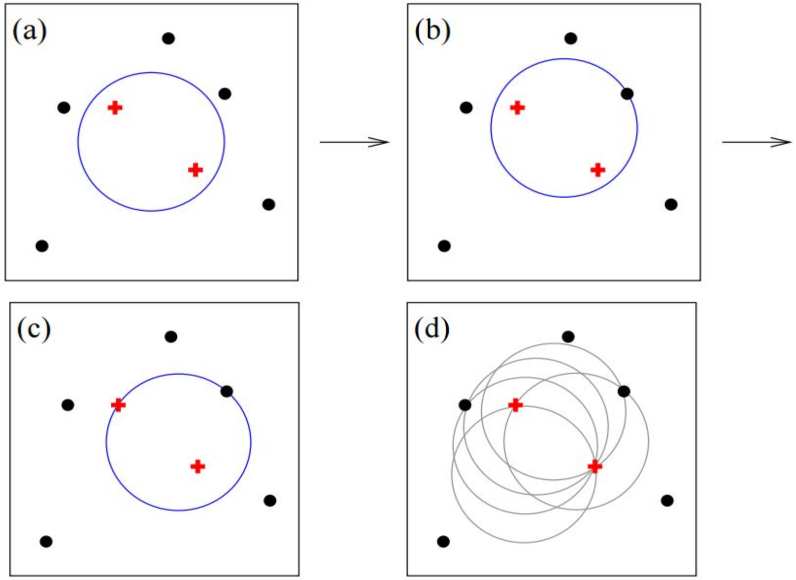
Soft emission shouldn't change jets!

- Identical procedure at parton and hadron level to compare experimental measurements to theory calculations. Defined at any order of perturbation theory.
- Minimal sensitivity to hadronization, underlying event, and detector pile-up
  These effects aren't modelled very well
- Applicable at detector level and easy to calibrate with well-defined jet area
- Not computationally expensive, especially in a high multiplicity environment

Two mainstream approaches:

Cone Algorithms

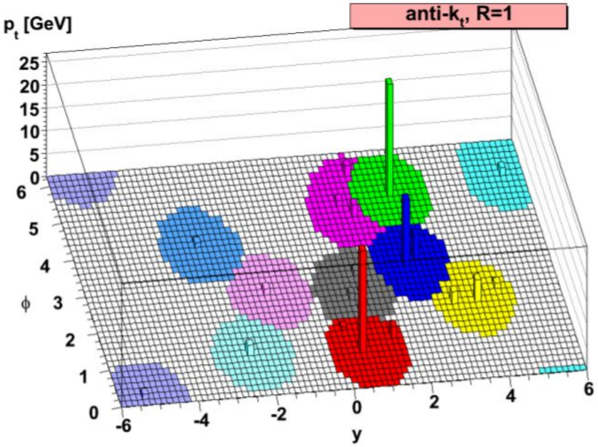
- Midpoint Cone, Iterative Cone, SISCone
- Typically not infrared or collinear safe
  - *Split-merge step (IR unsafe)*
  - *Iterate over seeds (collinear unsafe)*
- Complex, several non-physical parameters
- Previously favoured at hadron colliders
- Strongly disfavoured by theorists



Sequential Clustering Algorithms

- kT, Cambridge/Aachen, Anti-kT
- Infrared and collinear safe
- Clean and simple algorithms
- Favoured by theorists
- Previously not widely used at hadron colliders due to computational performance and nontrivial access to jet area.

  No longer issues!



Sequential Clustering Algorithms

Use the following distance measure between two constituents ***i*** and ***j***

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \qquad \Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$$

And the distance between constituent ***i*** and the beam:

$$d_{iB} = k_{ti}^{2p}$$

Isn't this algorithm O(N^3)?

Compute all distances d_{ij} and d_{iB} and find the smallest

If d_{ij} is the smallest, then combine consituents i and j

If d_{iB} is the smallest, then call i a jet and remove it from list of constituents

Repeat until list of remaining constituents is empty

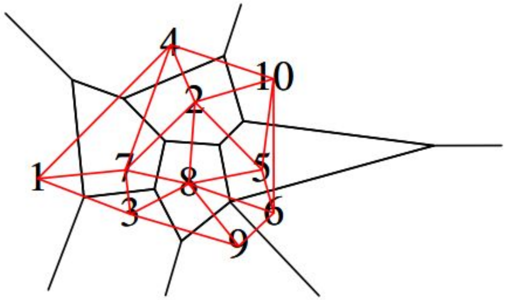
Parameter R scales d_{ij} relative to d_{iB} such that all final jets are at least separated by R

Parameter p sets the trade-off between momentum and spatial differences

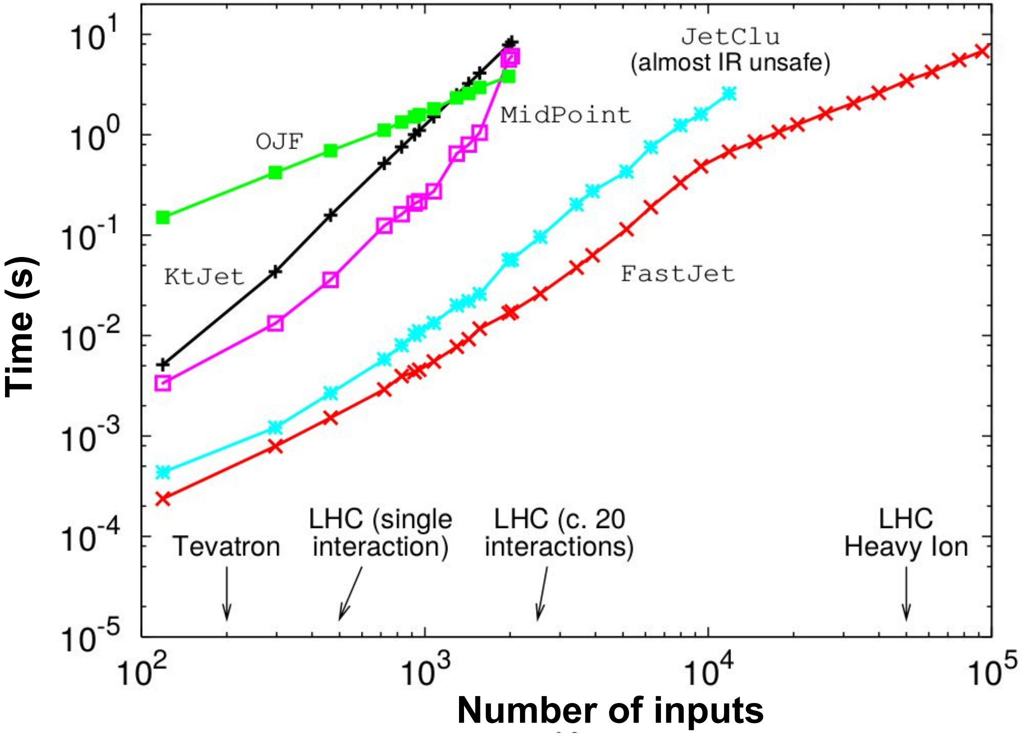
p=1 gives k_t , p=0 gives Cambridge/Aachen, and p=-1 gives anti-k_t

FastJet

- C++ library implementing all widely used jet algorithms
- Huge performance improvement for sequential clustering over previous implementations O(N ln(N))

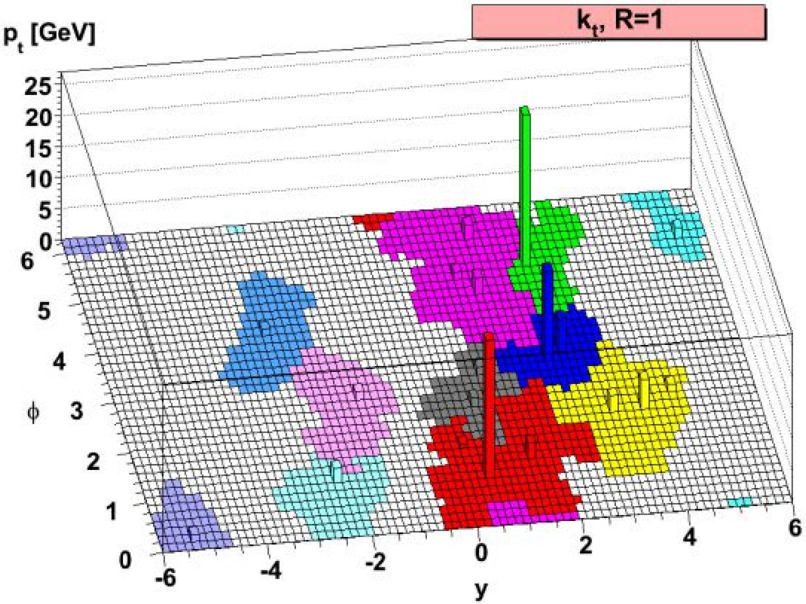


- Introduced concept of jet area, important in addressing pileup and underlying event contributions

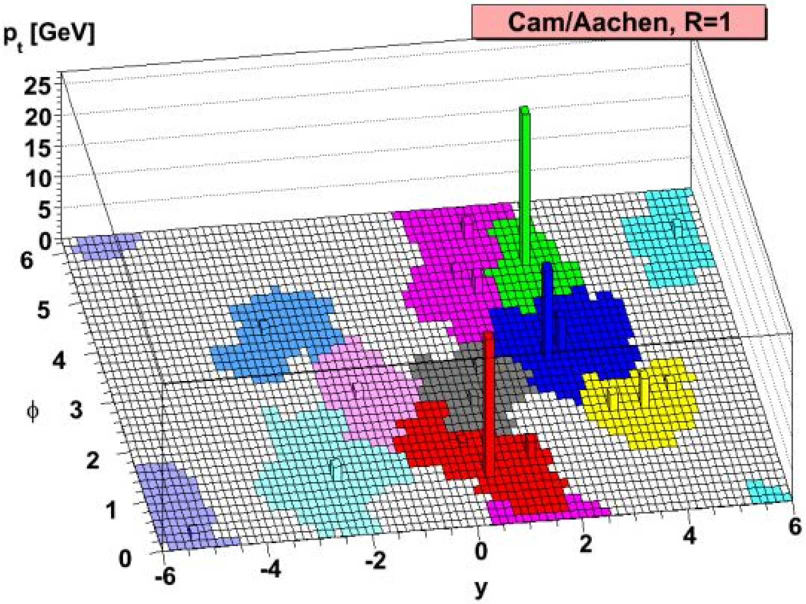


http://fastjet.fr/

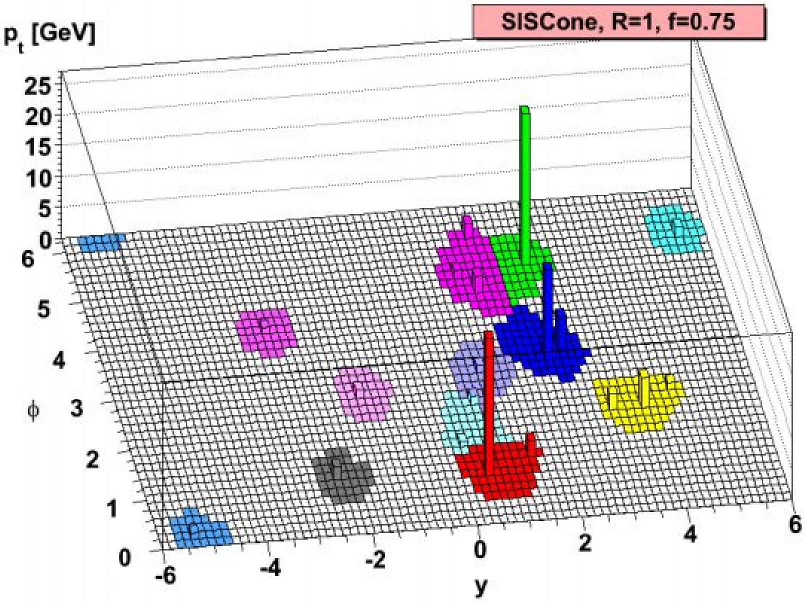
https://arxiv.org/abs/hep-ph/0512210



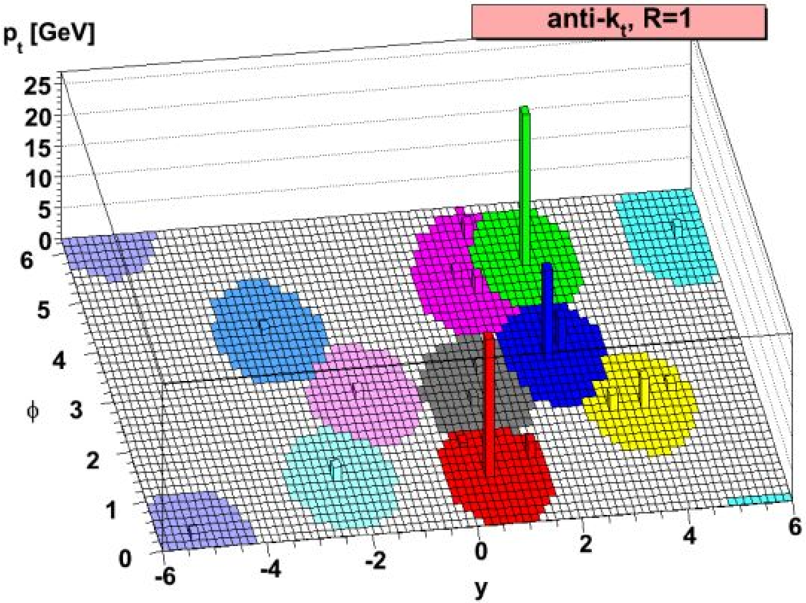
k$_t$, R=1



Cam/Aachen, R=1



SISCone, R=1, f=0.75



anti-k$_t$, R=1

https://arxiv.org/abs/0802.1189

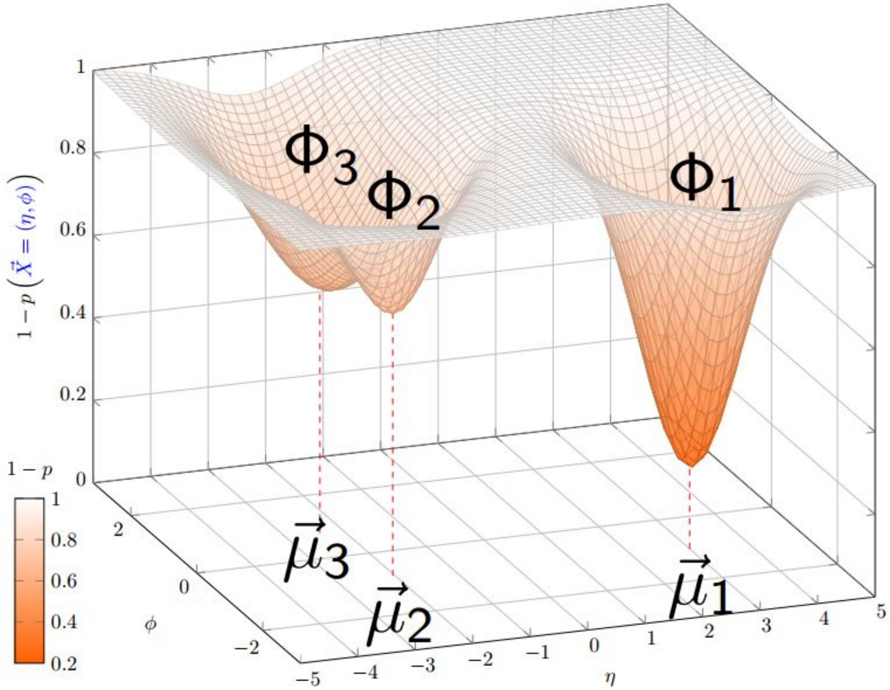
The Future: Fuzzy Jets?

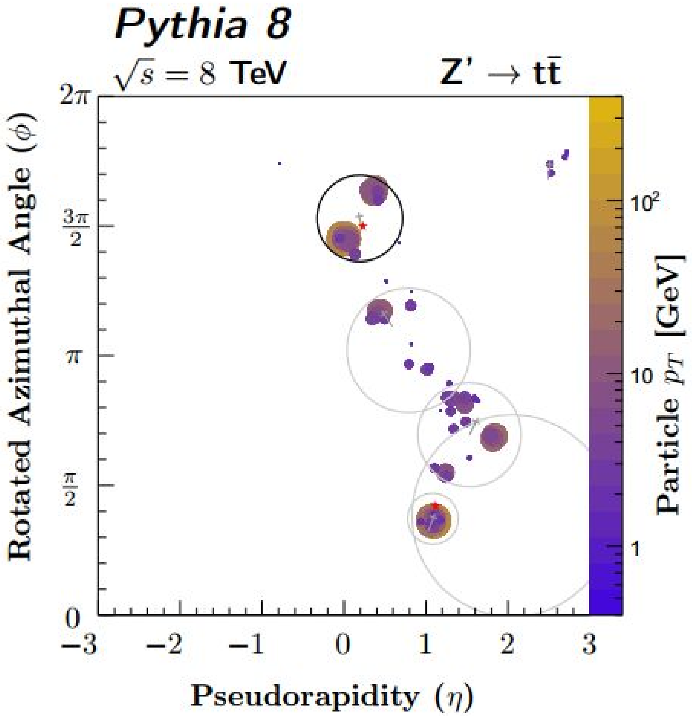
View jet clustering as an unsupervised machine learning task

Gaussian mixture models…

Clustered using maximum likelihood

Dynamically determine jet size



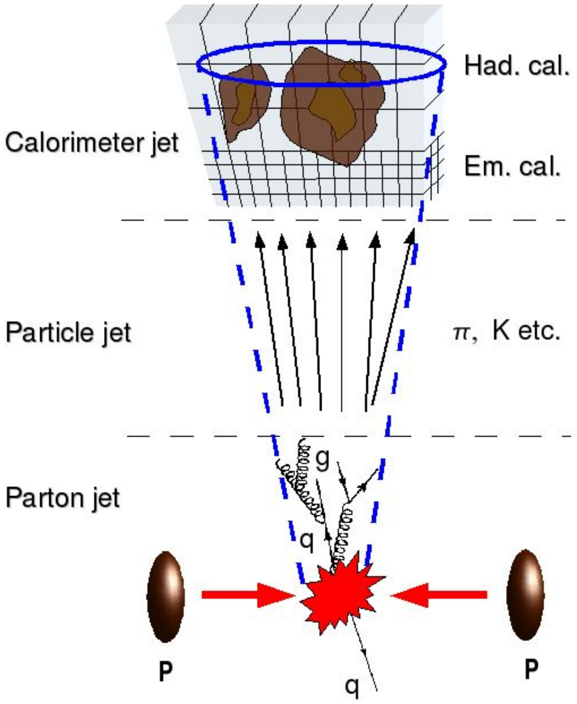


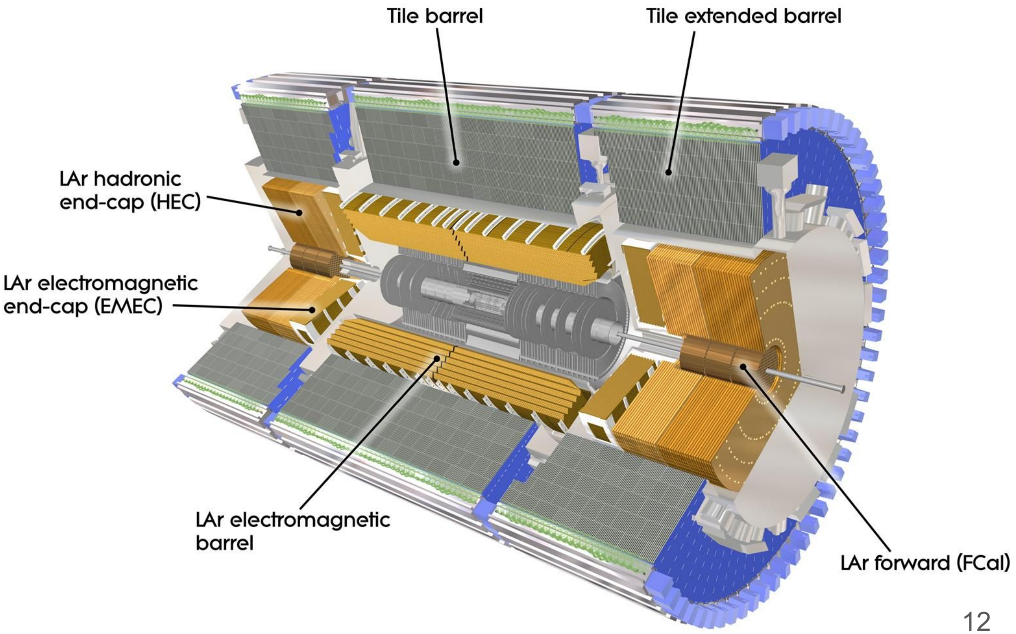
https://arxiv.org/abs/1509.02216

Jet Corrections

Need a calibration to "unfold" the detector and provide the best particle level estimate for a jet at detector level. Must correct for:

- **Non-compensation:** e / h ≠ 1
- **Dead material:** the detector is not hermetic or uniform
- **Clustering inefficiencies:** pileup contamination and electronic noise
- **Jet algorithm inefficiencies:** dependence on jet size parameter R



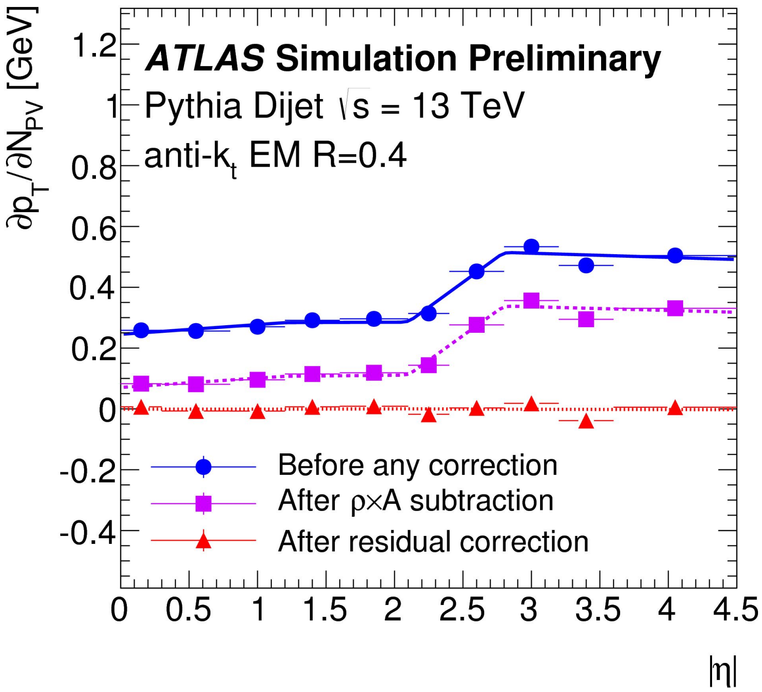


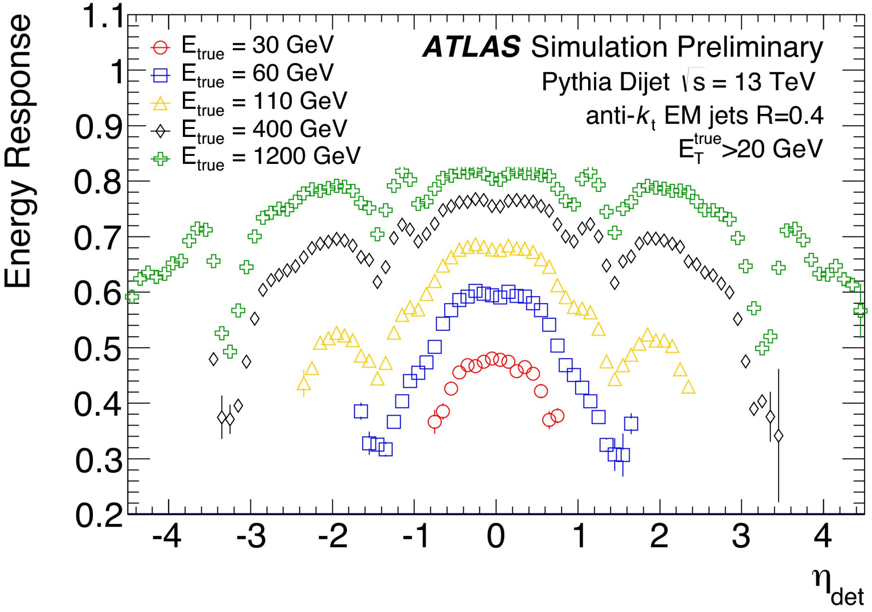
Jet Corrections

Origin Correction
Correct jet to point to the primary vertex instead of the centre of the detector

Pile-up Correction
Reduce pile-up effects by using average energy density across jet area

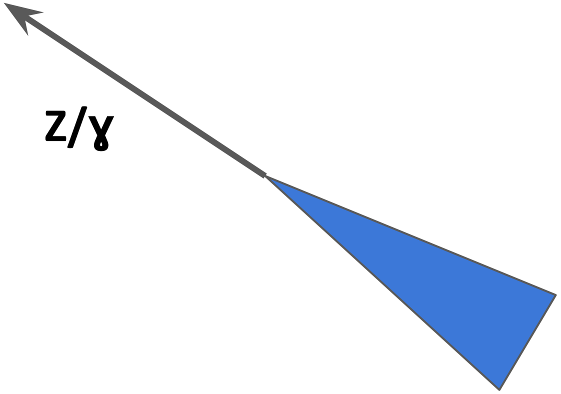
MC energy scale and η
Correct reco-truth relative differences in simulation
Correct η bias

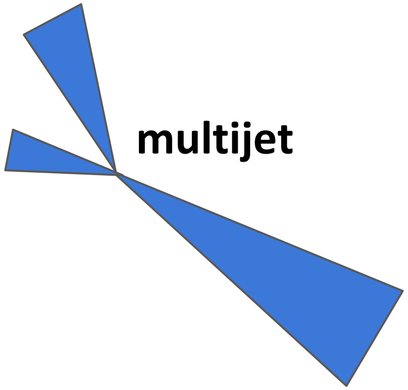


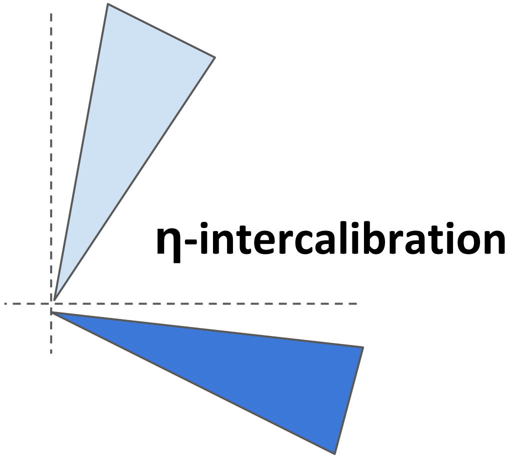


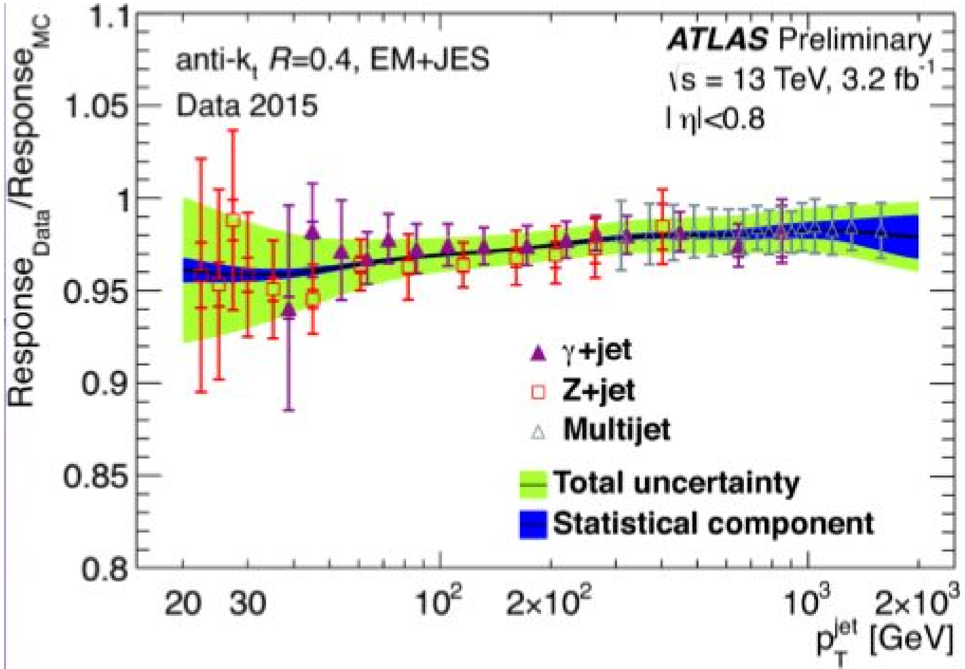
Jet Corrections: In Situ

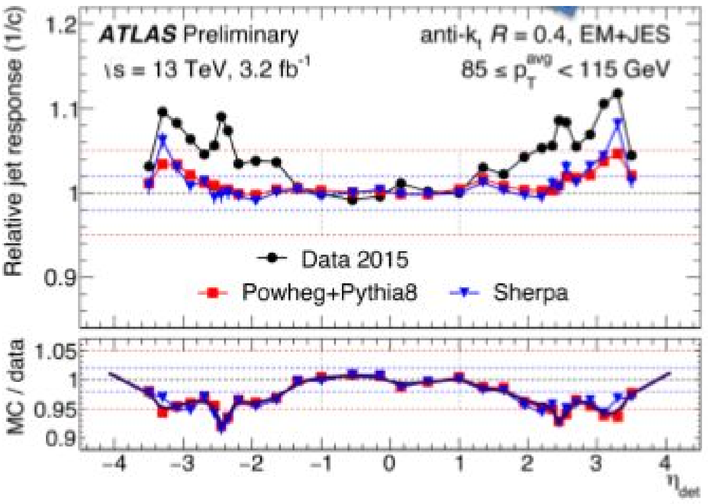
Correct jet p_T in data using well-calibrated reference objects





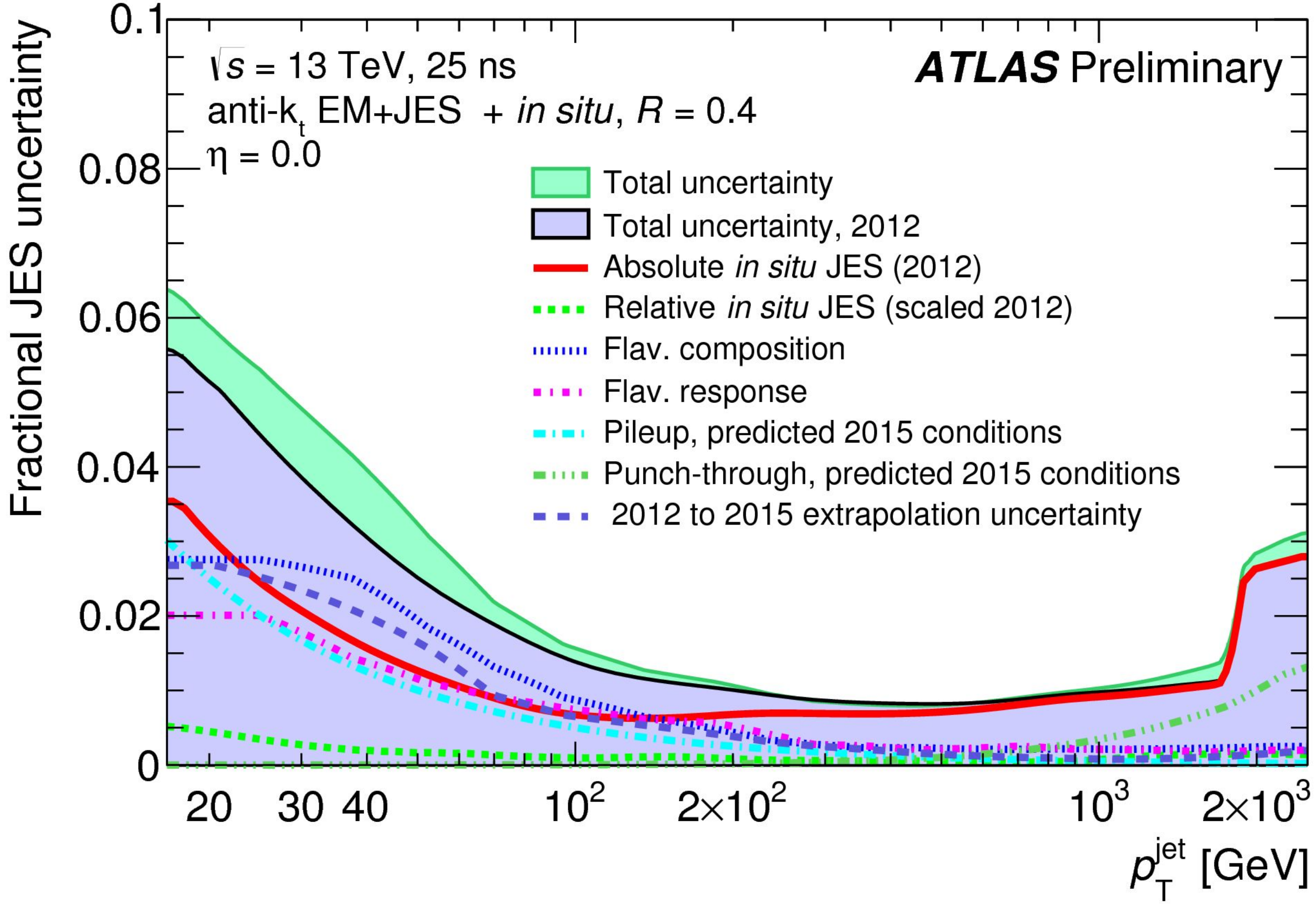




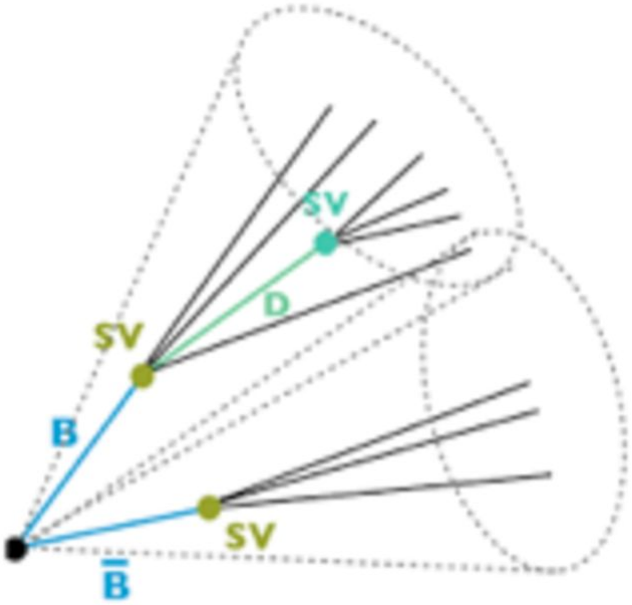


Jet Corrections: Uncertainties

In the end we have 77 independent sources of systematic uncertainty!



b-jet Tagging



Standard Model and new physics measurement need efficient techniques to identify jets containing b-flavoured hadrons

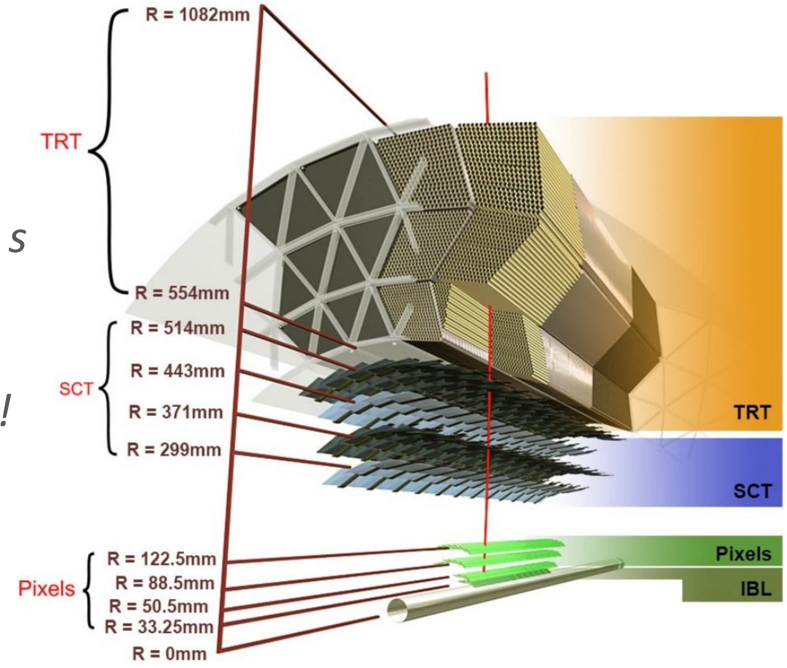
Top decays, $H \to bb$, …

Typical b-hadron lifetime $\boldsymbol{\tau \approx 1.5}$ ***ps***

So for $\boldsymbol{p_T = 50}$ ***GeV*** and $\boldsymbol{m = 5}$ ***GeV*** the mean flight path length is roughly:

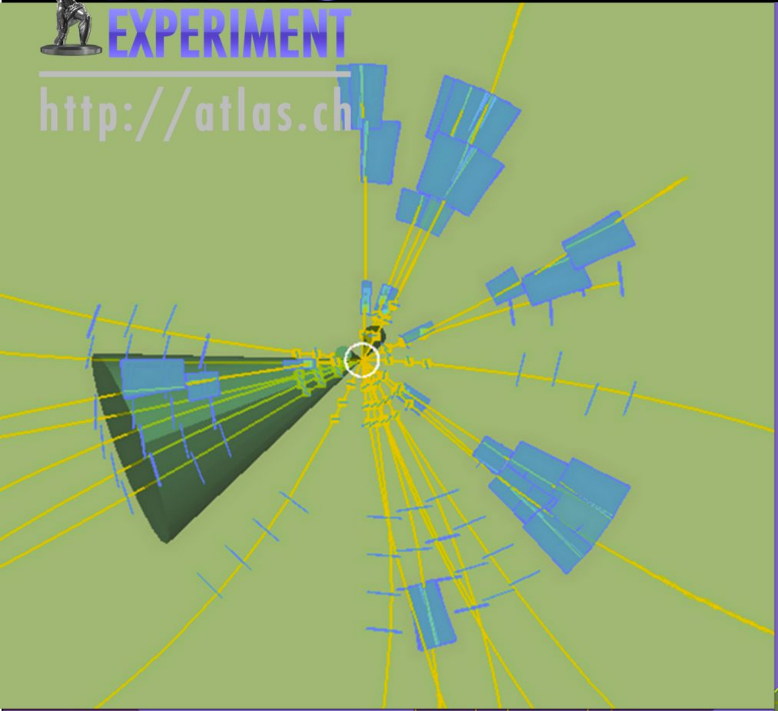
$$\langle L\rangle \approx \beta\, \gamma\, c\, \tau = 0.99 \times 10 \times 3 \times 10^{8}\ m/s \times 1.5 \times 10^{-12}\ s$$
$$\approx 4\ mm$$

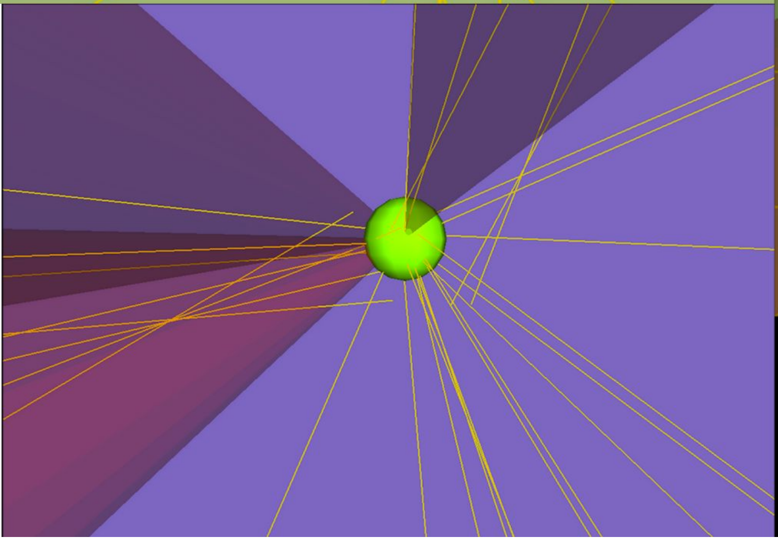
Can be measured with ATLAS resolution of 20 µm!

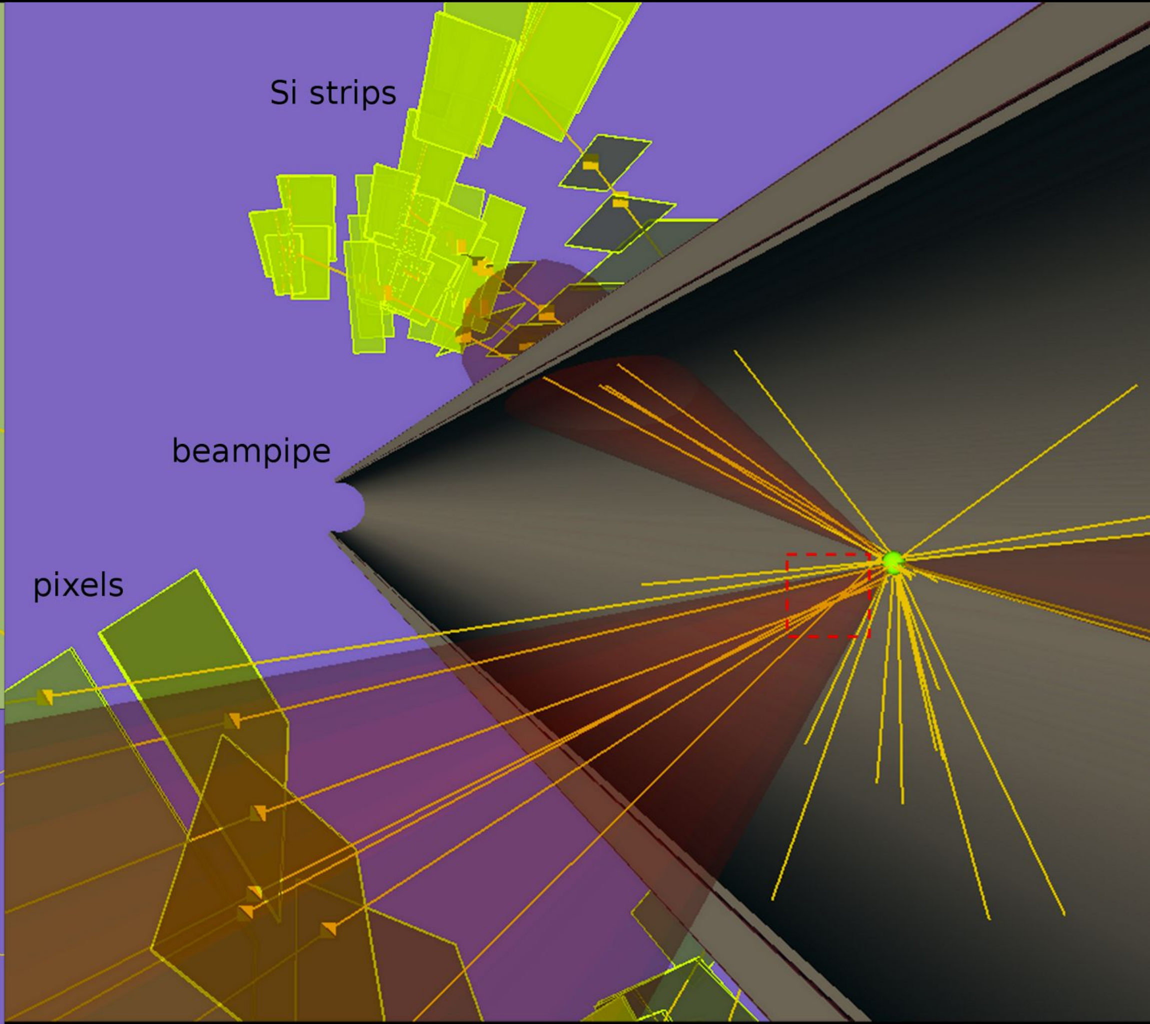


Run 152166
Event 817271

b-tagged jet in 7 TeV collisions





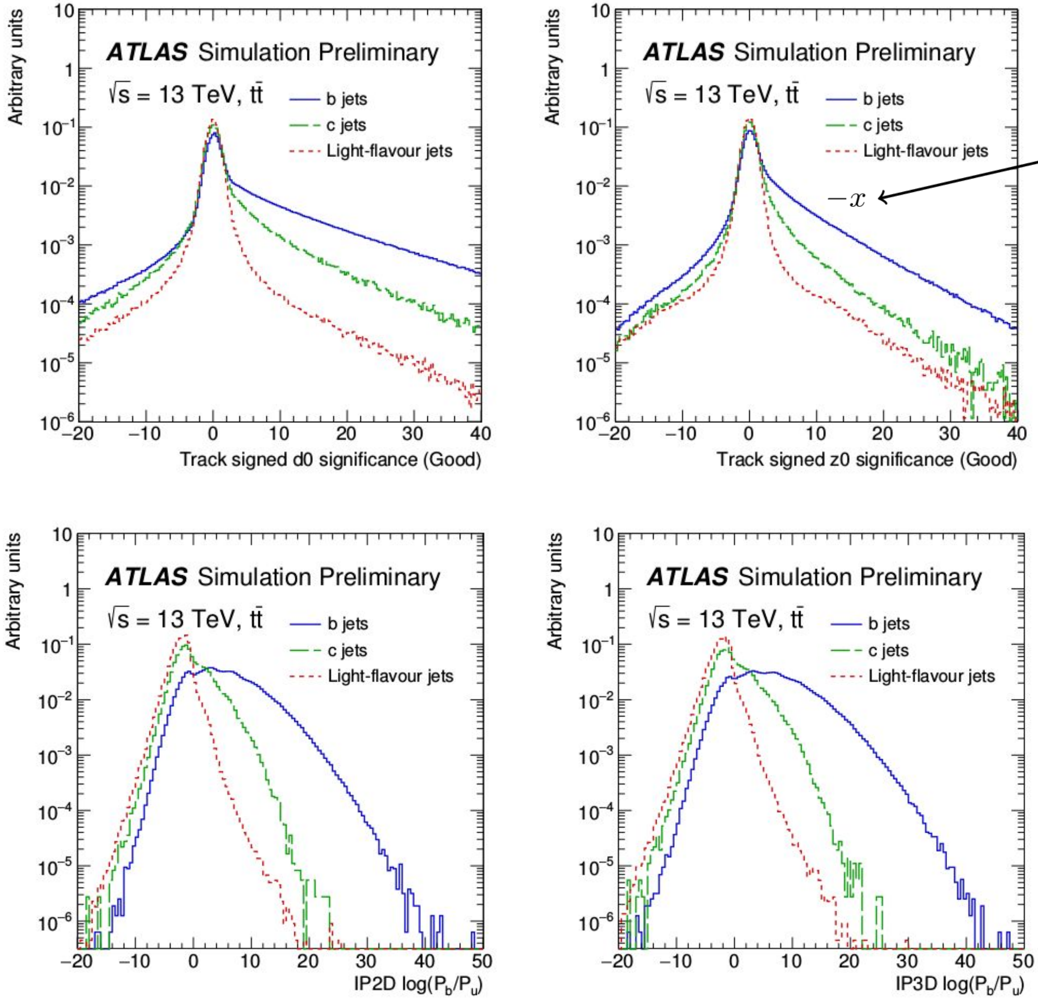


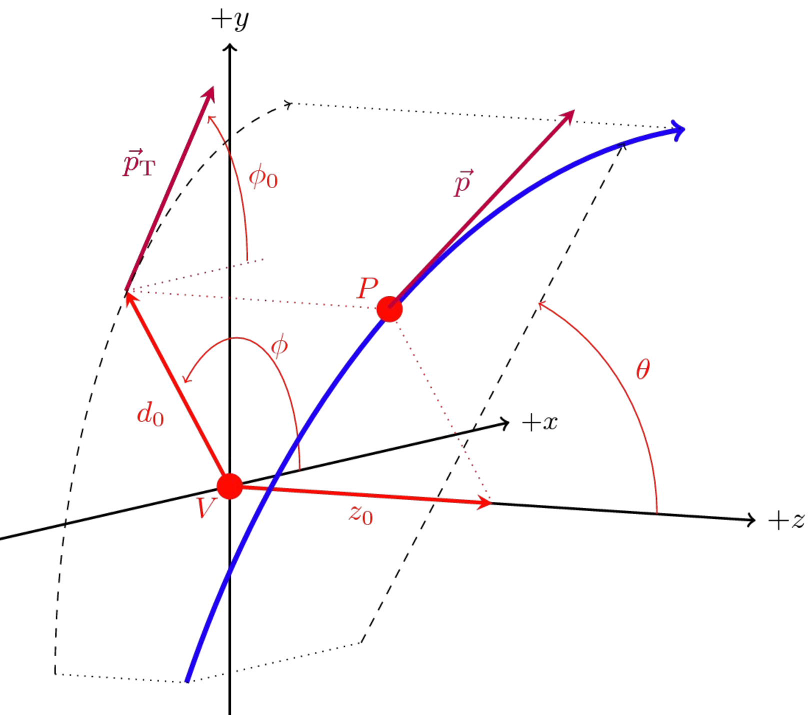
p_T^{jet}=19 GeV (measured at electromagnetic scale)

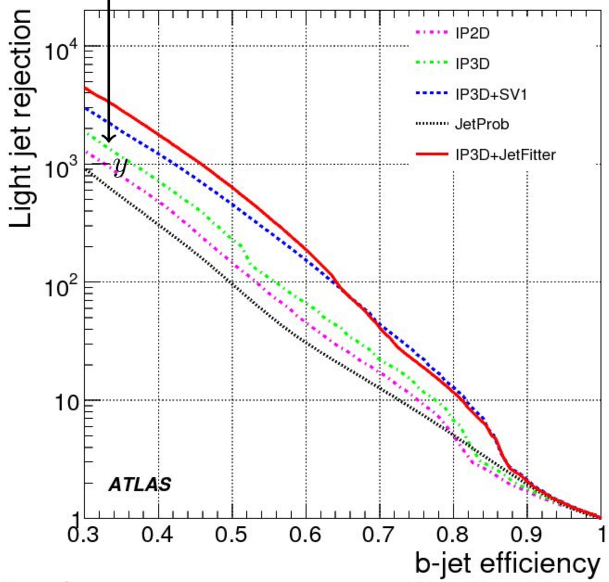
4 b-tagging quality tracks in the jet

b-tagging Algorithms

Look at track impact parameters: d_0 and z_0





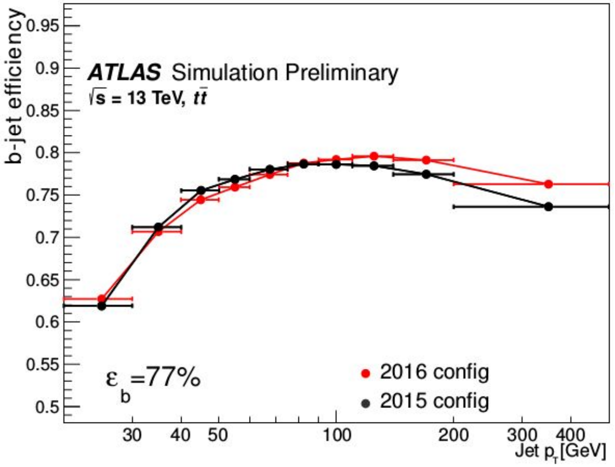


$$\sum_{i=1}^{N} \frac{\log p_b}{\log p_u}$$

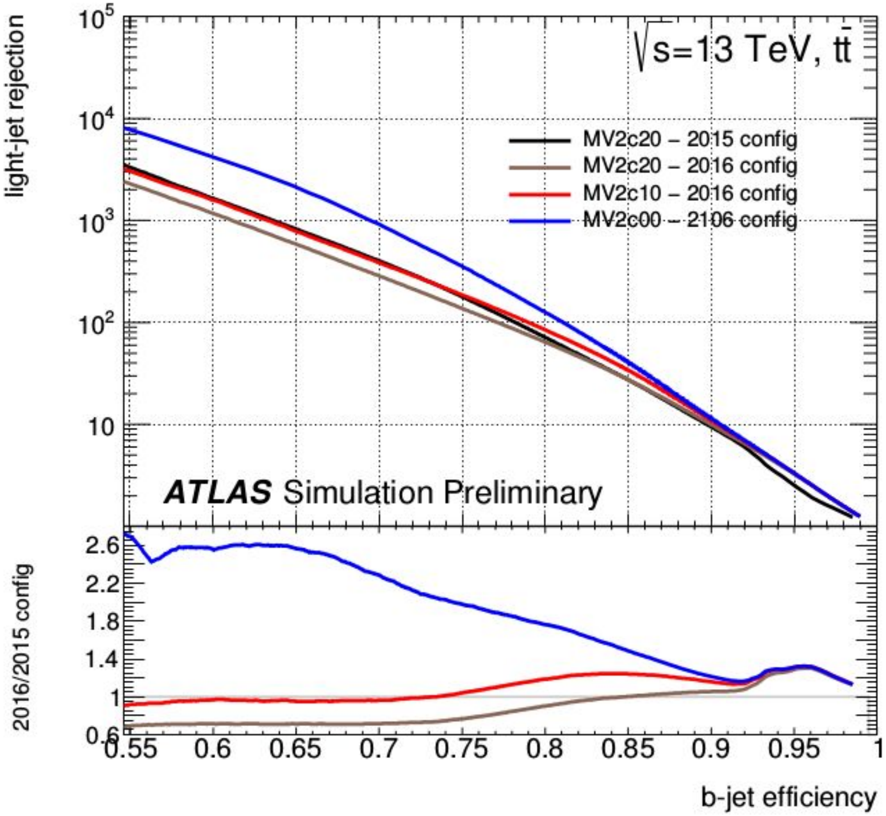
b-tagging Algorithms

Advanced techniques:

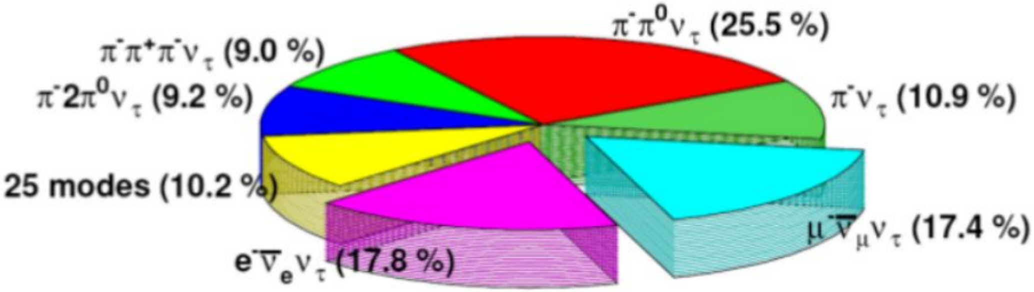
- ★ Decay chain multi-vertex reconstruction
- ★ Multivariate machine learning methods



Input	Variable	Description
Kinematics	$p_T(jet)$	Jet transverse momentum
	$\eta(jet)$	Jet pseudo-rapidity
IP2D, IP3D	$\log(P_b/P_{\text{light}})$	Likelihood ratio between the *b*- and light jet hypotheses
	$\log(P_b/P_c)$	Likelihood ratio between the *b*- and *c*-jet hypotheses
	$\log(P_c/P_{\text{light}})$	Likelihood ratio between the *c*- and light jet hypotheses
SV	$m(\text{SV})$	Invariant mass of tracks at the secondary vertex assuming pion masses
	$f_E(\text{SV})$	Fraction of the charged jet energy in the secondary vertex
	$N_{\text{TrkAtVtx}}(\text{SV})$	Number of tracks used in the secondary vertex
	$N_{\text{2TrkVtx}}(\text{SV})$	Number of two track vertex candidates
	$L_{xy}(\text{SV})$	Transverse distance between the primary and secondary vertices
	$L_{xyz}(\text{SV})$	Distance between the primary and secondary vertices
	$S_{xyz}(\text{SV})$	Distance between the primary and secondary vertices divided by its uncertainty
	$\Delta R(\text{jet}, \text{SV})$	ΔR between the jet axis and the direction of the secondary vertex relative to the primary vertex
Jet Fitter	$N_{\text{2TrkVtx}}(\text{JF})$	Number of 2-track vertex candidates (prior to decay chain fit)
	$m(\text{JF})$	Invariant mass of tracks from displaced vertices assuming pion masses
	$S_{\text{xyz}}(\text{JF})$	Significance of the average distance between the primary and displaced vertices
	$f_E(\text{JF})$	Fraction of the charged jet energy in the secondary vertices
	$N_{\text{1-trk vertices}}(\text{JF})$	Number of displaced vertices with one track
	$N_{\geq\text{2-trk vertices}}(\text{JF})$	Number of displaced vertices with more than one track
	$N_{\text{TrkAtVtx}}(\text{JF})$	Number of tracks from displaced vertices with at least two tracks
	$\Delta R(\vec{p}_{\text{jet}}, \vec{p}_{\text{vtx}})$	ΔR between the jet axis and the vectorial sum of the momenta of all tracks attached to displaced vertices

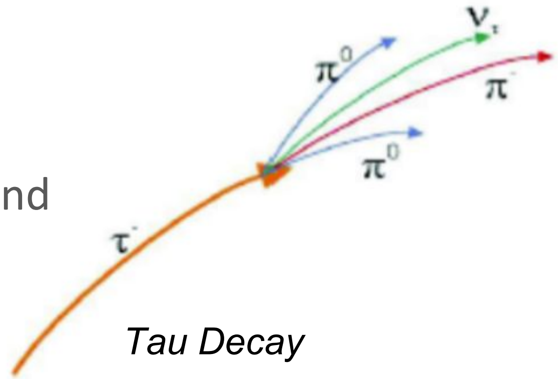


Tau Decays

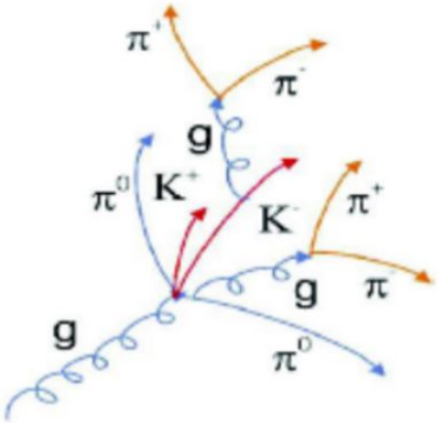


Tau Characteristics

- ★ m_τ = 1.8 GeV
- ★ Lifetime: $c\tau$ = 87 µm
- ★ 65% of taus decays hadronically
- ★ Hadronic decays are mostly collimated collections of neutral and 1 or 3 charged pions “prongs”. Some rare decays include kaons.



Tau Decay



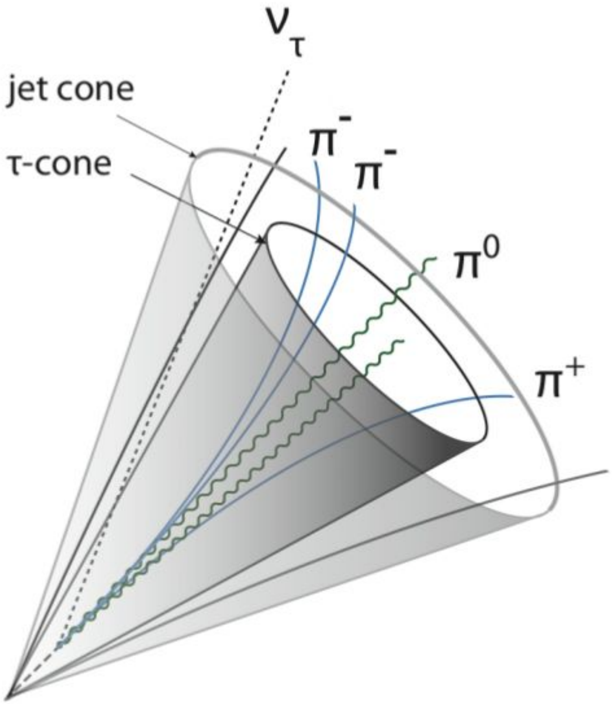
Typical Jet

Experimentally, taus are characterised by:

- ➢ Few tracks
- ➢ Leading track and displaced secondary vertex
- ➢ Narrow jet with large EM component (1-prong decays) from $\pi^0 \to \gamma\gamma$

Requires good performance from calorimeters and tracking systems

Tau Reconstruction

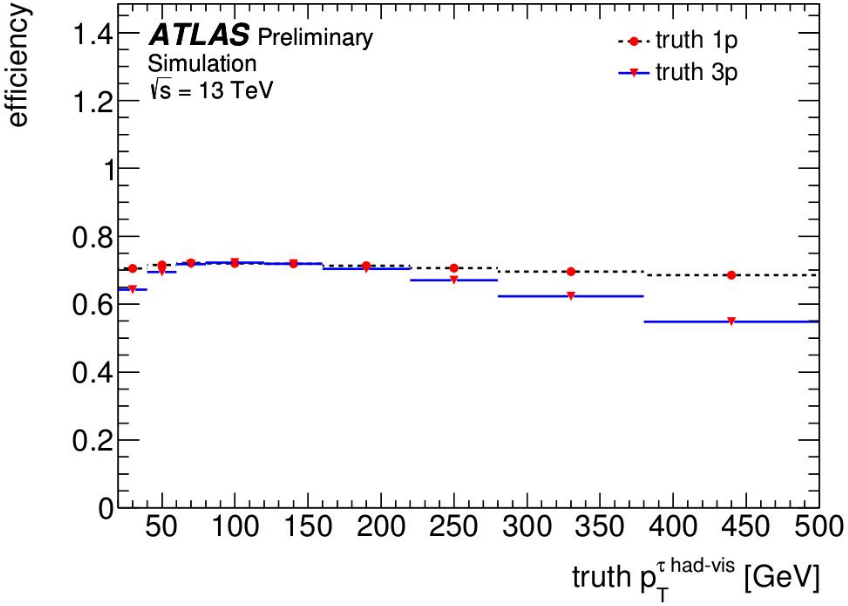


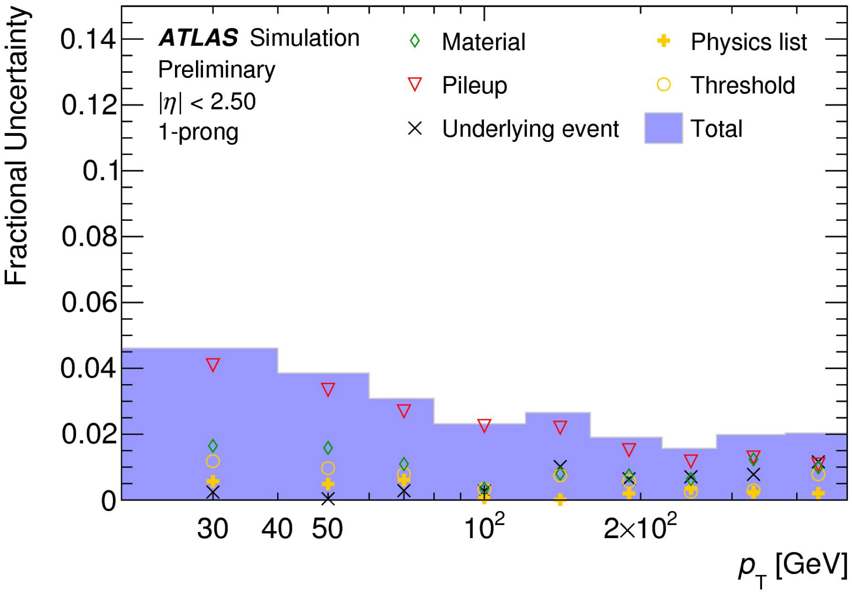
Tau candidates are seeded by calorimeter anti-k_T $R = 0.4$ calorimeter jets

Tracks with $p_T > 1$ GeV in $\Delta R < 0.2$ around the axis of the jet are associated to the candidate

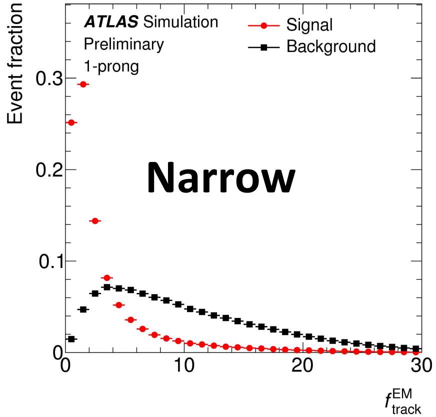
These tracks are used to identify the tau's production vertex which may not be the primary vertex of the event

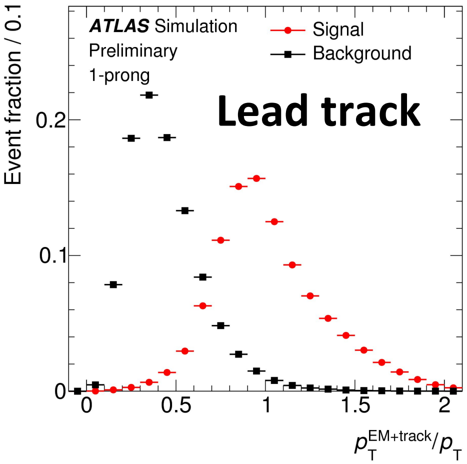
Impact parameters are recalculated w.r.t. the tau vertex

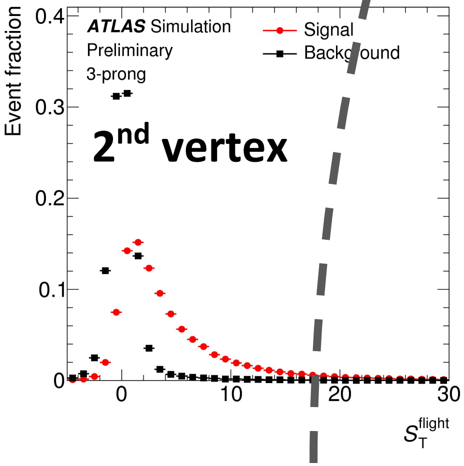


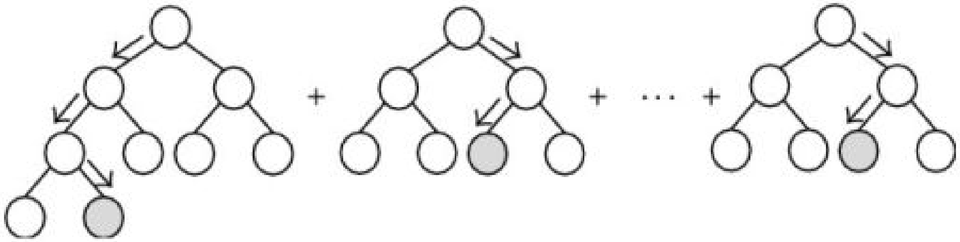


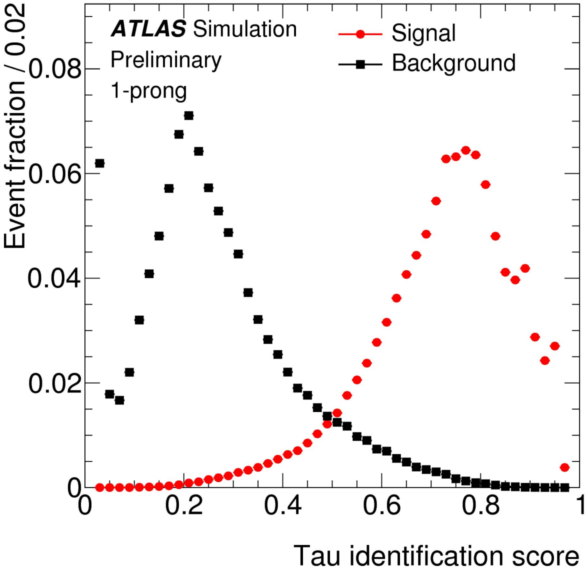
Tau Identification

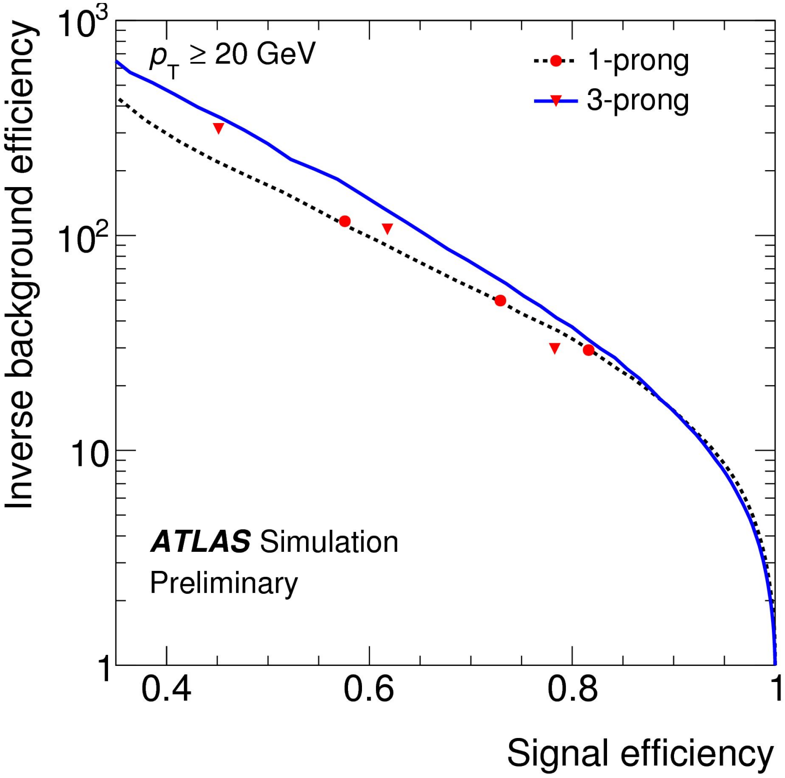


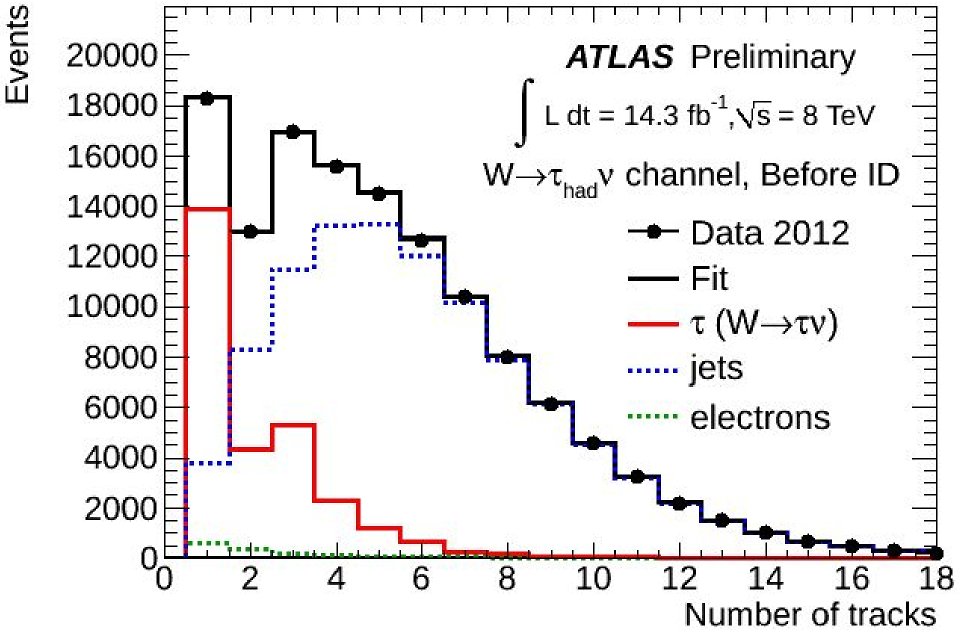


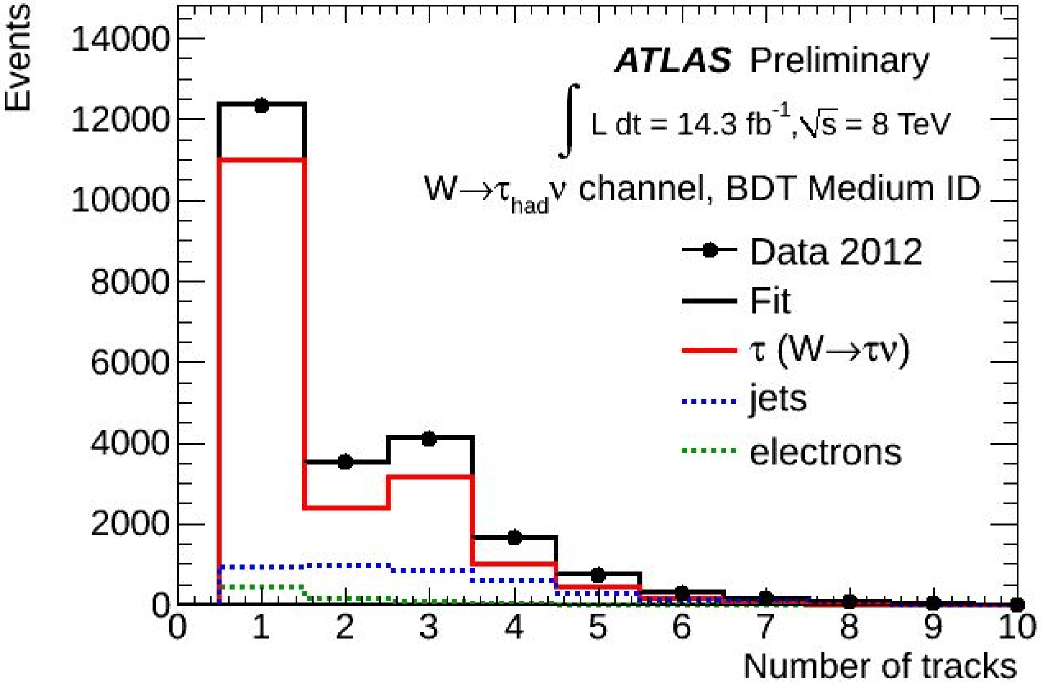




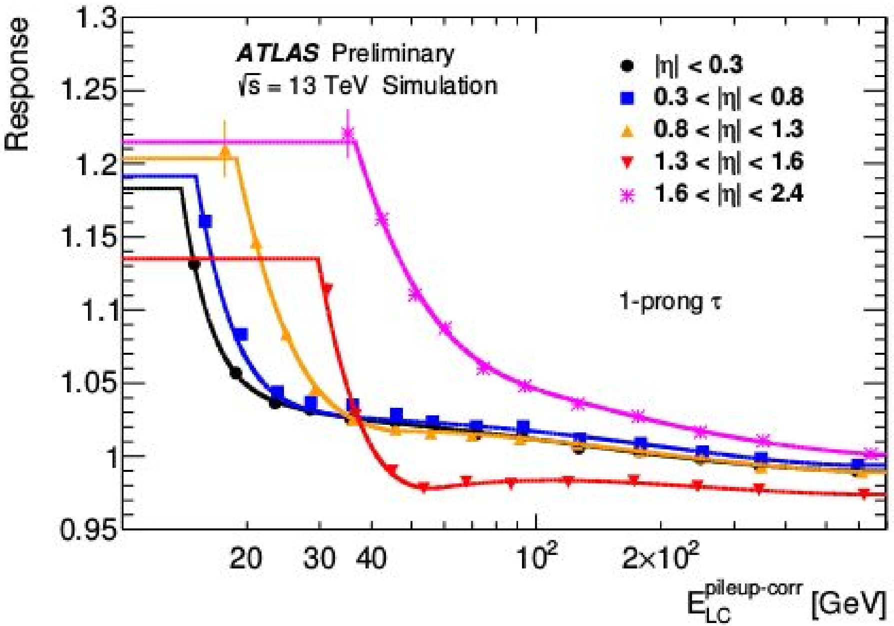


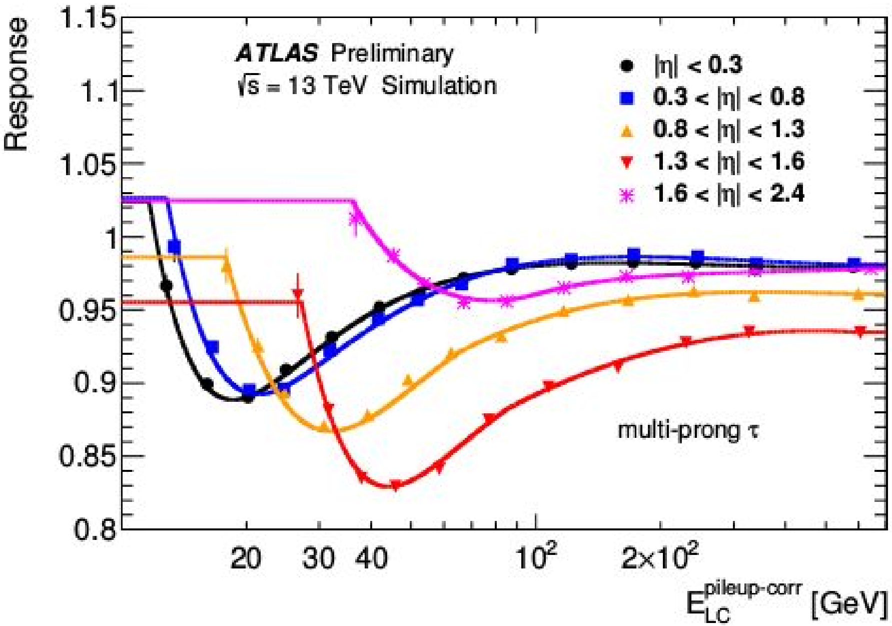



ATLAS Preliminary
∫ L dt = 14.3 fb⁻¹, √s = 8 TeV
W→τ_had ν channel, Before ID
Data 2012
Fit
τ (W→τν)
jets
electrons
Events
0
2000
4000
6000
8000
10000
12000
14000
16000
18000
20000
0
2
4
6
8
10
12
14
16
18
Number of tracks


ATLAS Preliminary
∫ L dt = 14.3 fb⁻¹, √s = 8 TeV
W→τ_had ν channel, BDT Medium ID
Data 2012
Fit
τ (W→τν)
jets
electrons
Events
0
2000
4000
6000
8000
10000
12000
14000
0
1
2
3
4
5
6
7
8
9
10
Number of tracks

Tau Energy Calibration

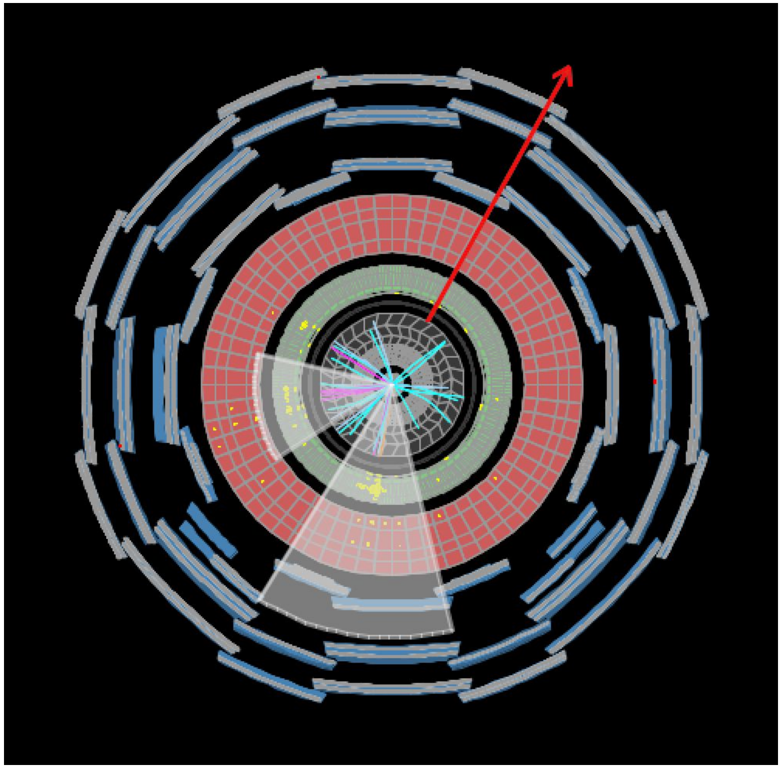




Due to their particular mix of charged and neutral pions in a narrow cone, taus have their own energy calibration

$$E_{\text{calib}} = \frac{E_{\text{LC}} - E_{\text{pileup}}}{\mathcal{R}\left(E_{\text{LC}} - E_{\text{pileup}}, |\eta|, n_p\right)}$$

Missing Transverse Momentum



Transverse momentum of collision products should sum to zero!

Any imbalance is called “missing transverse momentum”: E_T^{miss}

Weakly-interacting stable particles in the final state. Neutrinos or BSM particles.
Dark matter?

Fake E_T^{miss} can also be created by interacting SM particles that escape detector acceptance, or are poorly reconstructed!

$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss}, e} + E_{x(y)}^{\text{miss}, \gamma} + \boxed{E_{x(y)}^{\text{miss}, \tau}} + \boxed{E_{x(y)}^{\text{miss, jets}}} + E_{x(y)}^{\text{miss}, \mu} + \boxed{E_{x(y)}^{\text{miss, soft}}}$$

$$E_{\text{T}}^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + \left(E_y^{\text{miss}}\right)^2}$$

$$\phi^{\text{miss}} = \arctan\left(E_y^{\text{miss}} / E_x^{\text{miss}}\right)$$

Missing Transverse Momentum

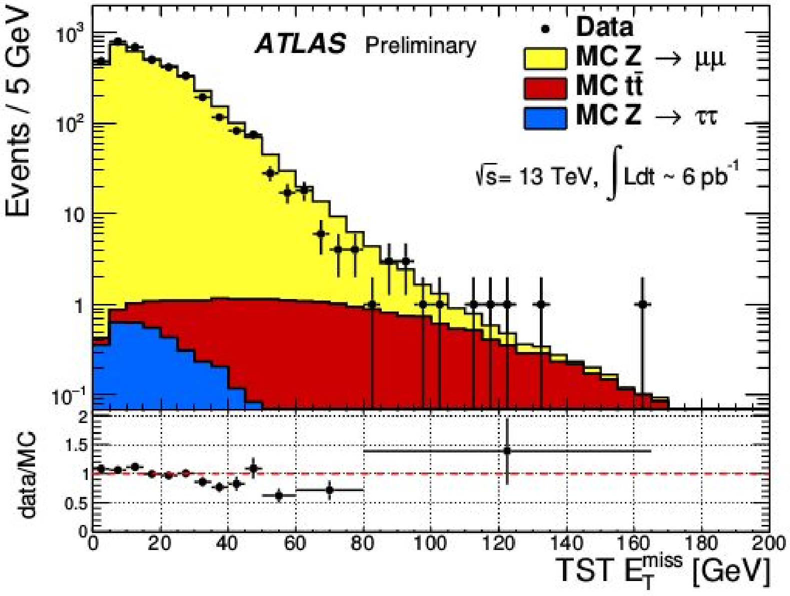
$$E_{x(y)}^{\text{miss}} = E_{x(y)}^{\text{miss}, e} + E_{x(y)}^{\text{miss}, \gamma} + \boxed{E_{x(y)}^{\text{miss}, \tau}} + \boxed{E_{x(y)}^{\text{miss, jets}}} + E_{x(y)}^{\text{miss}, \mu} + \boxed{E_{x(y)}^{\text{miss, soft}}}$$

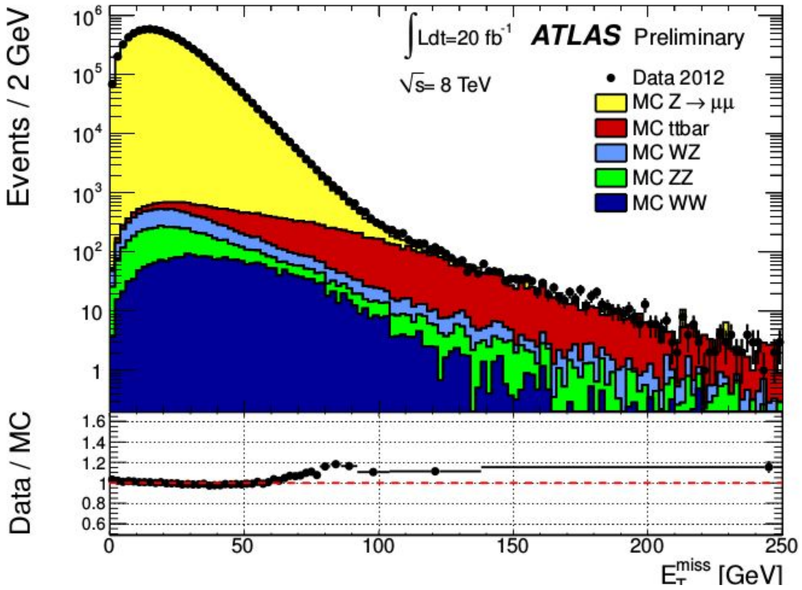
Reconstructed tau candidates passing the “medium” identification threshold and with $p_T > 20$ GeV and not in any detector transition region.

Anti-k_T jets (R=0.4) with calibrated $p_T > 20$ GeV. Also required to have a jet-vertex tagger value > 0.64 where jets are within tracker coverage.

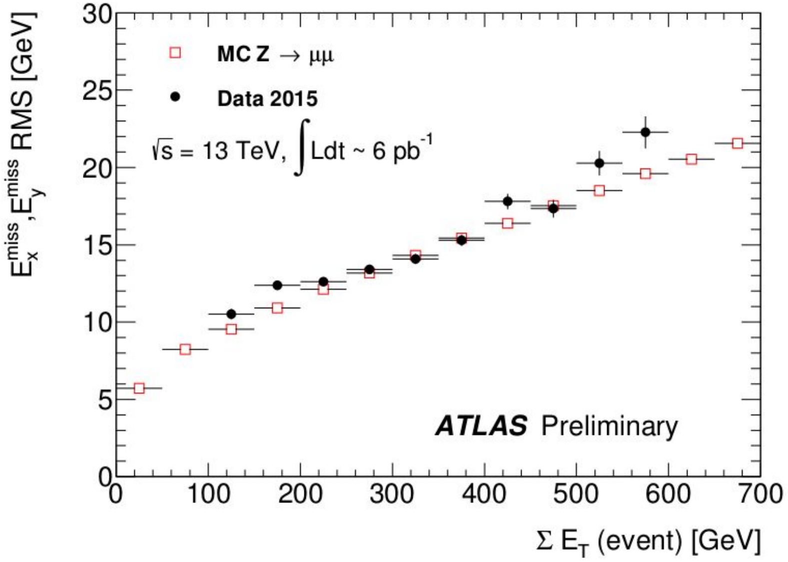
Tracks associated to the primary vertex but not matched to any object already included in the E_T^{miss}

Missing Transverse Momentum Performance

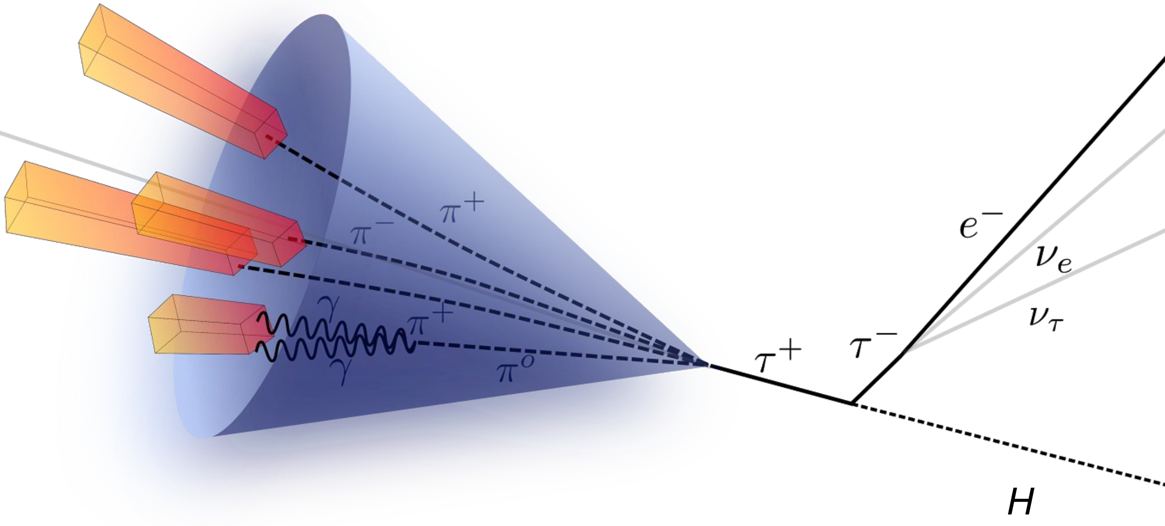




E_T^{miss} provides a good measure of overall performance!



Missing Transverse Momentum: Di-tau mass



Di-tau system is underconstrained with hadronic/leptonic decays producing 1 or 2 neutrinos

But we can use E_T^{miss} constraints!

Still 6 - 8 unknowns…

$$E_x^{\mathrm{miss}} = p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \cos\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \cos\phi_{\mathrm{mis}_2}$$

$$E_y^{\mathrm{miss}} = p_{\mathrm{mis}_1} \sin\theta_{\mathrm{mis}_1} \sin\phi_{\mathrm{mis}_1} + p_{\mathrm{mis}_2} \sin\theta_{\mathrm{mis}_2} \sin\phi_{\mathrm{mis}_2}$$

$$M_\tau^2 = m_{\mathrm{mis}_1}^2 + m_{\mathrm{vis}_1}^2 + 2\sqrt{p_{\mathrm{vis}_1}^2 + m_{\mathrm{vis}_1}^2}\sqrt{p_{\mathrm{mis}_1}^2 + m_{\mathrm{mis}_1}^2} - 2p_{\mathrm{vis}_1} p_{\mathrm{mis}_1} \cos\alpha_{vm_1}$$

$$M_\tau^2 = m_{\mathrm{mis}_2}^2 + m_{\mathrm{vis}_2}^2 + 2\sqrt{p_{\mathrm{vis}_2}^2 + m_{\mathrm{vis}_2}^2}\sqrt{p_{\mathrm{mis}_2}^2 + m_{\mathrm{mis}_2}^2} - 2p_{\mathrm{vis}_2} p_{\mathrm{mis}_2} \cos\alpha_{vm_2}$$

Use maximum likelihood approach to estimate neutrino momenta