Precision jet calibrations. Towards a b-jet energy scale and the impact in Top Physics

Miguel Jiménez

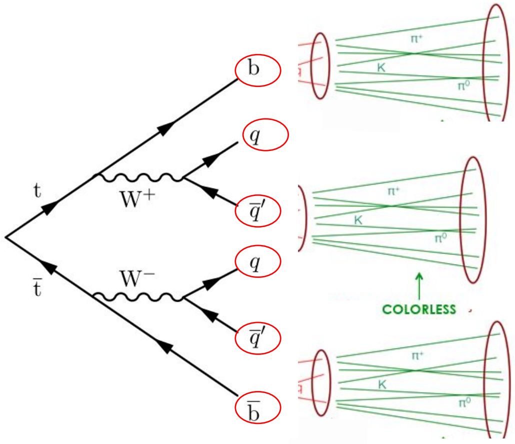
19-21st November

CPAN 2025, Valencia

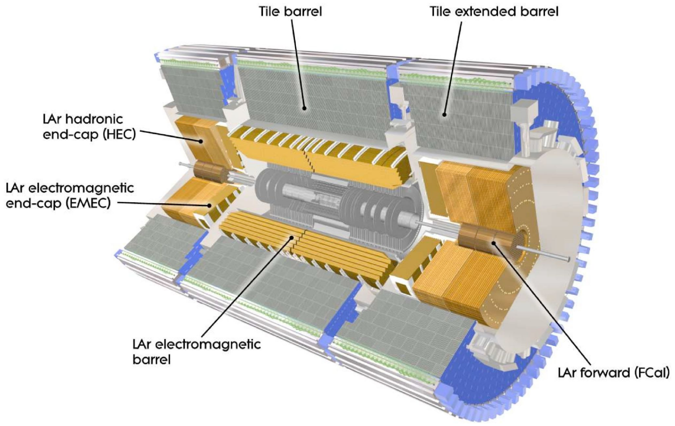
Supervisors: Marcel Vos and Davide Melini
b-jet task effort team

What is a jet ?

$$\mathcal{L}_{\mathrm{QCD}} = \sum_{f=1}^{N_F} \bar{q}_{f,i}\left(i\gamma^{\mu}(D_{\mu})_{ij} - m_f\delta_{ij}\right)q_{f,j} - \frac{1}{2}\mathrm{tr}(G^{\mu\nu}G_{\mu\nu})$$

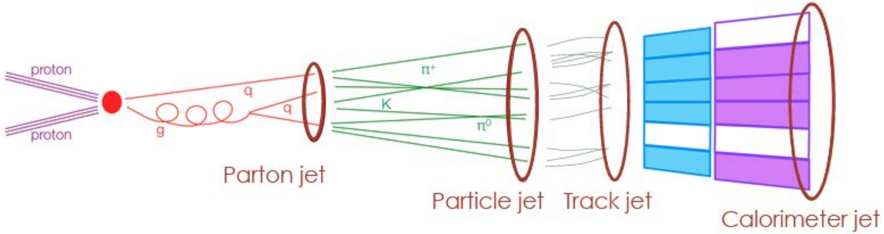


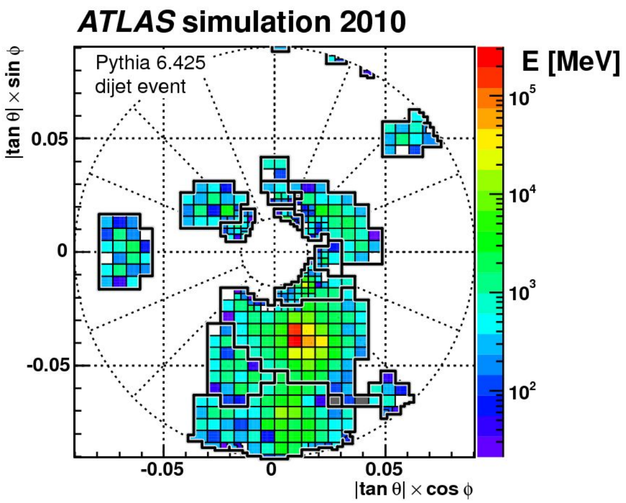
Eur. Phys. J. C 77 (2017) 490



This is nice but… how do I actually characterize this?

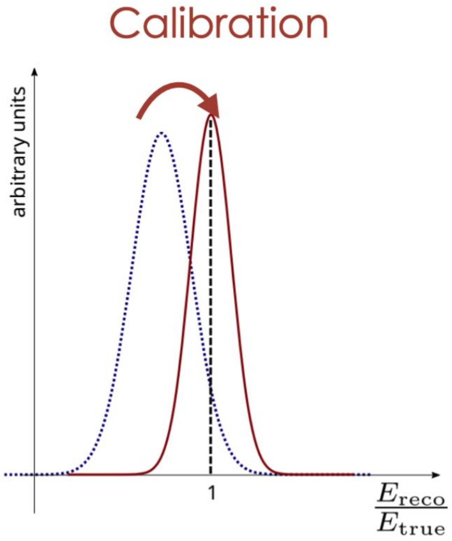
In short: Jets represent a collimated shower of energetic hadrons.



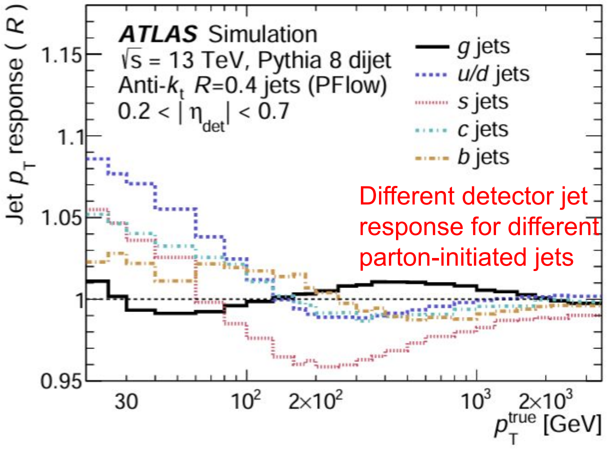


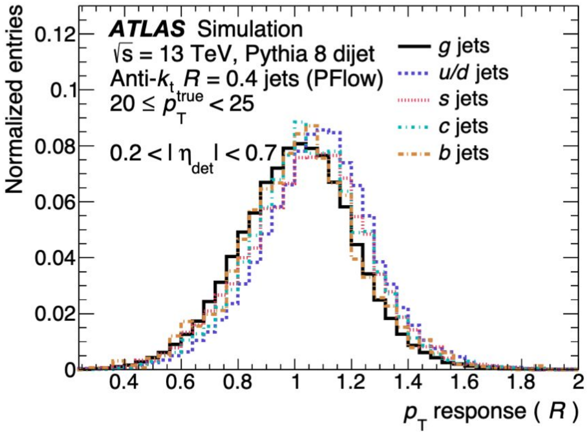
Jet calibrations → What & Why

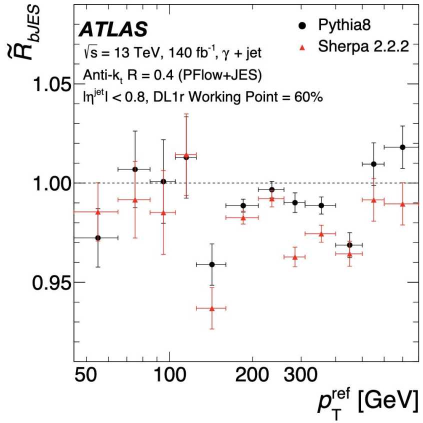
https://arxiv.org/abs/2303.17312



$$R(E_{\mathrm{jet}}^{\mathrm{truth}}, \eta_{\mathrm{det}}) = E_{\mathrm{jet}}^{\mathrm{reco}} / E_{\mathrm{jet}}^{\mathrm{truth}}$$



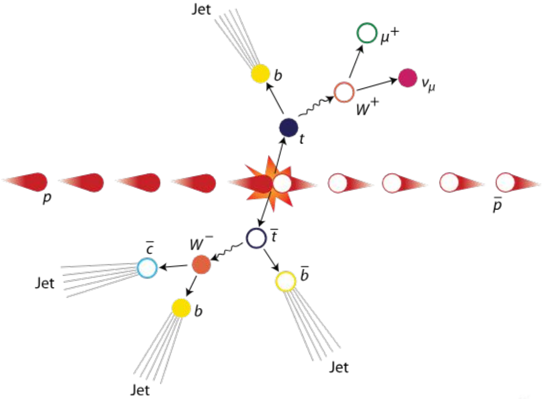




Corrections of the Inclusive Jet Energy Scale factor respect to the b-Jet Energy Scale were found to be of the order >2-3% (depending on photon pT)

Jet calibrations in Top Physics → Does it really matter ?

- Most of the **top quark mass measurements** (and other particles) have as main uncertainties features related to Jet calibrations.



JES Uncertainty Component	Δm_{top} [GeV] Dilepton	ℓ+Jets
JES_EffNP_Model1	0.254	-0.014
JES_Pileup_RhoTopology	0.201	-0.057
JES_EtaInt_Modelling	0.110	-0.004
JES_Flavor_PerJet_GenShower_HF	0.110	0.122
JES_Flavor_PerJet_Hadronization	0.002	-0.162
Total	**0.414**	**0.220**

Template top mass (dilepton+ljets)

Source	Uncertainty [GeV]
JES	± 0.29
Radiation (ISR and FSR)	± 0.17
Colour reconnection (CR1 and CR2)	± 0.15
JES heavy flavour	± 0.14
Parton shower and hadronisation model	± 0.14
JER	± 0.10
MC statistics	± 0.08
Underlying event	± 0.08
Recoil	± 0.07
Fit closure	± 0.07
Background modelling	± 0.05
Matrix element matching ($p_T^{hard} = 1$)	± 0.04
b-tagging	± 0.04
Higher-order corrections	± 0.02
E_T^{miss}	± 0.02
Pileup	± 0.01
JVT	± 0.01
PDF	± 0.01
Leptons	± 0.01
Luminosity	< 0.01
Total statistical	± 0.27
Total systematic	± 0.46
Total	± 0.53

Uncertainty category	Uncertainty impact [GeV] LHC	ATLAS	CMS
b-JES	0.18	0.17	0.25
b tagging	0.09	0.16	0.03
ME generator	0.08	0.13	0.14
JES 1	0.08	0.18	0.06
JES 2	0.08	0.11	0.10
Method	0.07	0.06	0.09
CMS b hadron $\mathcal{B}$	0.07	—	0.12
QCD radiation	0.06	0.07	0.10
Leptons	0.05	0.08	0.07
JER	0.05	0.09	0.02

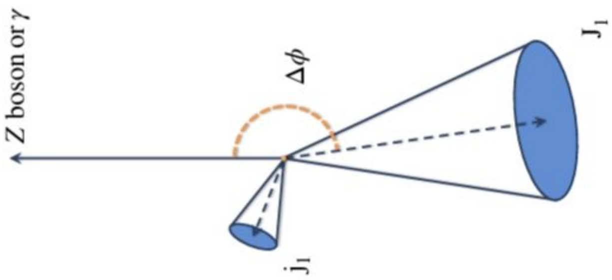
top quark mass with with a high transverse momentum

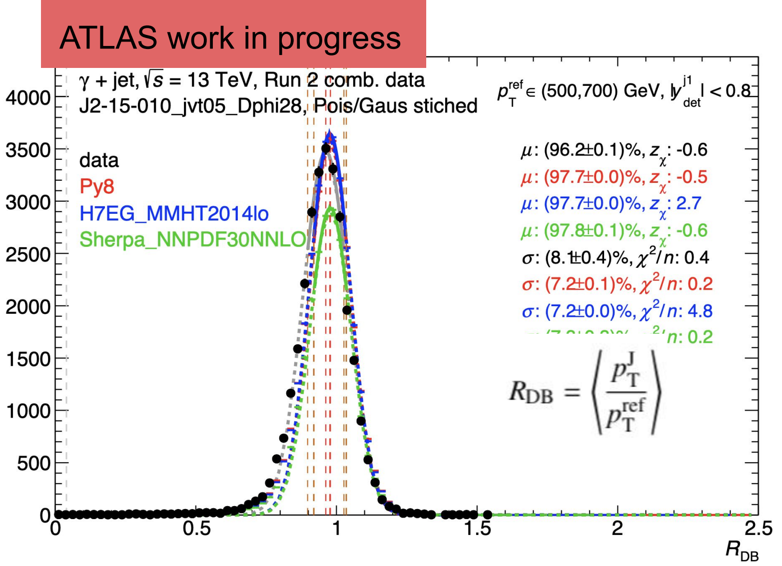
Combination of measurements ATLAS & CMS 7 TeV

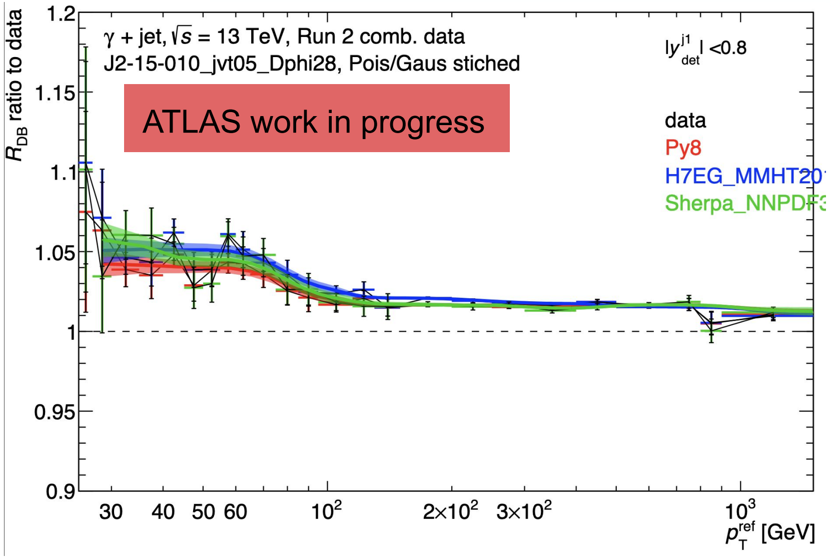
In-situ Direct Balance technique

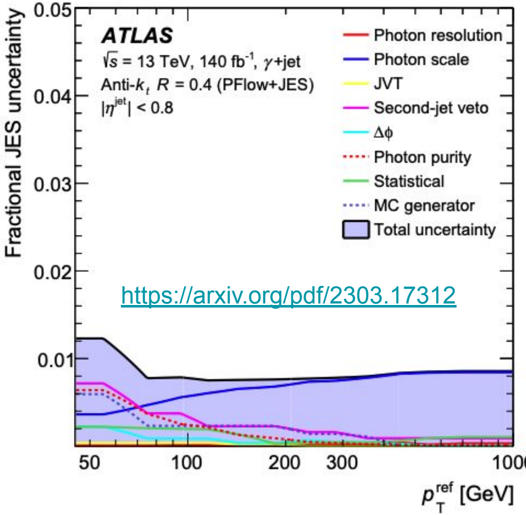
$$p_T^{ref} = p_T^Z \left|\cos(\Delta\phi)\right|$$

Z boson or photon







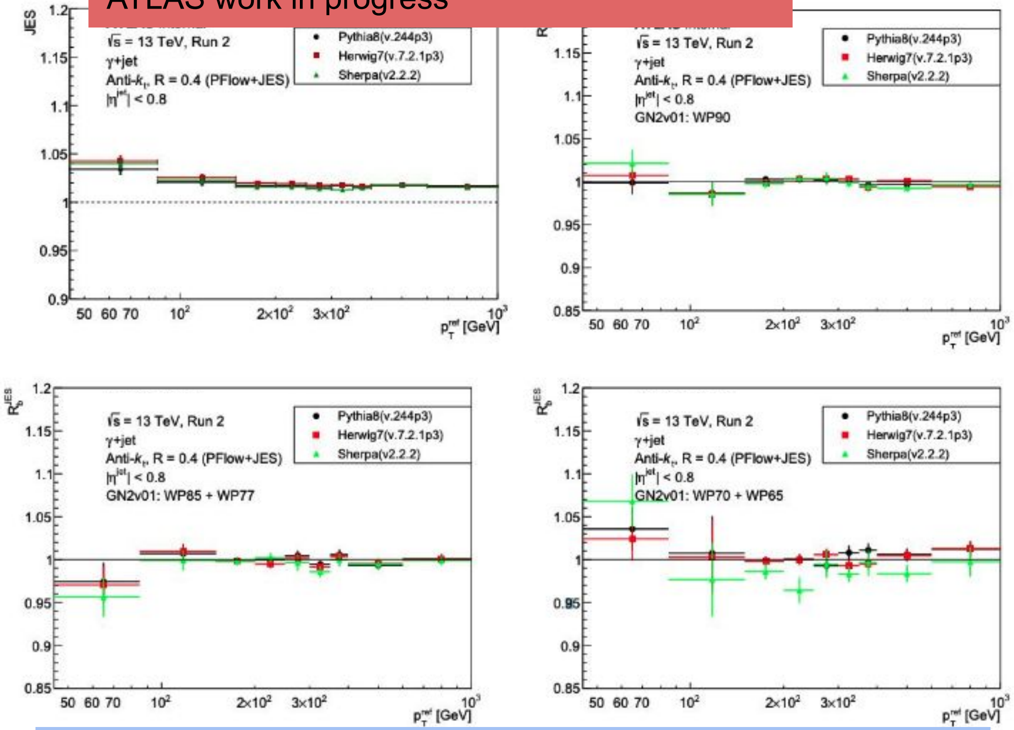


https://arxiv.org/pdf/2303.17312

Inclusive measurements of the Jet Energy Scale historically have as leading uncertainties the photon scale and topology (~1% depending on photon pT)

The actual work: cJES and bJES calibration factors

ATLAS work in progress



GN2 Working Point	c-quark	light-quark	b-quark	gluon
Inclusive	11.7%	68.5%	1.9%	17.8 %
90%	44.6%	39.9%	2.8%	12.7%
85%+77%	44.3%	32.4%	10.1%	13.08%
70%+65%	8.5%	10.3%	63.8%	17.4%

Table 25: Flavour composition 150-1000 GeV. Generator: Pythia8

GN2 Working Point	c-quark	light-quark	b-quark	gluon
Inclusive	13.2%	69.5%	2.0%	15.13 %
90%	47.7%	38.03%	3.0%	38.03%
85%+77%	48.7%	28.7%	11.7%	10.8%
70%+65%	7.8%	6.9%	68.3%	16.88%

Table 26: Flavour composition 150-1000 GeV. Generator: Herwig7

The potential precision we aim to achieve is the per mill level, can be statistically feasible, but still need to handle properly systematics

The actual work: cJES and bJES systematic uncertainties

ATLAS work in progress

Leading Uncertainties	bJES/JES (65% ‖ 70% WP)	cJES/JES (77% ‖ 85% WP)	cJES/JES (90% WP)
stat. uncertainty	0.0018	0.0012	0.0009
MC stat. unc. (Py8)	0.0016	0.0009	0.0008
photon scale unc.	0.0015	0.00027	0.00023
veto cut unc.	0.006	0.0009	0.00024
total uncertainty	0.008	0.0027	0.0014

- The systematic uncertainties go down to per mill level as expected, i.e the **photon scale**.
- There is **topology uncertainties** that are giving rise to larger values, a further investigation might be needed
- But the overall uncertainty lies below 1%, which is the main goal of the calibration strategy.

Conclusions

- Jets play a central role in High Energy collisions where hadrons are involved. An efficient way of calibrating them is crucial for the physics program.
- The “flavour” of the jets proved to be important in the regime of precision physics, 2-3% correction for some bins in photon pT.
- In this new study we are developing a b-jet data-driven calibration with a precision of few per mille, adapting direct balance methods for gamma/Z+jet channel.
- These improved calibrations will constrain the Jet Energy scale and the b-Jet energy scale, which are found to be the leading ones in some physics analysis, like in Top Physics.

Thank you!

Backup

How to actually build a jet (in ATLAS)

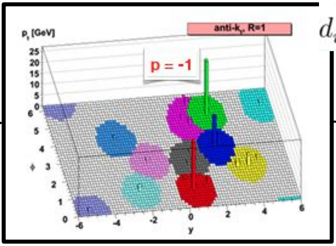
Jet inputs

Topo-cluster

Tracks

→

Jet finding



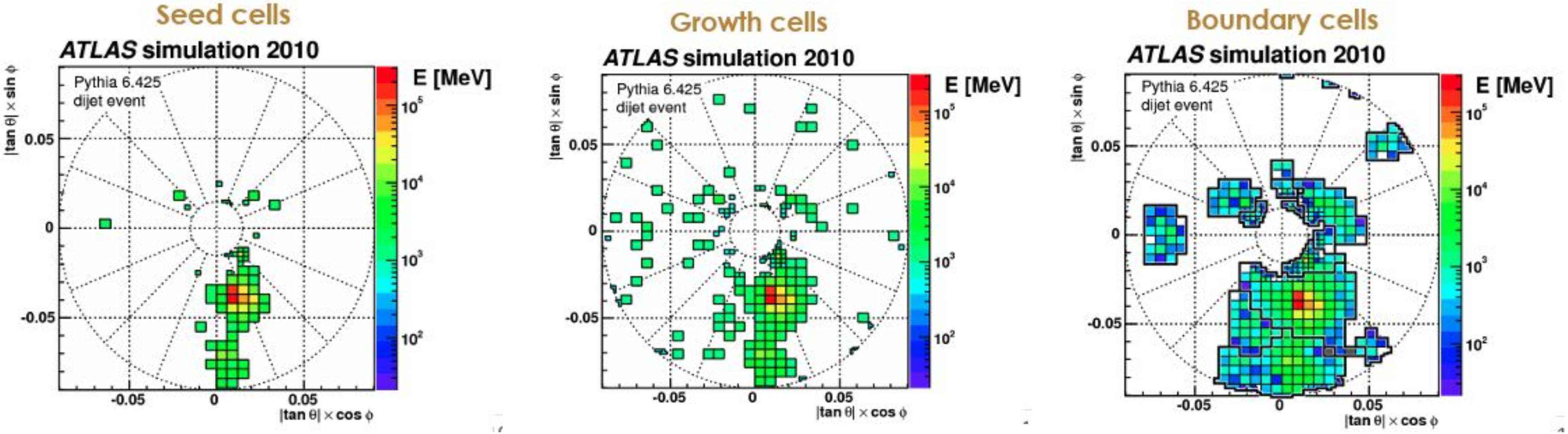
$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p}) \frac{\Delta R_{ij}^2}{R^2}$$

$$d_{iB} = k_{t,i}^{2p}$$

→ Reconstructed jet

Creation of topo-clusters

- Clusters are seeded by cells with **large energy over noise ratio** |ζ|>4 (highly significant)
- **Expanded on neighbouring cells**. All Neighbors with |ζ|>2 are added
- All neighbouring cells are added **regardless of the significance** |ζ|>0
- Final cluster splitting step **breaks up large topo-clusters with multiple local maxima**



Dilia Portillo's slides

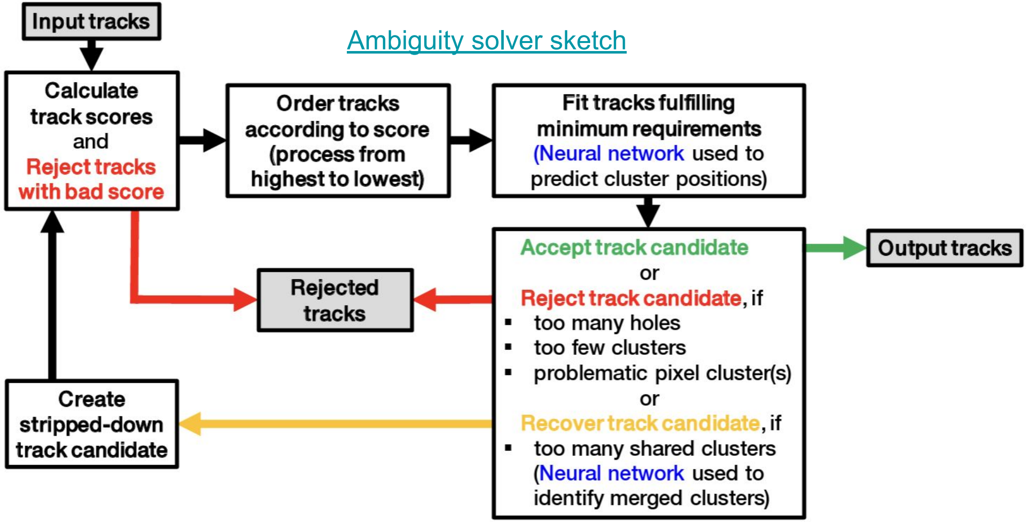
How to actually build a jet (in ATLAS): second round

Creation of tracks

On top of topo-clusters you can also use the track information to build jets

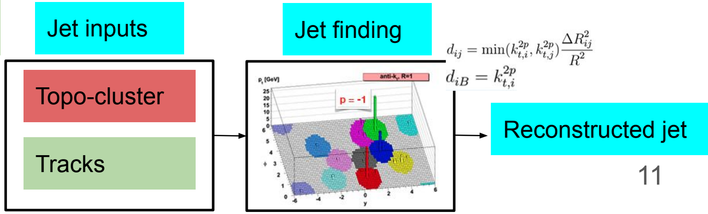
- They are built up with information from the **SCT and pixel detectors**.
- A connected component analysis (CCA) **groups pixels and strips in a given sensor**, where the deposited energy yields a charge above threshold, **with a common edge or corner into clusters**
- From these clusters, three-dimensional measurements referred to as **space-points** are created. The assembling is different for SCT and the pixel detector.
- **Track seeds are formed** from sets of three space-points
- A combinatorial **Kalman filter** is then used to build track candidates from the chosen seeds. Track candidates go through an ambiguity solver and quality cuts to choose those ones with the highest scores.
- Finally, a fit is performed on these tracks

Ambiguity solver sketch



Jet definitions

- Particle Flow Objects (PFO)
- Track Calo Clusters (TCC)
- Unified Flow Objects (UFO)

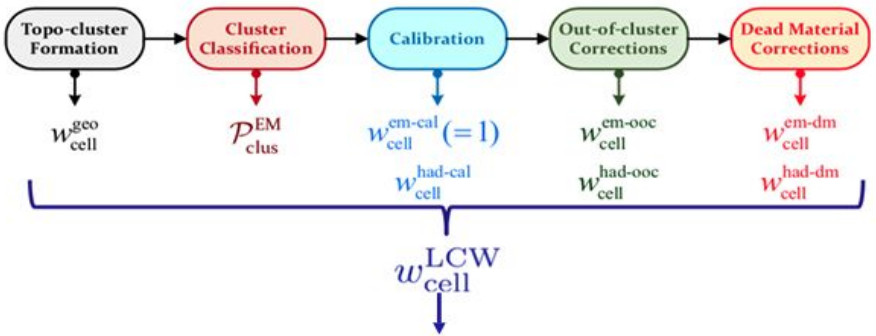


$$d_{ij} = \min(k_{t,i}^{2p}, k_{t,j}^{2p})\frac{\Delta R_{ij}^2}{R^2}$$
$$d_{iB} = k_{t,i}^{2p}$$

How to actually build a jet (in ATLAS): bonus

The topo-cluster can enter the jet finding algorithm either calibrated or uncalibrated

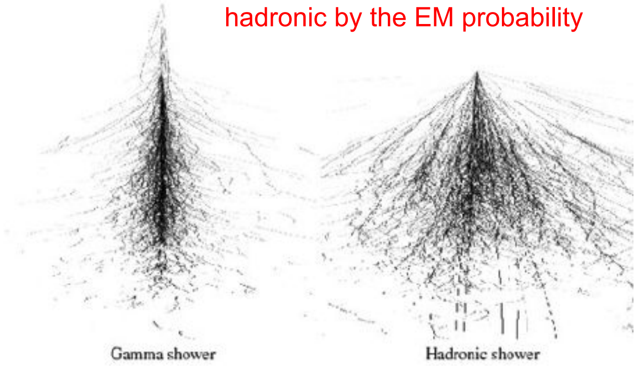
- **Electromagnetic Scale** (EM-Topo): same scale as the cells. Used for **small-R jets.**
- **Local cell weighted (LCW) scale** (LC-Topo): Topo-clusters calibrated based on their properties. Used for **large-R jets** .

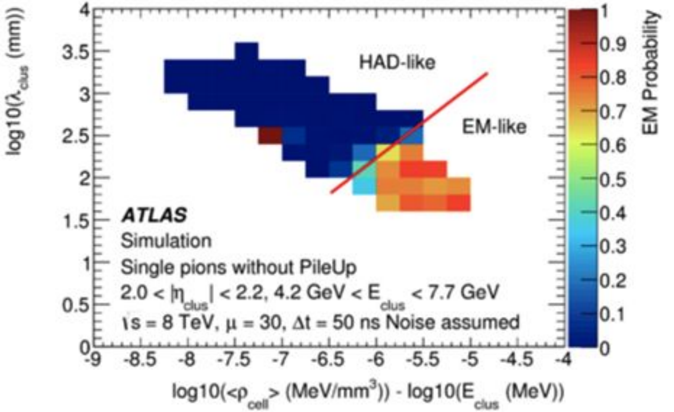


$$E^{\text{LCW}}_{\text{clus}} = \sum_{i \in \text{cluster}} w^{\text{LCW}}_{\text{cell},i} \; E^{\text{EM}}_{\text{cell},i}$$

Weights are then assigned to account for: Differences in detector response due to calorimeter non-compensation, energy falling in unclustered cells, energy deposited in inactive (dead) regions of the detector

Topo-clusters are identified as either electromagnetic or hadronic by the EM probability



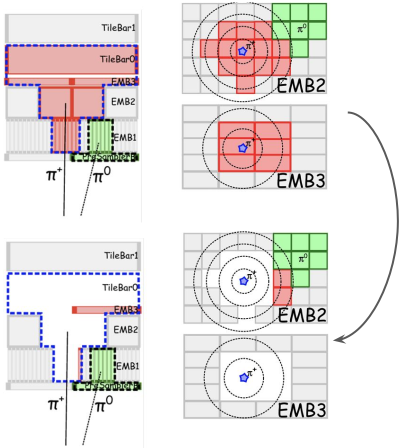


Few words about jet definitions (in ATLAS)

lower pT small-R jets

Particle Flow Objects (PFlow)

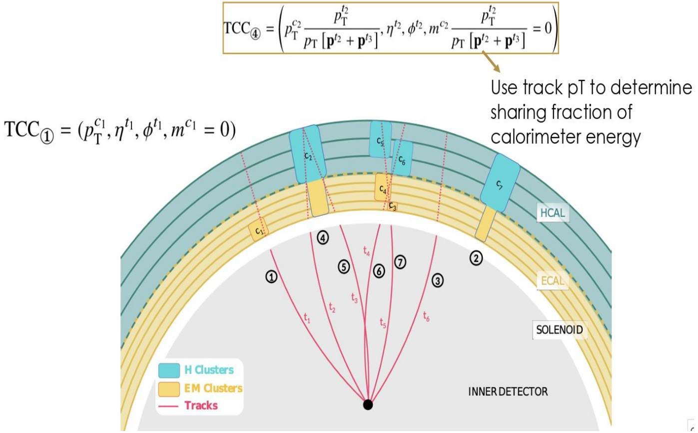
- The tracks are matched to one or several topo-clusters.
- Then track energy deposit is subtracted from topo-clusters together with the remnant cluster energy.
- The output of the algorithm are tracks, updated topo-clusters (with the subtraction applied) or just the same topo-cluster if no matching with tracks was found.



boosted environments (Large-R jets)

Track Calo Clusters (TCC)

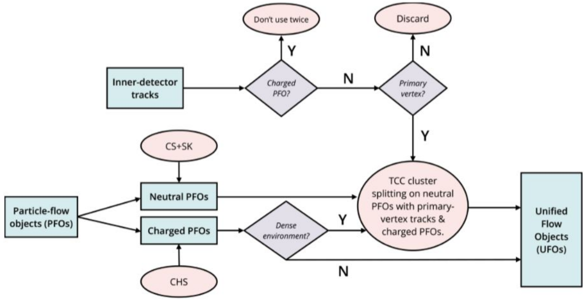
- Match tracks to topo-clusters.
- A 4-vector is built up of matched objects, with dedicated calculation of energy sharing between tracks and calorimeter deposits.
- Take spatial information from tracks (phi, eta) and calorimeter information with energy deposition.
- Improve angular resolution of high pT particles



best of two (?)

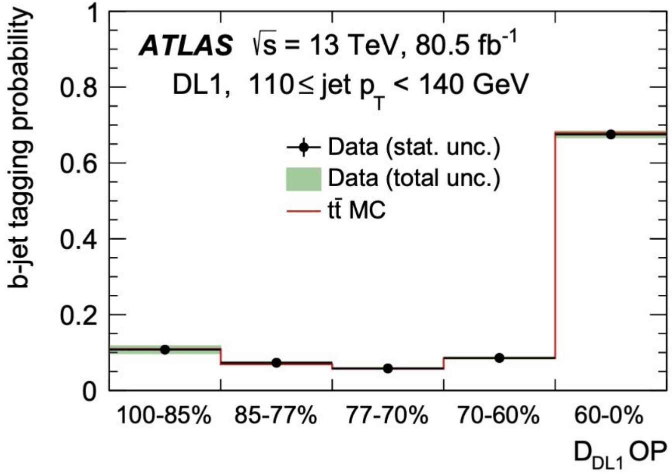
Unified Flow Objects (UFO)

- UFO algorithm uses PFlow and TCC inputs to build the jet

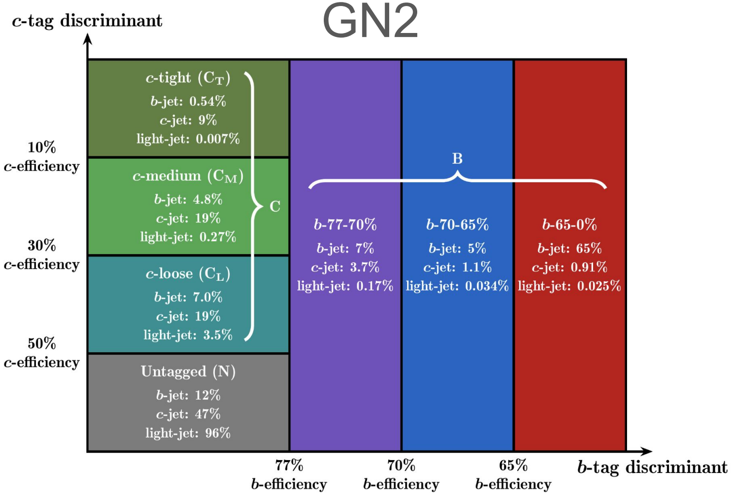


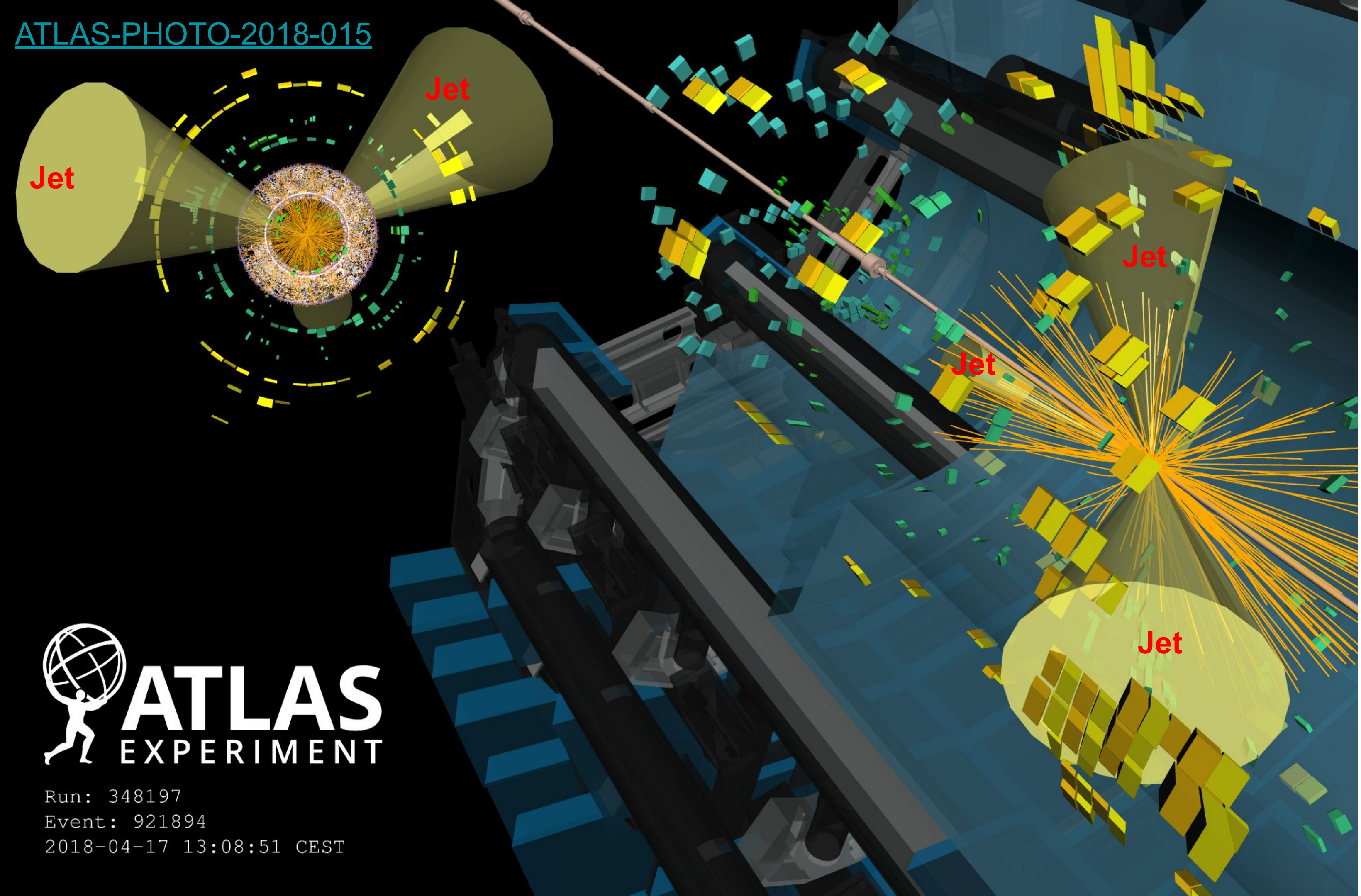
A word on pseudocontinuous btagging (https://ftag.docs.cern.ch/calibrations/cdi/pcbt/)

Below you can see an example PCBT tag-weight distribution (left) for a particular p_T bin. At the analysis level, you can utilize the tag-weight distribution, for example, at the **exclusive** working-point of $70 - 60\%$. What this *means* is that you will tag a jet if it's tagger-score falls in that particular PCBT bin. In the fixed-cut working-point, one would use an **inclusive** working-point of 70%, which would also tag a jet if its' score falls in the $60 - 0\%$ bin. **Note: you should be able to reconstruct the inclusive efficiency for a particular fixed-cut working-point by summing over the PCBT efficiency bins that are below that working-point.**



GN2




ATLAS-PHOTO-2018-015
Jet
Jet
Jet
Jet
Jet
ATLAS
EXPERIMENT
Run: 348197
Event: 921894
2018-04-17 13:08:51 CEST