Jet reconstruction (experimental)
Part 1

Steven Schramm
University of Geneva

PREFIT School, DESY
March 10, 2020

Jet reconstruction: an experimental perspective

UNIVERSITÉ DE GENÈVE

- Gavin has already told you about jets and the algorithms we use to build them
- Experimentally, jets are a tool to represent hadronic showers in a detector
- We primarily use "small-R" jets to represent quarks and gluons
  - We typically use $R = 0.4$ anti-k_t jets to do this (both ATLAS and CMS)
- In some cases, we use "large-R" jets to represent more complex objects: W, Z, H, top
  - ATLAS uses $R = 1.0$ anti-k_t jets, while CMS uses $R = 0.8$ anti-k_t jets

- The main difference from what Gavin discussed is the importance of the detector
  - We work with "reconstructed" jets, where we run jet building on detector objects
  - The vast majority of our time goes to understanding the resulting detector features
- These slides will have an ATLAS bias, but the concepts are generally applicable

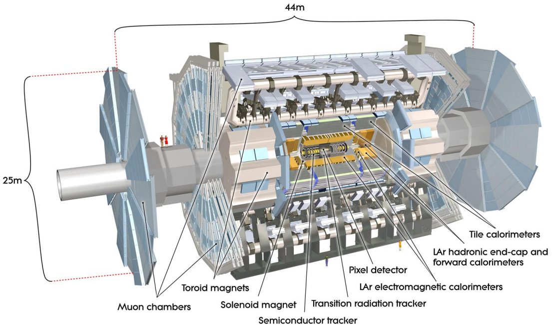
Jet reconstruction steps

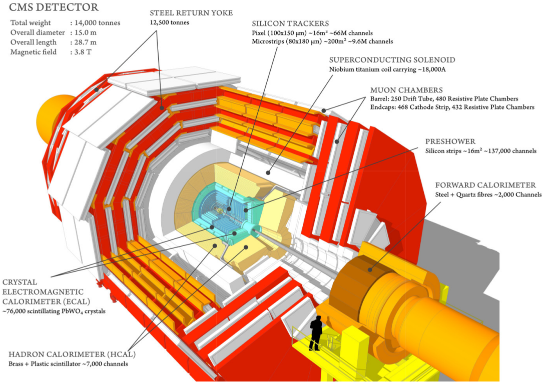
1. Detectors and pileup
   - How jets are observed and the big LHC challenge
2. Building robust inputs to jet reconstruction
   - These are the four-vectors that the jet algorithms are run on
3. Calibrating the resulting jets
   - Trying to remove experimental effects and make reconstructed jets look like truth jets
4. Correcting data vs simulation differences
   - Accounting for mismodelling and defining uncertainties on jet quantities

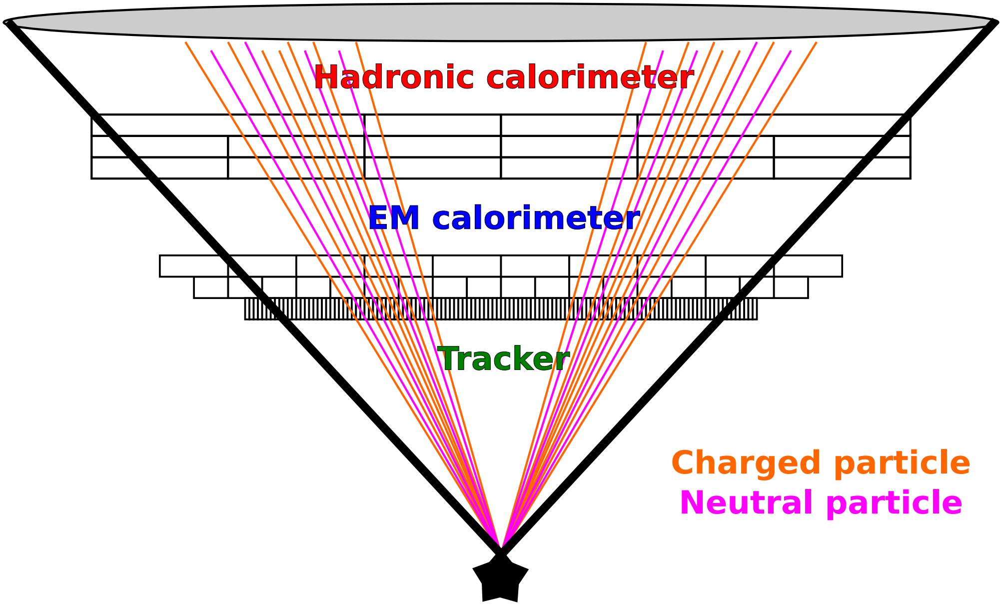
The detectors

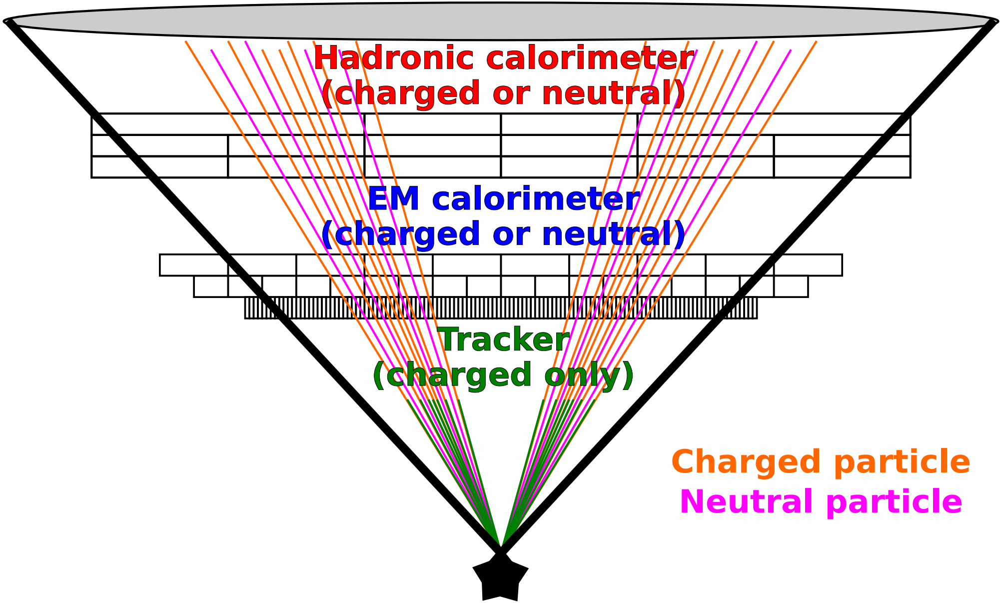
UNIVERSITÉ DE GENÈVE

- From a high-level view, they are very similar detectors
  - Inner tracker, electromagnetic calorimeter, hadronic calorimeter, then the muon system
- Modern experimental jet usage relies heavily on the tracker and both calorimeters

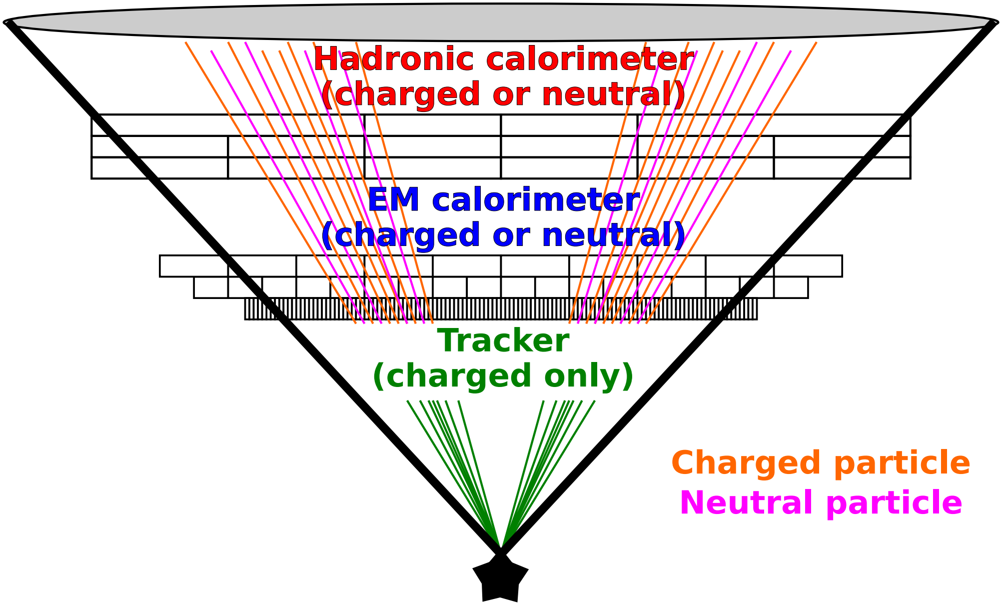


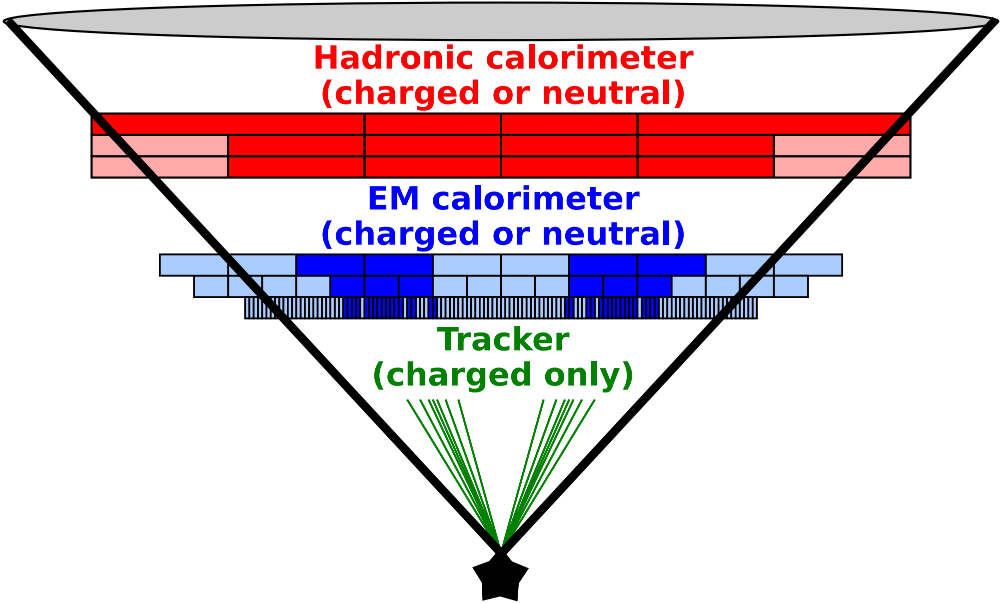



Hadronic calorimeter
EM calorimeter
Tracker
Charged particle
Neutral particle

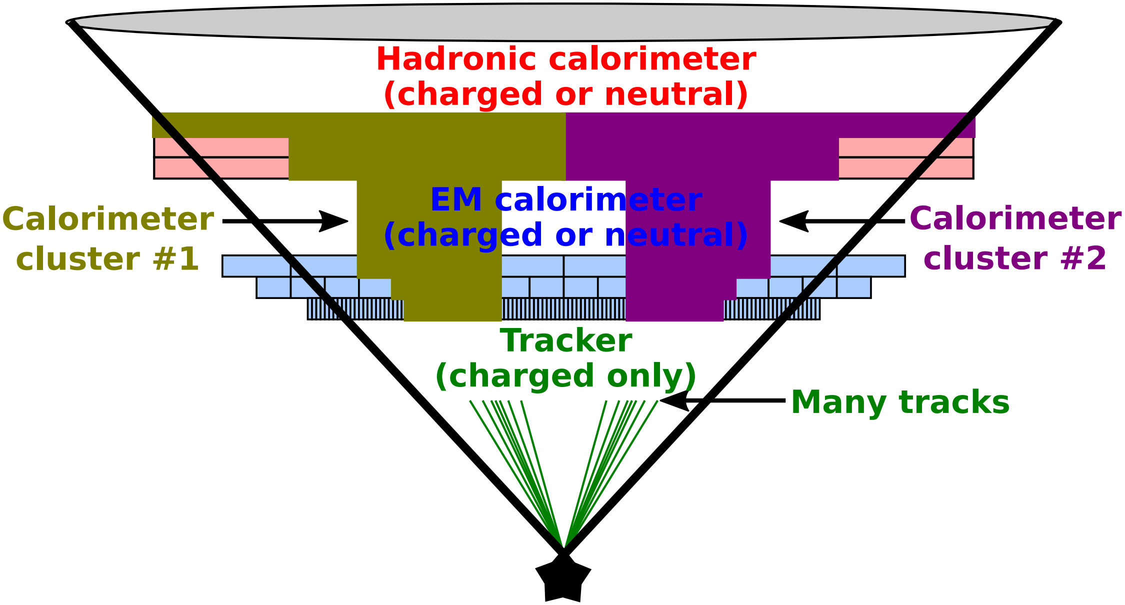


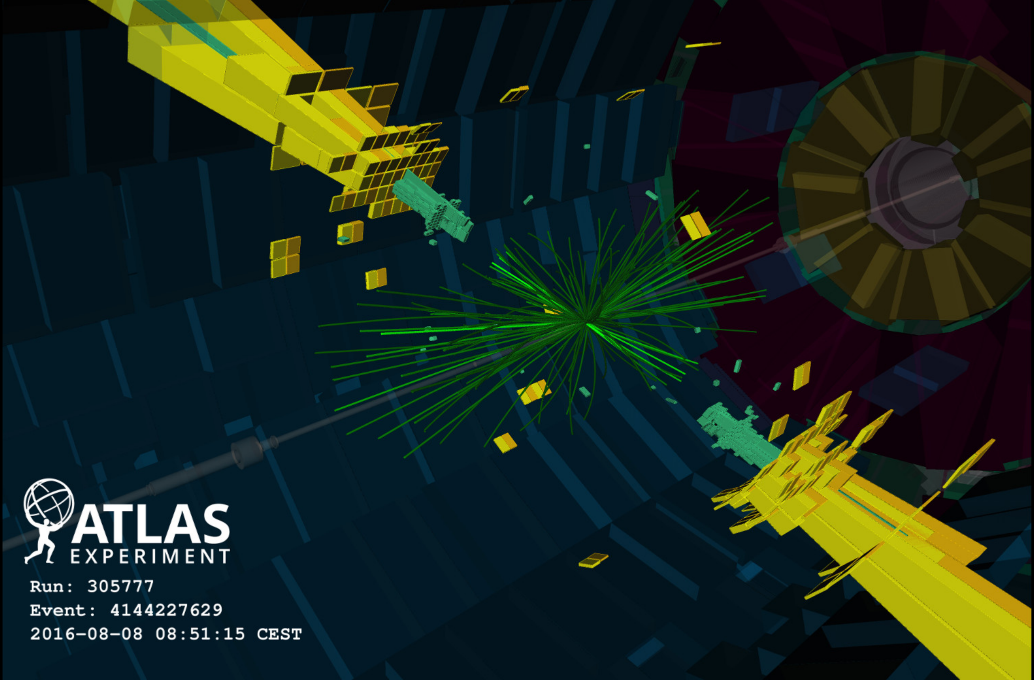
Hadronic calorimeter
(charged or neutral)
EM calorimeter
(charged or neutral)
Tracker
(charged only)
Charged particle
Neutral particle


Hadronic calorimeter
(charged or neutral)
EM calorimeter
(charged or neutral)
Tracker
(charged only)
Charged particle
Neutral particle



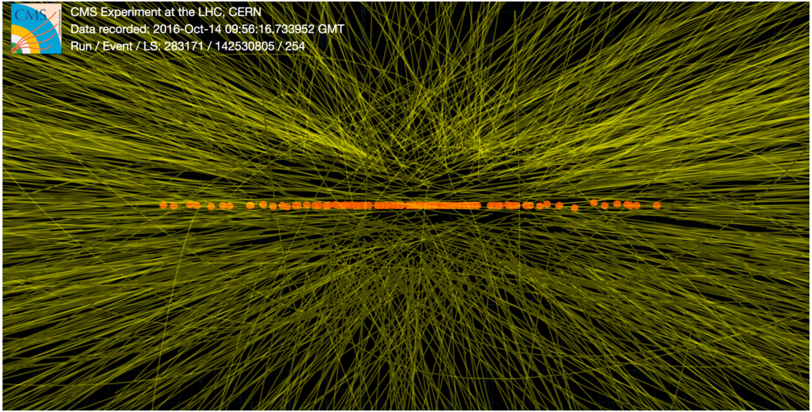
Hadronic calorimeter
(charged or neutral)
EM calorimeter
(charged or neutral)
Tracker
(charged only)


Hadronic calorimeter
(charged or neutral)
EM calorimeter
(charged or neutral)
Calorimeter
cluster #1
Calorimeter
cluster #2
Tracker
(charged only)
Many tracks


ATLAS
EXPERIMENT
Run: 305777
Event: 4144227629
2016-08-08 08:51:15 CEST

The LHC and pileup

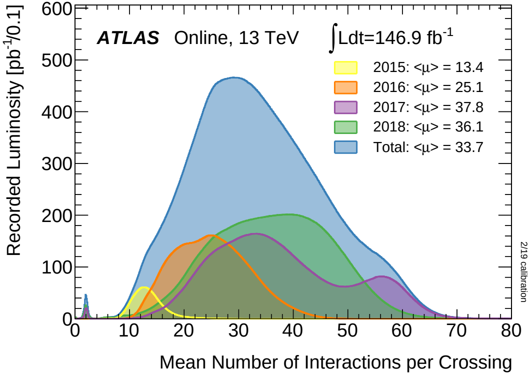
- The last slides were a very simple picture of what happens, showing a single collision
- In reality, we now have **many** simultaneous proton-proton collisions
  - Below is real data from a high-pileup run in 2016, each point is an independent collision



Pileup in Run 2 (2015–2018)

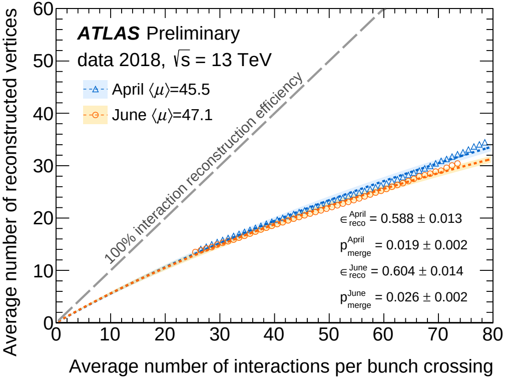
UNIVERSITÉ DE GENÈVE

- In Run 2, the average number of collisions per beam crossing was ~ 34
  - This is typically called μ
- Pileup is key to the huge LHC dataset
  - More simultaneous collisions $=$ higher probability of a rare event
  - More rare events $=$ improved measurement of rare processes (Higgs), increased sensitivity to new physics
- This number is expected to grow to an average of 60 for Run 3 (2021–2024)



The number of vertices

- Tracking detectors can be used to identify individual collision vertices
- Not all vertices can be reconstructed
  - Too low p_{T} to be observed
  - No charged part. in tracker acceptance
  - Vertices can merge (overlap)
  - Other effects are also possible
- In ATLAS, $\sim$60% vertex reco efficiency
- The number of vertices is called N_{PV}



Types of pileup effects

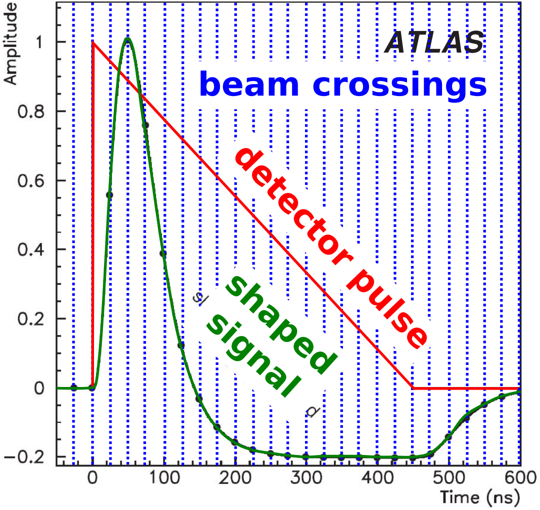
UNIVERSITÉ DE GENÈVE

- We have now seen two different observables which quantify pileup
  - μ: the **average** number of collisions in a small window of time
  - N_{PV}: the observed number of collisions in a **single event**
- There are also two types of pileup effects in the detector
  - In-time pileup: simultaneous collisions, related to N_{PV}
  - Out-of-time pileup: remnants of signals from previous or subsequent collisions, related to μ
- In-time pileup is pretty clear from the picture with all of the vertices
  - Easy to confuse overlapping contributions from different collisions
- Out-of-time pileup is a bit less obvious: why should previous/subsequent collisions matter?

Calorimeter read-out time

UNIVERSITÉ DE GENÈVE

- What if detector readout time is longer than the space between beam crossings?
  - Previous/subsequent events can overlap with the current event
- Space between LHC collisions: 25 ns
- ATLAS LAr calorimeter readout: 450 ns
  - That's 18 beam crossings!
  - Reduce impact using shaping function
- Most common example at the LHC due to having such a long readout time
  - Other ATLAS/CMS calorimeters less impacted due to shorter readout times



Detectors and pileup

- In a detector, jets are a complex mixture of visible signals
  - Individual charged particles appear in the tracker
  - Charged and neutral particles appear together (but not distinguishable) in the calorimeter
- Pileup makes the situation more complex
  - Typically quantified using N_{PV} (in-time) and μ (out-of-time)
- Signals from different processes can overlap in the calorimeter (in-time pileup)
- Some calorimeters also have very long readout times and are thus sensitive to previous/subsequent events (out-of-time pileup)

Building inputs to jet reconstruction

Building inputs to jet reconstruction

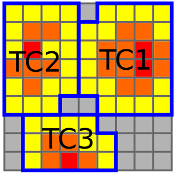
UNIVERSITÉ DE GENÈVE

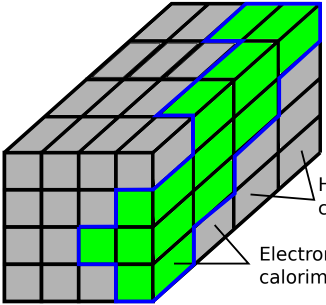
- Jets are comprised of $\sim 2/3$ charged particles and $\sim 1/3$ neutral particles
  - The tracker only sees the charged particles, while the calorimeter sees both types
  - Energy deposits in the calorimeter are thus the key to jet reconstruction
- However, as we just saw, the calorimeter can be susceptible to pileup
  - The amount of pileup sensitivity depends on the specific calorimeter
  - In general, calorimeters are always more sensitive to pileup than the tracker
- Calorimeter objects are designed to reduce pileup effects, within limits
  - Ultimately, we need to combine information from trackers and calorimeters

Calorimeter objects

UNIVERSITÉ DE GENÈVE

- Clusters of calorimeter energy are the classic object used for jet reconstruction
  - ATLAS approach: "topo-clusters"
  - Allows for dynamic sizes, unlike towers
  - Use each cell's expected noise to determine whether to include it or not
  - $\sigma = \sigma_{\text{electronic}} \oplus \sigma_{\text{pileup}}$, $\sigma_{\text{pileup}} \gg \sigma_{\text{electronic}}$
- Nonetheless, pileup still impacts clusters
  - If pileup and hard scatter directly overlap, their energies are combined, changing the measured cluster energy
  - Pileup can also create separate clusters

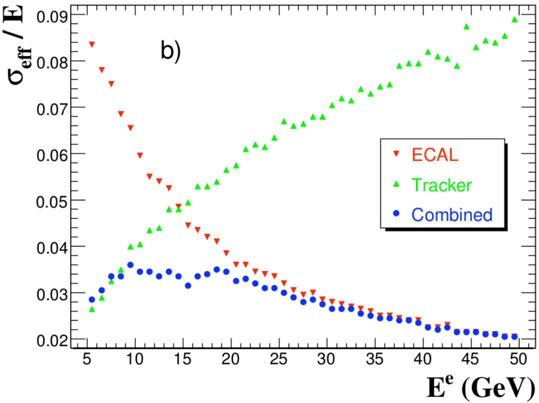




Using tracks for jet reconstruction

UNIVERSITÉ DE GENÈVE

- Tracks are "inherently" pileup stable
- As mentioned, roughly 2/3 of a jet is comprised of charged particles
  - Tracks are thus only 2/3 of the energy
  - Alone, not great for jet reconstruction
- Tracks also have inverted measurement sensitivity compared to the calorimeter
  - Tracks are best at low p_T
    - Ability to measure track curvature degrades with p_T
  - Calorimeter is best at high p_T
    - Fluctuations less relevant at high p_T
- Maybe we can combine tracks and calo?



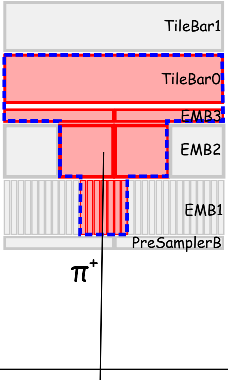
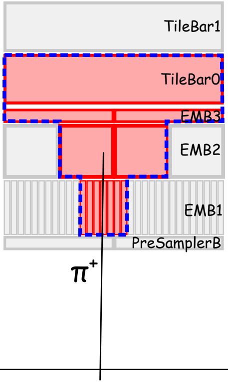
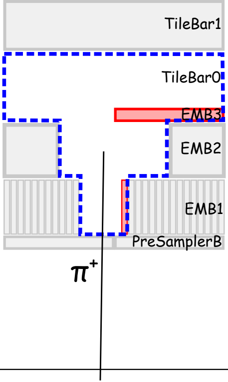
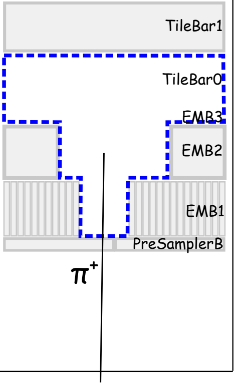
Particle flow

UNIVERSITÉ DE GENÈVE

- We saw on the last slide that trackers and calorimeters can be complementary
  - Trackers are pileup-robust, while calorimeters are less so
  - For charged particles, the tracker is best at low p_{T} and calo is best at high p_{T}
  - For neutral particles, we always need the calorimeter
- We can thus consider mixtures of tracks and calorimeter clusters
  - This requires *matching* tracks to clusters so we don't double-count energy
- Once we have matched track(s) to a cluster, check if their measurements are consistent
  - Consistent: the cluster likely corresponds to only charged particle(s)
  - Inconsistent: the cluster likely corresponds to overlapping charged and neutral particles
- This process is typically called "particle flow" (PFlow)
  - Particle flow is **very** detector-specific, and each collaboration has their own algorithm
  - However, the general idea and approach follows the above mentality

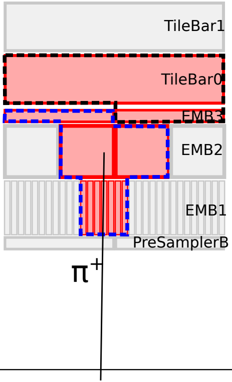
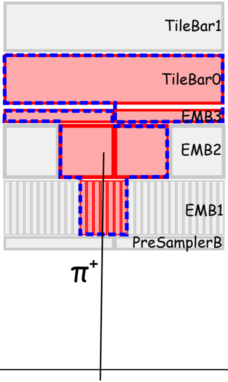
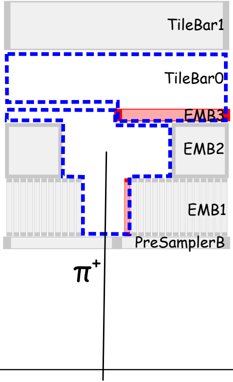
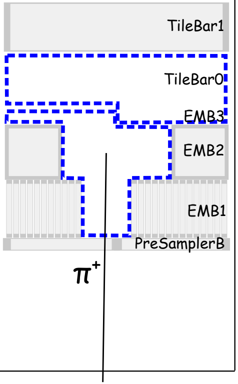
Particle flow examples: 1 particle, 1 cluster

UNIVERSITÉ DE GENÈVE

	Track/topo-cluster matching	Split shower recovery	Cell subtraction	Remnant removal
1 particle, 1 topo-cluster				

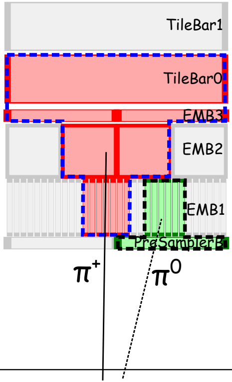
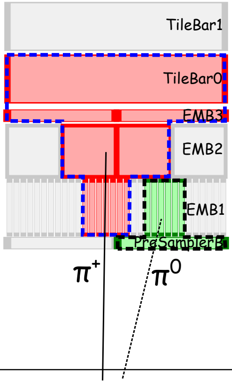
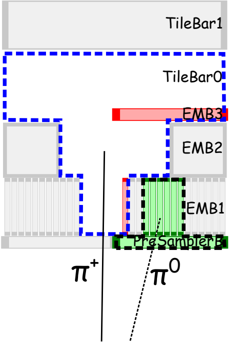
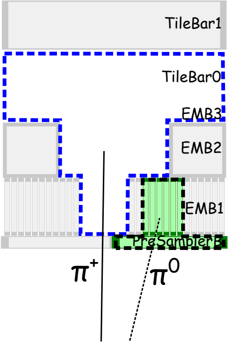
Particle flow examples: 1 particle, 2 clusters

UNIVERSITÉ DE GENÈVE

<table>
<tr><th></th><th>Track/topo-cluster matching</th><th>Split shower recovery</th><th>Cell subtraction</th><th>Remnant removal</th></tr>
<tr><td>1 particle,
2 topo-clusters</td><td>
</td><td>
</td><td>
</td><td>
</td></tr>
</table>

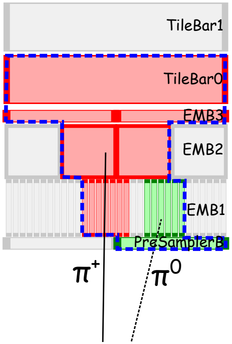
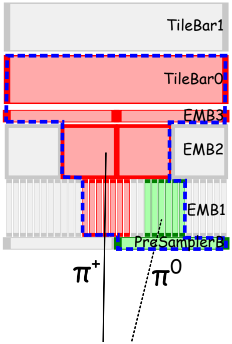
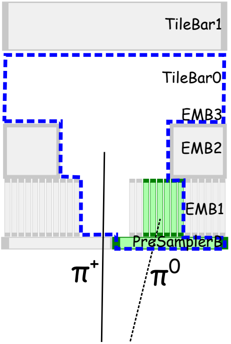
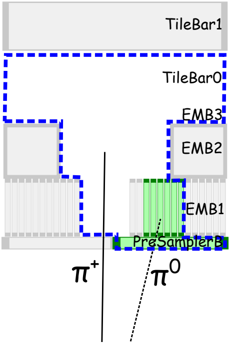
Particle flow examples: 2 particles, 2 clusters

UNIVERSITÉ DE GENÈVE

	Track/topo-cluster matching	Split shower recovery	Cell subtraction	Remnant removal
2 particles, 2 topo-clusters				

Particle flow examples: 2 particles, 1 cluster

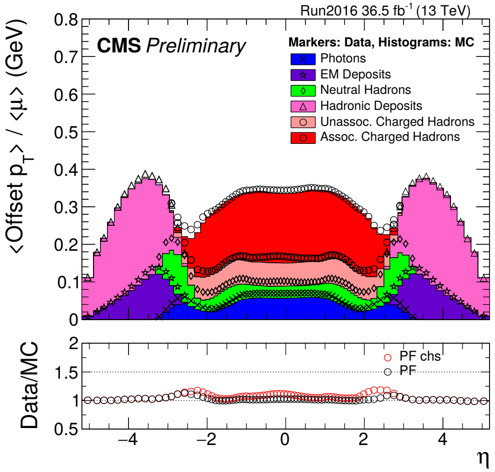
UNIVERSITÉ DE GENÈVE

	Track/topo-cluster matching	Split shower recovery	Cell subtraction	Remnant removal
2 particles, 1 topo-cluster				

Benefits of particle flow

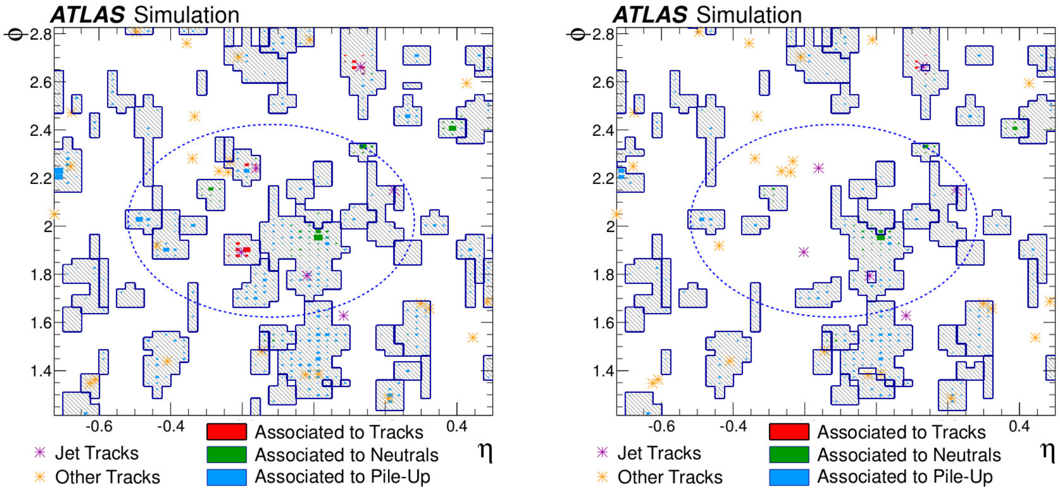
UNIVERSITÉ DE GENÈVE

- In the absence of pileup, PFlow already improves low p_T jet performance
  - Better measurement of low p_T particles
  - Exploits the complementary nature of two different types of detectors
- PFlow also naturally suppresses pileup
  - Matching of tracks to clusters identifies the originating vertex
  - Charged PFlow objects from other vertices can be naturally removed
  - This is “Charged Hadron Subtraction”



Before and after Charged Hadron Subtraction (CHS)

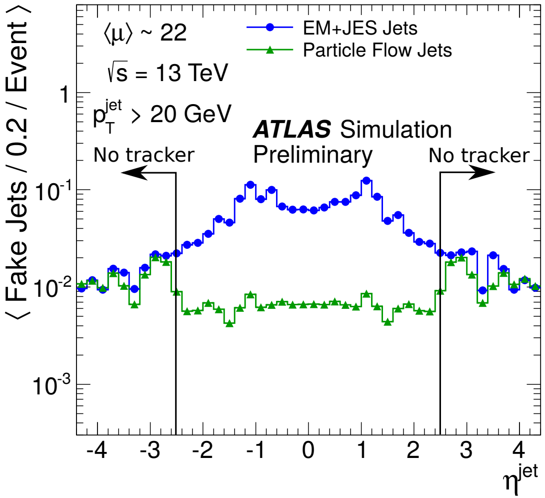
UNIVERSITÉ DE GENÈVE



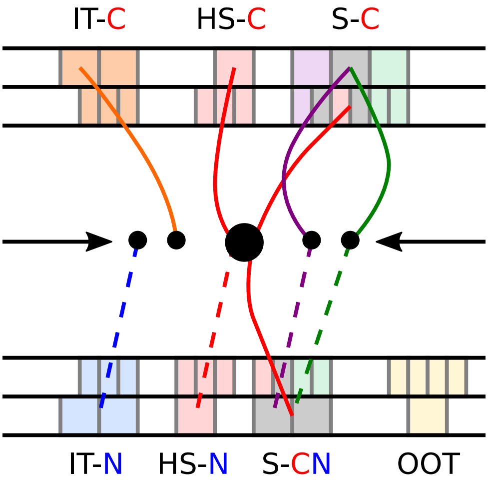
The impact of Charged Hadron Subtraction

UNIVERSITÉ DE GENÈVE

- Consider the average number of pileup jets (reco jets not matched to truth jets)
  - PFlow should reduce pileup jets
- Clear benefit using PFlow instead of calorimeter-only jets within the tracker
  - Tracker acceptance: $|\eta| < 2.5$
  - Mixture of improved jet p_T measurement and pileup suppression
  - Note that the plot is log-scale



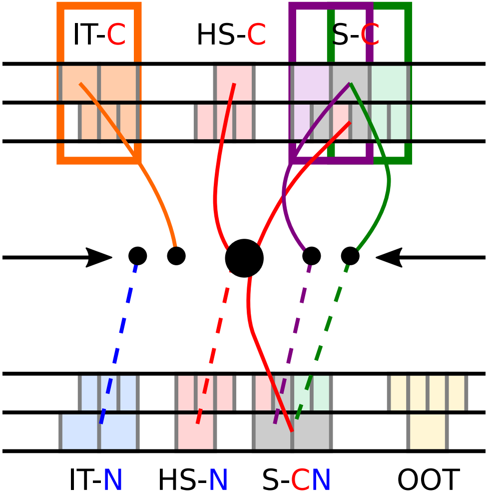
Pileup and jet inputs: different possibilities



HS = hard scatter
IT = in-time pileup
S = stochastic pileup
OOT = out-of-time pileup
-
C = charged energy flow
N = neutral energy flow

What about Charged Hadron Subtraction (CHS)?

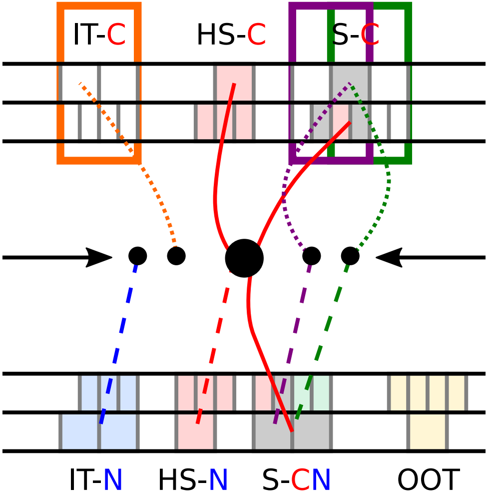
UNIVERSITÉ DE GENÈVE



HS = hard scatter
IT = in-time pileup
S = stochastic pileup
OOT = out-of-time pileup
-
C = charged energy flow
N = neutral energy flow

Removed using CHS (energy deposits)

Pileup and jet inputs: post-CHS

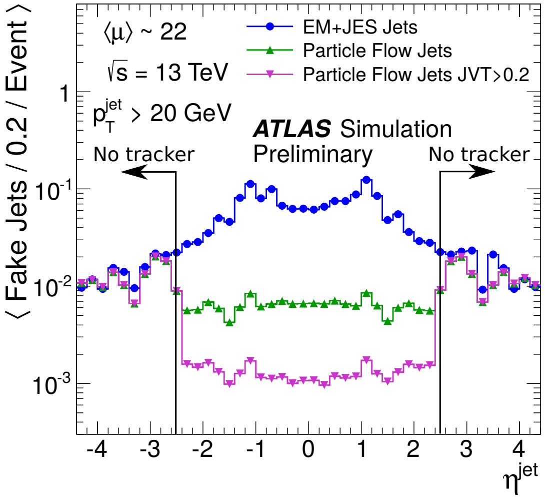


HS = hard scatter
IT = in-time pileup
S = stochastic pileup
OOT = out-of-time pileup
-
C = charged energy flow
N = neutral energy flow

Removed using CHS (energy deposits)

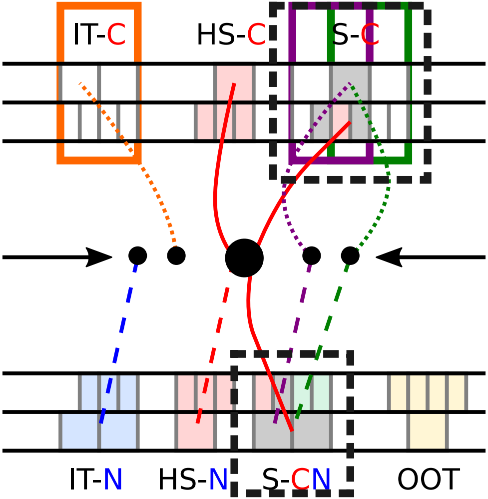
Moving beyond Charged Hadron Subtraction (CHS)

- The CHS technique has already been very useful, but what more can we do?
  - Stochastic overlap of multiple charged signals is challenging to suppress
- Instead of considering only the input objects, consider the collective effects
  - Example: JVF = Jet Vertex Fraction
  - Sum p_T of all tracks matched to jets
  - Cut on fraction of momentum from the vertex of interest vs all vertices
- Plots shows JVT, similar idea to JVF
  - Also uses expectation of $\sum \text{track}/p_T^{\text{jet}}$
  - Helps with stochastic neutral pileup



Pileup and jet inputs: adding JVT

UNIVERSITÉ DE GENÈVE



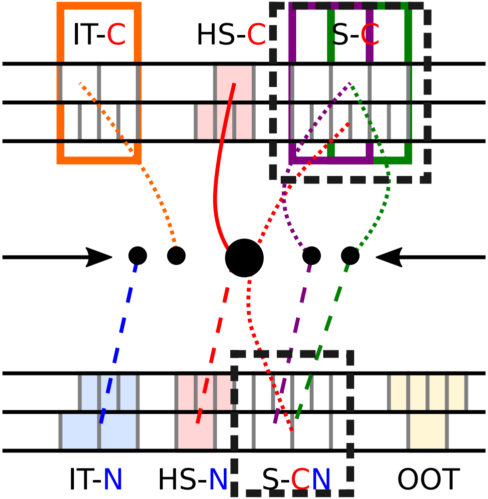
HS = hard scatter
IT = in-time pileup
S = stochastic pileup
OOT = out-of-time pileup
-
C = charged energy flow
N = neutral energy flow

Removed using CHS (energy deposits)

Removed using JVT (entire jets)

Pileup and jet inputs: post-JVT

UNIVERSITÉ DE GENÈVE



HS = hard scatter
IT = in-time pileup
S = stochastic pileup
OOT = out-of-time pileup
-
C = charged energy flow
N = neutral energy flow

Removed using CHS (energy deposits)

Removed using JVT (entire jets)

What types of pileup are we left with?

UNIVERSITÉ DE GENÈVE

- CHS and JVT are a powerful pair of tools to suppress pileup
  - CHS removes individual in-time pileup deposits from charged particles
  - JVT helps suppress stochastic (overlapping object) pileup:
    - $\sum$track(vertex of interest)/$\sum$track(all vertices): suppresses overlapping in-time charged pileup
    - $\sum$track/$p_{\mathrm{T}}^{\mathrm{jet}}$: suppresses overlapping out-of-time pileup and in-time neutral pileup
- JVT is not a perfect discriminant: some good hard-scatter jets are also removed

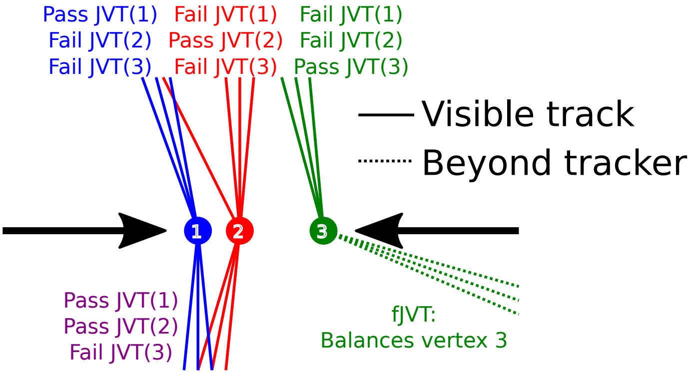
- What remains is jets built of exclusively out-of-time pileup and neutral pileup
  - For in-time pileup, this is quite rare: the vast majority of jets have tracks
  - Independent out-of-time jets are mostly suppressed by calorimeter pulse shapes or jet cleaning

- That, and anything outside of the tracker; we've only discussed areas with tracks!

What about the forward region?

UNIVERSITÉ DE GENÈVE

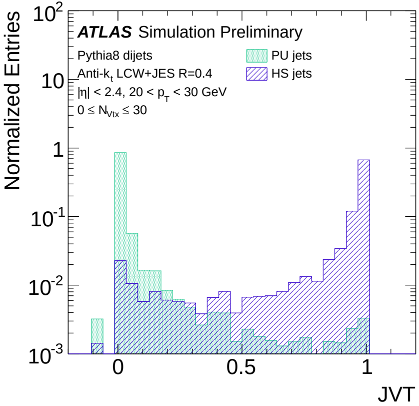
- Where we don't have tracks, our ability to reject pileup is dramatically reduced
- Instead, use tracks in central region to determine momentum imbalance of each vertex
  - Pileup vertices are mostly dijet events, where one jet may be central and one forward
  - If a forward jet resolves the vertex imbalance, the jet is likely from that vertex

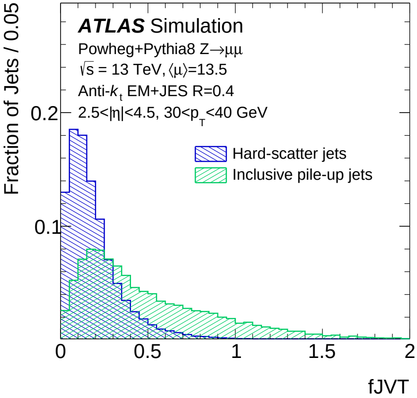


Comparing JVT and fJVT

UNIVERSITÉ DE GENÈVE

- Clearly **much better** separation between hard-scatter and pileup within the tracker
  - However, vertex balance information can reject some forward pileup jets

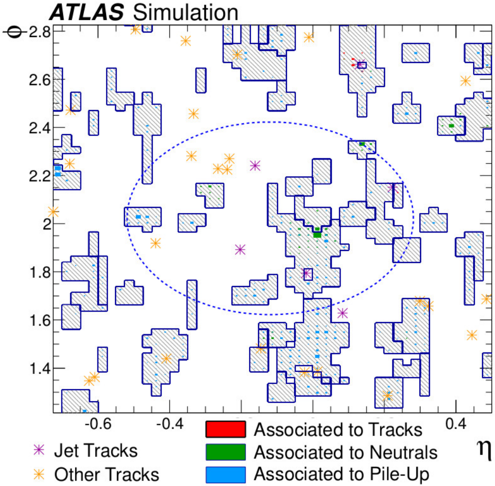




Inputs to jet reconstruction

UNIVERSITÉ DE GENÈVE

- Discussed clusters, tracks, and PFlow
  - PFlow exploits the complementary nature of trackers and calorimeters
  - Also suppresses pileup via CHS
- Saw how to remove pileup jets with JVT
  - However, it can remove good jets too
- Note: we only removed individual energy deposits from charged in-time pileup
  - Energy deposits from neutral in-time pileup remain within hard-scatter jets
  - Mitigating the impact of this neutral energy is left to jet calibration



Calibrating the resulting jets

Jet reconstruction

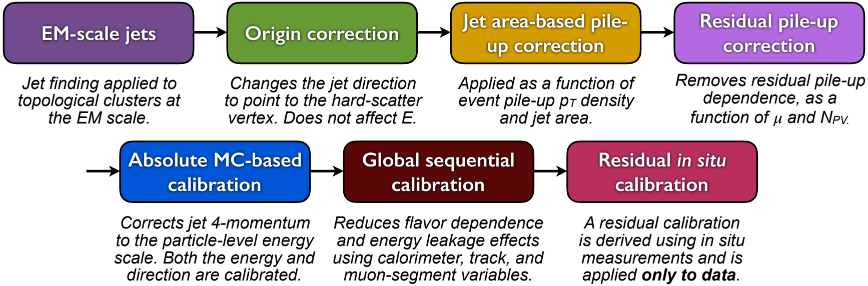
UNIVERSITÉ DE GENÈVE

- Now that we have finalized our input objects, we can run jet reconstruction
  - Inputs: either calorimeter clusters or PFlow objects
  - Jet algorithm: anti-k_t, meaning we start with the most energetic inputs
  - Jet radius parameter: $R = 0.4$, as we're focusing on small-R jets today

- Recall that unlike electrons and muons, jets are not a physical particle
  - They are a useful tool to represent the hadronization of quarks and gluons
  - They are only defined by the algorithm used to build them
- As such, the "correct" jet energy depends on the algorithm
  - Need to run the same algorithm on both truth and reco jets

Jet calibration

UNIVERSITÉ DE GENÈVE



- Jet calibration consists of many different steps, addressing different experimental effects
  - Aiming to get back to the truth jet scale, as there is no absolute reference scale for jets
- Most calibration steps thus involve comparing reco and truth jets in MC
  - The final box (*in situ* calibration) is the only one where data is involved [next section]

Optimization metrics

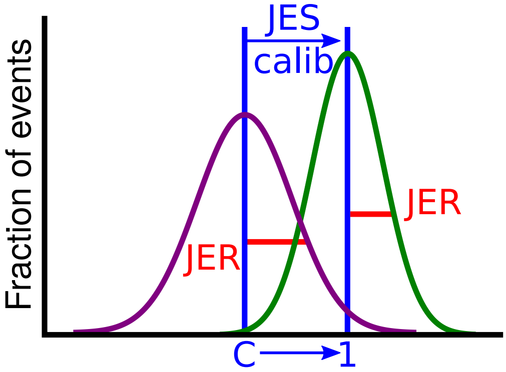
UNIVERSITÉ DE GENÈVE

- In jet calibration, we are always comparing to the relevant truth jet
  - This requires matching reconstructed jets to their originating truth jet
  - This is typically done using $\Delta R = \sqrt{(\eta_{\text{truth}} - \eta_{\text{reco}})^2 + (\phi_{\text{truth}} - \phi_{\text{reco}})^2}$, such as $\Delta R < 0.3$
- Once a reco jet is matched to its truth jet, two quantities may be calculated:
  - The response, $X^{\text{reco}}/X^{\text{true}}$, where X is most commonly E or p_{T} (variables with a scale)
  - The difference, $X^{\text{reco}} - X^{\text{true}}$, where X is most commonly η or ϕ (variables without a scale)

- The response in particular is very important, and it is typically a Gaussian distribution
  - The mean of the Gaussian is referred to as the scale (typically Jet Energy Scale, JES)
  - The width of the Gaussian is called the resolution (typically Jet Energy Resolution, JER)

JES and JER

- JES = the mean of the Gaussian
  - Ideal: a value of 1, so the average reco jet is at the truth energy scale
  - “Easy” to make this happen, just multiply reco jet energy by 1/response
  - Complications come from this quantity varying dramatically across detector
- JER = the width of the Gaussian
  - Ideal: a delta-function, so every reco jet is exactly at the truth jet energy
  - In practice, this will never happen
  - Reality: as narrow as possible

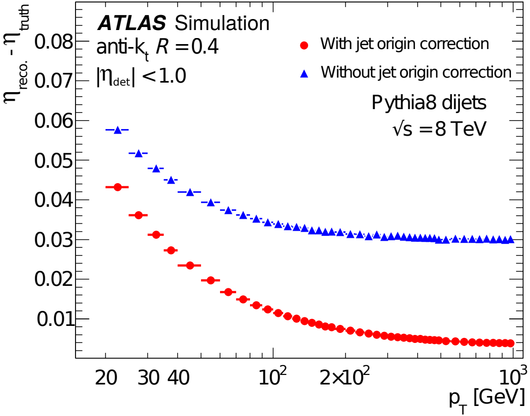


$$\text{Response} = \frac{E^{\mathrm{reco}}}{E^{\mathrm{true}}}$$

Origin correction

UNIVERSITÉ DE GENÈVE

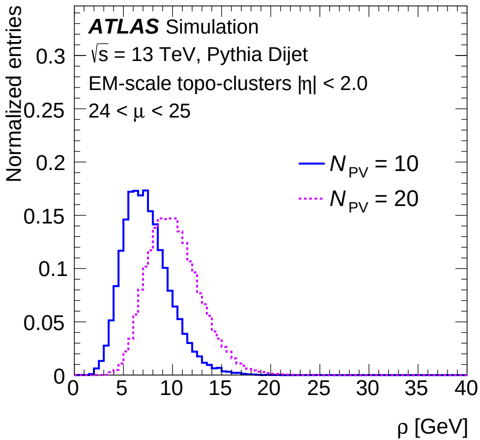
- First calibration step in the past
  - Correct the jet axis to point to the primary vertex of interest
  - Currently applied directly to clusters
- By default, jets/clusters point to geometric center of the detector
  - Pointing at the vertex improves the subsequent angular precision
- Large gains for jet η
  - Negligible impact for jet ϕ due to geometric considerations



Pileup subtraction

UNIVERSITÉ DE GENÈVE

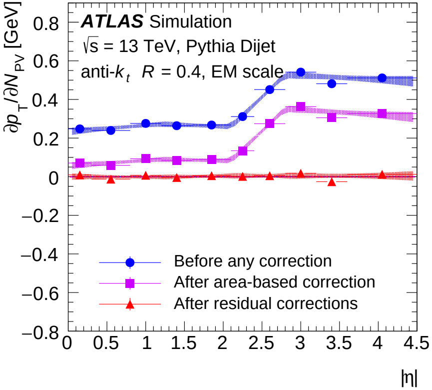
- CHS only removes charged pileup energy
  - Pileup corrections are used to account for neutral pileup energy expectation
- $p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco}} - \rho \times \mathcal{A} - \alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \mu$
  - $p_{\mathrm{T}}^{\mathrm{reco}}$: starting jet energy
  - ρ: per-event energy density, expected amount of pileup energy per unit area
  - $\mathcal{A}$: the area of the jet in question
  - α and β: constants, derived to correct the mean of a dataset (not per-event)
- ρ is calculated as the median of $R = 0.4$ k_t jet energy density, built for $|\eta| < 2.0$

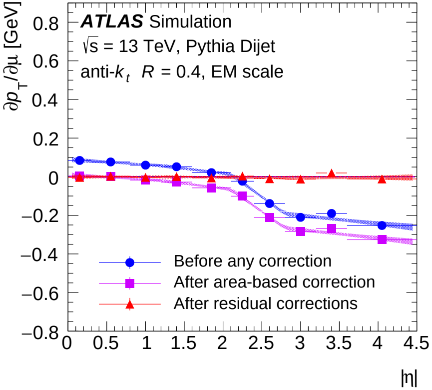


Pileup correction: dependence after each step

UNIVERSITÉ DE GENÈVE

- Plots are shown per unit of N_{PV} (left) or μ (right), after different PU correction steps
  - $\rho \times \mathcal{A}$ does reasonably well in the central region where energy density is roughly constant
  - Really need α and β to help the more complex forward region (very different energy density)

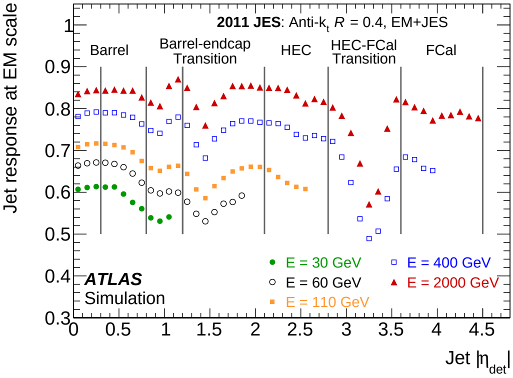




MC JES: ATLAS, EM jets

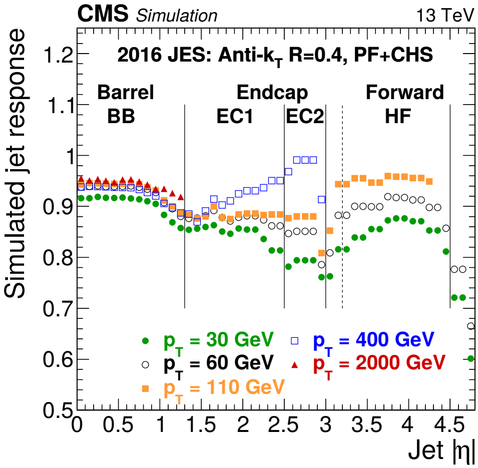
UNIVERSITÉ DE GENÈVE

- A large fraction of the jet energy is not visible in the calorimeter
  - Inactive material, noise thresholds, etc
- The MC JES is the main step that corrects for this missed energy
- Calorimeter structure is clearly visible
  - More inactive material in transition regions $\Longrightarrow$ more lost energy
- Also shows clearly increased sensitivity (higher response) for high energy jets
  - The more energy there is, the better the calorimeter can measure it



MC JES: CMS, PFlow jets

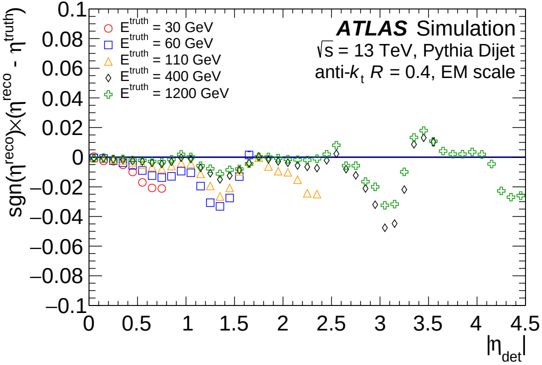
- CMS is a different detector
  - Structures show up in different places
- This plot is also showing PFlow jets
  - Last slide was calorimeter-only jets
  - PFlow improves the response dramatically at low p_T (the tracker sees the full charged energy)
    - Energy spread reduced for $|\eta| < 2.5$
- Despite differences, intent is the same
  - Correct for energy in the truth jet which was not observed by the detector



MC JES: η correction

UNIVERSITÉ DE GENÈVE

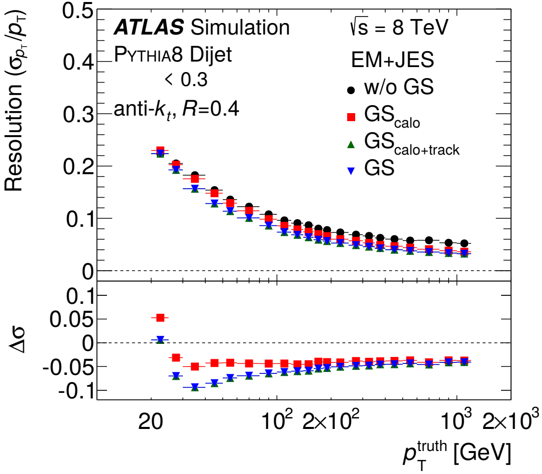
- Additional correction to jet η done together with the MC JES
  - Jet axis is defined as energy-weighted centroid $\Longrightarrow$ jets in a low-response area appear in adjacent regions
- A shift in η corrects for this bias
  - Results in large differences in calorimeter transition regions



Global Sequential Calibration (GSC)

UNIVERSITÉ DE GENÈVE

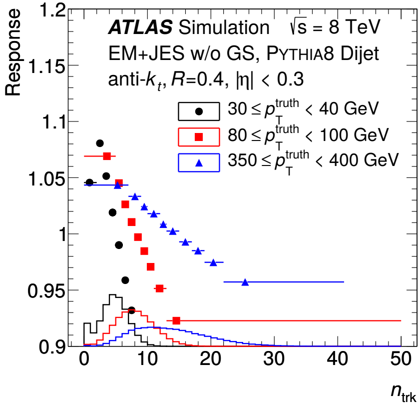
- MC JES is complete, we're nearly done
  - The scale is now correct
  - Let's improve the resolution
- Recall: resolution is the response width
- How can we narrow the width?
  - If we can find variables that the response depends on, we can exploit them to narrow the peak
- Looking for sub-populations of jets
  - Sub-populations = offset Gaussians
  - Adjusting their scales to match $\implies$ improved resolution [aligned Gaussians]

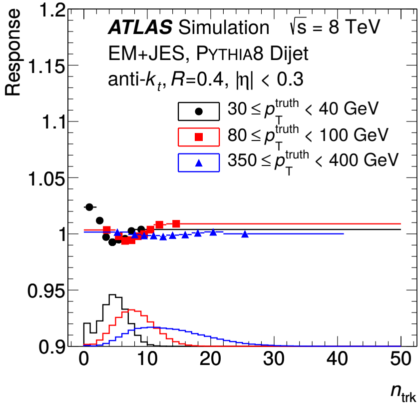


GSC: N_{trk} dependence, before and after

UNIVERSITÉ DE GENÈVE

- One example variable is the number of tracks matched to the jet, N_{trk}
  - Clear dependence of the jet response on N_{trk}, which we exploit (left vs right)

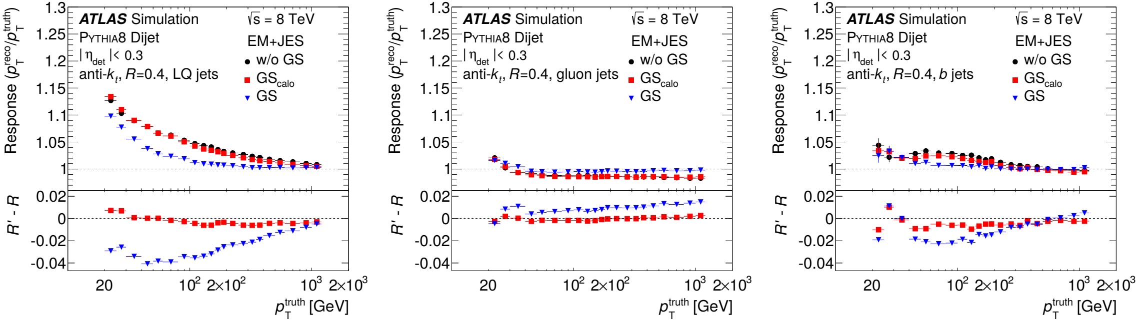




GSC: flavour dependence, before and after

UNIVERSITÉ DE GENÈVE

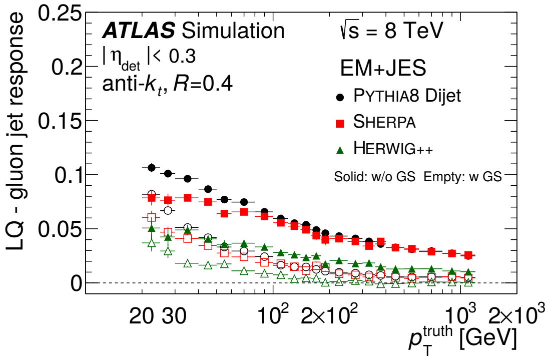
- N_{trk} is only one of the several variables used in the GSC
  - It is however likely the most important: correlated with quark vs gluon response differences
- GSC significantly improves light quark (left), gluon (middle), and b-quark (right) response



GSC: flavour differences

UNIVERSITÉ DE GENÈVE

- Reducing the differences between jet flavours is extremely important
  - Uncertainties are derived in light-quark-dominated topologies
  - Extrapolation from light quarks to gluons is a leading uncertainty
- By reducing light quark vs gluon differences, the GSC also significantly reduces jet uncertainties
  - We'll see this later



Jet calibration steps

UNIVERSITÉ DE GENÈVE

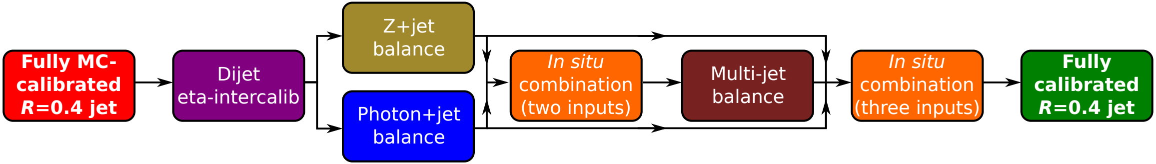
- We have now seen the many steps of the jet calibration sequence
  - Each step is designed to mitigate experimental effects
- Origin correction points the jet axis back towards the truth jet (improves jet η)
- Pileup corrections reduce dependence on μ and N_{PV}
  - This is how we mitigate the impact of (primarily) neutral pileup particles in our jets
- MC JES corrects for energy not observed in the detector
  - Largest correction, accounts for detector geometry, dead material, etc
- Global Sequential Correction reduces dependence on auxiliary variables
  - Improves the jet resolution and reduces flavour differences (and thus uncertainties)

- This has all been done using only MC, now let's look at data!

Correcting data vs simulation differences

In situ calibrations

UNIVERSITÉ DE GENÈVE

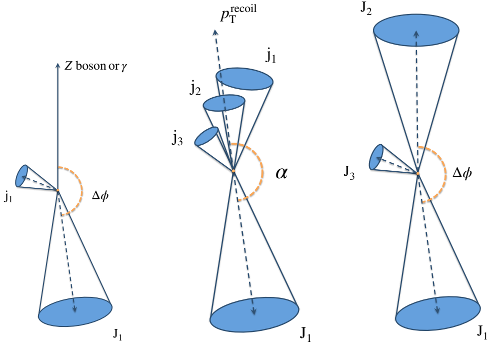


- Looking at data requires having an independent object that can be used as the reference
  - That is the "tag" (reference), and then we work to measure the "probe" (jet under study)
- The *in situ* JES calibration is comprised of several sequential steps
  - Each one is used to evaluate possible differences between data and MC
  - This evaluation includes both a calibration factor and associated uncertainties
- We are only evaluating differences between data and MC, **not** the *absolute* scale in data
  - We know our MC is at the "truth scale", so we correct our data to match our MC (!)

Jet energy scale: balance techniques

UNIVERSITÉ DE GENÈVE

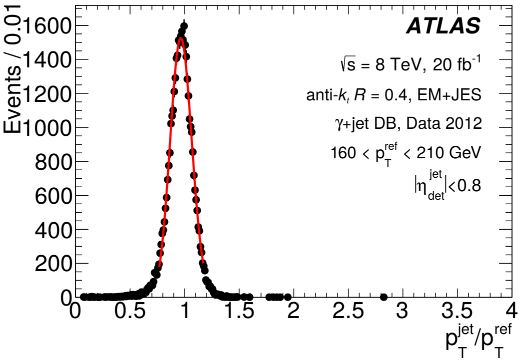
- All of the *in situ* techniques rely on having a well-calibrated reference
- Idea: exploit well-balanced events
  - $\mathcal{R} = \langle p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\mathrm{ref}} \rangle \approx 1$, binned in $p_{\mathrm{T}}^{\mathrm{ref}}$
  - $\mathcal{R}^{\mathrm{MC}} / \mathcal{R}^{\mathrm{data}}$ gives the calibration factor
- We transfer uncertainties from the reference object to the jet under study



Jet energy scale: balance techniques

UNIVERSITÉ DE GENÈVE

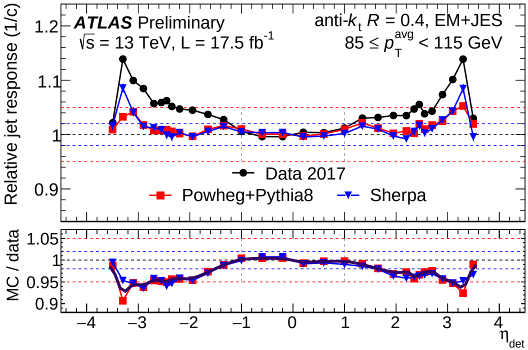
- All of the *in situ* techniques rely on having a well-calibrated reference
  - Example: a photon in this plot
- Idea: exploit well-balanced events
  - $\mathcal{R} = \langle p_{\mathrm{T}}^{\mathrm{jet}} / p_{\mathrm{T}}^{\mathrm{ref}} \rangle \approx 1$, binned in $p_{\mathrm{T}}^{\mathrm{ref}}$
  - $\mathcal{R}^{\mathrm{MC}} / \mathcal{R}^{\mathrm{data}}$ gives the calibration factor
  - In this case, $p_{\mathrm{T}}^{\mathrm{ref}} = p_{\mathrm{T}}^{\gamma}$
- We transfer uncertainties from the reference object to the jet under study
  - Photons have very small uncertainties, so we use them to study balancing jets



Jet energy scale: dijet η-intercalibration

UNIVERSITÉ DE GENÈVE

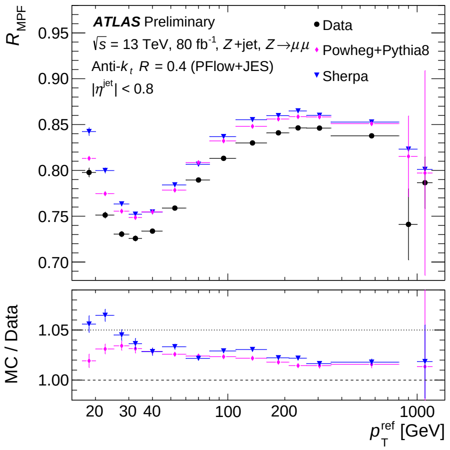
- The first technique balances central jets (reference) against more forward jets
  - Flattens the JES across the detector
  - Next techniques correct the global offset using central reference region
  - Reference region starts with $|\eta| < 0.8$
- Forward jet response is challenging for most MC generators
  - One of the leading uncertainty sources for jets beyond the tracker, $|\eta| \gtrsim 2.5$



Jet energy scale: $Z\ell\ell$+jet

UNIVERSITÉ DE GENÈVE

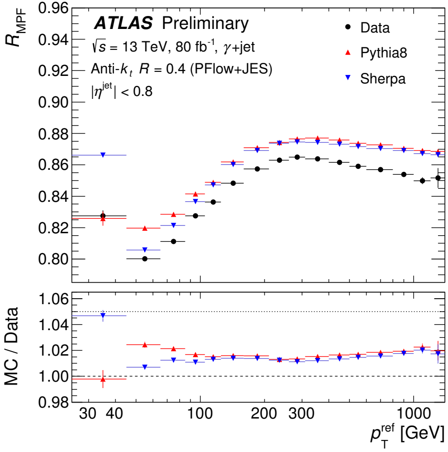
- $Z\ell\ell$+jet is the first method aimed at correcting the global data vs MC offset
  - Reference object: $Z \to \mu\mu$ or $Z \to ee$
- Can access the lowest jet p_T range
  - Best method up to $\sim$500 GeV of jet p_T



Jet energy scale: γ+jet

UNIVERSITÉ DE GENÈVE

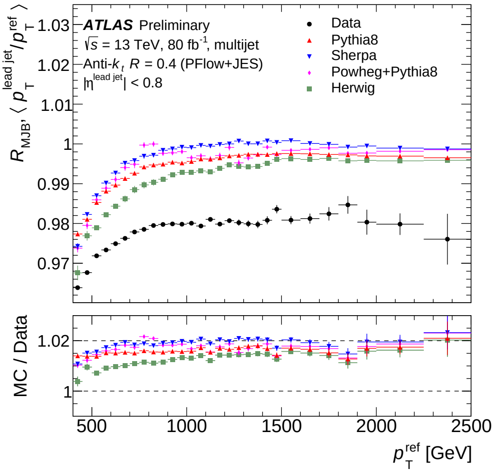
- γ+jet is the second method aimed at correcting the global data vs MC offset
  - Reference object: photon
- Starts at slightly higher p_T than Z+jet
  - Complementary up to ∼500 GeV of p_T
  - Best method up to ∼1000 GeV of p_T



Jet energy scale: multi-jet balance

UNIVERSITÉ DE GENÈVE

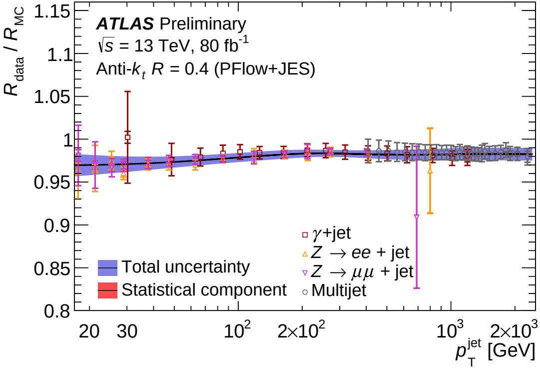
- The multi-jet balance (MJB) is the third method aimed at correcting the offset
    - Reference object: system of at least two lower-p_{T} jets that have already been calibrated by $Z\ell\ell$+jet and γ+jet
- Starts at high p_{T}, where Z/γ+jet approaches start to run out of stats
    - Exploits the huge QCD multijet xsec
    - Best method from 1000–2500 GeV

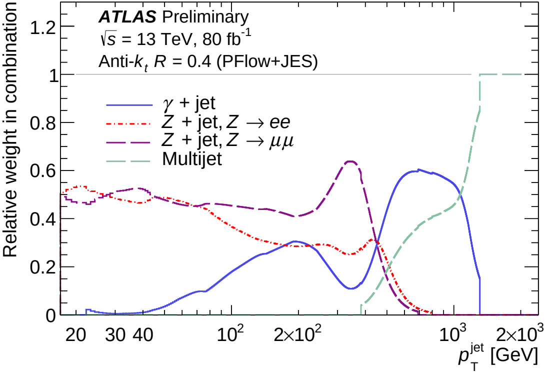


Jet energy scale: combining methods

UNIVERSITÉ DE GENÈVE

- The different *in situ* global offset corrections are then combined
  - $Z\ell\ell$+jet, γ+jet, and MJB are consistent where overlapping and are complementary vs p_T
- Dijet η-intercalibration corrects the forward region with respect to this global offset

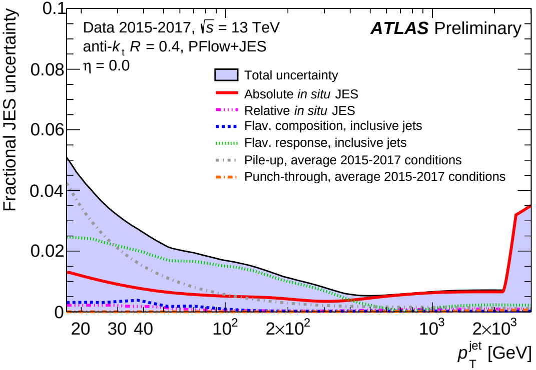


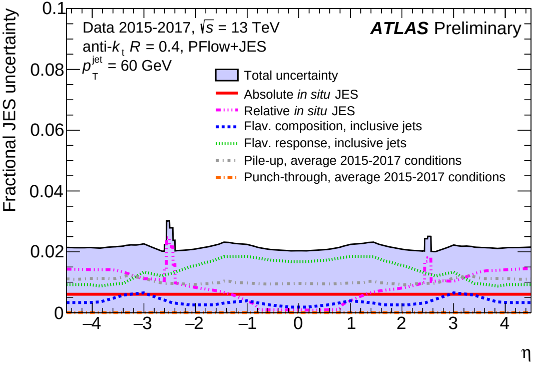


Jet energy scale: uncertainties

UNIVERSITÉ DE GENÈVE

- The *in situ* uncertainties are shown for the global offset and the forward jet correction
- Additional MC-based uncertainties for flavour (two sources, reduced by GSC) and pileup
- Results in > 100 sources of JES systematic uncertainties

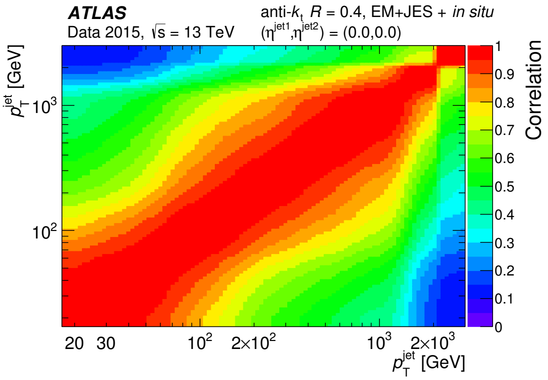


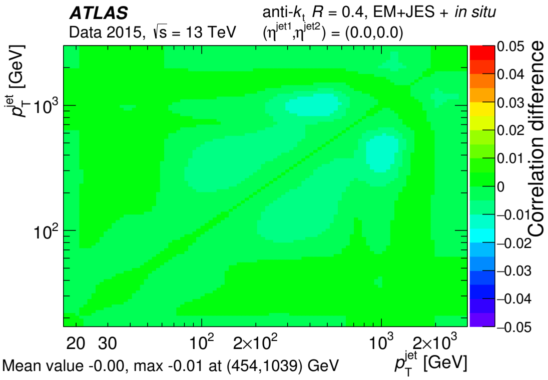


Jet energy scale: reductions

UNIVERSITÉ DE GENÈVE

- Retaining all of the >100 sources of JES systematic uncertainties is usually not necessary
- Instead, derive "eigenvector decompositions" (linear combinations preserving correlations)
  - Result: 30–40 JES uncertainties, depending on required usage, with negligible correlation loss

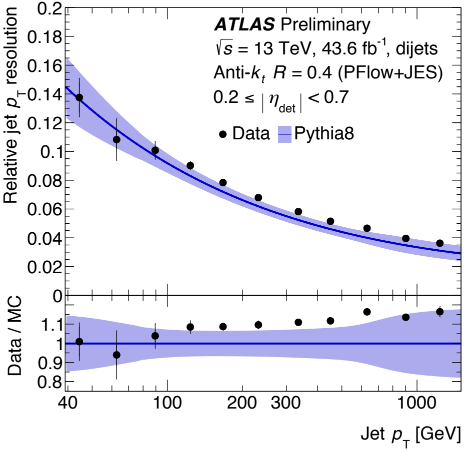




Jet energy resolution: smearing jets

UNIVERSITÉ DE GENÈVE

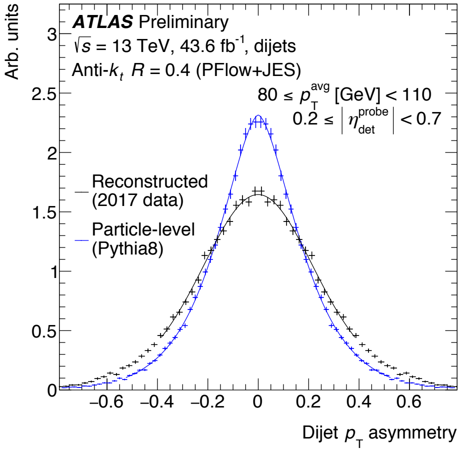
- To correct the JER, we "smear" jets to make the response distribution wider
  - Need to match widths of two Gaussians without changing central value
  - Do this by adding a new Gaussian of the same central value with a width equivalent to quadrature difference
    - $\sigma_2 = \sigma_1 \oplus \sigma_{\text{smear}} \therefore \sigma_{\text{smear}} = \sigma_2 \ominus \sigma_1$
- In contrast with JES, the *in situ* JER corrects MC to match data
  - Resolution in MC typically better than in data, and we can't anti-smear data



Jet energy resolution: removing the truth width

UNIVERSITÉ DE GENÈVE

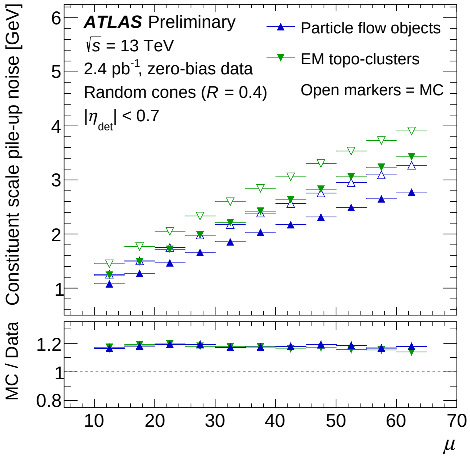
- The *in situ* JER is aimed at improving data/MC differences
  - We have to be careful to not fold in MC/MC generator differences
  - Want to correct for detector effects, not truth-level differences
  - Generators can disagree significantly
- JER thus subtracts truth jet widths
  - Resulting measurement is of the reco. resolution (not the full jet resolution)
  - Thus applicable to various generators



Jet energy resolution: random cones

UNIVERSITÉ DE GENÈVE

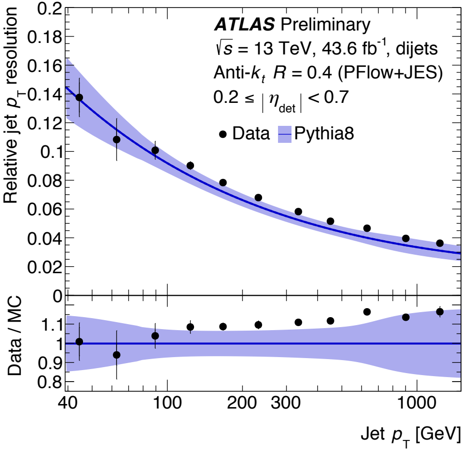
- JER has an expectation: $\frac{N}{E} + \frac{S}{\sqrt{E}} + C$
  - N = noise term, mostly pileup
  - S = sampling term, random showers
  - C = constant term, det. limitations
- We can measure N directly
  - Use completely random data
  - Measure energy in random cones, defined to be back-to-back (balancing)
  - Imbalances quantify the amount of ambient noise present in the detector
- Still need to measure S and C



Jet energy resolution: dijet asymmetry

UNIVERSITÉ DE GENÈVE

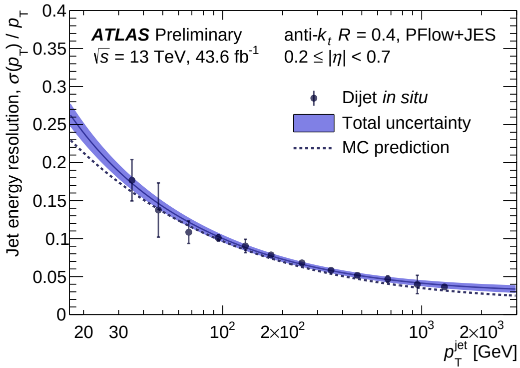
- This plot is actually a data/MC comparison of dijet events
  - Jets in dijet events should be balanced
  - Asymmetry quantifies the resolution of the measurement, compare data/MC
- Determine asymmetry for many p_{T} bins
- Fit the full (N,S,C) resolution function to the resulting distribution
- Result: a weak measurement of N (low stats), but good control of S and C

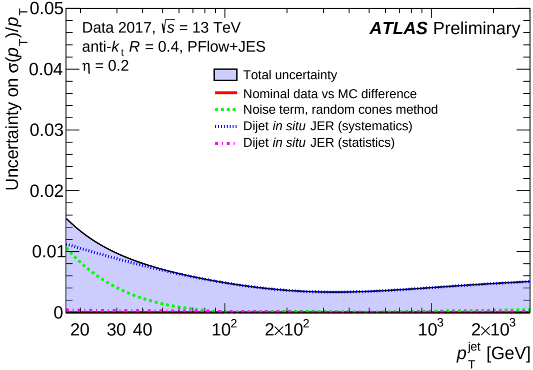


Jet energy resolution: results

UNIVERSITÉ DE GENÈVE

- Combine the random cone and dijet asymmetry measurements, using (N,S,C) fit
  - Result gives a precise measure of the JER across the full kinematic regime





In situ calibration steps

UNIVERSITÉ DE GENÈVE

- We compare data and MC in well-understood topologies for both the JES and JER
- JES: correct data to match MC, as MC remains our truth-jet reference
  - Use dijet events to correct forward jets with respect to central jets
  - Use $Z\ell\ell$+jet, γ+jet, and multi-jet balance to correct central jets
  - Combine all of the methods to obtain a final *in situ* JES calibration
  - Additional uncertainties for other effects (flavour, etc) are added
- JER: correct MC to match data, as we can't anti-smear the data
  - Use primarily the balance of dijet events to constrain the JER
  - JER accounts for only reconstruction effects: the truth asymmetry is removed

- Usually, the JES is more of a limiting factor than the JER in analysis usage

Summary

Summary

UNIVERSITÉ DE GENÈVE

- Experimentally, jet reconstruction is mostly about dealing with the detector
  1. Devising robust input objects that exploit both trackers and calorimeters (PFlow)
  2. Mitigating pileup effects, both before (CHS) and after (JVT/fJVT) jet reconstruction
  3. Calibrating reconstructed jets back to the truth jet scale, thus reducing detector effects
  4. Comparing jets between data and MC, to correct for detector simulation imperfections
- Today we have gone into these topics in detail for $R = 0.4$ anti-k_t jets
  - Today's exercise will allow you to compare truth, calo, and track jets
  - You will thus be looking at pileup and detector response plots
- On Thursday, we will switch focus to large-R jets and identifying the origin of such jets