Inclusive dijet cross-section measurements at the ATLAS experiment

Ota Zaplatílek [1]

[1] Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University in Prague

JCF Summer Workshop, June 16, 2022

Outline

- motivation for jet measurements
- ATLAS detector
- main dijet variables
- ATLAS MC
- ATLAS Data
- one-dimensional inclusive dijet measurement
- two-dimensional inclusive dijet measurement
- three-dimensional inclusive dijet measurement
- dedicated studies

QCD Lagrangian, quarks, gluons and jets

- QCD Lagrangian with quarks and gluons

$$\mathcal{L}_{QCD} = \sum_q \left[i\bar{\psi}_q \gamma^\mu \left(\partial_\mu - i g_s \frac{\lambda^a}{2} A^a_\mu(x) \right) \psi_q - m_q \bar{\psi}_q \psi_q \right] + \mathcal{L}_{gauge}$$

$$\mathcal{L}_{gauge} = -\frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu}$$

$$= -\frac{1}{4} A^a_{\mu\nu} A^{a\mu\nu} - \frac{1}{2} g_s f^{abc} (\partial_\mu A^a_\nu - \partial_\nu A^a_\mu) A^{b\mu} A^{c\nu} - \frac{1}{4} g_s^2 f^{abc} f^{ajk} A^b_\mu A^c_\nu A^{j\mu} A^{k\nu} \tag{1}$$

Quarks and gluons are not directly observed in experiments; Experimenters usually observe the secondary hadronic particles produced in the same direction after the quark/gluon fragmentation. Such secondary particles are found in collimated showers, so-called jets.

Motivation for jet measurements

jets can be matched to quarks/gluons

$$\mu_F^2 \frac{q_i(x,\mu_F)}{\mu_F^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dy}{y} P\left(\frac{x}{y}\right) q_i(x,\mu_F)$$

Theoretical motivation:

- structure of proton
- parton distribution functions $q_i(x, \mu_F)$
- splitting functions $P_{ij}\left(\frac{x}{y}\right)$
- running $\alpha_s(\mu_R^2)$
- test of asymptotic freedom of QCD
- potential New Physics at TeV scale

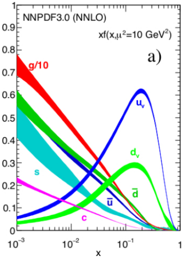
$$\mu_R^2 \frac{\partial \alpha_s}{\partial \mu_R^2} = -(b_0\alpha_s^2 + b_1\alpha_s^3 + b_2\alpha_s^4 + \ldots)$$

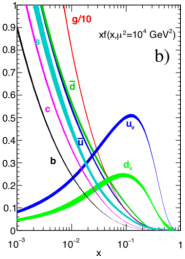
$$\alpha_s(Q^2) = \frac{\alpha_s(Q_0^2)}{1 - B_s \cdot \alpha_s(Q_0^2)\ln\frac{Q^2}{Q_0^2}}$$

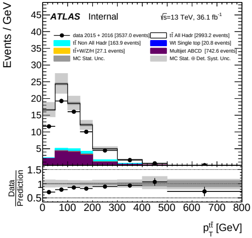
Experimental motivation:

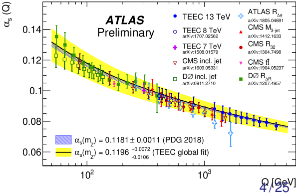
- large theoretical unc. coming from QCD
- background for most of Standard Model measurements

$$B_s = -\frac{11N_C - 2N_f}{6\pi}$$

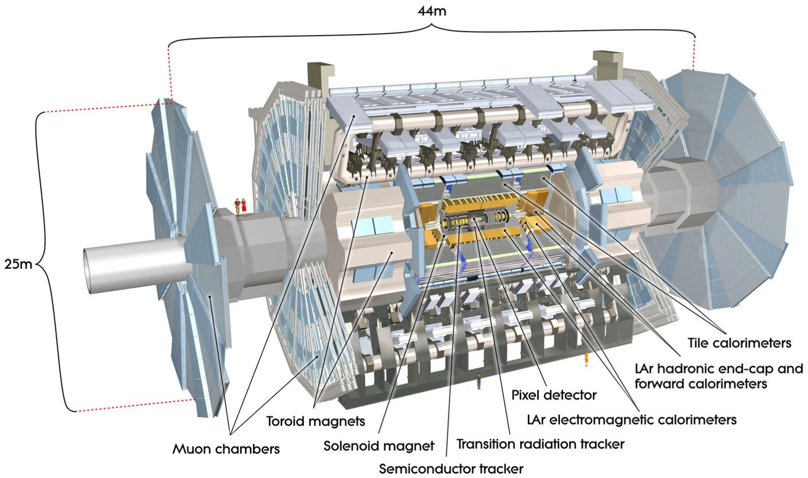




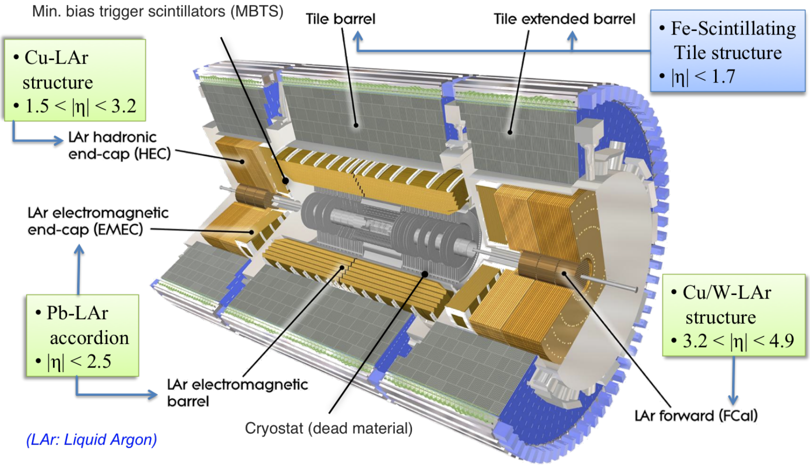




ATLAS detector



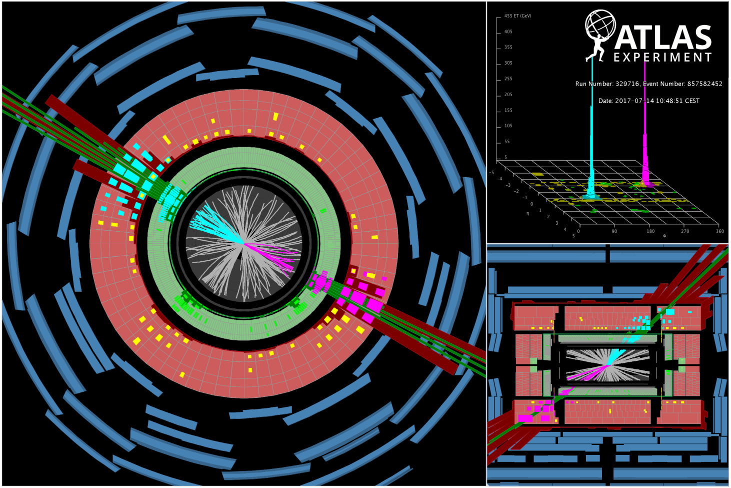
ATLAS calorimeter system



- essential for jet measurements
- topological jets built from topological clusters in the calorimeter

Event display

- dijet event, proton–proton collision, $\sqrt{s} = 13$ TeV, ATLAS data 2017
  - transverse plane to beam direction (left)
  - longitudinal Z-Y plane (right-bottom)
  - $\eta - \phi$ plane (right-top)



- exclusive dijet event
  - $p_{T,1} = p_{T,2} = 2.9$ TeV, $y_1 = -1.2$, $y_2 = 0.9$
  - $m_{jj} = 9.3$ TeV, $y^\star = 1.05$, $y_{boost} = 0.15$

Dijet variables

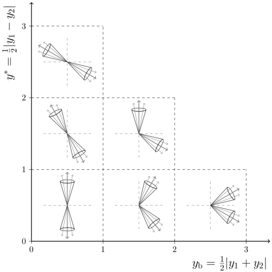
leading jet:

- transverse momentum $p_{T,1}$
- rapidity y_1

subleading jet:

- transverse momentum $p_{T,2}$
- rapidity y_2

inclusive dijet defined as a vector sum of leading and subleading jets in the event with at least two jets
most important variables for dijet analysis: m_{jj}, $y^{\star}$, y_{boost}



- $y^{\star} = \frac{1}{2}\,|y_1 - y_2|$
  - forward vs. central jets
  - sensitive to New Physics
  - New Physics models predict different shape; more dijets with low $y^{\star}$

- $y_{boost} = \frac{1}{2}\,|y_1 + y_2|$
  - same-side vs. opposite-side jet event
  - sensitive to pdf
  - low y_{boost} for approx. equal momentum fractions of incoming partons x_1, x_2
  - high y_{boost} for very asymmetric momentum fractions of incoming partons x_1, x_2; one very high, one very low

Selection criteria for inclusive dijet measurements

- proton–proton collisions at $\sqrt{s} = 13$ TeV
- anti-k_T $R = 0.4$ calibrated topological jets

Simplified dijet selection:

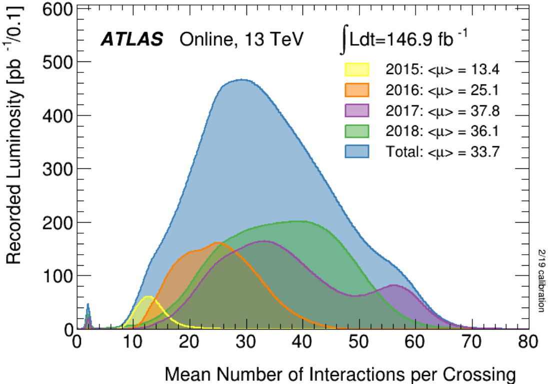
1. identify well reconstructed events (data quality)
2. identify well reconstructed jets (jet quality)
3. identify inclusive di-jet using dijet selection criteria

Selection criterion	Applied condition
jet multiplicity n_{jets}	$n_{\text{jets}} \geq 2$
leading jet $p_{T,1}$	$p_{T,1} > 75$ GeV
subleading jet $p_{T,2}$	$p_{T,2} > 75$ GeV
leading jet $\|y_1\|$	$\|y_1\| < 3.0$
subleading jet $\|y_2\|$	$\|y_2\| < 3.0$
dijet $y^\star = \frac{1}{2}\|y_1 - y_2\|$	$y^\star < 3.0$
dijet $H_T^{\text{two}} = p_{T,1} + p_{T,2}$	$H_T^{\text{two}} > 200$ GeV

Table: Per–dijet selection criteria.

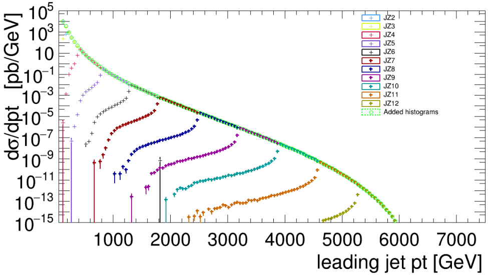
ATLAS MC

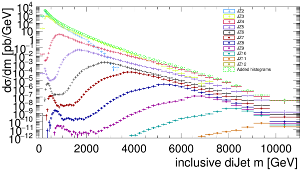
- jet/dijet cross-section falls steeply over 10 orders of magnitude
- MC generated in multiple $\hat{p}_T$ slices
  (13 slices for nominal MC Pythia 8)
- various pile-up conditions in LHC run2 phase:
  - MC16a (Data 2015+16)
  - MC16d (Data 2017)
  - MC16e (Data 2018)
  - MCFullRun2 as MC16a+d+e (Data 2015+16+17+18)

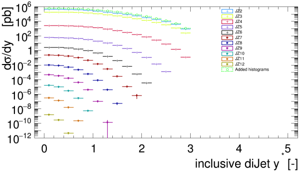


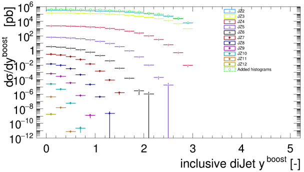
ATLAS MC Pythia 8

- Pythia 8 as nominal LO MC
- final MC distribution as a sum of all relevant JZX slices





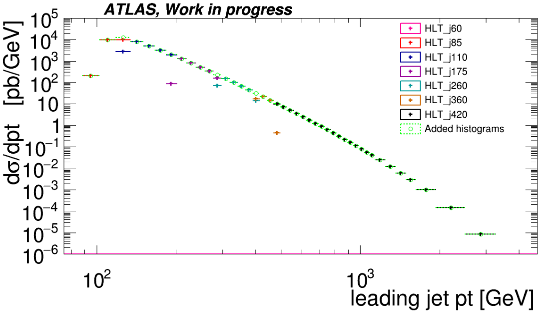




- low JZX slices dominate to low jet p_T, angular variables in the whole range
- high JZX slices dominate high jet p_T
- high dijet mass events originate from high p_T jets (high JZX) or low p_T jets with high angular separation (low JZX)
- MC16a+d+e need in case of full LHC run 2 measurement

ATLAS Data and Triggers

- single jet p_T High Lever Triggers[1]
- trigger combination using so-called leading jet p_T trigger strategy
- final distribution is reached as a sum of all relevant trigger distributions

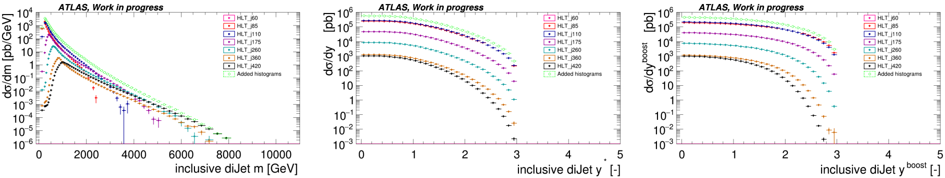


example for leading jet pt cross-section using the ATLAS data 2018

[1] e.q. `HLT_jXXX` trigger fires if the event includes at least one jet with $p_T > XXX$ GeV, the event is recorded with predefined prescale.

ATLAS Data and Triggers

- low (prescaled) trigger dominates to angular distributions like $y^{\star}$, y_{boost}
- highest (unprescaled) trigger dominates to high jet p_{T} and high m_{jj}
- final distribution is reached as a sum of all relevant trigger distributions



some more examples for one-dimensional dijet cross-section using the ATLAS data 2018

- the ATLAS data of of 2015,16,17,18 need in case of full LHC run 2 measurement years
- applied triggers differs in various years of data acquisition

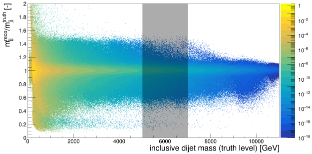
Resolution studies

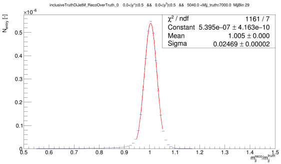
- finite detector resolution effects
- dijet mass resolution[2] $\sigma\left(\frac{m_{\mathrm{jj}}^{\mathrm{reco}}}{m_{\mathrm{jj}}^{\mathrm{truth}}}\right)$
- 2D resolution histogram in a given range of $y^{\star}$, y_{boost}, m_{jj} (left)
- 1D projection to a given range of m_{jj}^{truth}, Gaussian fit, sigma extraction (middle)
- 1D response histogram as a iteration over all possible ranges of m_{jj}^{truth} (right)

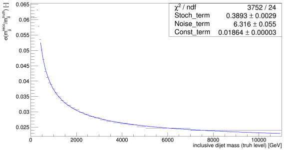
- 1D fit of response function:

$$\sigma(X) = \frac{S}{X} + \frac{N}{\sqrt{X}} + C.$$

  - X as variable of interest, m_{jj}
  - S as Stochastic term
  - N as Noise term
  - C as Constant term







It has been done for various of $y^{\star}$ and y_{boost} bins.
Considering 6 bins in $y^{\star}$ and y_{boost}, there are 21 response functions.

[2] Dijet mass resolution as a Gaussian sigma of $m_{\mathrm{jj}}^{\mathrm{reco}}/m_{\mathrm{jj}}^{\mathrm{truth}}$ histogram in a given range of $y^{\star}$, y_{boost}.

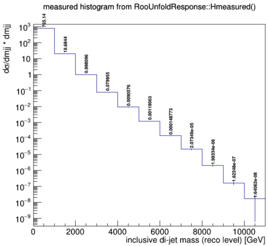
Folding and Unfolding

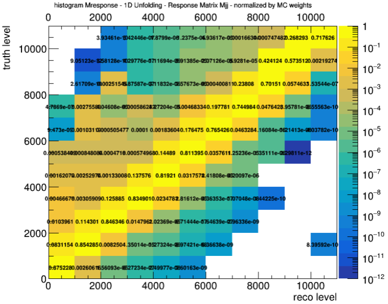
- each measurement suffers by detector effects:
  - detector resolution
  - finite acceptance
  - finite efficiency
- reco level distribution (measured distribution) is not the same as truth level distribution (generated distribution)
- reco level distribution is folded (convoluted) truth level distribution with the detector resolution; inverse operation is called unfolding (deconvolution)

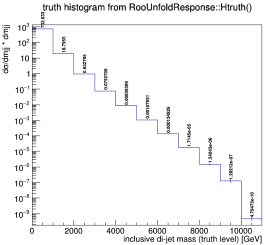
Folding equation:

$$\nu_i = \sum R_{ij}\mu_j$$

- ν as reco level distribution (left)
- μ as truth level distribution (right)
- R as response matrix derived from full detector simulation using Geant4 (middle)

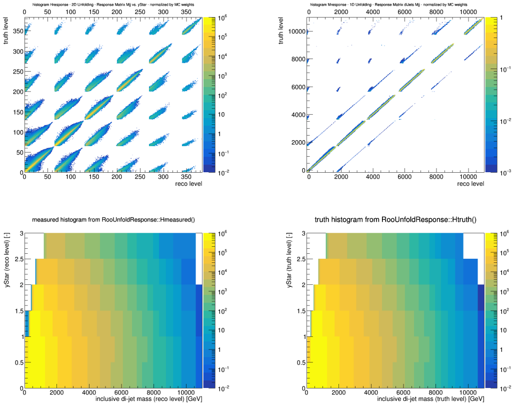






RooUnfoldResponse 2D - m_{jj} vs. $y^\star$

- Unfolding and Response studies within `RooUnfold` package
- 1D, 2D, 3D measurements possible (but not optimized by default)
- example for 2D measurement of incl. dijet mass and $y^\star$



- response histogram[3] (top-left)
- response matrix[4] (top-right)
- reco level distribution (bottom-left)
- truth level distribution (bottom-right)

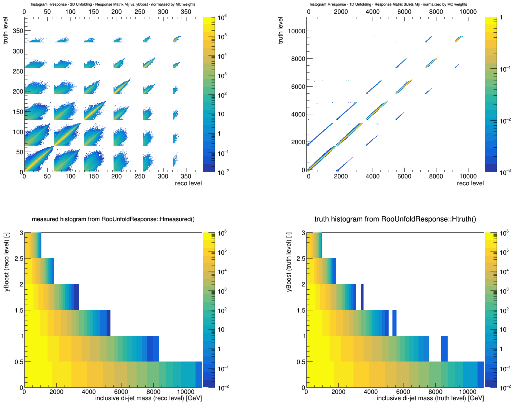
---

[3] number of reconstructed events in reco bin *i* and generated in truth bin *j*

[4] conditional prob. to measure an event in reco bin *i* if it was generated in truth bin *j*

RooUnfoldResponse 2D - m_{jj} vs. y_{boost}

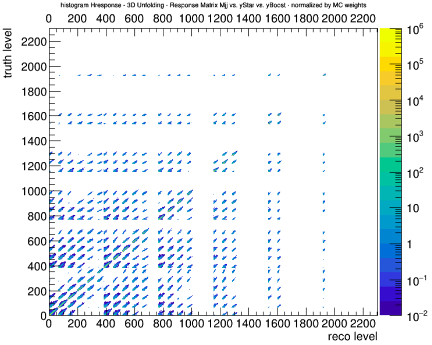
- very similar study for m_{jj} vs. y_{boost}
- see migration between various bins of m_{jj} and various bins of y_{boost}

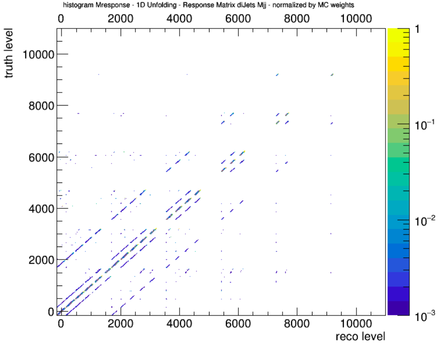


- multidimensional measurements should include full information about the event migration (not like one-dimensional measurement in several windows for your second variable)

RooUnfoldResponse 3D - m_{jj} vs. $y^\star$ vs. y_{boost}

- potential 3D measurement of m_{jj} vs. $y^\star$ vs. y_{boost} attach CPU and memory limitations
- propagation of syst. unc. through unfolding as a challenge (we have $\approx$ 1400 syst. unc. for the ATLAS jets)





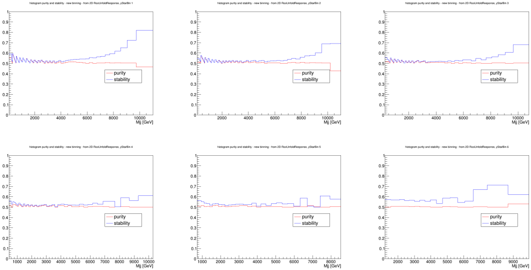
Bin Optimization

- fine response histogram, fine reco level distribution and fine truth level distribution for bin optimization; considering high enough statistic MC
- useful variables purity and stability:

$$\text{stability} = \frac{\text{\# of expected events passed selection criteria in the reco bin i and the same truth bin i}}{\text{\# of expected events passed selection criteria in the truth bin i}}$$

$$\text{purity} = \frac{\text{\# of expected events passed selection criteria in the reco bin i and the same truth bin i}}{\text{\# of expected events passed selection criteria in the reco bin i}}$$

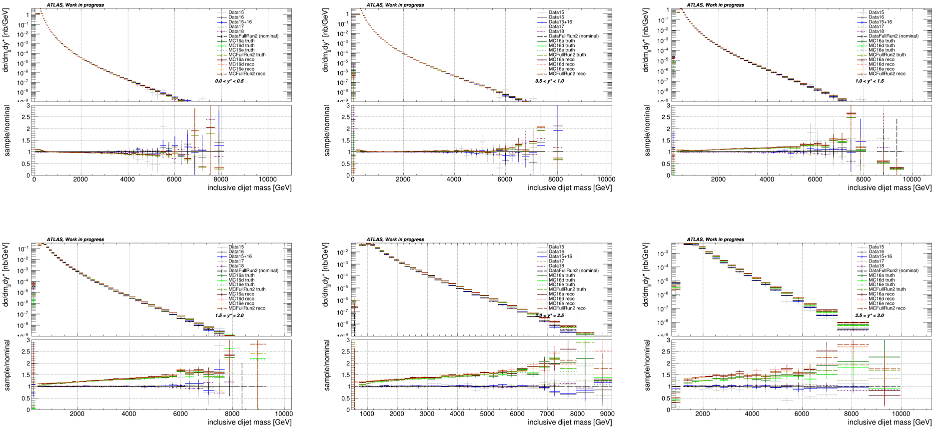
- algorithm: sum neighboring bins in response histogram, truth level distribution and reco level distribution consistently, until the purity is high enough
- example for 2D measurement of incl. dijet mass vs. $y^{\star}$, using 6 equidistant bins of $y^{\star}$ and required at least 50% purity and stability



- last bins included/excluded according stat. unc. in the ATLAS data
- improved χ^2 / *nfd* with new binning expected
- similar studies also for 1D, 2D, 3D measurements of incl. jets and incl. dijets

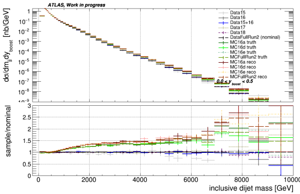
Data to MC - dijet - m_{jj} vs. $y^{\star}$

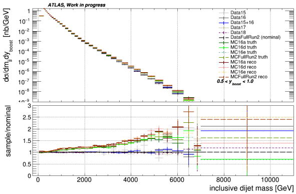
- comparison for ATLAS Data and LO MC Pythia 8
- considering:
  - various years of data acquisitions
  - various MC campaigns
  - data (Blackish), MC truth (Greenish), MC reco (Redish)

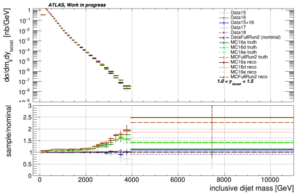


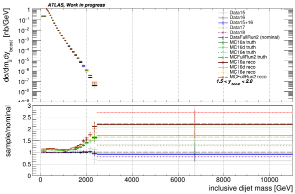
Data to MC - dijet - m_{jj} vs. y_{boost}

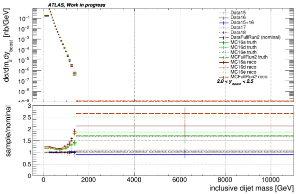
- comparison for ATLAS Data and LO MC Pythia 8
- considering:
  - various years of data acquisitions
  - various MC campaigns
  - data (Blackish), MC truth (Greenish), MC reco (Redish)

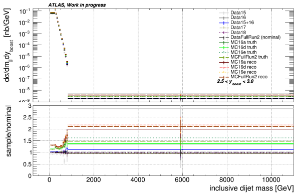






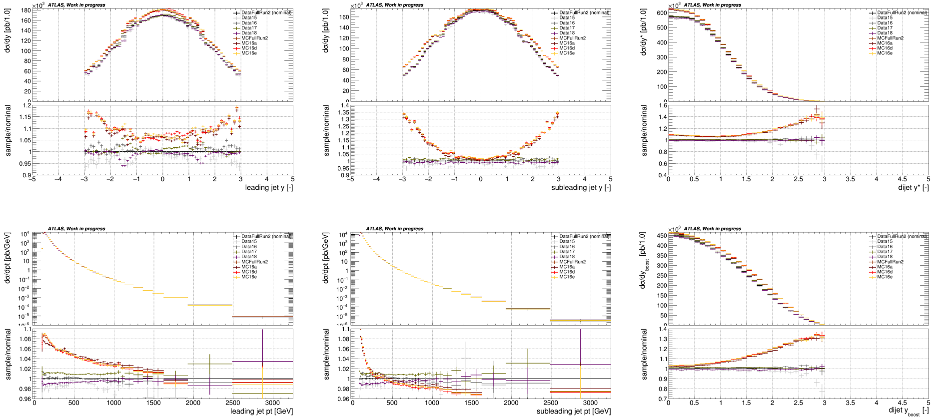






Data to MC - leading and subleading jets

- Data to MC differences due to LO MC Pythia 8 missmodelling of forward jets

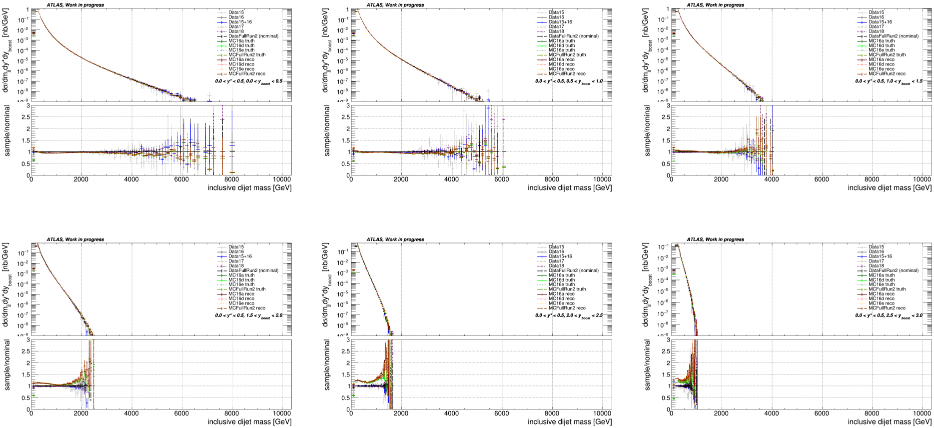


LO MC Pythia 8:

- decent description of central (sub)leading jet
- poor description of forward jets
- the differences are propagated from leading and subleading jets to dijets by construction
- higher order MC needed (NLO, NNLO) for Data to MC comparison
- still acceptable for binning, response, unfolding studies etc.

Data to MC - dijet - m_{jj} vs. $y^{\star}$ vs. y_{boost}

- similar study for 3D measurement of m_{jj} vs. $y^{\star}$ vs. y_{boost}
- 21 plots in total due to kinematics restriction of $y^{\star}$ and y_{boost}
- here $0.0 < y^{\star} < 0.0$ and all relevant y_{boost} (6 plots)
- other remaining plots available in back-up



Summary

- inclusive dijet measurements using the ATLAS experiment in progress
- for the first time
  - two-dimensional measurement for m_{jj} vs. $y^\star$ with 2D unfolding[5]
  - two-dimensional measurement for m_{jj} vs. y_{boost}
  - possible three-dimensional measurement m_{jj} vs. $y^\star$ vs. y_{boost}

- following plans:
  - finish multidimensional unfolding
  - extraction of systematic unc.
  - propagation systematic unc. through unfolding
  - preparation of NLO (NNLO) MC

[5] previously it has been done 1D measurements in windows of $y^\star$ ignoring migration in $y^\star$ bins

Thank you for your attention.

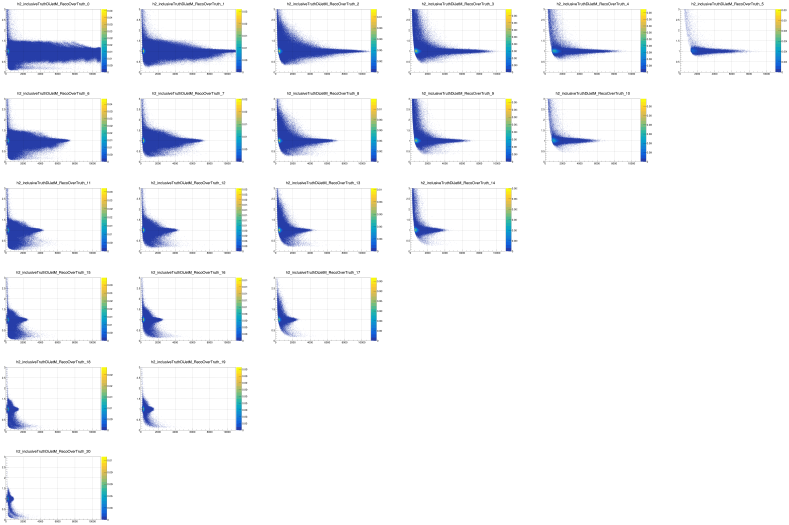
BACK-UP

links:

- ATLAS detector - figure taken from here.
- ATLAS calorimeter system - figure taken from here.
- ATLAS Event display - dijet event - figure taken from here.
- ATLAS Pile-up - taken from here
- ATLAS alpha s - taken from here
- $y^{\star}$ and y_{boost} plane - taken from here.

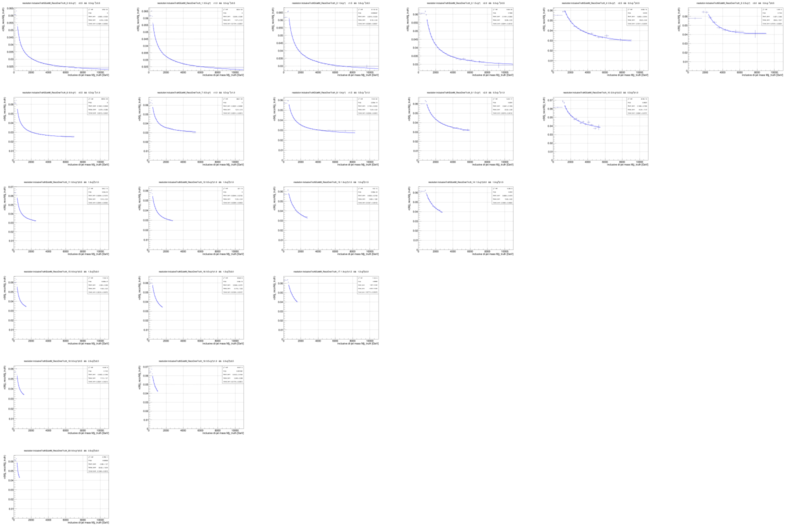
Dijet mass resolution for all 21 $y^\star$ and y_{boost} intervals

- considering 6 equidistant bins of $y^\star$ and y_{boost} from 0.0 to 3.0, then there are 21 2D response histograms
- various line - various $y^\star$
- various column - various y_{boost}



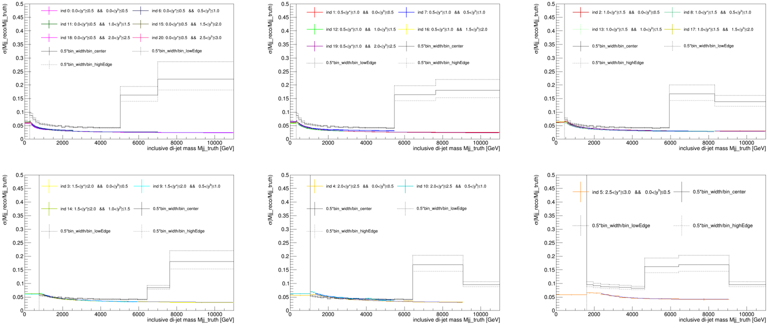
Dijet mass resolution for all 21 $y^{\star}$ and y_{boost} intervals

- considering 6 equidistant bins of $y^{\star}$ and y_{boost} from 0.0 to 3.0, then there are 21 1D response histograms
- various line - various $y^{\star}$
- various column - various y_{boost}



Resolution studies - results

- dijet mass resolution $\sigma\left(\frac{m_{jj}^{reco}}{m_{jj}^{truth}}\right)$ in $y^\star$ and y_{boost} bins
- comparison for 21 dijet mass resolutions; fixed $y^\star$ bin and all relevant y_{boost} bins in each plot below



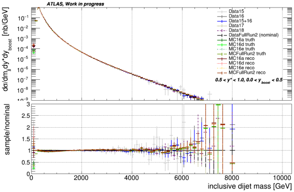
- comparable resolutions for various y_{boost} in a given $y^\star$ bin (color lines)
- comparison also for bin-like resolution[6] - (black lines)

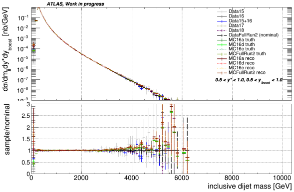
There is a place for improvement, possible twice as many bins as in previous measurements.

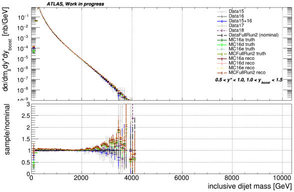
[6] historically the bin-width was chosen as a twice of the resolution

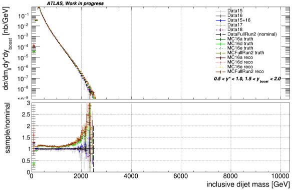
Data to MC - dijet m_{jj} vs. $y^\star$ vs. y_{boost}

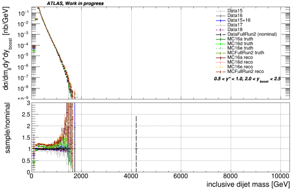
- similar study for 3D measurement of m_{jj} vs. $y^\star$ vs. y_{boost}
- 21 plots in total due to kinematics restriction of $y^\star$ and y_{boost}
- here $0.5 < y^\star < 1.0$ and all relevant y_{boost} (5 plots)





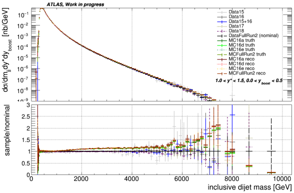


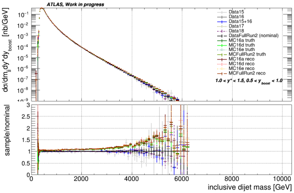


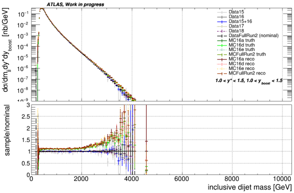


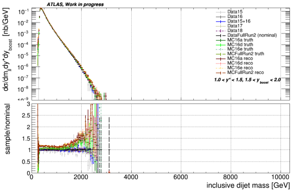
Data to MC - dijet m_{jj} vs. $y^{\star}$ vs. y_{boost}

- similar study for 3D measurement of m_{jj} vs. $y^{\star}$ vs. y_{boost}
- 21 plots in total due to kinematics restriction of $y^{\star}$ and y_{boost}
- here $1.0 < y^{\star} < 1.5$ and all relevant y_{boost} (4 plots)



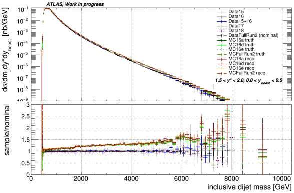


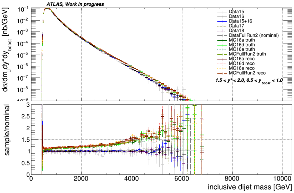


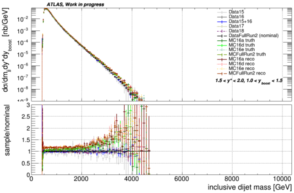


Data to MC - dijet m_{jj} vs. $y^{\star}$ vs. y_{boost}

- similar study for 3D measurement of m_{jj} vs. $y^{\star}$ vs. y_{boost}
- 21 plots in total due to kinematics restriction of $y^{\star}$ and y_{boost}
- here $1.5 < y^{\star} < 2.0$ and all relevant y_{boost} (3 plots)

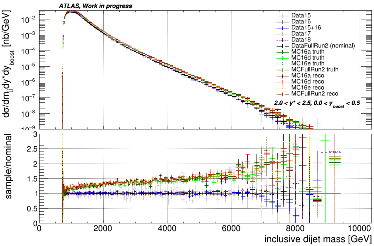


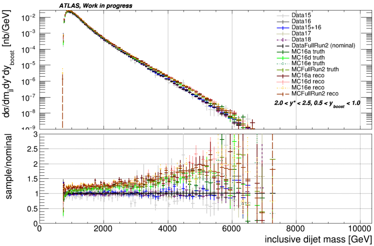


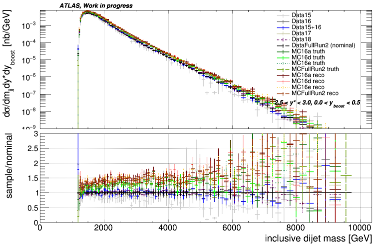


Data to MC - dijet m_{jj} vs. $y^{\star}$ vs. y_{boost}

- similar study for 3D measurement of m_{jj} vs. $y^{\star}$ vs. y_{boost}
- 21 plots in total due to kinematics restriction of $y^{\star}$ and y_{boost}
- fist line: $2.0 < y^{\star} < 2.5$ and all relevant y_{boost} (2 plots)
- second line: $2.5 < y^{\star} < 3.0$ and all relevant y_{boost} (1 plot)







Selection criterion		Applied condition
MC cleaning - ptAvg cleaning	(MC only)	$\frac{p_{\mathrm{T}}^{\mathrm{avg,reco}}}{p_{\mathrm{T,\,1}}^{\mathrm{truth}}} < 1.4$
jet cleaning - `Tight0Loose1`	(Data and MC, reco level only)	true
data quality - LAr, Tile, Core and Scintillator	(Data only, reco level only)	true

Table: Per–event selection criteria.