



ATLAS Note

KIP-2016



Draft version 0.1

1

2 Jet Observables using Subjet-assisted Tracks

3

The ATLAS Collaboration¹, Oleg Brandt^a, Sascha Dreyer^a, Fabrizio
4 Napolitano^a

5

^a*Heidelberg University*

6

5th December 2016

7

This note presents the details of the Monte-Carlo studies on the subjet-assisted observables for groomed large-radius jet. In particular the observables for the Energy Correlation Functions and n-Subjettiness variables used by the ATLAS collaboration, C_2 , D_2 , τ_{21} and τ_{32} are discussed using subjet-assisted tracks; the mass observable constructed with this technique, m^{TAS} , is presented and discussed with a modified four-momentum prescription. In all the variables studied, large improvement have been found using this novel techniques, the first ones evaluating in terms of QCD event rejection in W/Z boson, top quark and Higgs boson tagging; the second one in terms of precision reconstruction of the large-radius jet mass.

15

© 2016 CERN for the benefit of the ATLAS Collaboration.

Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.

16 **Contents**

17 **1 Introduction**

18 Jets are collimated streams of particles resulting from quarks and gluons fragmentation and hadronization.
19 The distribution of energy inside a jet contains information about the initiating particle. When a massive
20 particle such as a top quark, Higgs boson or W/Z bosons is produced with significant Lorentz boost and
21 decays into quarks, the entire hadronic decay may be captured inside a single jet. The mass of such jets
22 (jet mass) is one of the most powerful tools for distinguishing massive particle decays from the continuum
23 multijet background; the Energy Correlation Functions and n-Subjettiness C_2 , D_2 , τ_{21} and τ_{32} provide an
24 ad-hoc tool pupusely developed for the multijet background and constitue a fundamental part of many for
25 boson taggers. This note documents the so-called subjet-assisted techniques with the ATLAS detector
26 *insref*. The track-assisted subjet mass m^{TAS} definition is presented and confronted with the standard
27 development in ATLAS, m^{comb} and m^TA . Energy Correlation Functions and n-Subjettiness with the
28 modified subjet-assisted technique is presented and confronted with the standard one in ATLAS. The note
29 ends with conclusions and future outlook in *insref*.

30 **2 ATLAS detector**

31 ATLAS (A Toroidal ApparatuS) is a multi-purpose particle detector with nearly 4π coverage in solid angle.
32 A lead/liquid-argon sampling electromagnetic calorimeter is split into barrel ($|\eta| < 1.5$) and end-cap ($1.5 < |\eta| < 3.2$) sections. A steel/scintillating-tile hadronic calorimeter covers the barrel region ($|\eta| < 1.7$)
33 and two end-cap copper/liquid-argon sections extend to higher pseudo-rapidity ($1.5 < |\eta| < 3.2$). Finally,
34 the forward region ($3.1 < |\eta| < 4.9$) is covered by a liquid-argon calorimeter with Cu (W), absorber in the
35 electromagnetic (hadronic) section. Inside the calorimeters there is a 2 T solenoid that surrounds an inner
36 tracking detector which measures charged particle trajectories covering a pseudo-rapidity range $|\eta| < 2.5$
37 with pixel and silicon micro-strip detectors (SCT) and additionally which covers the region $|\eta| < 2.0$ with
38 a straw-tube transition radiation tracker (TRT). Outside the calorimeter there is a muon spectrometer: a
39 system of detectors for triggering up to $|\eta| < 2.4$ and precision tracking chambers up to $|\eta| < 2.7$ inside a
40 magnetic field supplied by three large superconducting toroid magnets.
41

42 A breakdown of the ATLAS sub-detector performance is shown in Table ??.

43 **3 Monte Carlo Samples**

44 *refrseMT* The samples used are divided into two main groups: SM background and beyond SM signal.
45 The SM background includes the QCD multijet samples, produced with a falling p_T spectrum. The
46 beyond SM signals are $W' \rightarrow WZ \rightarrow q\bar{q}'q\bar{q}$, $Z' \rightarrow t\bar{t}$ (top quarks considered in the full hadronic channel
47 ($t \rightarrow W(\rightarrow q\bar{q}')b$)) and RS-Graviton $\rightarrow hh \rightarrow b\bar{b}bb\bar{b}$, i.e. final states have only jets in all the samples.
48 The details of the samples are given in Table 1; the masses considered span from 0.5 to 5 TeV to improve
49 and diversify the kinematic space covered.

ATLAS	Description and performance
Magnetic field	2 T solenoid; 0.5 T toroid barrel and 1 T toroid end-cap
Tracker	Inner detector: IBL, Silicon pixel and strips, TRT $\sigma_{p_T}/p_T \simeq 5 \times 10^{-4} p_T \otimes 1\%$
EM calorimeter	EMB, EMEC and pre-sampler (Liquid Argon and lead) $\sigma_E/E \simeq 10\%/\sqrt{E} \otimes 0.7\%$
Hadronic calorimeter	Tile (Fe and scintillating tiles) and HEC (Cu and LAr) $\sigma_E/E \simeq 50\%/\sqrt{E} \otimes 3\%$
Muons	Inner detector and muon spectrometers $\sigma_{p_T}/p_T \simeq 2\% \text{ at } 50 \text{ GeV}$ $\sigma_{p_T}/p_T \simeq 10\% \text{ at } 1 \text{ TeV}$
Trigger	L1 and HLT (L2 and EF) Rates from ~ 40 MHz to ~ 75 kHz (L1) and to ~ 200 Hz (HLT)

Process	ME Generator & Fragmentation	ME PDFs	UE Tune	Resonance Masses
QCD multijet	Pythia 8	NNPDF23LO	A14	N/A
$W' \rightarrow WZ$	Pythia 8	NNPDF23LO	A14	1.5, 2.5, 3, 4, 5 TeV
$Z' \rightarrow t\bar{t}$	Pythia 8	NNPDF23LO	A14	1.5, 1.75, 2.5, 3, 4, 5 TeV
$G_{RS} \rightarrow hh(\rightarrow b\bar{b})$	Pythia 8	NNPDF23LO	A14	0.5, 1, 1.5, 2, 2.5, 3 TeV
$W' \rightarrow WW$ with $m_{\tilde{W}} = m_t$	Pythia 8	NNPDF23LO	A14	1.5, 2.5, 3, 4, 5 TeV

Table 1: Overview of the Monte Carlo Samples used. The first line shows QCD standard model process, the second, the third and the forth the beyond SM samples considered; the last line the “massive W/Z ” sample.

50 A set of kinematic distributions for the W' is shown in Figure ??: on the left the p_T distribution where
51 the kinks correspond to the Jacobian peak of the mass considered and the η distribution on the right. The
52 green dots represent the distribution before the selection, which is $p_T > 250$ GeV and $|\eta| < 2.0$ and the red
53 dots after this selection. This selection typical for many searches for BSM physics. All the other samples
54 and the background can be found in the Appendix. In what follows, it will also be used the nomenclature
55 *boosted W/Z* for the W' sample, *boosted tops* for the Z' sample, *boosted Higgs* for the G_{RS} sample and
56 *massive W* for the $W' \rightarrow \tilde{W}\tilde{W}$ with $m_{\tilde{W}} = m_t$. *refrseMT*

57 4 Object Definition

58 This section gives an overview of the objects used for the subjet-assisted variables, which are the large-
59 radius jet mass, the Energy Correlation Functions and the n-Subjettiness, as they are used within ATLAS,

60 the next section will give the details of the modified approach of the subjet-assisted techniques.

61 **4.1 Large-radius jet mass definitions**

62 Large-radius jet, or large- R jets are jets constructed with a radius parameter of the reclustering algorithm
 63 much bigger than the standard 0.4; within ATLAS the size of large- R jets is 1.0 for anti- k_t and 1.2 for C/A
 64 (the area of C/A is $\sim 20\%$ smaller than anti- k_t).

65 It is worth noting that, for a standard anti- k_t 0.4 jet the active area [**antiktalgo**] is $A_{\text{anti}-k_t} = \pi R^2 \simeq 0.5$,
 66 while it is $\simeq 3.14$ for 1.0 jet, i.e. around six times bigger.

67 Already from this “geometrical” point of view, the necessity of further techniques can be understood: the
 68 effect of soft radiation contamination from Pile-Up (PU) and Underlying Event (UE) will be in this case
 69 six times bigger and spoil the efficiency of the jet mass measurements.

70 **4.1.1 Substructure: Grooming Techniques**

71 This section is based on the 7 TeV article on jet Substructure [**substructure1**]. In order to use large- R
 72 jets, it is necessary to gain additional information on the interior of these objects, i.e. using techniques
 73 that exploit its substructure allowing a jet-by-jet discrimination of the energy deposit most likely coming
 74 from the hard-scattering to other soft radiation.

75 A common feature in substructure is the use of *sub-jet*, i.e. jets obtained from a parent jet (e.g. the large- R
 76 jet), using its constituent but running the jet reclustering algorithm with a smaller radius parameter; in one
 77 large- R jet, typically there are two or more sub-jets depending on the originating process and its p_T .

78 Techniques have been developed, both using sub-jets or directly constituents of a jet, which are referred
 79 to as *grooming* algorithms.

80 Grooming algorithms are designed to retain the characteristic substructure within such a large- R jet while
 81 reducing the impact of the fluctuations of the parton shower and the UE, thereby improving the mass
 82 resolution and mitigating the influence of pile-up.

83 The grooming algorithms presented here are the most important ones in ATLAS: the *Trimming*; other
 84 used as well, the *Split-Filtering* and the *Pruning* can be found in the Appendix.

85 **4.1.2 Trimming**

86 The trimming algorithm is the most important in ATLAS and the one mainly used in the work presented in
 87 this thesis. It takes advantage of the fact that contamination from soft radiation has a much lower p_T with
 88 respect to the hard-scattering component. Therefore uses a transverse momentum balance to distinguish
 89 among those. The algorithm works on a two-dimensional parameter space: R_{sub} and f_{cut} . The steps are
 90 as follows:

- 91 • k_t algorithm (but of course other choices are also possible) is used to create sub-jets with a smaller
 92 radius R_{sub} , aiming at separating the soft radiation from the hard one in different sub-jets. Typical
 93 choices are 0.2 and 0.3 (0.2 is used as standard);

- 94 • for each sub-jet, the ratio f_{cut} of its p_T with the parent jet p_T^{jet} is calculated: if then this ratio is
 95 below a certain value, the sub-jet is removed. Standard choice is $f_{cut} = \frac{p_T}{p_T^{jet}} = 0.05$;
 96 • the sub-jets which survived this procedure are the only one which compose the trimmed jet.

97 The trimming procedure is also explained in Figure 1, an example of performance in simulation with
 98 standard parameters is shown in Appendix (Figure ??).

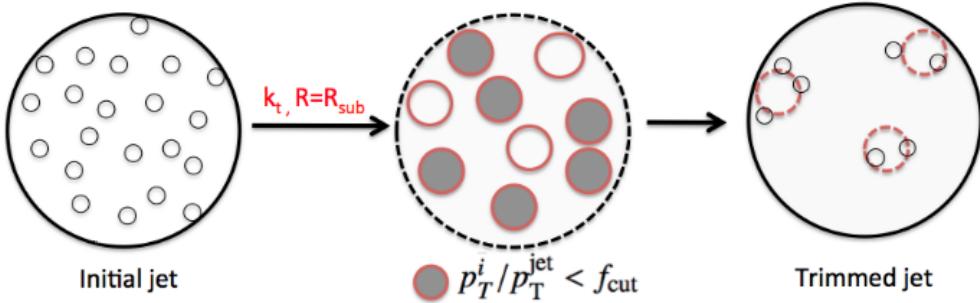


Figure 1: Schematic of the trimming algorithm.

99 4.1.3 Calorimeter Mass

100 Once the collection of constituents from the large- R jet is groomed, it is possible to use them for the
 101 measure of physical related properties such as the jet mass, since the possible sources of soft radiation
 102 from PU and UE have been reduced.

The *calorimeter mass* or m^{calo} is a widely used variable which takes as input the topo-cluster information. Given that each topo-cluster i has a 3D information on the energy deposit, E_i , the mass can be simply calculated from 4-vector properties:

$$m^{calo} = \sqrt{\left(\sum_{i \in J} E_i\right)^2 - \left(\sum_{i \in J} p_{T,i}\right)^2}$$

103 where J labels the Large- R jet.

104 4.1.4 Track Mass

105 This section briefly presents the tracks and their relation with the large- R jet's properties. There are
 106 significant advantages and few disadvantages of their usage for precise jet mass reconstruction, which are
 107 inherited both from the detector experimental properties and from the underlying physical processes.

108 First of all the excellent performance of track reconstruction and angular separation at low p_T is intrinsically
 109 better than the calorimeter one (see the Chapter 2. and Table ??). The second main advantage is that tracks
 110 can be associated with the primary vertex, thus simply excluding those from PU or other beam-induced
 111 soft radiation background (this is not the case for the UE).

112 The requirement made on tracks to achieve optimal performance are grouped into two categories, the
 113 quality of the track, i.e. if it was fully reconstructed from the detector and separated from others with no
 114 ambiguities, and the association conditions with the primary vertex:

- 115 • $p_T^{track} > 400$ MeV;
- 116 • $|\eta| < 2.5$;
- 117 • Maximum 7 hits in the Pixel and STC sub-detectors;
- 118 • Maximum 1 Pixel hole;
- 119 • Maximum 2 silicon holes;
- 120 • Less than 3 shared modules;
- 121 • Maximum 2 mm of displacement along beam axis (z_0) from the primary vertex;
- 122 • Maximum 2.5 mm of distance in x-y plane from the primary vertex and point of closest approach
 123 (d_0).

124 Given the set of tracks which pass this selection, the mass m^{track} is calculated summing up the 4-momenta
 125 of those tracks which are ghost associated to the groomed jet.

126 Apart from this benefits which derive from the tracker system, there is also an important disadvantage
 127 which comes from the underlying physics: it is completely blind to the electrically neutral component
 128 (mostly π^0) of the jet. As seen in Figure 2, the track mass (red distribution) is not only shifted towards
 129 lower values than the calorimeter mass (green distribution), but its width also degrades.

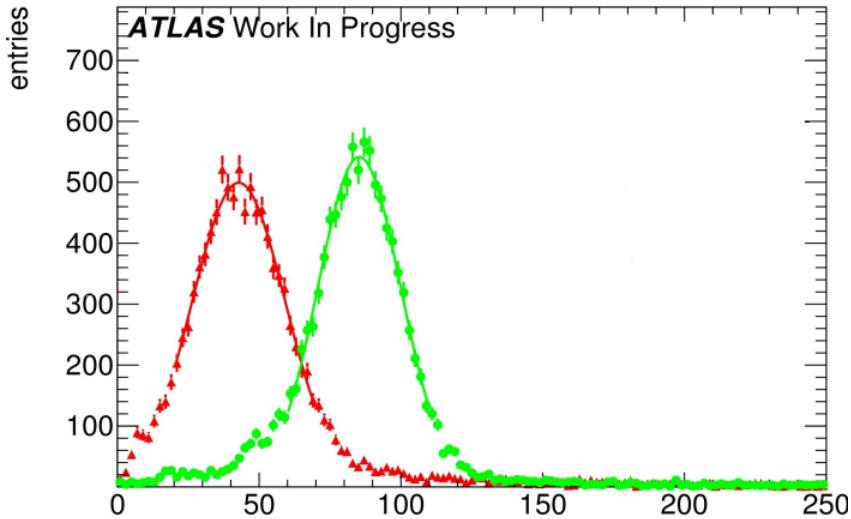


Figure 2: Mass distribution boosted W/Z : in green the m^{calo} and in red the m^{track} .

130 Tracks could be used either for independent mass reconstruction (and in this section is shown how this is
 131 not the case), or, most importantly, as an ulterior information to the calorimeter measurement.

132 **4.1.5 Performance Figure of Merit (FoM)**

133 Since we already introduced the calorimeter and track mass, a concrete, quantitative feature has to be
 134 defined in order to understand which observable is “better”, in the sense that we would prefer one or the
 135 other according to this criterion. This is often referred to as *Figure of Merit* or simply FoM.

136 There are few ways to look at the FoM: one can e.g. naively think about the mean of the mass distribution,
 137 since closer values of the mean to the e.g. W or Z mass (if we are speaking about W/Z decays), indicate a
 138 more correct mass reconstruction. However, this does not take into account the width of this distribution,
 139 as a large width spoils the reconstruction in terms of percentage of jets misreconstructed. Moreover, the
 140 mean is not as important since it can be rescaled to the desired value in a calibration procedure.

141 **4.1.6 Gaussian Fit**

142 The important feature to keep in mind, in fact, is the underlying physics which brings us to calculate the
 143 mass of a jet. In figure 3 this is made clear: if the width of the invariant mass distribution of the jet is
 144 smaller (highlighted), it allows a bigger background rejection, here shown as the QCD dijet, and a higher
 145 signal efficiency, by means of a simple mass requirement.

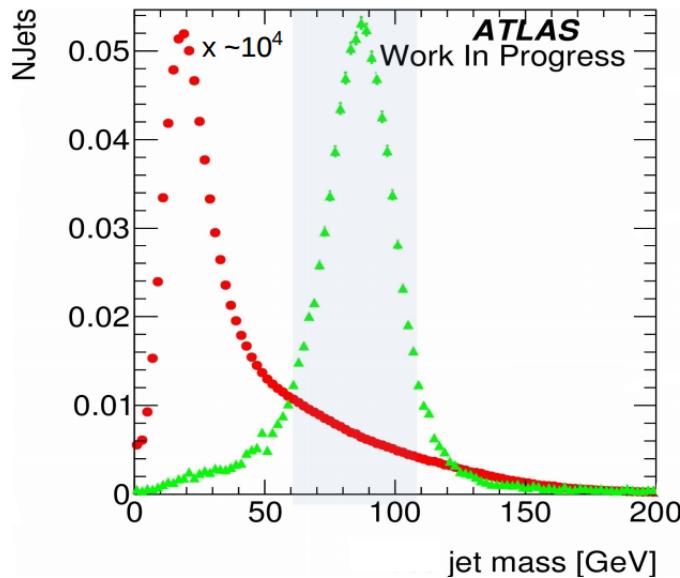


Figure 3: Mass distributions: in red the QCD dijet background rescaled, in green the W/Z from the W' sample.
 Highlighted the width of the W/Z distribution.

146 The width σ of the distribution, which can be obtained from a fit to the Gaussian core, is already a valid
 147 FoM, which has an underlying physical feature. Moreover, in order to be independent from the mean of
 148 the distribution, the width can be divided by the mean itself. This was in fact the FoM which was used
 149 at the beginning of the work for this thesis, since it provided a simple and fast solution. However, special
 150 care must be used both in the procedure of fitting Gaussian cores of responses, since they are asymmetric,
 151 and to how the tails are treated.

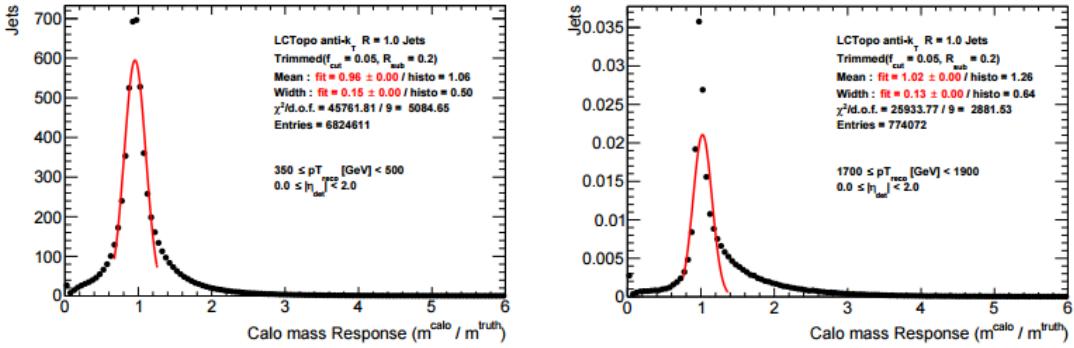


Figure 4: Mass Response distributions for the QCD multijet for various p_T ranges: on the right the failure of the Gaussian fit shows the limitation of this approach to evaluate the Figure of Merit. On the plot the fit parameters and transverse momentum ranges.

The situation is depicted e.g. in Figure 4, where a mass response is shown for calorimeter mass for QCD multijet: here the presence of a right-handed tail which enhances going from low to high transverse momenta makes the Gaussian fit clearly not the tool which provides the stability needed. The ideal tool should take care of managing the presence of at least tails outside the Gaussian core and should converge to the intuition of the standard deviation for a perfect Gaussian distribution. The closest tool to this idea was found to be the *InterQuantile Range*, which was therefore preferred and presented in the next section.

4.1.7 InterQuantile-Range

Another way to look at the mass FoM is half of the 68% of the InterQuantile range (IQnR) (here defined such as it corresponds to a sigma of a “perfect” Gaussian distribution: $q84\% - q16\%$ where $q84\%$ is the 84th percentile and $q16\%$ is the 16th, not to be confused with the InterQuartile Range (IQR) which is the $q75\% - q25\%$ and does not correspond to the sigma) divided by the Median ($\frac{1}{2} \times 68\% \text{ IQnR}/\text{median}$). It provides stability and high sensitivity to left-hand-side and right-hand-side tails.

Another important FoM, used for the work in this thesis, is the response distribution: given the reconstructed mass (calorimeter, track or whichever method) one can compare it to its *truth* mass (m^{truth}), computed from the particle at MC level before the interaction with the detector:

$$R_m = \frac{m^{reco}}{m^{truth}}$$

Standard descriptor of the FoM e.g. in [art35] and here is the IQnR of the R_m .

In Figure 5 a mass response for a single range of transverse momentum is shown, for the calorimeter mass. On the plot the contours of a standard deviation and of $q16\%$ and $q84\%$ are drawn with dashed and solid lines, respectively, showing the difference induced by the tail. This sort of plot is the key when looking quantitatively to the observable performance and can be found in the Appendix for each of the process studied in every p_T range considered. In this chapter will be shown, however, the quantity which describes this FOM, the IQnR, as a function of p_T , in order to get an understanding of the behavior in the entire spectrum and assure the exclusion of local sub-optimalities.

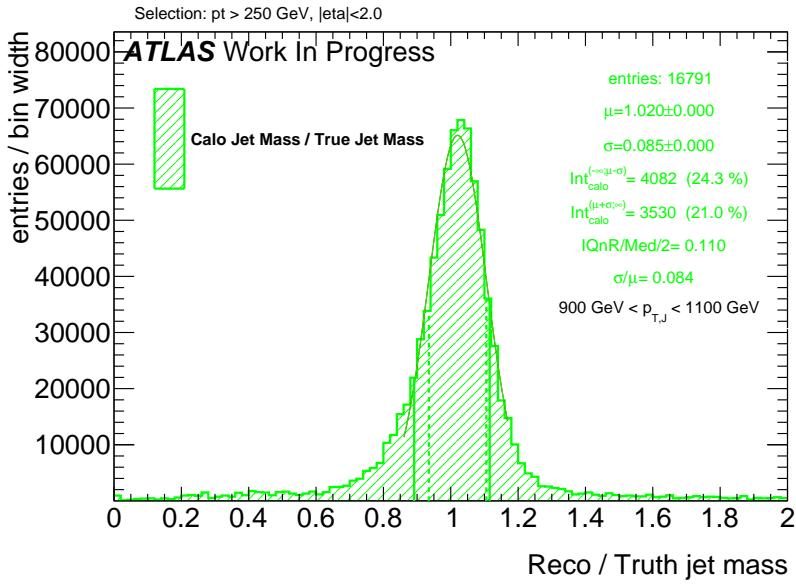


Figure 5: Calorimeter mass response plot for boosted W/Z . On the plot, right, are shown: the number of entries, the mean and the width of the fit to the Gaussian core, the integral from 0 to $\mu - \sigma$ and the one from $\mu + \sigma$ to $+\infty$, the values $\frac{1}{2} \times 68\%$ IQnR/median and σ/μ . On the distribution the dashed vertical lines represent the points $\mu - \sigma$ and $\mu + \sigma$ and the solid lines represent the $q16\%$ and $q84\%$.

175 4.2 Energy Correlation Functions

176 4.3 n-Subjettiness

177 5 Track-assisted subjet mass

178 The track-assisted subjet mass takes inspiration from the simpler development which is already implemented
179 within ATLAS, the track-assisted mass which is described briefly below for completeness.

180 5.1 Track-Assisted Mass (m^{TA})

181 The main limitation of the calorimeter mass comes from the angular resolution of the topo-clusters, which,
182 for extreme kinematic regimes, start approaching each other at the point that they hit the granularity of
183 the detector. The main advantage is that on the contrary the relative energy resolution increases at higher
184 energies.

185 The tracks instead have a very good angular resolution, but p_T relative resolution degrades linearly with
186 the transverse momentum.

187 One could then think about creating a variable which exploits the advantages of both and minimizes the
188 disadvantages. As seen, the track mass is missing the neutral component, i.e. each measurement is missing

189 the fraction $\frac{\text{neutral}+\text{charged}}{\text{charged}}$, but it could be corrected on a jet-by-jet basis: this leads to the definition of
 190 the *track-assisted mass* (m^{TA}):

$$m^{TA} = \frac{p_T^{\text{calo}}}{p_T^{\text{track}}} \times m^{\text{track}} \quad (1)$$

191 It can be intuitively understood as follows: the term m^{track} has the superior angular resolution, but misses
 192 the neutral component; the ratio $p_T^{\text{calo}}/p_T^{\text{track}}$, representing exactly the $(\text{neutral} + \text{charged})/\text{charged}$ ratio, “restores” the correct value of the mass back to $\text{charged} + \text{neutral}$.

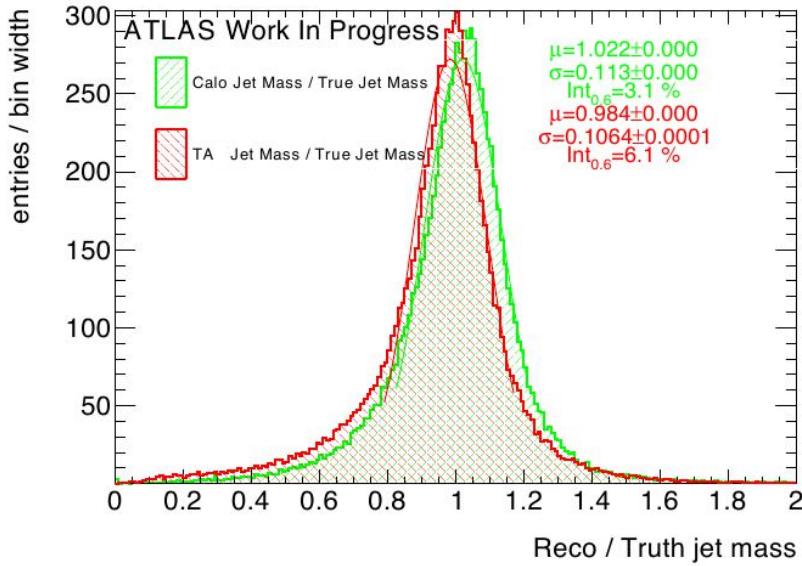


Figure 6: Track-assisted mass response plot for boosted W/Z : in green the calorimeter mass, in red the track-assisted mass. On the right are shown properties of the fit to the Gaussian core; it can be seen that the width of the m^{TA} distribution is smaller, and the mean is slightly below the calorimeter mass.

193

194 From Figure 6 the comparison of the track-assisted mass and the calorimeter mass; the width of the
 195 distribution is smaller, making this observable a good candidate for usage.

196 5.2 Advantages and Limitation of m^{TA}

197 The m^{TA} has a good handle on boosted W/Z , looking at all the transverse momentum spectrum for these
 198 results.

199 Another big advantage which supports the use of the track-assisted mass is the relatively small uncertainties:
 200 in Figure 7 the comparison of m^{calo} (left) and m^{TA} (right) fractional uncertainties on the JMS, shows how
 201 the tracking uncertainties are much smaller because of the ratio $m^{\text{track}}/p_T^{\text{track}}$. On the right plot the black
 202 line indicates the JMS fractional uncertainty for the m^{calo} , and is always above the m^{TA} . Of course this
 203 introduces another argument in the development of new techniques, which is to look for a good balance
 204 between performance and small uncertainties: a perfect observable in terms of behavior which has very
 205 big uncertainties is not really useful.

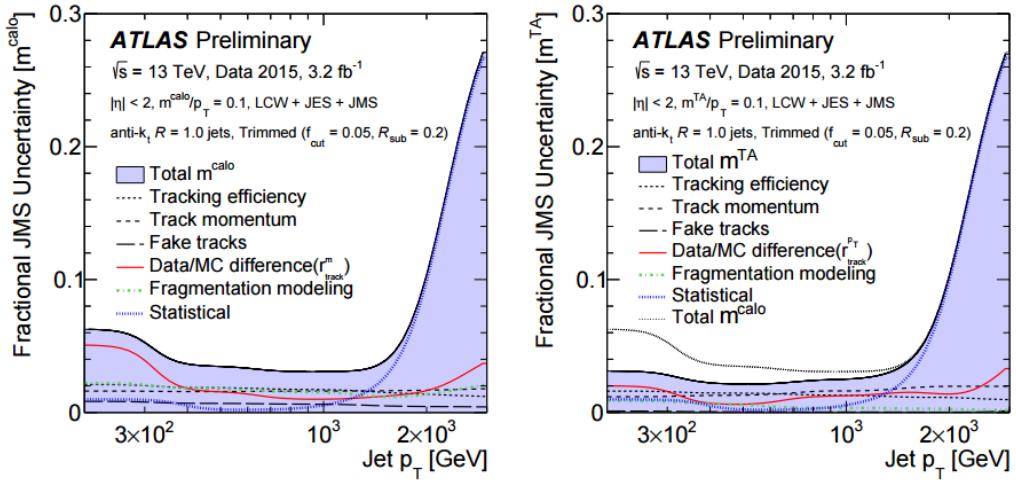


Figure 7: Comparison of the uncertainties for m^{calo} , on the left, and m^{TA} , on the right the rise on the high jet p_T is due to statistics. From the [art35].

When looking in the extreme kinematic regime, at very high p_T , as in the top plot in Figure 8, the m^{TA} shows its real strength, achieving much smaller value of the IQnR. However, there are some severe limitations which are worth noting, especially looking at the performance in different regions of transverse momentum: this is shown in the bottom plot of Figure 8, where at a low p_T it exhibits a much worse behavior.

5.2.1 Performance in $W \rightarrow q'\bar{q}$ Decays

The performance in all the bins of p_T can be studied looking at Figure 9; these plots have as horizontal axis the transverse momentum and as vertical one the value of the $\frac{1}{2} \times 68\%$ IQnR/median calculated from the correspondingly response. For W/Z jets, there is a crossing point around $p_T \sim 1$ TeV, which can be understood as the point in which the two sub-jet present start merging (sub-jet multiplicity shown in Figure 32 in Appendix).

5.2.2 Performance in $t \rightarrow q'\bar{q}b$ Decays

For top quarks the situation is much different: with respect to W/Z jets, in fact, there are two main disparities: on one side, the mass of the top quark is much higher than the one of the electroweak bosons, hence making the separation $\Delta R = \frac{2m}{p_T}$ bigger; on the other side, the decay is not anymore two-prong (two-sub-jet-like) but rather a three-prong (three-sub-jet-like) decay, one from the b -jet and the other two from the W decay. m^{TA} is here never performing better than m^{calo} , as can be seen e.g. in Figure 9, right.

5.2.3 Performance in $h \rightarrow b\bar{b}$ Decays

For boosted Higgs the m^{calo} outperforms the m^{TA} in the spectrum of transverse momentum. Although the decay is two-pronged, the mass of the Higgs is higher than the electroweak bosons, moreover another

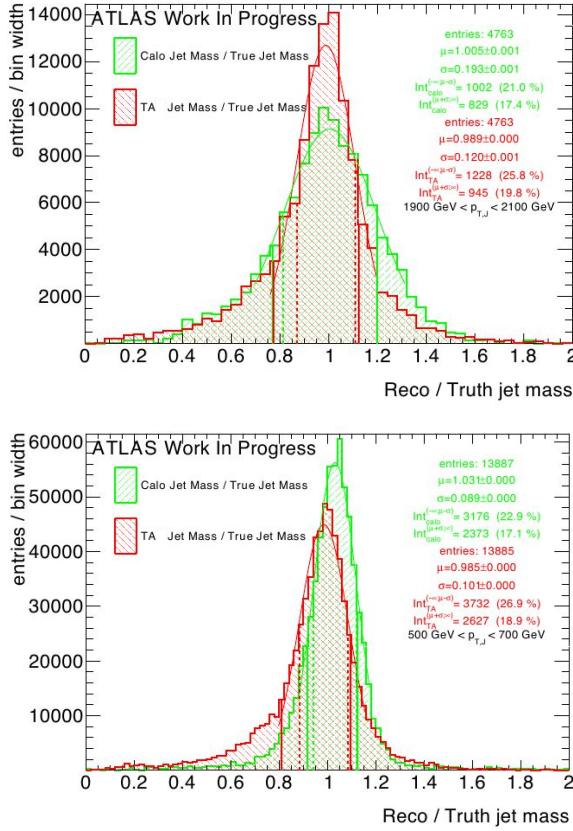


Figure 8: Mass response plots for selected ranges of p_T : on the bottom, a “low” range, $500 \text{ GeV} < p_T < 700 \text{ GeV}$, on the top an high p_T , $1900 \text{ GeV} < p_T < 2100 \text{ GeV}$. A difference in performance can be clearly seen.

227 difference lays in light quarks initiated jets and heavy quarks initiated ones, like the b-quarks from Higgs
 228 decay.

229 5.3 The Track-Assisted Sub-jet Mass (m^{TAS})

230 In this section the main outcome of the work of this thesis is presented: the *track-assisted sub-jet mass*
 231 (m^{TAS}). The main idea takes inspiration from the track-assisted mass: if one can use the tracks to exploit
 232 the better angular resolution and correct the missing neutral component jet-by-jet, there is an additional
 233 information that can be used. The neutral fraction, in fact, varies stochastically not only per-jet basis, but
 234 even per-sub-jet basis, since each sub-jet is originated from a different quark. Correcting the missed neutral
 235 component per-sub-jet, it should perform better already at an intuitive level, as it accesses information
 236 from the jet substructure. There are few question in the definition of this mass observable, whose answers
 237 are in the next section:

- 238 • Regarding the inputs:
 239 – How to select the set of tracks to be used?
 240 – Which kind of sub-jet should be used?

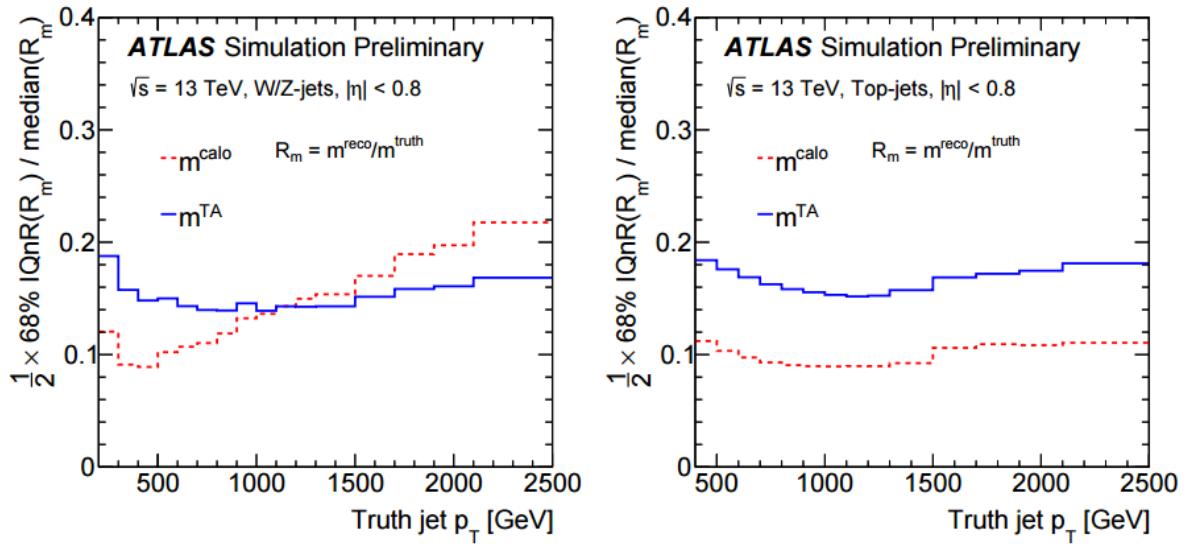


Figure 9: The comparison between the performance of m^{calo} and m^{TA} for W/Z jets (left) and top jets (right); on the x-axis the transverse momentum and on the y-axes the $\frac{1}{2} \times 68\%$ IQnR/median of the mass distribution, from [art35]. A better observable has lower values on the y-axis.

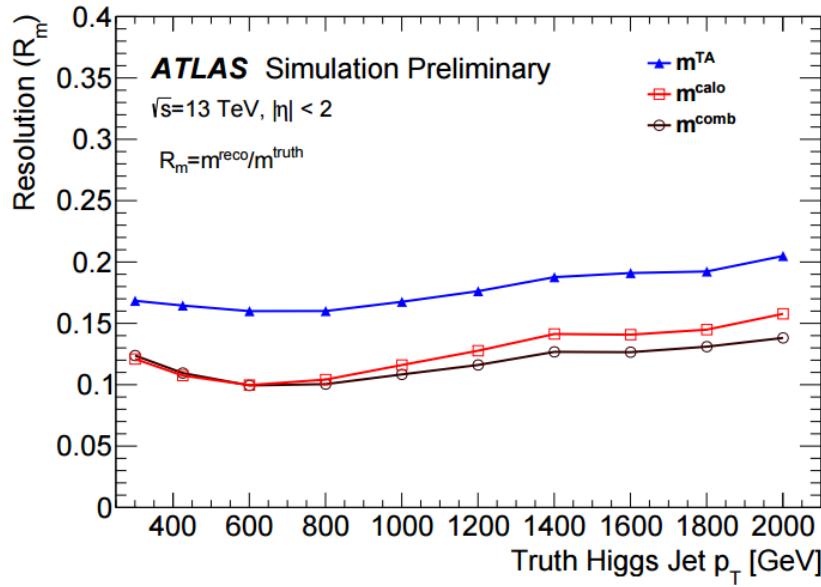


Figure 10: Performance of the m^{TA} with the boosted Higgs sample; the m^{TA} is the blue line, the m^{comb} will be described later in this chapter. From [art39]. The FoM here is the resolution of the Response.

- Regarding the procedure
 - How to associate the tracks to a sub-jet?
 - How to correct for the missed neutrals on a sub-jet basis?
 - How to add everything back together?

²⁴⁵ Those details are given in the next subsection.

²⁴⁶ 5.4 Observable Definition: Inputs

²⁴⁷ There are two inputs to the m^{TAS} : the tracks and the sub-jets. The definition of the standard inputs are
²⁴⁸ give here; alternative approaches are given in subsection 5.15.

²⁴⁹ 5.4.1 Tracks

²⁵⁰ Only the tracks that satisfy the quality criteria and primary vertex association, described in the previous
²⁵¹ section 4.1.4, are used. The tracks taken additionally are required to be ghost associated to the sub-jets
²⁵² of the groomed jet; namely only the sub-jets which survived the trimming procedure and are described in
²⁵³ the next subsection. Ghost association provides a one-to-one correspondence to the sub-jets set, and was
²⁵⁴ therefore chosen and preferred to other kind of assignments.

²⁵⁵ 5.4.2 Sub-jets

²⁵⁶ The choice of sub-jets must follow a simple requirement: of course we want to take those which most likely
²⁵⁷ come from the hard-scattering. This means that the choice of taking them after grooming is forced.

²⁵⁸ As grooming technique used, the trimming was preferred as being the standard in ATLAS and the most
²⁵⁹ flexible one for optimization studies.

²⁶⁰ The standard version of the trimming uses the k_t reclustering algorithm with radius of 0.2, with the
²⁶¹ transverse momentum ratio f_{cut} at 5%.

²⁶² As shown later, this is also the optimal configuration for sub-jets.

²⁶³ 5.5 Observable Definition: Procedure

²⁶⁴ Having tracks and sub-jets now well defined, we can describe the recipe to produce the m^{TAS} . For brevity
²⁶⁵ we will call the sub-jets SJ in the formulae below.

²⁶⁶ As said, the tracks are the one ghost-associated to the sub-jets; however, tracks which fall inside the area
²⁶⁷ of the large- R jet, but not inside the sub-jets area, are still much probably coming from the hard-scattering.
²⁶⁸ They are then associated again to the closest sub-jets via ΔR association.

²⁶⁹ Each sub-jet will have at this point some tracks associated via ghost-association and some other via ΔR
²⁷⁰ (which are maximally 5%). We call this set of tracks, a “custom” Track-Jet or TJ.

²⁷¹ At this point, the one-to-one correspondence is still preserved (for each SJ there is one and only one TJ),
²⁷² and we can move on correcting the neutral fraction.

²⁷³ Getting inspired from the formula $m^{TA} = p_T^{calo}/p_T^{track} \times m^{track}$, we would like to replicate this at sub-jet
²⁷⁴ level, i.e.

$$m^{TAS} = \sum_{SJ} \frac{p_T^{SJ}}{p_T^{TJ}} \times m^{TJ}$$

275 Since now we are working inside the sub-jets we need to change the sub-jet's 4-vector itself and not only
 276 the mass: if we call p_μ^{TJ} the Lorentz vector of the track-jet,

$$p_\mu^{TJ} = \begin{pmatrix} m^{TJ} \\ p_T^{TJ} \\ \eta^{TJ} \\ \phi^{TJ} \end{pmatrix} \rightarrow p_\mu^{TA} = \begin{pmatrix} m^{TJ} \times \frac{p_T^{SJ}}{p_T^{TJ}} \\ p_T^{SJ} \\ \eta^{TJ} \\ \phi^{TJ} \end{pmatrix}$$

277 where p_μ^{TA} is the track-assisted sub-jet's 4-vector. If we label i the i -th track-jet of the N ones present in
 278 the large- R jet,

$$m^{TAS} = \sqrt{\left(\sum_i^N p^{TA} \right)_\mu \left(\sum_i^N p^{TA} \right)^\mu}$$

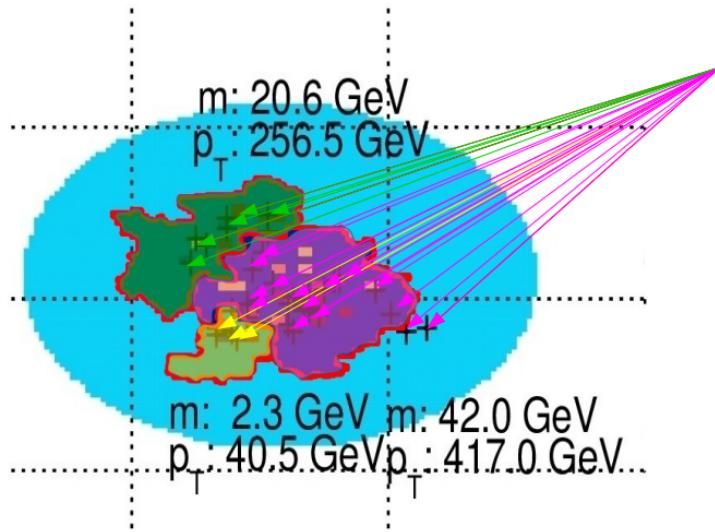


Figure 11: Pictorial event display showing the η ϕ region of a large- R jet, (in blue the catchment area of the anti- k_t) showing the different k_t sub-jets: they are highlighted in green, fuchsia and yellow. The associated track-jets (here as arrows pointing the calorimeter area) are colored with the same color of the correspondent sub-jet. Some tracks associated with ΔR procedure can be seen in the fuchsia sub-jet. The transverse momenta and mass values are also shown for the sub-jets.

279 An important remark is that, in the case of a large- R jet with only one sub-jet, the m^{TAS} has exactly
 280 the same definition of the m^{TA} . This implies, since the angular separation of the decay product scales
 281 inversely with p_T , that the performance should approach the one of the m^{TA} in the extreme kinematic
 282 regime. However, the space for improvement is precisely in the low-middle p_T regime, as seen in the m^{TA}
 283 section.

284 5.6 Performance in $W \rightarrow q' \bar{q}$ Decays

285 The boosted W/Z was the first one looked at, and with which the m^{TAS} was designed. The m^{calo} shows a
 286 fast deterioration of the performance at high p_T , and, as shown in the previous section, the m^{TA} prevents
 287 this deterioration but suffers at low transverse momenta ($p_T < 1$ TeV). The m^{TAS} has the same behavior in
 288 the extreme transverse momentum regime as the m^{TA} , since the sub-jet multiplicity peaks at one, where
 289 there are no differences between the two observables. In the low- p_T regime, on the contrary, it exploits
 290 the different charged to neutral fluctuation, achieving a better performance. This is shown in Figure 12 as
 291 a function of p_T : below ~ 1 TeV it achieves lower values of the IQnR converging from below to the m^{TA}
 292 as the number of sub-jets decreases to one.

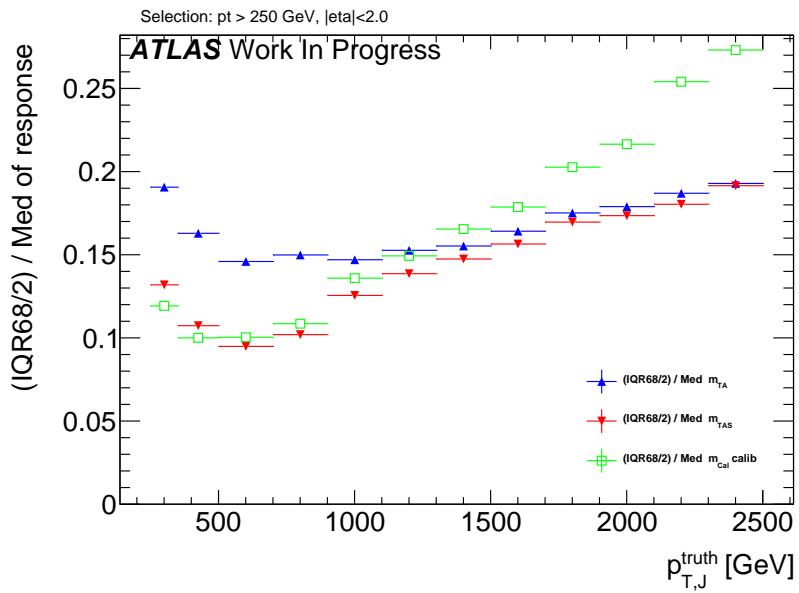


Figure 12: Performance of the m^{TAS} versus the m^{calo} and m^{TA} for the boosted W/Z sample.

293 5.7 Performance in $t \rightarrow q' \bar{q} b$ Decays

294 The boosted tops are shown on Figure 13; the m^{TAS} is comparable yet slightly worse than the m^{calo} in
 295 the low-middle p_T regime, while degrades at higher p_T approaching the m^{TA} , which is far beyond the
 296 track-assisted sub-jet mass in performance. As already noted, the worse performance can be ascribed both
 297 to the higher top-quark mass, and to its different and more complex decay topology.

298 5.8 Performance in $h \rightarrow b\bar{b}$ Decays

299 In the Randall-Sundrum graviton to di-Higgs to four b-quark, the performance is again problematic for the
 300 m^{TA} with respect to m^{calo} , which is far beyond the latter, while the performance of the m^{TAS} is partially
 301 similar to the boosted top-quark sample, but degrades much more in the extreme p_T regime, following the
 302 m^{TA} . Shown in Figure 14.

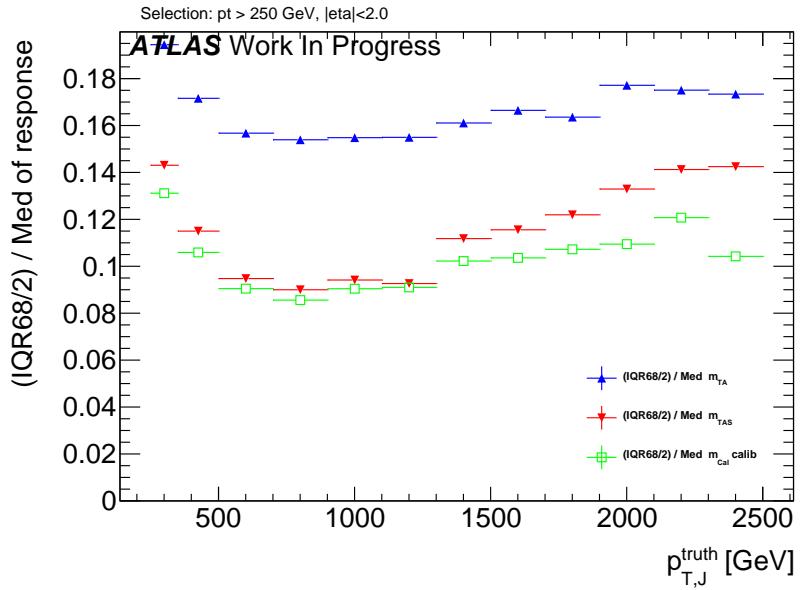


Figure 13: Performance of the m^{TAS} versus the m^{calo} and m^{TA} for the boosted top sample.

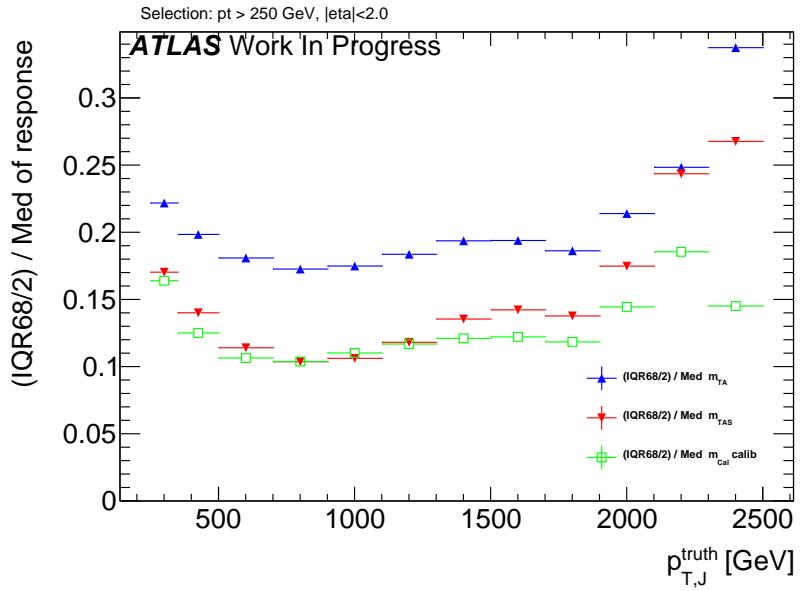


Figure 14: Performance of the m^{TAS} versus the m^{calo} and m^{TA} for the boosted Higgs sample.

5.9 Performance in QCD Multijet Events

The behavior of the QCD multijet sample is similar to the boosted W/Z sample, where the m^{TA} exhibits a crossing point in the middle-low regime $p_T \simeq 900 \text{ GeV}$ and proceeds with a better performance at high transverse momenta. Again the m^{TAS} follows this similarity showing no crossing point and an optimal overall behavior, both with respect to calorimeter- and track-assisted-based mass definition. On Figure 15.

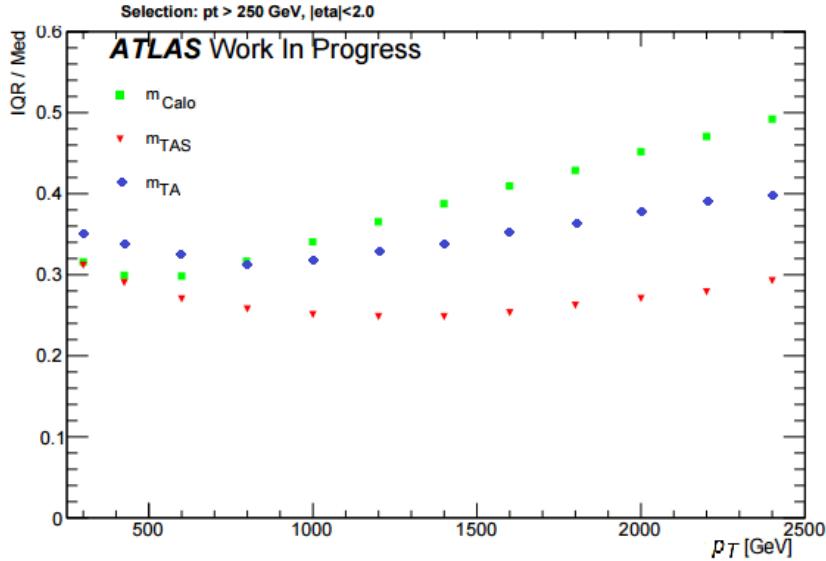


Figure 15: Performance of the m^{TAS} versus the m^{calo} and m^{TA} for the QCD multijet. Here shown IQR/Med not $\frac{1}{2} \times 68\%$ IQnR/median.

309 5.10 Performance in Massive $\tilde{W} \rightarrow q'\bar{q}$ Decays with $m_{\tilde{W}} = m_t$

310 The massive W sample is a special sample which was used to understand the behavior of the boosted
 311 tops, whether its worse resolution was coming from the higher mass of the top quark or from the more
 312 complex decay topology (three-pronged instead of two-pronged decay and b-quark presence). The sample
 313 is almost identical to the boosted W/Z one ($W' \rightarrow WZ$) but in this case the SM electroweak boson are
 314 set to have the mass of the top quark $m_{\tilde{W}} = m_t$. In fact, from the rule $\Delta R = 2m/p_T$, a bigger separation is
 315 expected between the quark from the hadronic decay. The comparison with m^{calo} is shown in Figure 16,
 316 together with the boosted top-quark for comparison. As seen here, the performance of the latter is clearly
 317 worse than the former, the trend is yet very similar. This difference is interpreted in terms of different
 318 and more complex topology and hence higher sub-jet multiplicity: in the three sub-jet structure, resolving
 319 accurately the components is more challenging.

320 5.11 Other Stability Quantifiers

321 The stability of the m^{TAS} was checked, although the IQnR is already a good quantifier of stability, explicitly
 322 for the mean of the mass response distribution and for the left-hand-side tail, as a function of the transverse
 323 momentum. This was an important check to assure the overall gaussianity of the final distribution in the
 324 whole spectrum of p_T , and suitability in regards of the calibration step, which is not discussed in this
 325 thesis.

326 The mean of the response distribution is shown for boosted W/Z decays in Figure 17, left; as seen here,
 327 despite being the mean constantly below the unity, its behavior is much more flat and independent of
 328 p_T , especially in the low-middle regime. This is surprising since the m^{calo} is already shown after the
 329 calibration step, which is not taken instead for the m^{TAS} . Conversely the left-hand-side tail of the mass
 330 response which is shown in the same figure, right, shows a more enhanced behavior than the m^{calo} , but

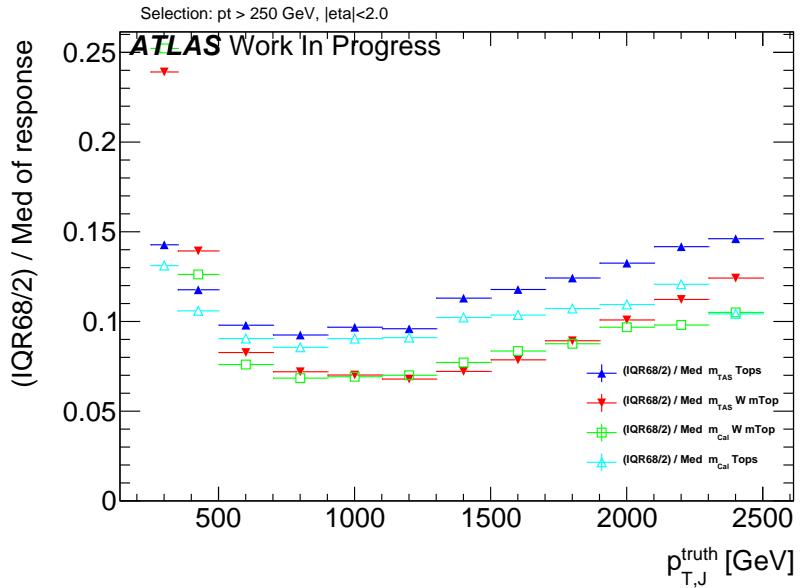


Figure 16: Performance of the m^{TAS} versus the m^{calo} for the massive W/Z (in red and green); shown on the same plot also the boosted top sample (in blue and light blue).

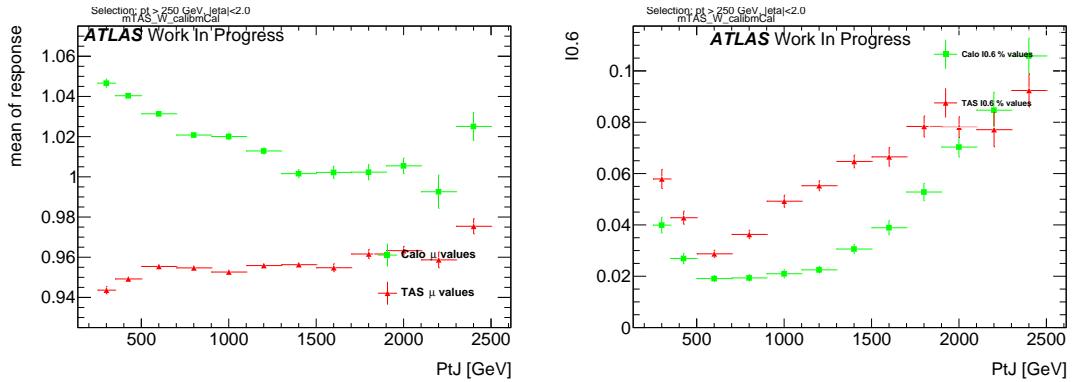


Figure 17: Stability quantifiers which were checked for the m^{TAS} : mean, on the left, and normalized left-hand side integral, on the right, of the mass response distribution. The mean is calculated from a Gaussian fit and the integral goes from 0 to 0.6.

331 still never reaches the 10%. Of course an enhancement of the tail causes a loss of gaussianity and a
 332 number of jets which are reconstructed with a lower mass than they should, but it is still comparable with
 333 the calorimeter mass.

334 Those quantifiers show analogous behavior for the other samples considered and those figures can be
 335 found in the Appendix.

336 5.12 Sub-jet Calibration

337 An additional attempt of calibrating the sub-jet was also tried and, although the results were not sub-
 338 stantially improved, it is presented in this section. This study was performed using only boosted W/Z

339 samples.

340 5.13 Preliminary Studies on Sub-jet Calibration

341 The first attempt in calibrating the sub-jets had as start a “perfect calibration”, which means using the
 342 truth-level information from the MC sample *before* the interaction with the calorimeter. Truth-level
 343 tracks are the particles in the jet which have an electric charge and are stable, truth-level sub-jets are
 344 all the particles, charged and not, which are ghost associated to the calorimeter sub-jets. There are few
 345 possibilities in doing so, here some nomenclature for this study will be introduced:

- 346 • m^{TAS} using truth-level sub-jets and tracks; normal tracks (with all detector effects) are used to assist
 347 the truth-level sub-jets;
- 348 • m^{TAS} using truth-level tracks and truth-level sub-jets; the truth-level tracks are used to assist the
 349 truth-level sub-jets;
- 350 • m^{calo} truth, calculated using only the truth sub-jets.

351 5.13.1 Perfect Calibration

352 The *perfect calibration* refers to the procedure of using m^{TAS} with truth-level sub-jets and track, i.e.
 353 looking at the best possible scenario with an ideal detector. The performance is of course expected to be
 354 optimal, because of the use of the truth-level. This step was necessary as feasibility study, to understand
 355 whether ulterior efforts in this direction were meaningful. The perfect calibration is shown in Figure
 356 18; since the performance exhibits room for big improvement below ~ 1 TeV and moderate to small
 357 improvement above this value, the second step of a simple calibration was tried.

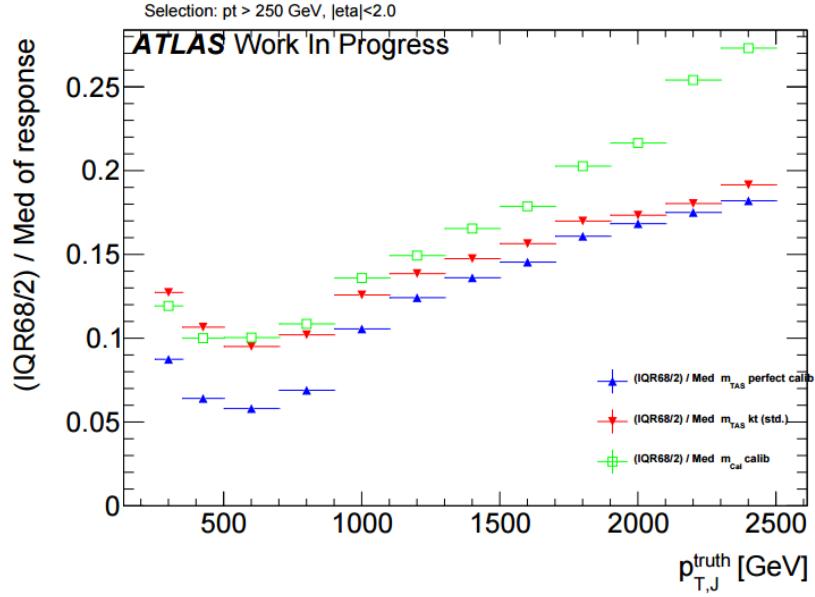


Figure 18: Performance of the perfect calibration. It shows room for improvement especially at low-middle p_T .

358 **5.13.2 Simple Sub-jet Calibration**

359 Following the example of calibration of jets in general, a simple approach to emulate this procedure was
360 tried, constructing in various bins of transverse momenta the responses of the sub-jet's energy to derive
361 the weights factors to be applied. The detailed procedure is as follows:

- 362 1. Responses in energy $R_E = E^{reco}/E^{truth}$ were built in several bins of p_T , spanning to the whole
363 transverse momentum range;
- 364 2. The mean μ_R of this response was calculated via a fit to the Gaussian core;
- 365 3. Those values (*scale factors*) were stored and applied again to the sub-jets before the computation of
366 the m^{TAS} via 4-momentum correction $E' = E/\mu_R$; the p_T (the value which only enters the m^{TAS}
367 variable) was changed then correspondingly to keep the sub-jet's mass constant.

368 This procedure was called *poor man's calibration* or PM calibration or *simple calibration*. A check on
369 the p_T response before and after calibration together with the mean of the entire Large- R jet response is
370 shown in Figure 38 and 39 in Appendix.

371 The results are on Figure 19; there are only marginal improvements in few ranges of low transverse
372 momentum where the scale factors are further away from unity, and the overall observable is not performing
373 better than the standard m^{TAS} . This is interpreted both in terms of a missing calibration as a function of
374 the η variables (having hence a befit from the crack region) and because the correction done on average
375 does not provide the sufficient handle in a jet-by-jet basis, especially when all the sub-jets are rescaled by
376 similar factors (which translates into a similarity of p_T s of the sub-jets, often the case for e.g. boosted
377 W/Z, less for boosted tops entirely contained in the large- R jet).

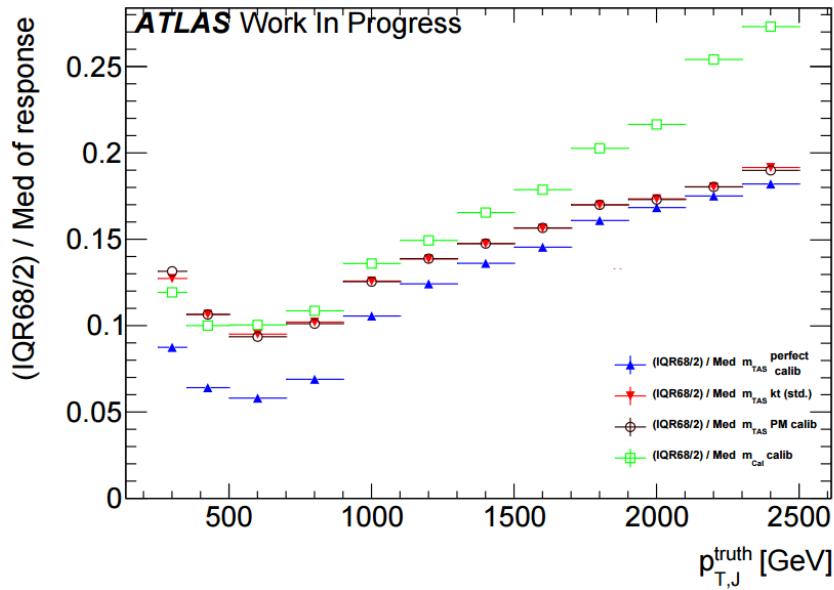


Figure 19: Performance of the poor man's calibration. The improvement is marginal throughout the entire transverse momentum space.

378 **5.14 Limitation of m^{TAS}**

379 The final effort to understand the various and competing effects, which take place in the m^{TAS} and which
 380 was inspired by the perfect calibration procedure, brought to a final study on the variable to understand
 381 the reason for the worsening of the resolution at high transverse momenta, using again the truth MC
 382 information.

383 The preliminary investigation in this direction was then the study on the track-resolution: since the track
 384 relative resolution of the transverse momentum is expected to worsen linearly with this variable, a response
 385 of the mass of the tracks was constructed, using the truth-level tracks.

386 The result is shown on Figure 20: for the samples considered, it shows a linear degradation of the mass of
 387 the tracks, both for massive and SM W/Z .

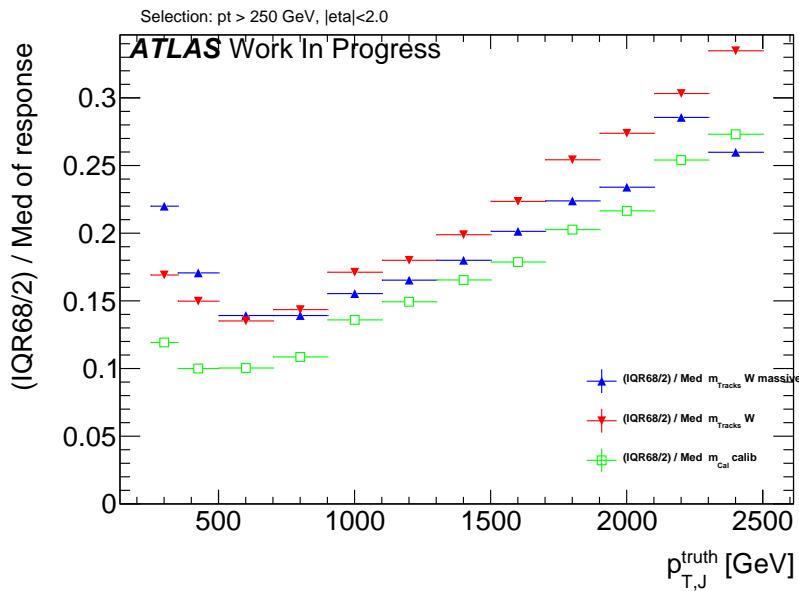


Figure 20: The performance of the track mass in blue and red for massive W sample and boosted W/Z respectively; for reference in green the calorimeter mass of the large- R jet.

388 The hypothesis of the degradation of the m^{TAS} driven by the tracks is also supported by the Figure 40
 389 in Appendix, where the truth-level tracks are used instead of real tracks to compute the variable; it can
 390 be seen the flat behavior at high p_T , hence ascribing the worsening of the resolution to tracks at higher
 391 transverse momenta.

392 A complete breakdown of the variable in terms of truth-level particles is given in Figure 21, where all the
 393 different components are separated. In particular the black dots show the m^{TAS} using truth-level sub-jets
 394 but real tracks for the track assistance procedure. Even combining this truth-level information, in fact, it
 395 shows a large worsening of the performance (truth-level sub-jets only are shown as blue dots).

396 Other results using truth-level information on boosted tops are shown and described in the Appendix.

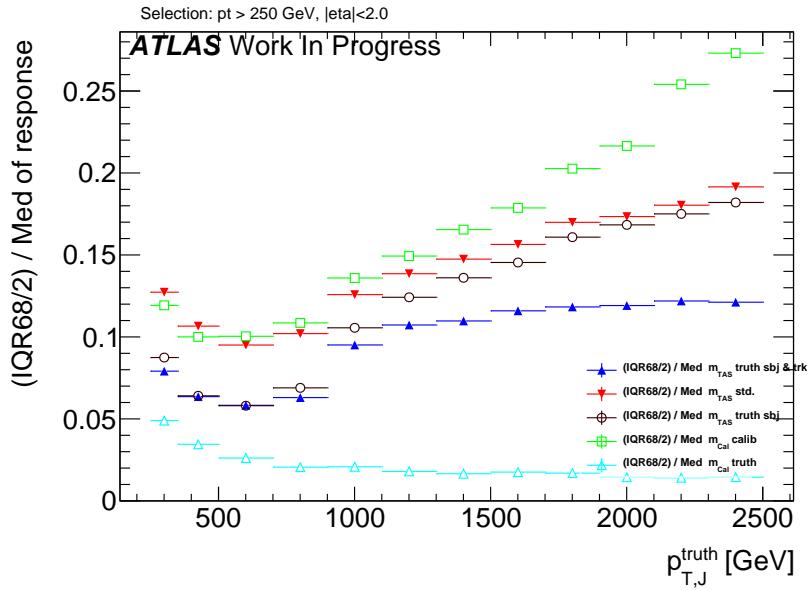


Figure 21: Breakdown of the m^{TAS} in its component using truth-level information for boosted W/Z decays.

397 5.15 Alternative Observable Definitions

398 There are quite a few ways to modify the track-assisted sub-jet mass; however, all the alternative approaches
 399 showed worse performance, and they are mentioned here for completeness only.

400 Alternatives considered were:

401 • for the tracks:

- 402 – use of tracks not as input directly, but only taking those belonging to anti- k_t reclustered
 403 track-jet with radius of 0.3 or 0.2;
- 404 – tighter or looser quality conditions were explored;
- 405 – tighter or looser primary vertex association requirement were explored.

406 • for the sub-jets:

- 407 – the trimming procedure was modified: various radii R_{sub} of the sub-jets were tested;
- 408 – the sub-jets were reclustered using not only the standard k_t , but also anti- k_t and C/A.

409 • for the procedure: different 4-momentum correction scheme was also explored.

410 The different reclustering algorithm choice has a deep impact and was studied in details, since it changes
 411 the topo-cluster added to the sub-jets and the tracks associated to them. The situation is depicted in the
 412 event-display in Figure 22; the display on the left shows the standard choice of k_t , the one on the right
 413 shows the modified approach anti- k_t .

414 In the Appendix, figure 33 34 35 the performance for boosted W/Z , tops and Higgs are shown, respectively.
 415 It can be seen that the k_t algorithm provides the best observable definition, in all the samples considered.
 416 However, the anti- k_t algorithm provides similar performances; this was an important check as the jet

417 calibration procedure currently going on in ATLAS, the *R-Scan* procedures includes the anti- k_t algorithm
 418 with radius of $R=0.2$ and aims at providing the calibration and uncertainties that could be used directly in
 419 the computation of the m^{TAS} .

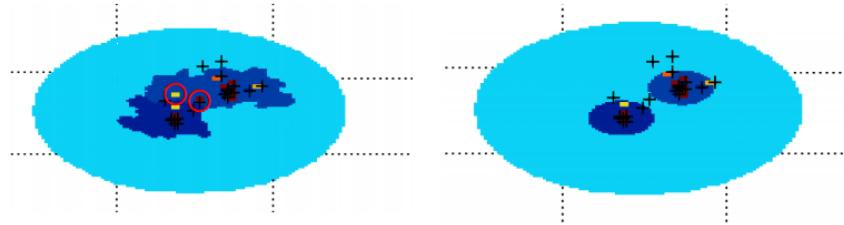


Figure 22: An example of event-display shows the differences in the reclustering algorithm used for the sub-jets: on the right k_t and on the left anti- k_t . Highlighted some constituents trimmed away with the second choice.

420 6 Combining the mass observables

421 Since the calorimeter large- R jet mass is not explicitly used in the track-assisted (sub-jet) mass, it may be
 422 possible to improve the performance creating a new observable which combines both mass definitions.

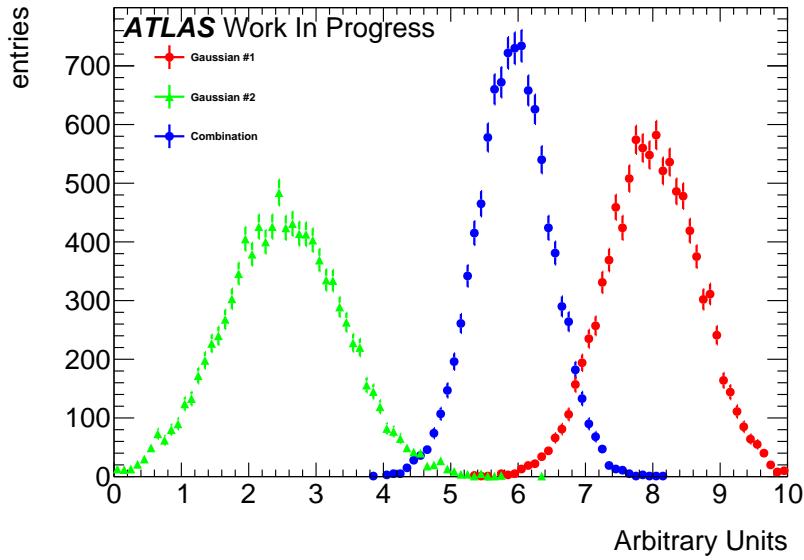


Figure 23: A toy example of the combination of two independent Gaussian observables, in red and green, and their combination, in blue. It can be seen that the combination has a smaller width.

423 This is true for both the m^{TA} and the m^{TAS} ; they are introduced in the next subsections. Provided that the
 424 two observables are nearly independent (correlation coefficient are $\sim 10\%$, see Figure 36 in the Appendix),

425 due to the Gaussian nature of the p_T and mass response, the optimal combination of the two is linear¹. An
426 example is provided in Figure 23.

427 **6.1 Combination $m^{TA} - m^{calo}$**

428 For the $m^{TA} - m^{calo}$ combination the observable are considered nearly independent, then

$$\begin{aligned} m^{comb} &= a \times m^{calo} + b \times m^{TA}, \\ a &= \frac{\sigma_{calo}^{-2}}{\sigma_{calo}^{-2} + \sigma_{TA}^{-2}} & b &= \frac{\sigma_{TA}^{-2}}{\sigma_{calo}^{-2} + \sigma_{TA}^{-2}} \end{aligned} \quad (2)$$

429 where σ_{calo} and σ_{TA} are the m^{calo} 's and m^{TA} 's resolution functions. The m^{comb} then is the $m^{TA} - m^{calo}$
430 combination.

431 **6.2 Combination $m^{TAS} - m^{calo}$**

432 There is a main difference between the m^{TAS} and m^{TA} when it comes to combination: since the m^{TAS} is
433 using sub-jet level information but m^{TA} not, the correlation with the m^{calo} is expected to be higher. This
434 can be seen e.g. in the plots in Figure 24 (additional plots shown in Figure 37 in Appendix), where the
435 correlation is not only higher for the simple W/Z and Higgs jets, but above 50% for tops. The assumption
436 of independent variables here falls, forcing a more complete approach. The Ansatz is to take into account
437 the correlation via the formula:

$$\begin{aligned} m_{TAS}^{comb} &= w \times m^{calo} + (1 - w) \times m^{TAS}, \\ w &= \frac{\sigma_{TAS}^2 - \rho \sigma_{calo} \sigma_{TAS}}{\sigma_{calo}^2 + \sigma_{TAS}^2 - 2\rho \sigma_{calo} \sigma_{TAS}} \end{aligned} \quad (3)$$

438 where now m_{TAS}^{comb} is the new $m^{TAS} - m^{TA}$ combination. This expression reduces then to the form:

$$\begin{aligned} m_{TAS}^{comb} &= a \times m^{calo} + b \times m^{TAS}, \\ a &= \frac{\sigma_{TAS}^2 - \rho \sigma_{calo} \sigma_{TAS}}{\sigma_{calo}^2 + \sigma_{TAS}^2 - 2\rho \sigma_{TAS} \sigma_{calo}} & b &= \frac{\sigma_{calo}^2 - \rho \sigma_{calo} \sigma_{TAS}}{\sigma_{calo}^2 + \sigma_{TAS}^2 - 2\rho \sigma_{TAS} \sigma_{calo}} \end{aligned} \quad (4)$$

439 which reduces to equation (2) after simple algebra for the case when $\rho = 0$. Of course, this value can be
440 set to the value of the specific sample considered, or to an average of 0.3 if one wants to give a definition
441 generally valid for all the cases considered; in this case, the performance would be slightly sub-optimal.

¹ If the joint distribution of the responses is Gaussian, then one can write their probability distribution function as $f(x, y) = h(x, y) \times \exp[A(\mu) + T(x, y)\mu]$, where x is the calorimeter-based jet mass response, y is the track-assisted jet mass response, μ is the common average response, and h, A, T are real-valued functions. This form shows that the distribution is from the exponential family and therefore T is a sufficient statistic. Since the natural parameter space is one-dimensional, T is also complete. Therefore, the unique minimal variance unbiased estimator of μ is the unique unbiased function of $T(x, y) = x/\sigma_x^2 + y/\sigma_y^2$. See e.g. Ref. [statistic] and [art35] for details.

442 6.2.1 Procedure

443 The procedure of producing the m_{TAS}^{comb} is defined as follows:

- 444 1. For the given sample, the m^{TAS} and m^{calo} are produced;
- 445 2. The mass responses are also produced for the given ranges of p_T ;
- 446 3. For each of these responses, the value of the IQnR as defined previously is calculated and stored;
- 447 4. The average correlation factor of 0.3 is assumed;
- 448 5. With the formula 3, m_{TAS}^{comb} is calculated using the m^{TAS} , m^{calo} and the values stored from before.

449 A remark on the procedure: the step 3. uses values of the IQnR because this was showed to be a more
450 robust way to look at the response and fit-independent. For step 4. the correlation factor was decided to
451 be an average of the samples considered.

452 Additionally, the IQnR weights are produced for each sample specifically. In order to give a sample-
453 independent definition of the m_{TAS}^{comb} , following also the procedure adopted for the m^{comb} , these weights
454 could be taken from a QCD multijet sample and applied indiscriminately to the particular case. Here of
455 course the performance would be again sub-optimal, since the variable was not developed in an ad-hoc
456 way.

457 Throughout the results presented in the following sections, both observables were calculated with ad-hoc
458 weights. Quantitative statements between them would still hold in the case of QCD weights. However,
459 when confronting e.g. m^{TAS} with them it has to be kept in mind that in this case their performance is
460 overestimated, since this choice, although being more general, would perform slightly worse.

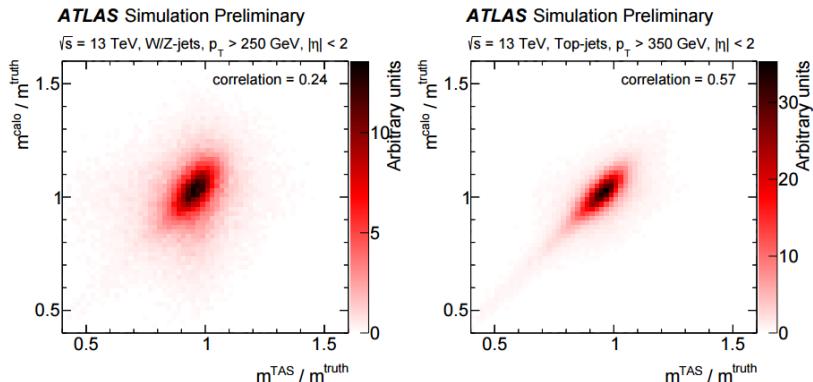


Figure 24: The calorimeter based jet mass response versus the track-assisted sub-jet mass response, on the left for boosted W/Z on the right for boosted tops.

461 6.3 Performance in $W \rightarrow q' \bar{q}$ Decays

462 On the boosted W/Zs sample, the performance of the m_{TAS}^{comb} outperforms all the other definitions
463 throughout all the transverse momentum space; on Figure 25 they are shown for reference together with
464 the m^{TAS} . It can be noted here that the track-assisted sub-jet mass, although being sub-optimal, has
465 comparable performance, yet presenting fewer complications due to the combination procedure.

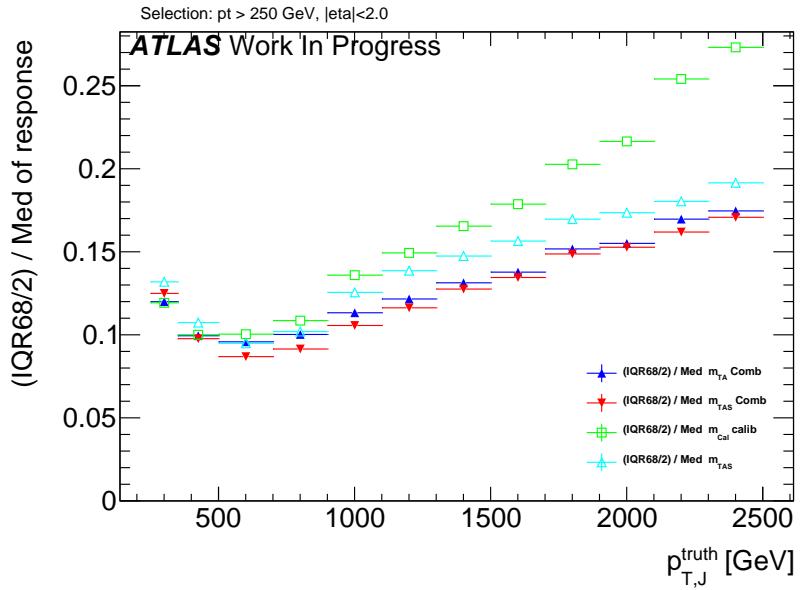


Figure 25: Performance of the combined mass on W/Z samples; here shown the two definitions of the combined mass, m^{comb} and m_{TAS}^{comb} , together with the calorimeter mass and the track-assisted sub-jet mass.

466 6.4 Performance in $t \rightarrow q' \bar{q} b$ Decays

467 The boosted top sample remains the most challenging one also with the combined mass; as seen on Figure
 468 26, the m^{comb} performs quite similarly to the calorimeter based mass definition, yet behaving considerably
 469 better than the m^{TAS} especially at high transverse momentum. The m_{TAS}^{comb} , however, outperforms all the
 470 other definitions, and shows its optimal observable strength at middle p_T i.e. in the range $1 < p_T < 1.6$
 471 TeV.

472 6.5 Performance in $h \rightarrow b\bar{b}$ Decays

473 Again, for the Higgs decay there are similarities as for the top sample; on Figure 27 the two definitions of
 474 the combined mass, together with the simpler m^{TAS} . Although this variable is lightly sub-optimal yet still
 475 comparable in the low to intermediate range in transverse momenta, where the tracks are driving a decrease
 476 in performance for the high to very-high p_T . The m_{TAS}^{comb} uses this advantage to achieve optimal behavior
 477 in the entire transverse momentum spectrum, outperforming both m^{calo} and m^{comb} almost everywhere.

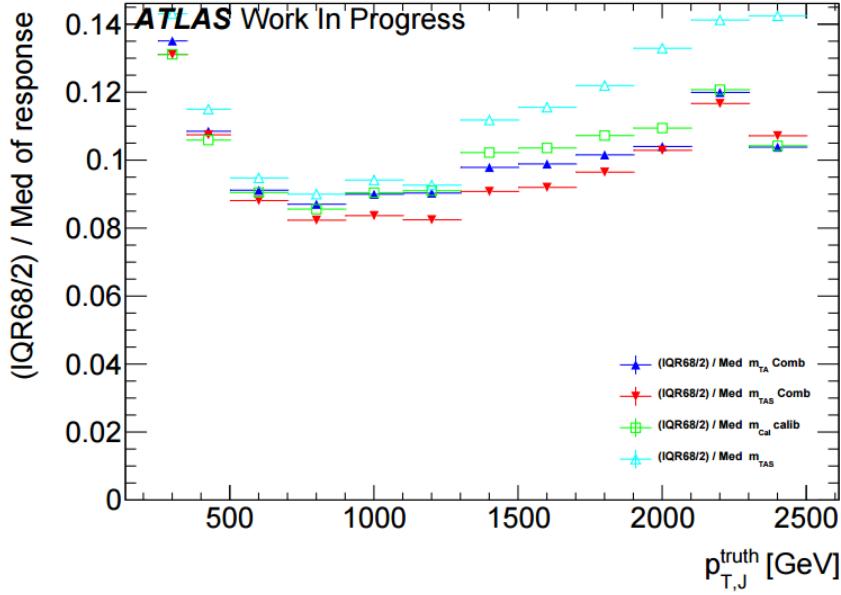


Figure 26: Performance of the combined mass on the top sample; here shown the two definitions of the combined mass, m^{comb} and m_{TAS}^{comb} , together with the calorimeter mass and the track-assisted sub-jet mass.

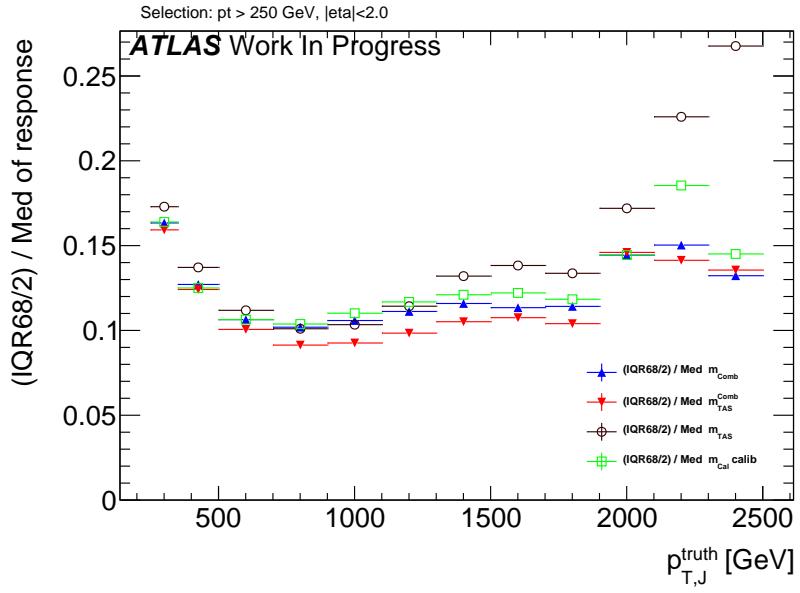


Figure 27: Performance of the combined mass on the Higgs decay; here shown the two definitions of the combined mass, m^{comb} and m_{TAS}^{comb} , together with the calorimeter mass and the track-assisted sub-jet mass.

478 7 Energy Correlation Functions and n-Subjettiness

479 8 Conclusions & Outlook

480 8.1 Jet mass observables

481 The m^{TAS} variable was developed for the large- R jet mass; it combines the information of the tracker- and
 482 calorimeter-system to achieve an higher precision in the jet mass reconstruction, correcting the missed
 483 neutral fraction which is absent in the tracker but not in the calorimeter. With respect to the m^{TA} , it
 484 applies this correction at sub-jet by sub-jet level and not at jet by jet level, therefore providing a more
 485 accurate reconstruction. It was shown in Monte Carlo simulation to be a very good observable confronting
 486 quantitatively with the other definitions which are either standard or in preparation, m^{calo} , m^{TA} and m^{comb} .
 487 In fact, it behaves better in terms of $\frac{1}{2} \times 68\%$ IQnR/median and all the other ways to look at the figure of
 488 merit, the mass response, for the boosted W/Z and QCD sample; is always better than the m^{TA} and similar
 489 to the m^{calo} for the boosted tops and Higgs. Moreover, it is a slightly worse observable than the m^{comb} , yet
 490 being comparable, and avoiding the development of ad-hoc weights. The optimal configuration of m^{TAS}
 491 is shown and confronted with different approaches, in particular in terms of different trimming procedure
 492 of the large- R jet to be used as an input. All the components of the observable have been studied with the
 493 use of truth Monte Carlo information without detector effect, in order to evaluate quantitatively its limits
 494 and strengths; the track p_T measure degradation was found to be the cause of the variable decreasing
 495 performance at higher transverse momenta.

496 The m_{TAS}^{comb} is the logical extension of the m^{TAS} , which improves by construction the results beyond the
 497 m^{calo} and the m^{TAS} , combining these two variables on the same way of the m^{comb} , but taking into account
 498 the higher correlation factor which is inherited from the sub-jet usage. Weights for its construction can
 499 be in both cases either derived specifically for the sample considered, or constructed on average with the
 500 QCD sample, in this case getting a sub-optimal performance. In all the cases studied, it has a better
 501 behavior than the m^{comb} , m^{calo} and m^{TA} .

502 For the very conclusion, both the variables constructed in the work of this thesis, m^{TAS} and m_{TAS}^{comb} ,
 503 exhibit a better performance of their counterparts, m^{TA} and m^{comb} , which are now ready to be use or in
 504 preparation within the ATLAS collaboration, and share the same advantages -and disadvantages. Further
 505 steps are necessary to get this observables to usage: calibration and uncertainties.

506 8.1.1 Outlook

507 The outlook of the m^{TAS} and m_{TAS}^{comb} variables follows two main scenarios, concerning the calibration and
 508 uncertainties determination which are necessary to get this observables ready to be used. The procedure
 509 involved are already fully understood, since the the same was applied or is being applied for the m^{TA} and
 510 m^{comb} .

511 For the simple scenario here the procedure that would take place is the direct Monte Carlo calibration
 512 of the m^{TAS} , aiming at correcting the reconstructed jet mass to the particle-level jet mass by applying
 513 the calibration factors derived from QCD multijet events, an analogous procedure to the one described in
 514 Section ?? for the jet energy scale.

- 515 The more complex scenario considers an additional calibration to the sub-jets with R=0.2, which is already
516 at an advanced stage within ATLAS for anti- k_t reclustering algorithm (it has a slightly worse performance
517 than k_t , as presented previously).
- 518 The uncertainties are expected to be similar to the one which were derived for the m^{TA} and which are
519 compared to the m^{calo} on Figure 7; the tracking uncertainties are smaller for the track-assisted mass
520 because of the ratio m^{track}/p_T^{track} and will be smaller as well for the track-assisted sub-jet mass since it
521 uses the same ratio.
- 522 In-situ uncertainties were derived for the m^{TA} with a sample of enriched top-quark; the same technology
523 used here can be applied to the m^{TAS} .
- 524 In the more complex scenario, the uncertainties could be derived for the sub-jets R=0.2 reclustered with
525 anti- k_t .

526 **8.2 Energy Correlation Functions and n-Subjettiness**

527 ***here sascha conclusions***

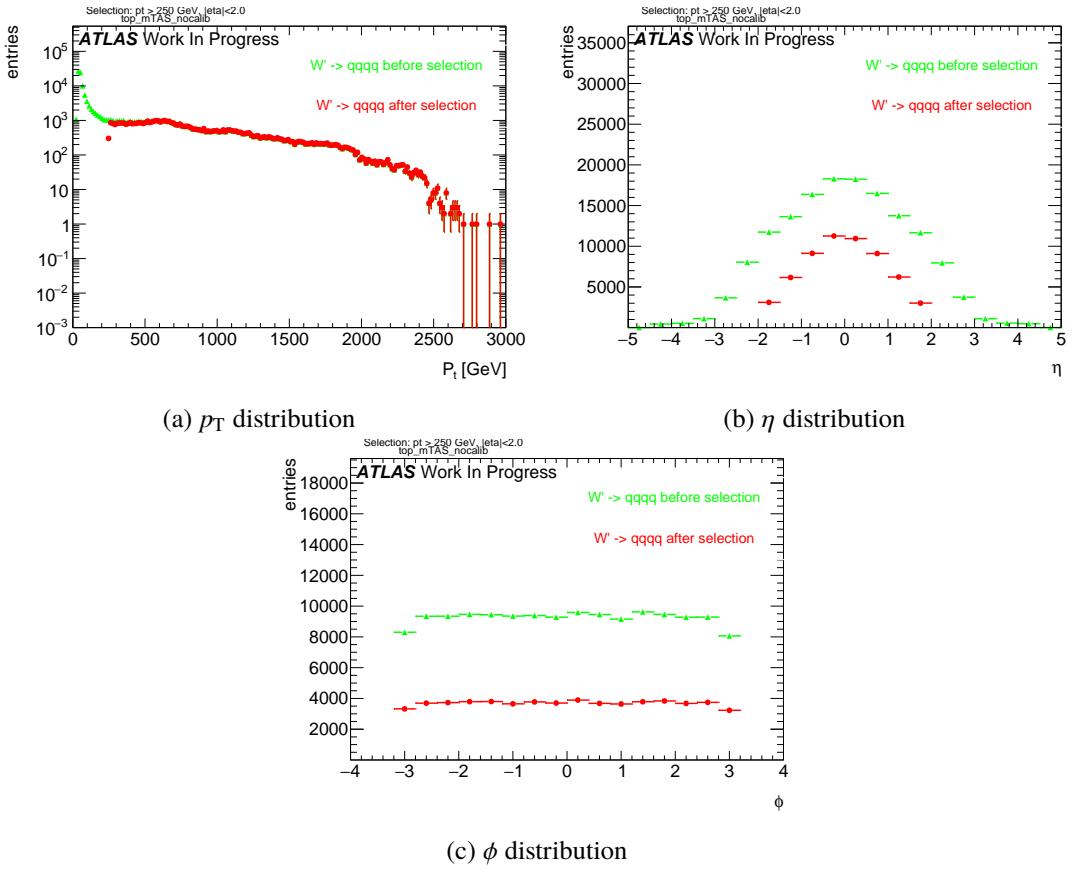


Figure 28: Boosted tops kinematic distribution.

528 Appendix

529 Jet Mass Observable Distribution Kinematic distribution for all the samples, p_T η and ϕ is shown.

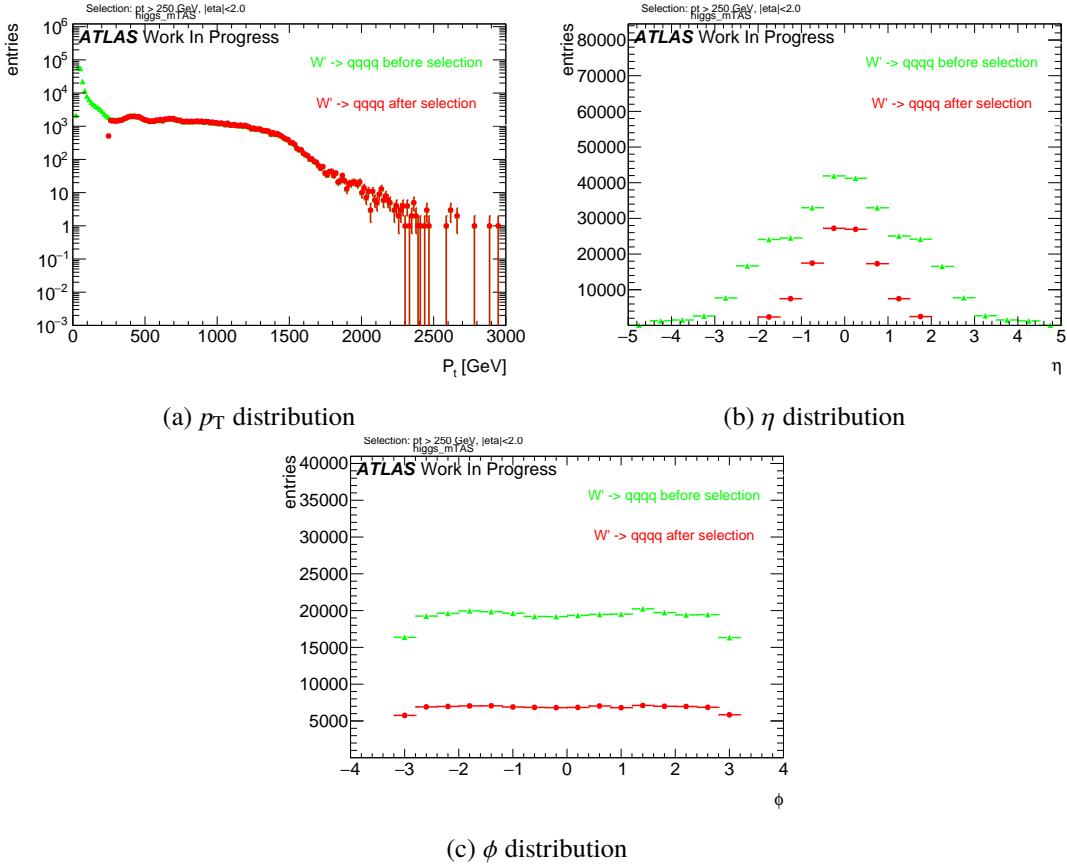


Figure 29: RS-Graviton kinematic distribution.

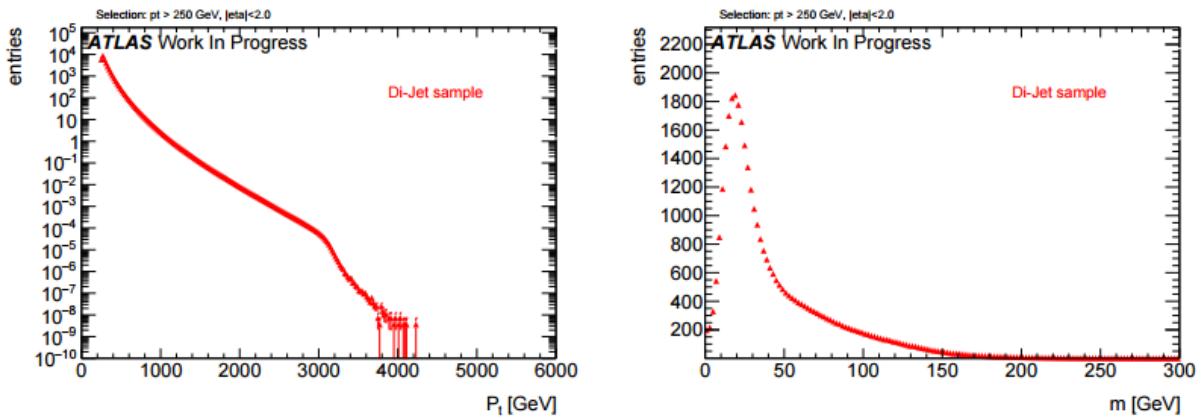


Figure 30: QCD dijet transverse momentum and mass distributions.

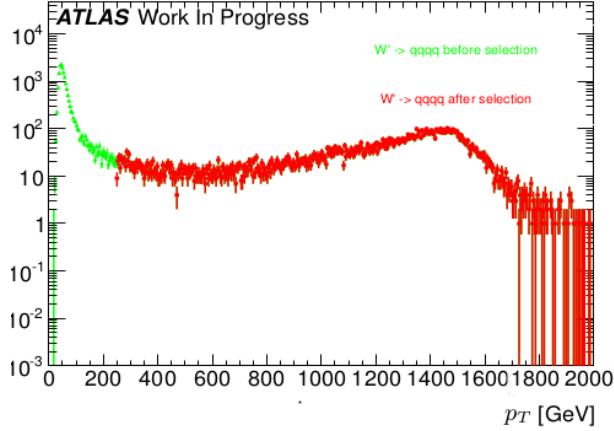


Figure 31: The p_T distribution of a 3 TeV resonance from the hadronically decaying W or Z , in logarithmic plot. As can be seen, the jacobian peak is around $p_T \simeq m_{W'}/2 \simeq 1.5$ TeV.

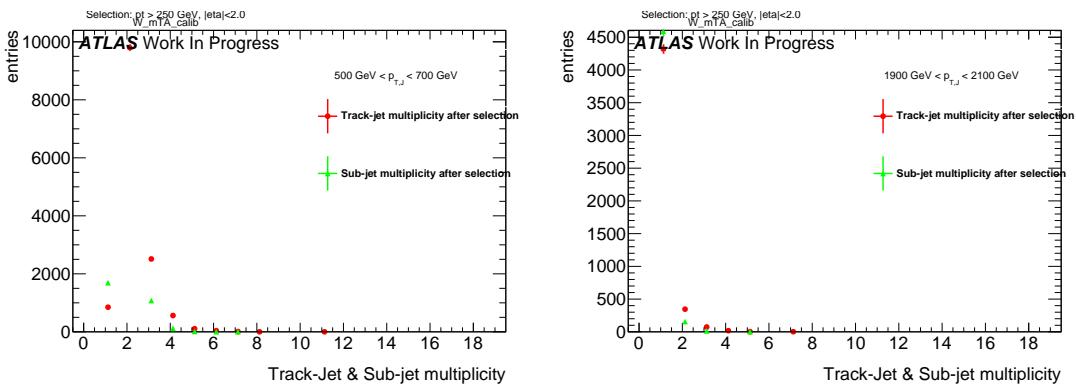


Figure 32: Sub-jet and Track-jet (jets created having tracks as input) multiplicity, for selected bins of transverse momentum.

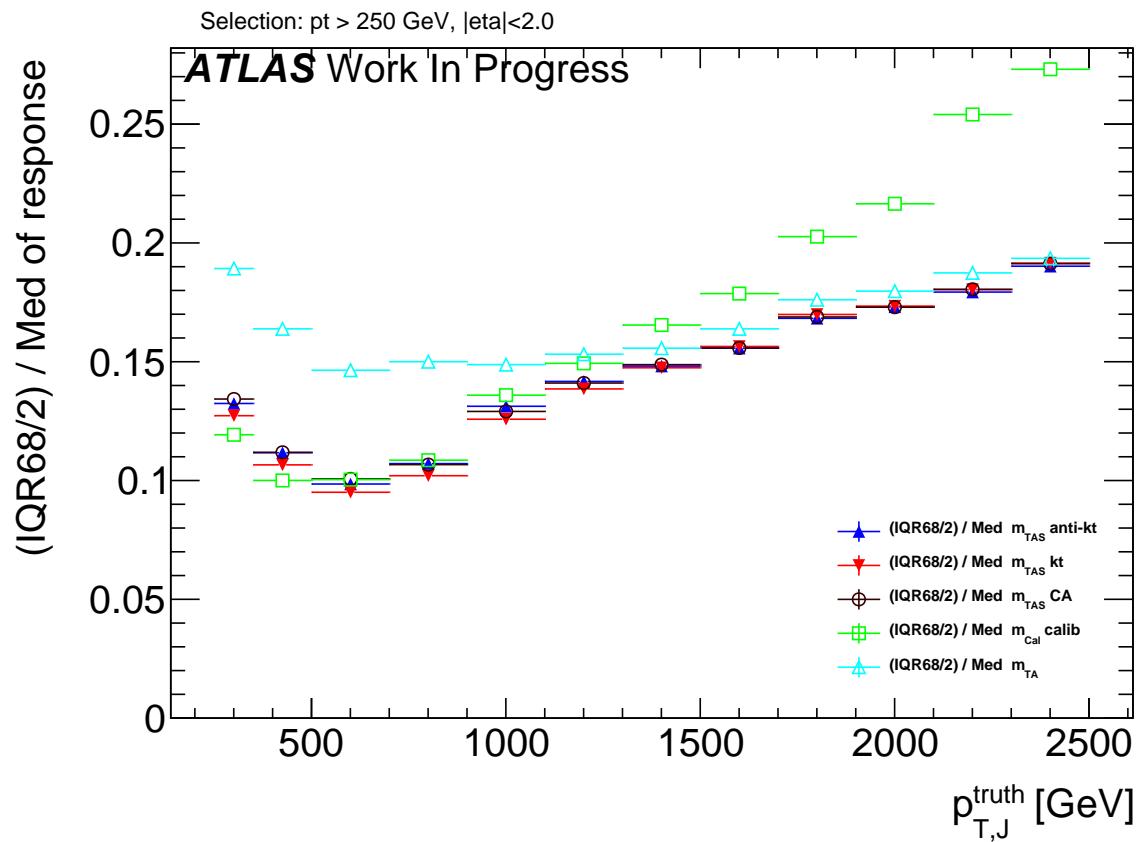


Figure 33: Performance of m^{TAS} with different reclustering algorithm for the sub-jets: anti- k_t , k_t and C/A. Boosted W/Z sample.

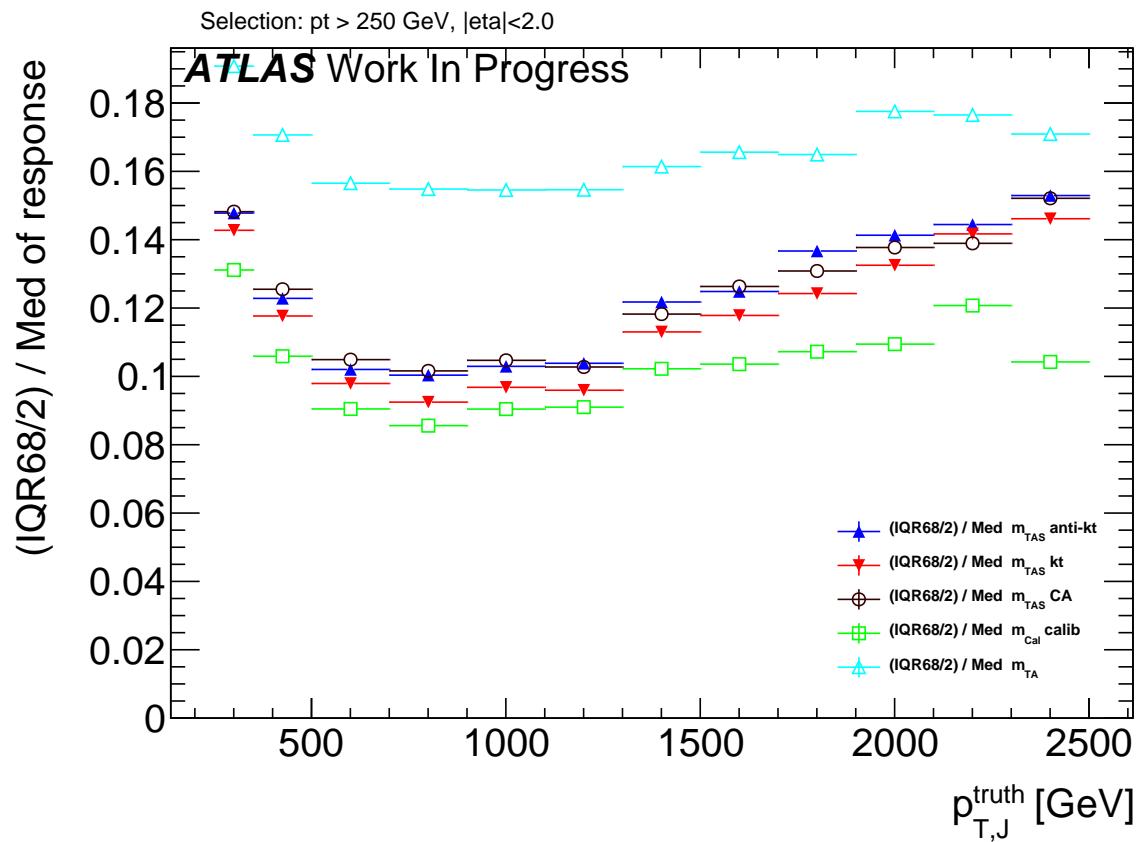


Figure 34: Performance of m^{TAS} with different reclustering algorithm for the sub-jets: anti- k_t , k_t and C/A. Boosted top sample.

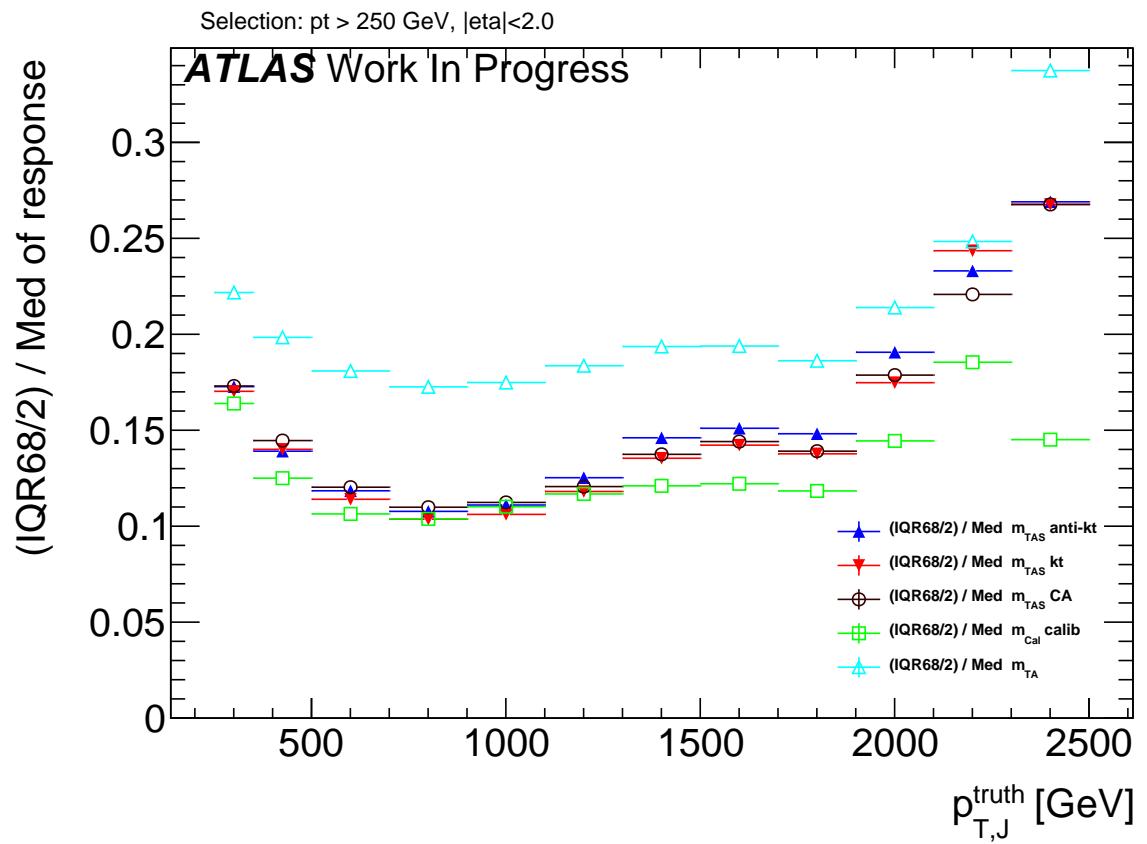


Figure 35: Performance of m^{TAS} with different reclustering algorithm for the sub-jets: anti- k_t , k_t and C/A. Boosted higgs sample.

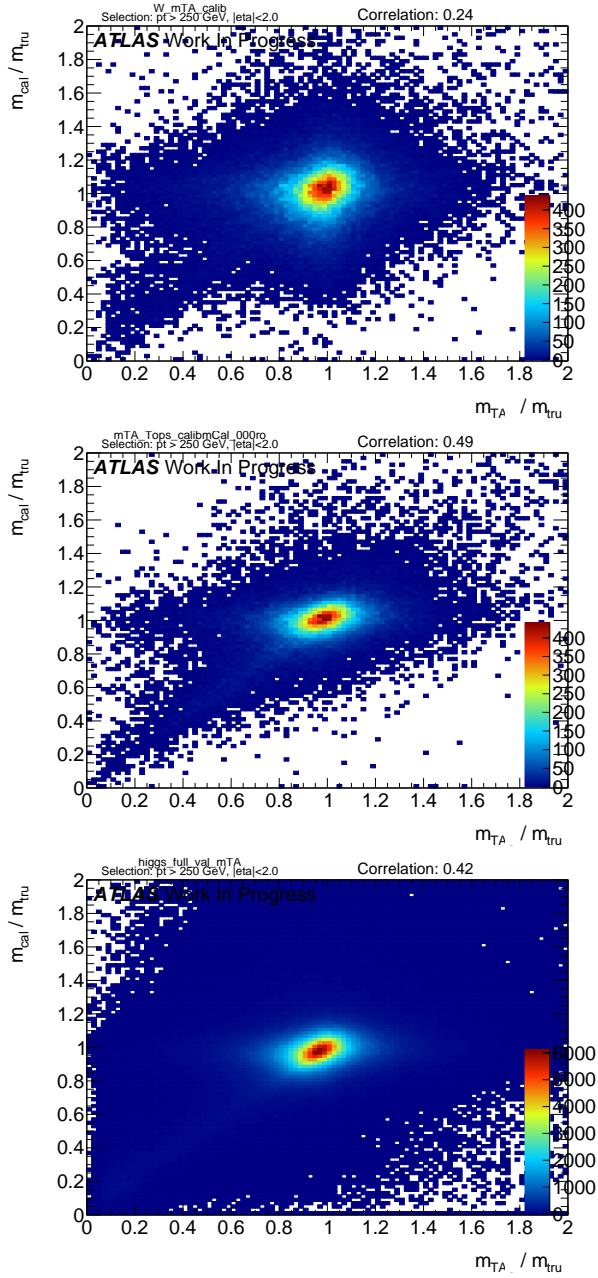


Figure 36: Calorimeter based jet mass response vs the track-assisted mass response for the three signal samples. Correlation coefficient is indicated on the top right.

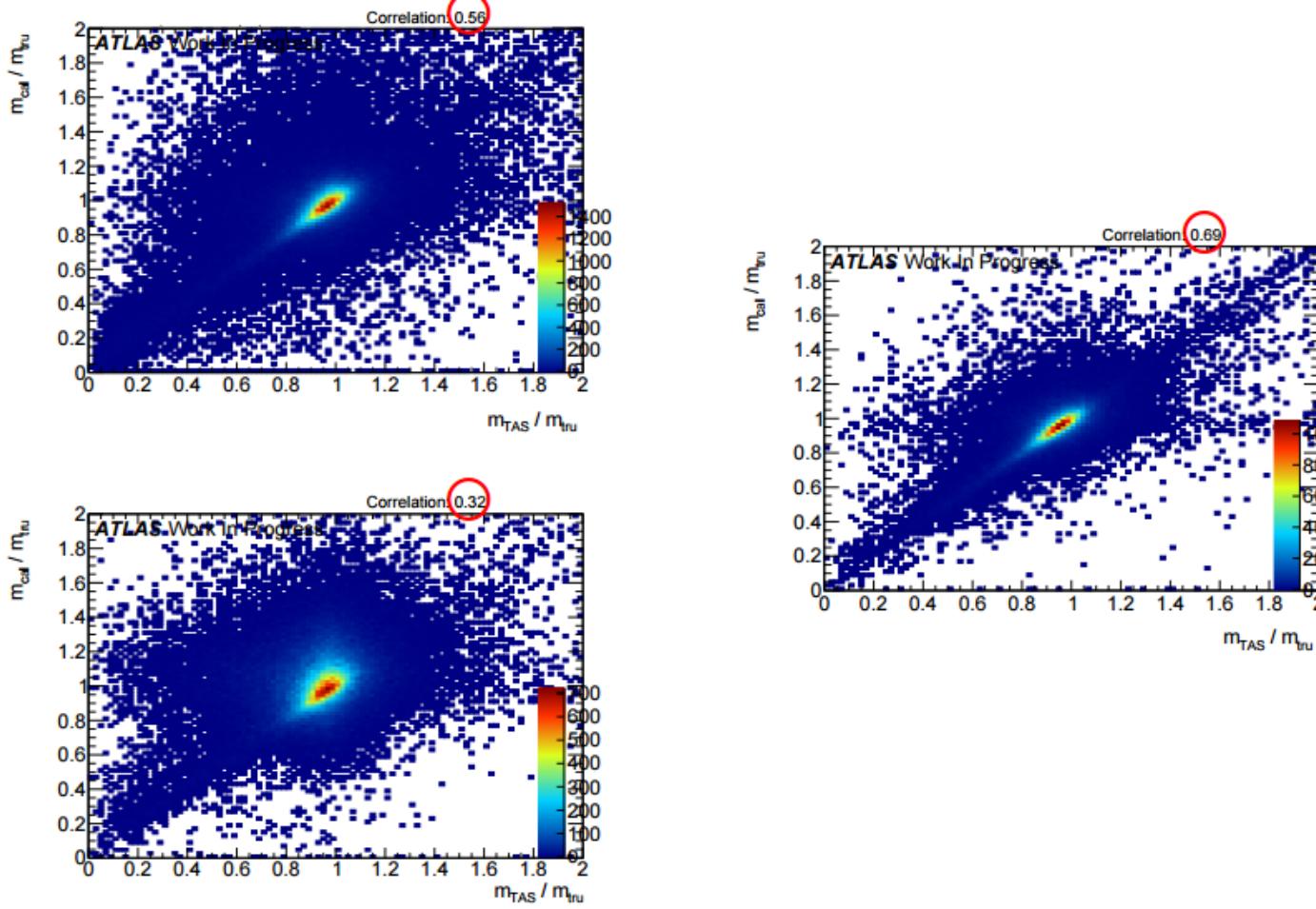


Figure 37: Calorimeter based jet mass response vs the track-assisted sub-jet mass response for the three signal samples. Correlation coefficient is indicated on the top right and highlighted. On the left, top, the higgs sample, bottom, the W/Z ; on the right the top-quark sample.

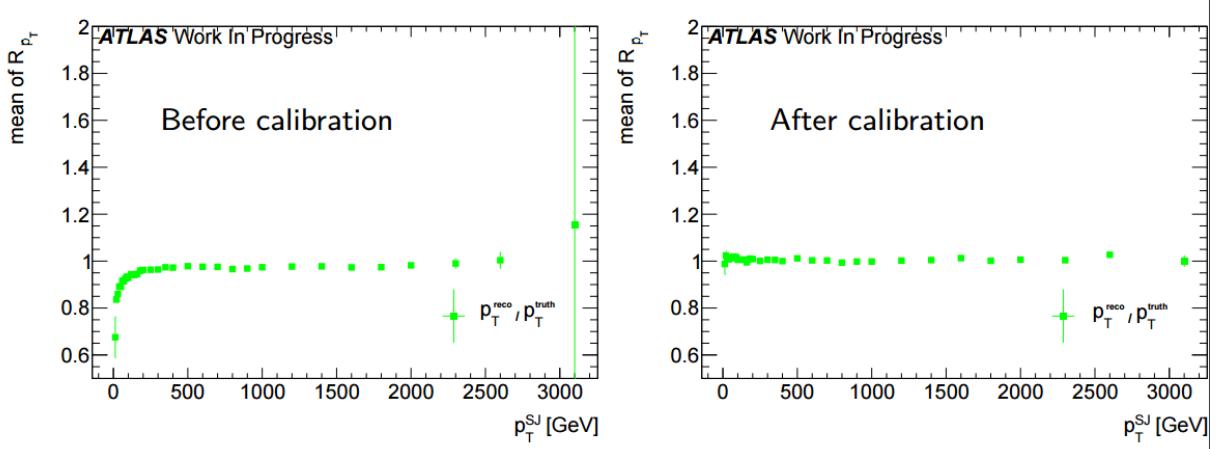


Figure 38: Poor's man calibration effect on mean of transverse momentum's response of the sub-jet, before, left, and after, right, the procedure.

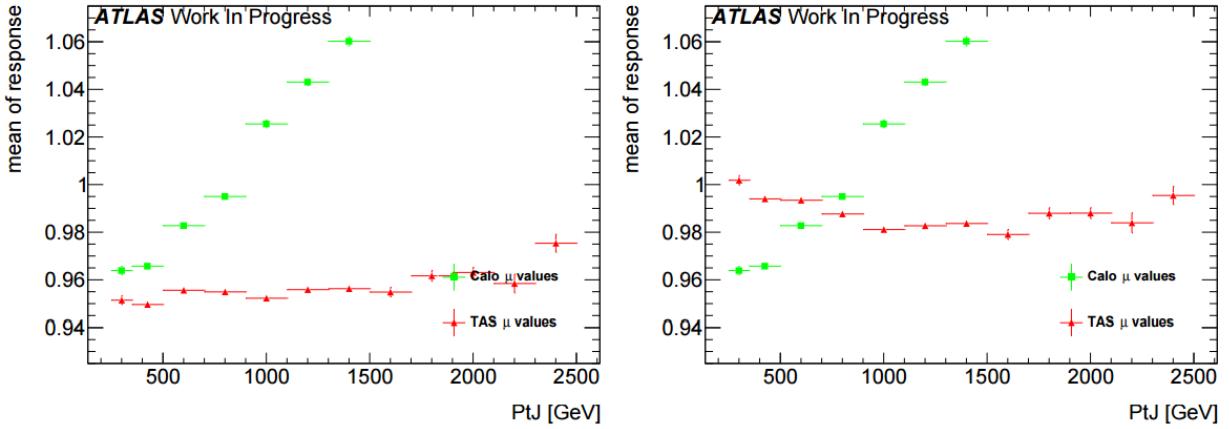


Figure 39: Poor's man calibration effect on the mean of the mass response of the large-R jet, before, left, and after, right, the procedure.

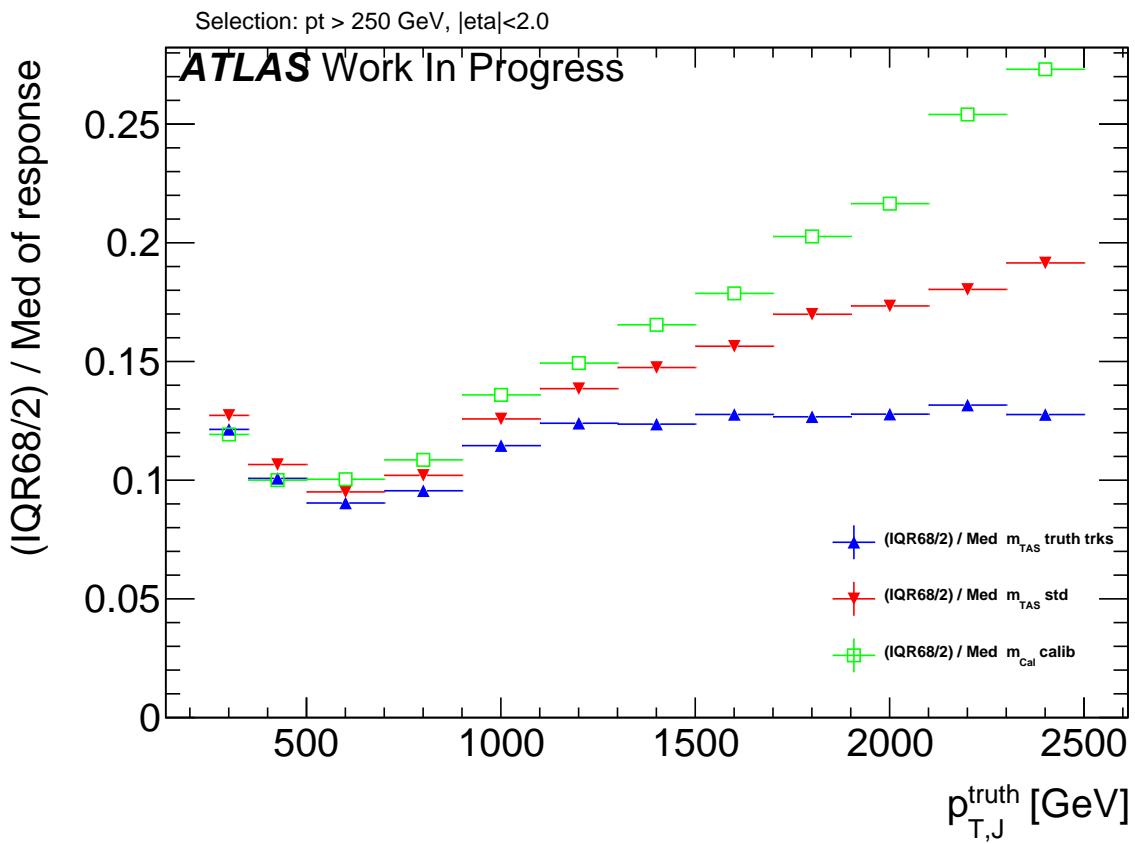
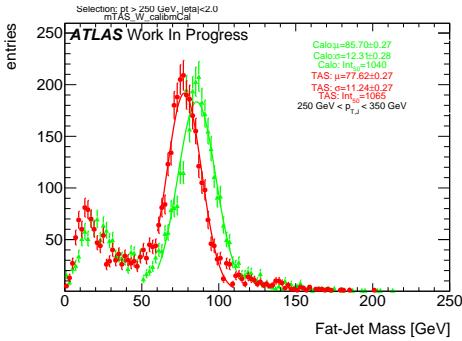
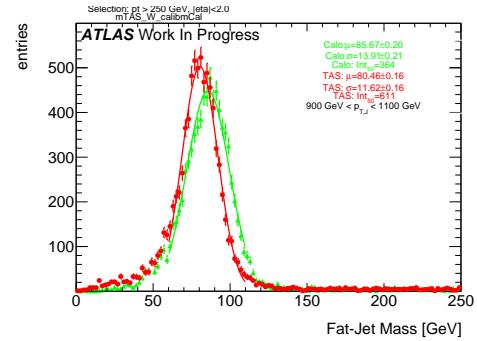
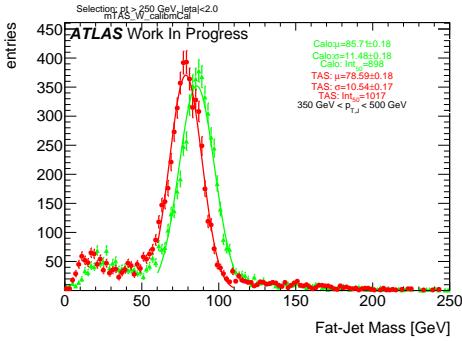
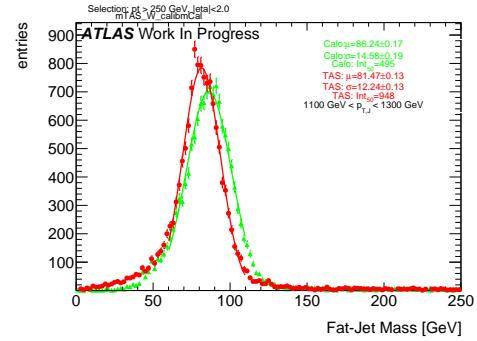
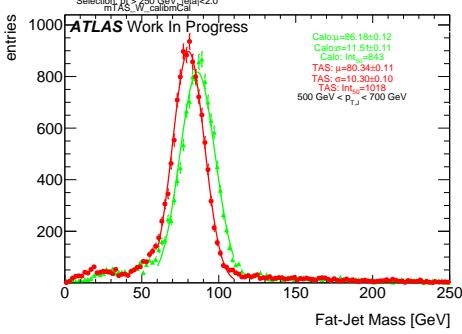
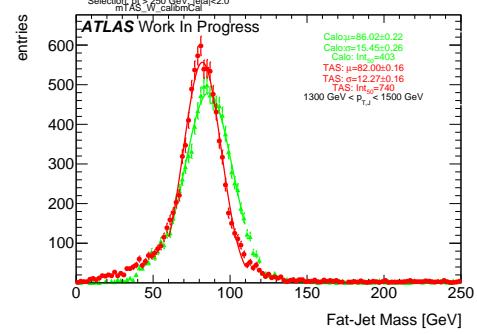
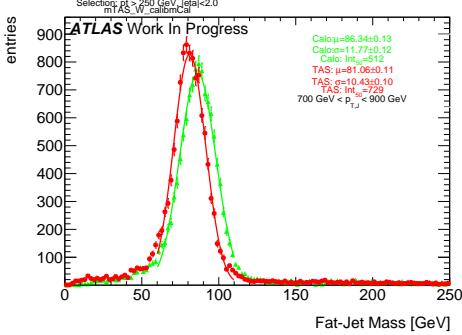
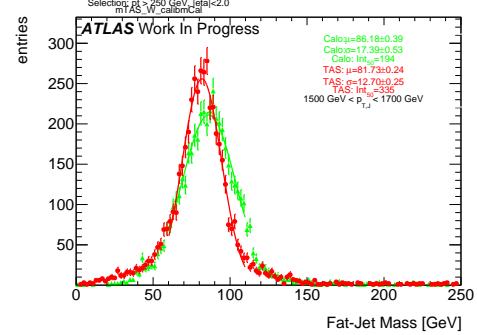


Figure 40: Comparison of the m^{TAS} and the same variable using truth-level information for the tracks.

530 **A m^{TAS} distributions, boosted W/Z**

Figure 41: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 45: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 42: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 46: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 43: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 47: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 44: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)Figure 48: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

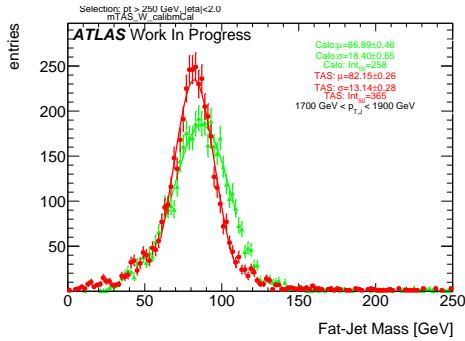


Figure 49: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

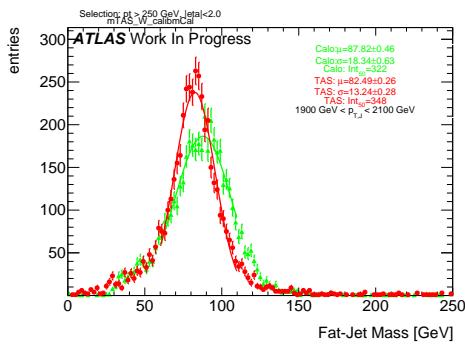


Figure 50: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

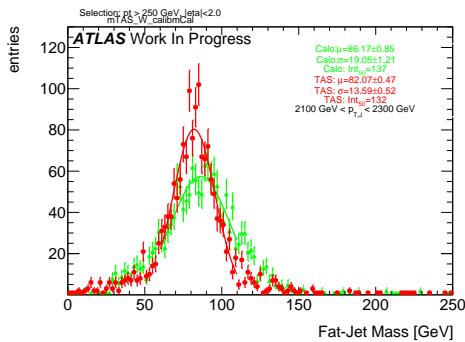


Figure 51: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

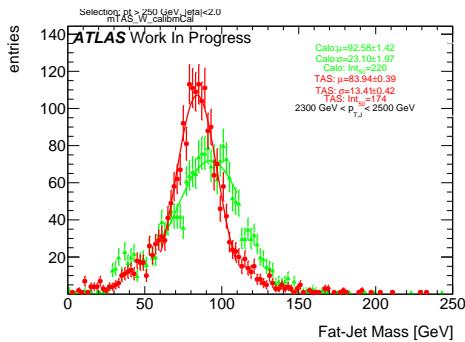


Figure 52: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

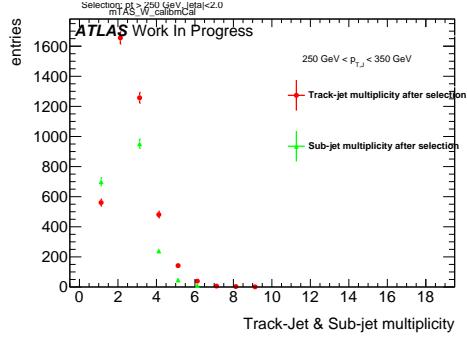


Figure 53: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

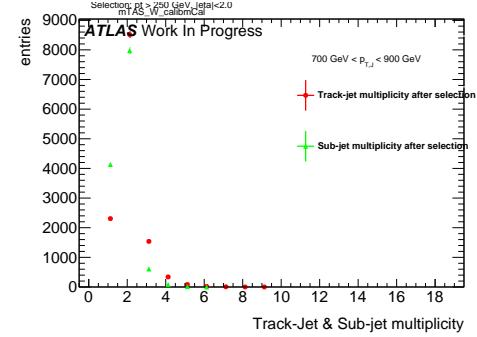


Figure 56: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

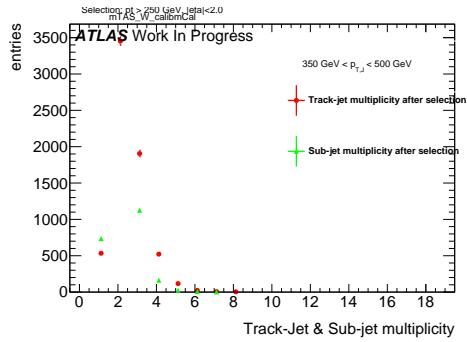


Figure 54: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

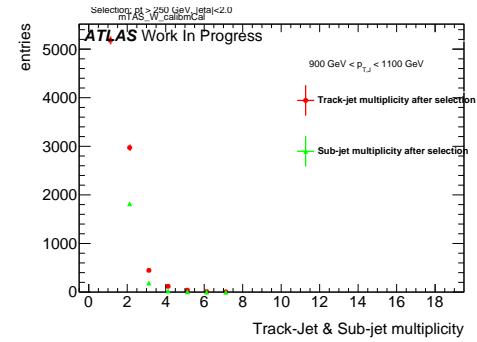


Figure 57: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

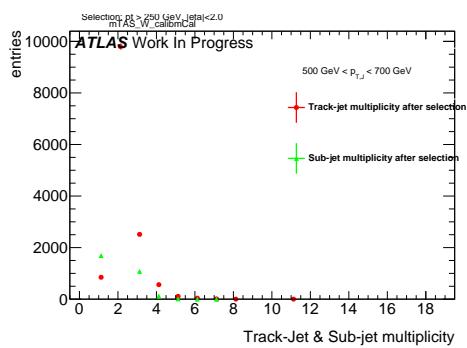


Figure 55: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

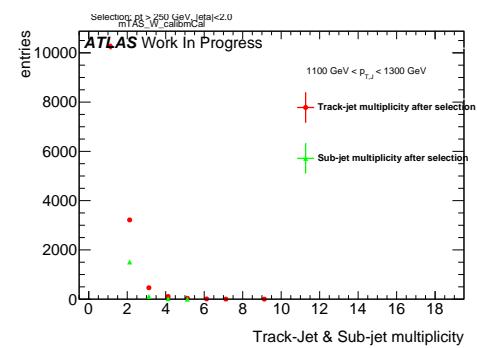


Figure 58: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

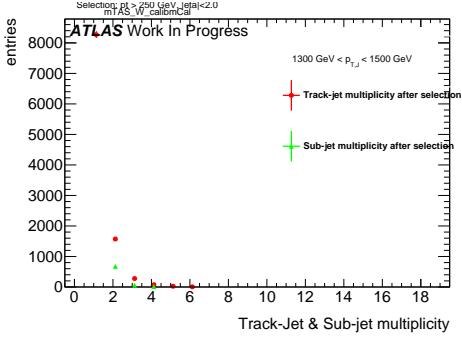


Figure 59: Track-jet $R=0.2$ and sub-jet multiplicity for p_T^J bin (indicated on plot)

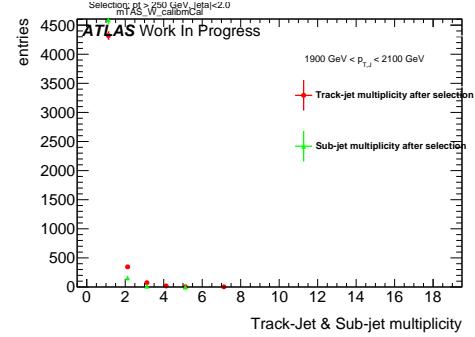


Figure 62: Track-jet $R=0.2$ and sub-jet multiplicity for p_T^J bin (indicated on plot)

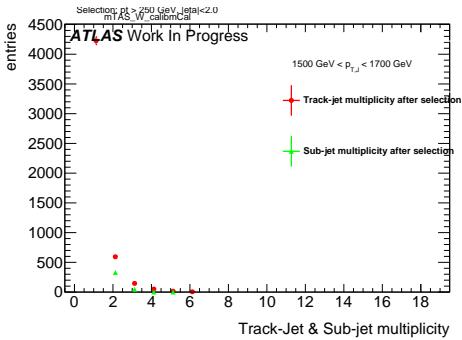


Figure 60: Track-jet $R=0.2$ and sub-jet multiplicity for p_T^J bin (indicated on plot)

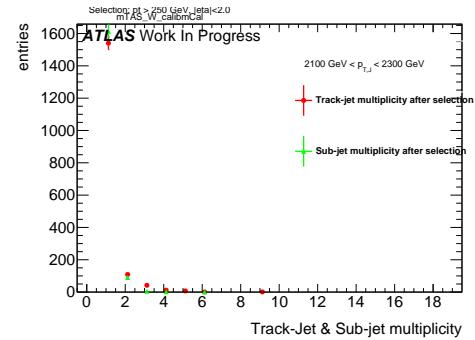


Figure 63: Track-jet $R=0.2$ and sub-jet multiplicity for p_T^J bin (indicated on plot)

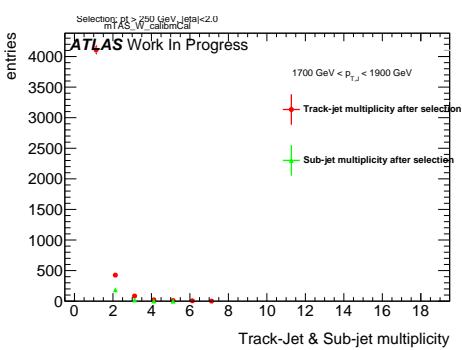


Figure 61: Track-jet $R=0.2$ and sub-jet multiplicity for p_T^J bin (indicated on plot)

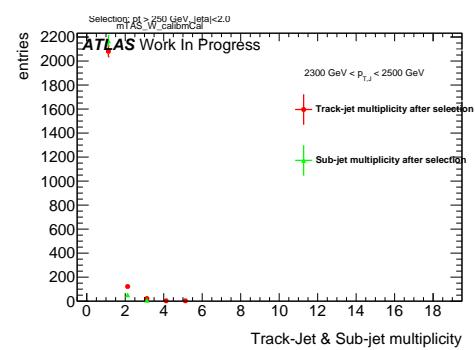


Figure 64: Track-jet $R=0.2$ and sub-jet multiplicity for p_T^J bin (indicated on plot)

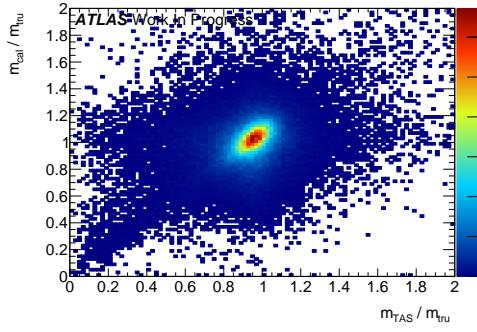
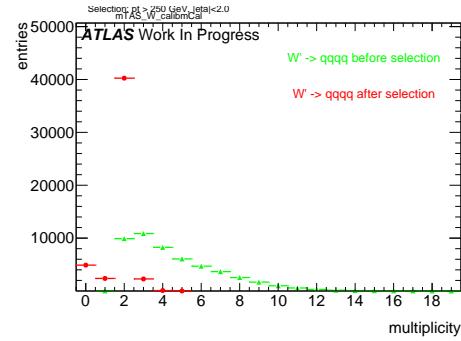
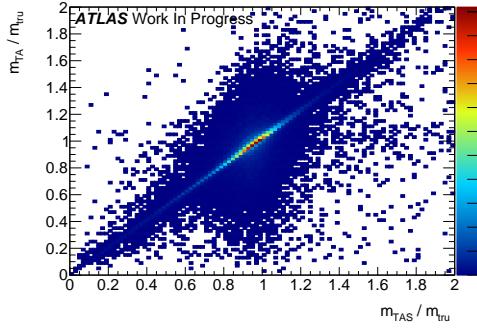
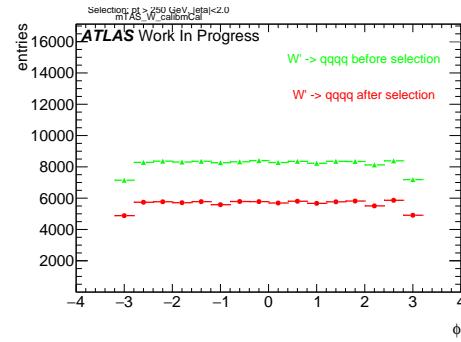
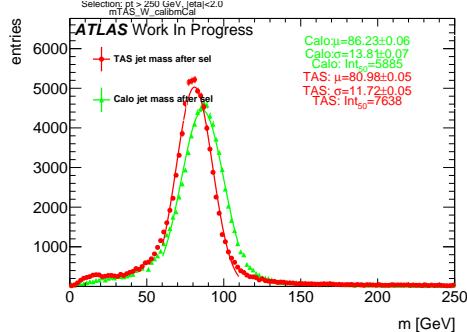
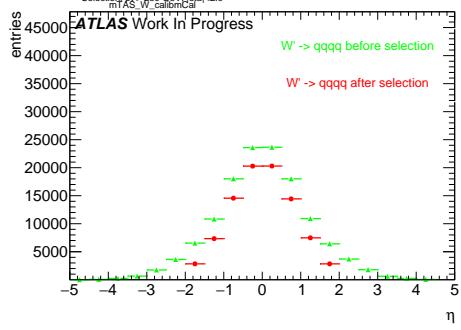
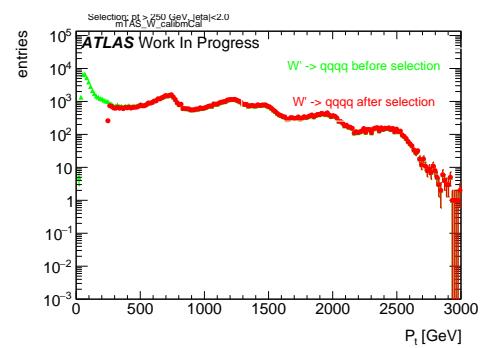
Figure 65: Scatter plot m^{TAS} versus m^{calo} responses

Figure 69: large-R jet Multiplicity, before and after selection

Figure 66: Scatter plot m^{TAS} versus m^{TA} responsesFigure 70: ϕ distribution of the large-R jet, before and after selectionFigure 67: m^{TAS} distribution in all the p_T binsFigure 68: η distribution of the large-R jet, before and after selectionFigure 71: p_T distribution of the large-R jet, before and after selection

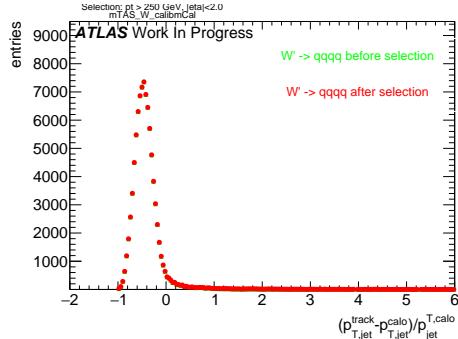


Figure 72: p_T resolution: $\frac{p_{T,jet}^{track} - p_{T,jet}^{calo}}{p_{T,jet}^{calo}}$, before and after selection

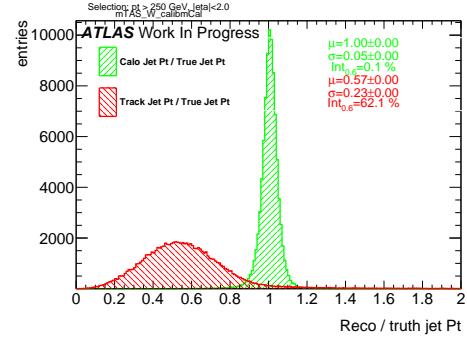


Figure 75: Transverse momentum response p_T^{Reco} / p_T^{Truth} for calorimeter and tracks

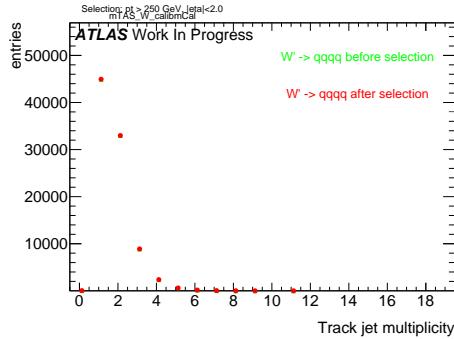


Figure 73: Multiplicity of track-jets R=0.2 per large-R jet

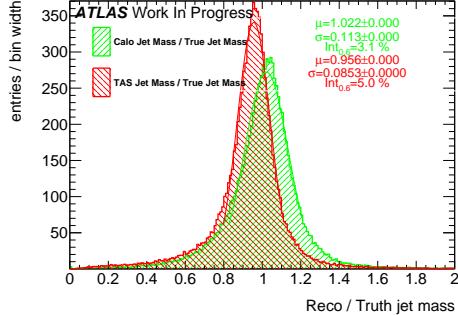


Figure 74: Response m^{Reco} / m^{Truth} for all the p_T bins

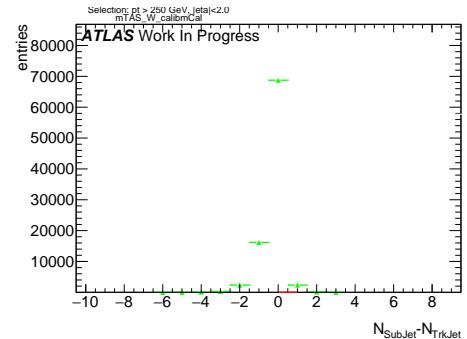


Figure 76: sub-jet - track-jet Multiplicity

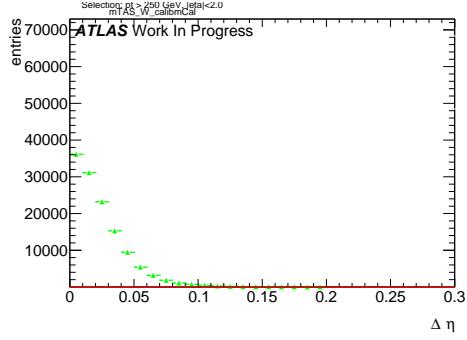


Figure 77: $|\eta_{\text{sub-jet}} - \eta_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

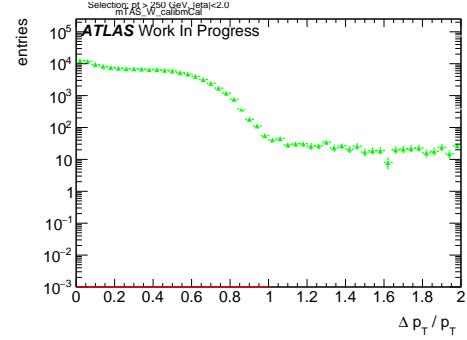


Figure 80: $|p_{T,\text{sub-jet}} - p_{T,\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

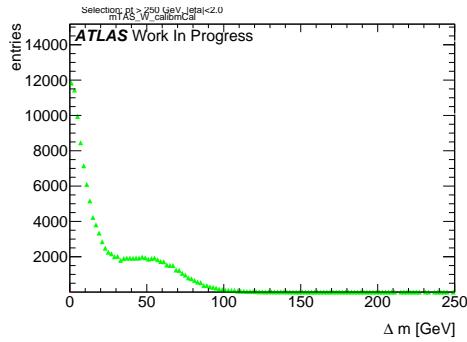


Figure 78: $|m_{\text{sub-jet}} - m_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

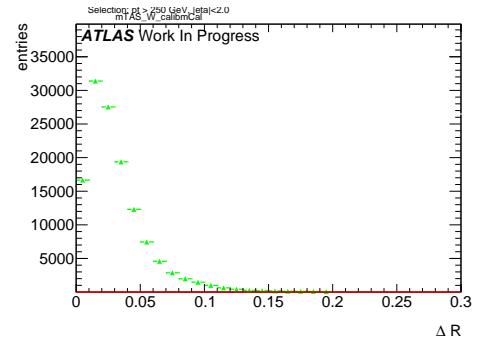


Figure 81: $|R_{\text{sub-jet}} - R_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

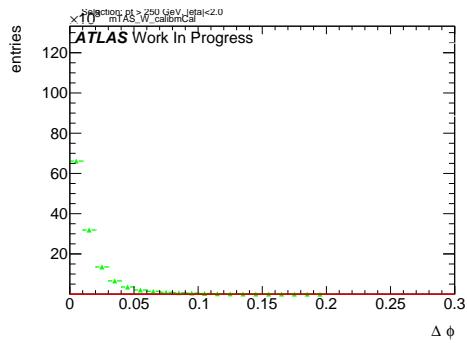


Figure 79: $|\phi_{\text{sub-jet}} - \phi_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

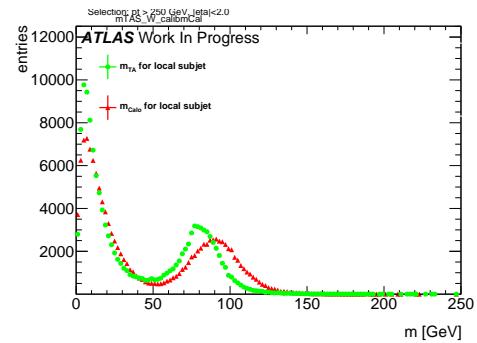


Figure 82: Mass distribution of the sub-jet, calorimeter and track-assisted

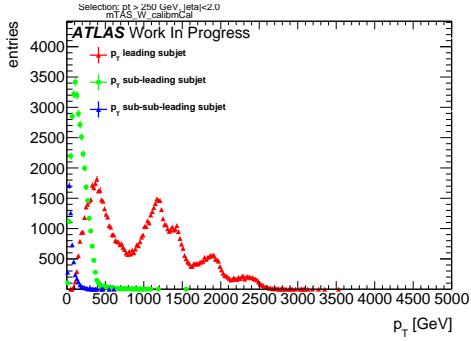


Figure 83: p_T distribution for leading, sub-leading and sub-sub-leading sub-jets

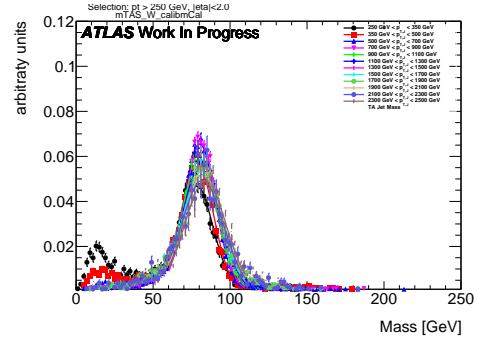


Figure 86: m^{TAS} for p_T^J bin, superimposed

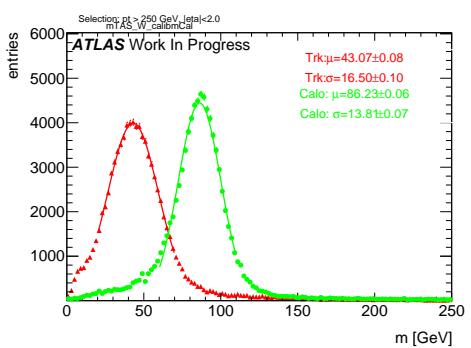


Figure 84: Mass distribution for calorimeter and tracks associated to the large- R jet

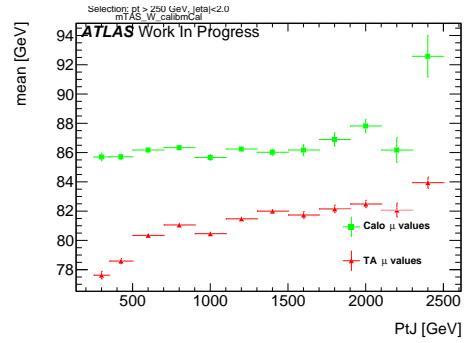


Figure 87: μ from fit of the mass distribution vs bin of p_T^J

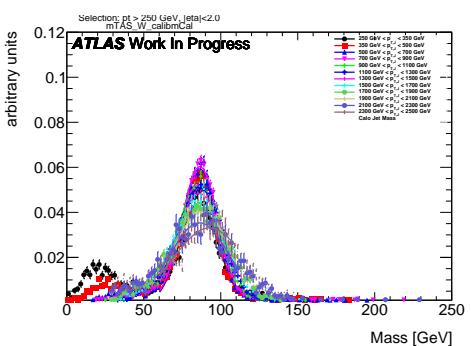


Figure 85: m^{calo} for p_T^J bin, superimposed

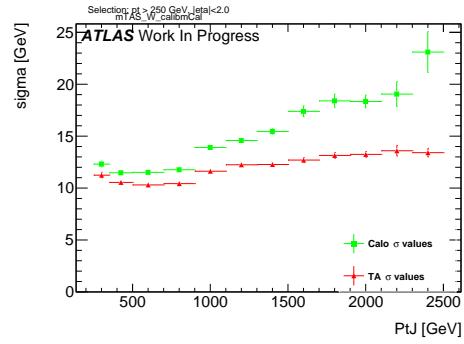


Figure 88: σ from fit of the mass distribution vs bin of p_T^J

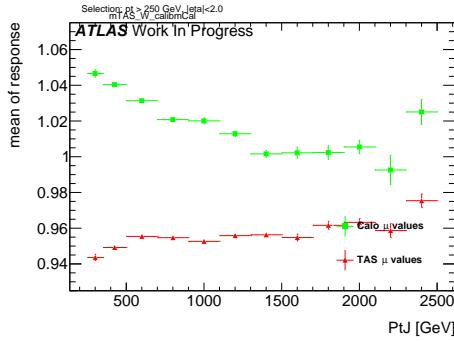


Figure 89: μ from fit of the mass Response vs bin of p_T^J

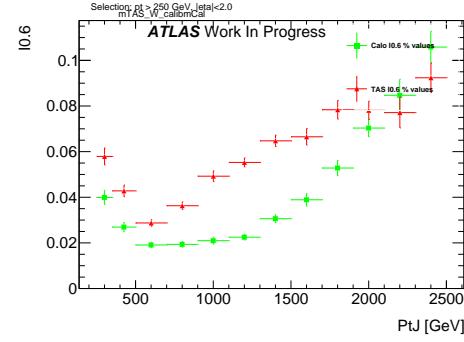


Figure 92: Left integral normalized, $\int_0^{0.6}$ of the mass response, vs bin of p_T^J

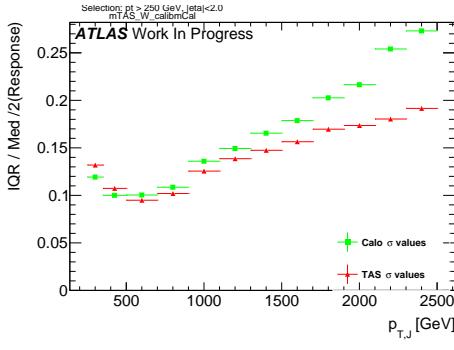


Figure 90: σ from fit of the mass Response vs bin of p_T^J

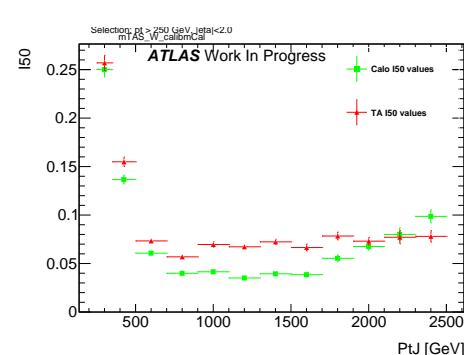


Figure 93: \int_0^{50GeV} from fit of the mass distribution vs bin of p_T^J (normalized)

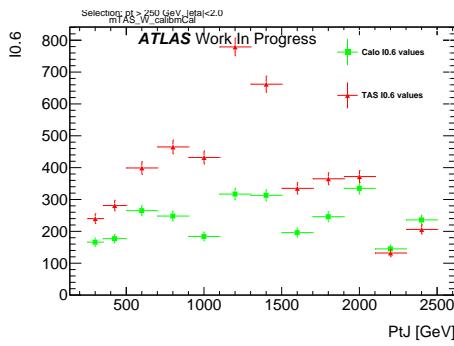


Figure 91: Left integral, $\int_0^{0.6}$ of the mass response, vs bin of p_T^J

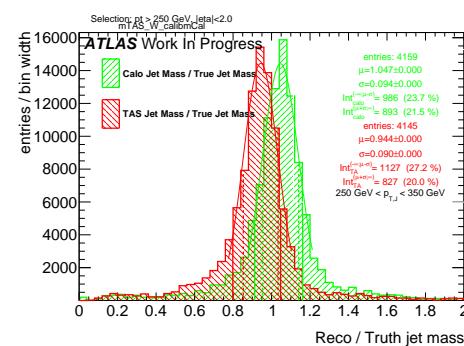


Figure 94: Response in bin of p_T^J (indicated on plot)

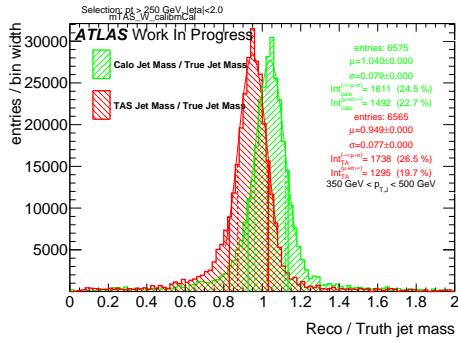


Figure 95: Response in bin of p_T^J (indicated on plot)

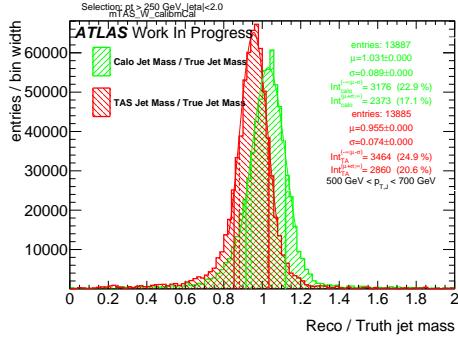


Figure 96: Response in bin of p_T^J (indicated on plot)

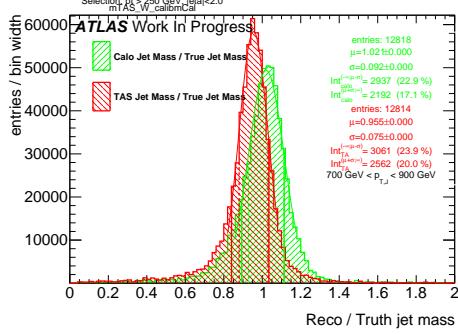


Figure 97: Response in bin of p_T^J (indicated on plot)

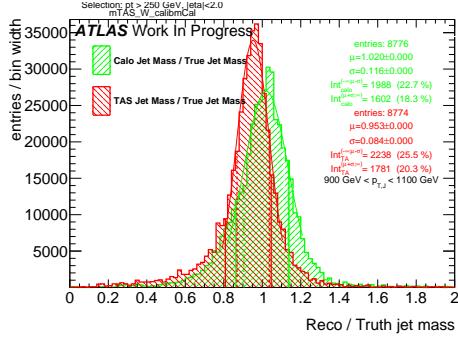


Figure 98: Response in bin of p_T^J (indicated on plot)

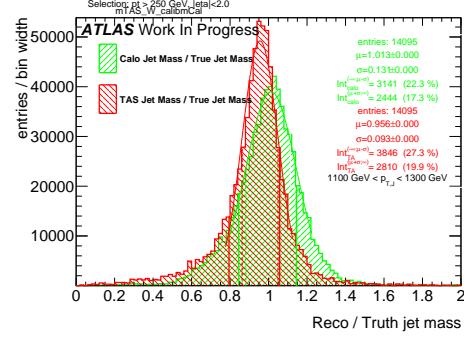


Figure 99: Response in bin of p_T^J (indicated on plot)

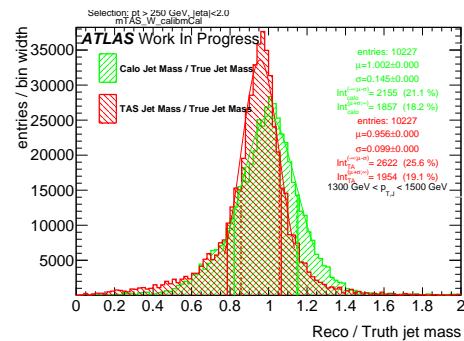
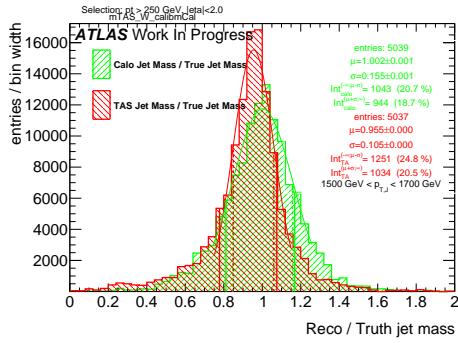
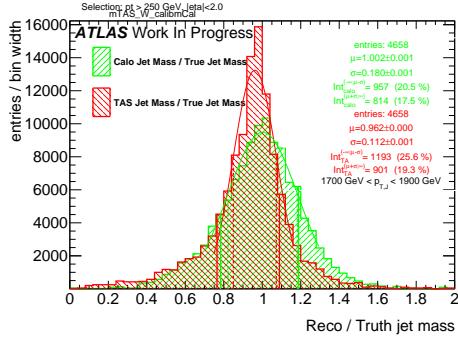
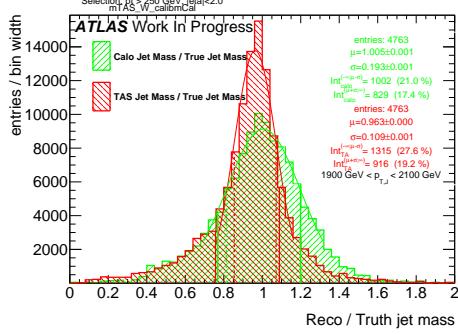
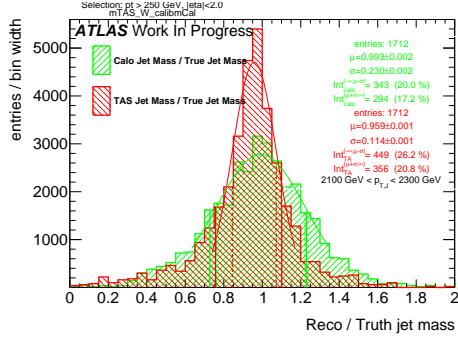
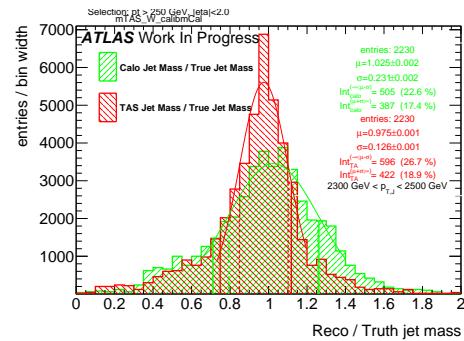


Figure 100: Response in bin of p_T^J (indicated on plot)

Figure 101: Response in bin of p_T^J (indicated on plot)Figure 102: Response in bin of p_T^J (indicated on plot)Figure 103: Response in bin of p_T^J (indicated on plot)Figure 104: Response in bin of p_T^J (indicated on plot)Figure 105: Response in bin of p_T^J (indicated on plot)

₅₃₁ **B m^{TAS} distributions, boosted tops**

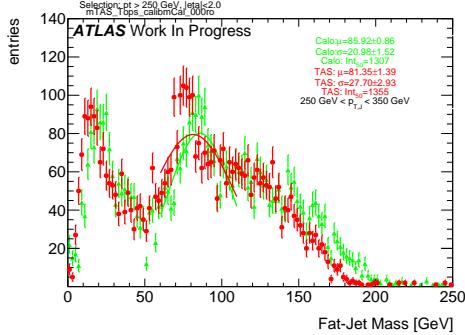


Figure 106: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

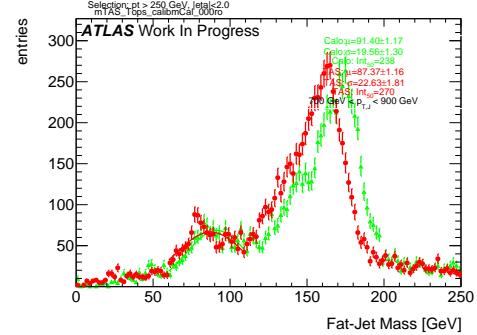


Figure 109: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

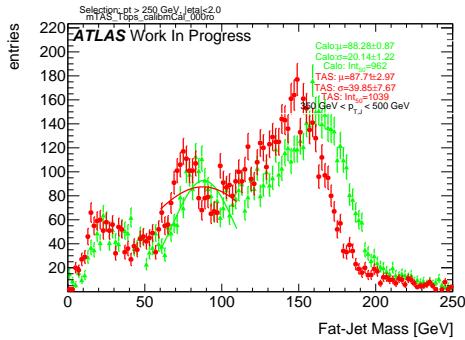


Figure 107: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

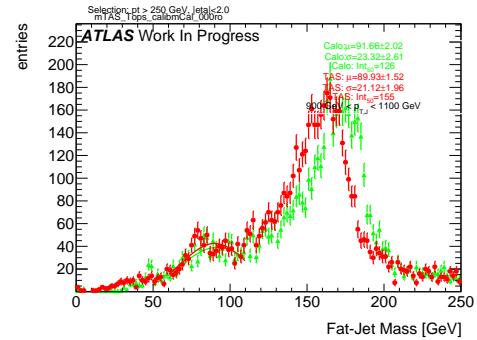


Figure 110: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

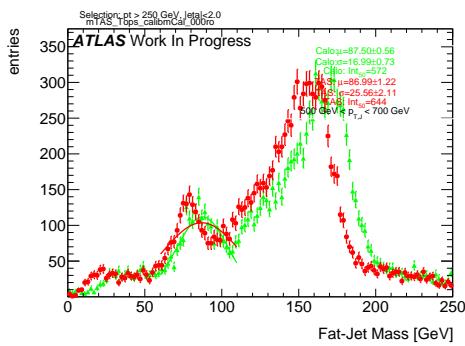


Figure 108: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

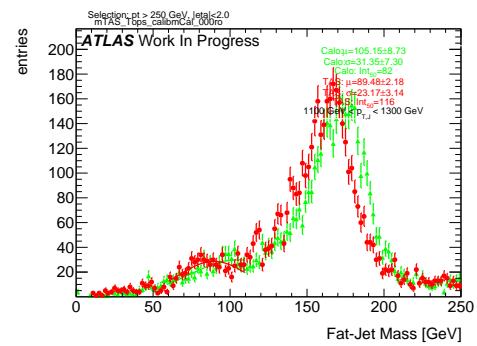


Figure 111: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

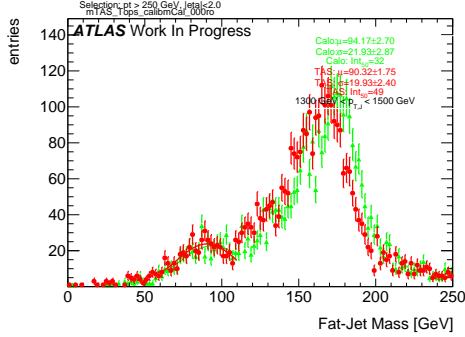


Figure 112: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

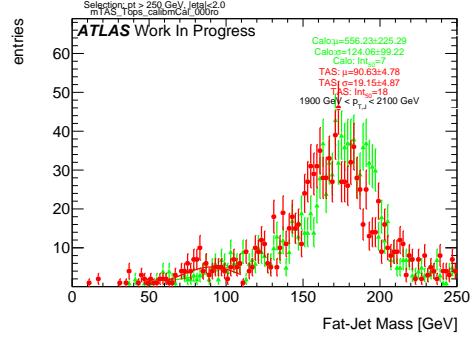


Figure 115: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

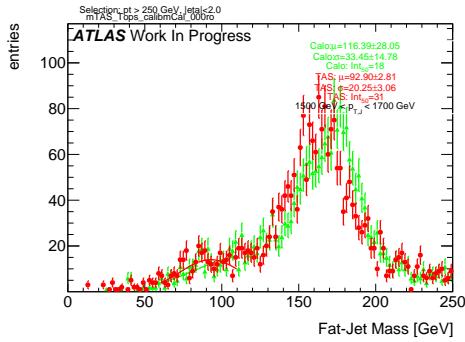


Figure 113: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

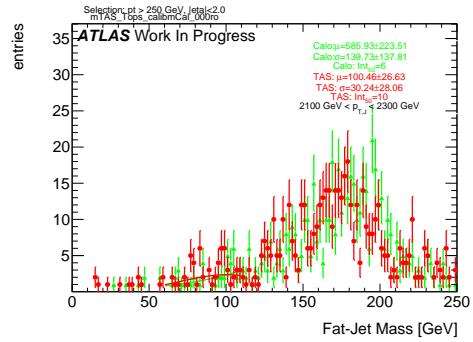


Figure 116: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

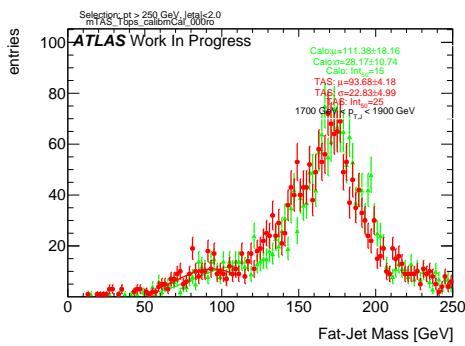


Figure 114: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

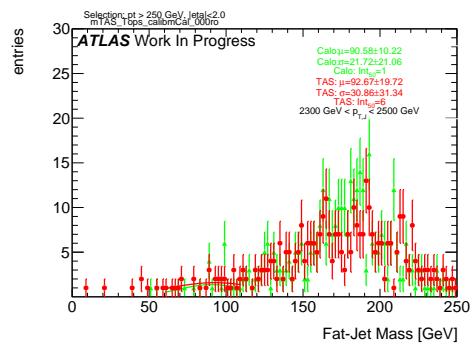


Figure 117: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

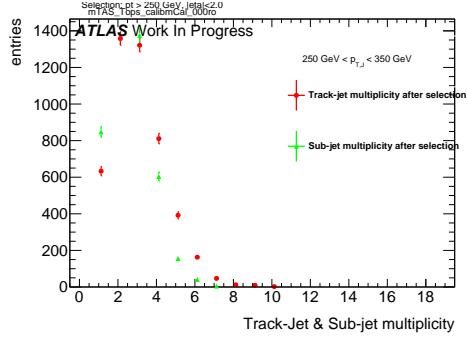


Figure 118: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

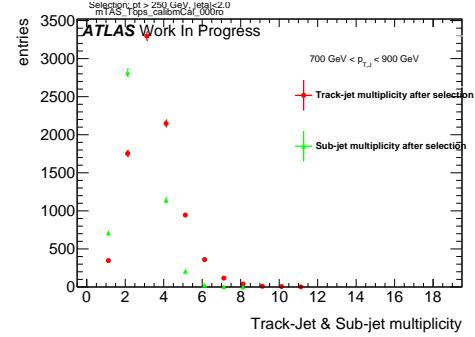


Figure 121: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

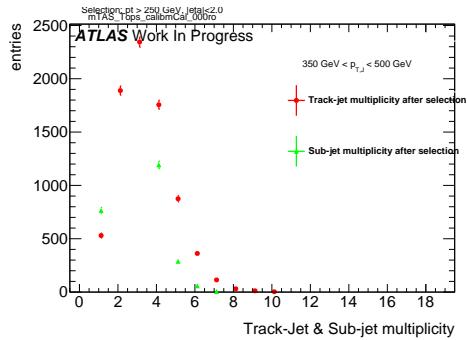


Figure 119: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

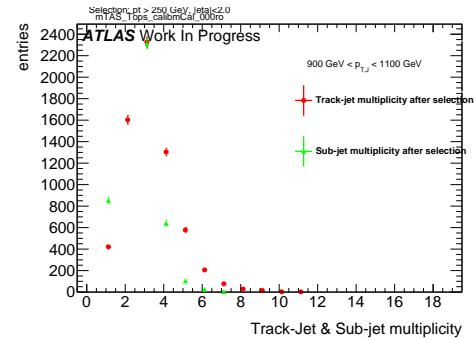


Figure 122: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

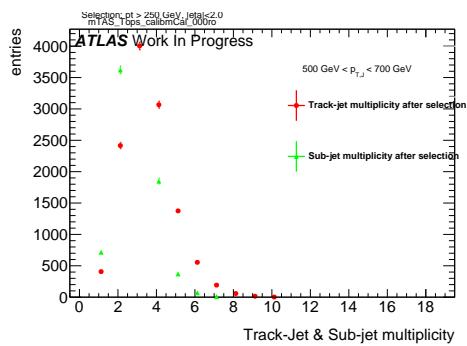


Figure 120: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

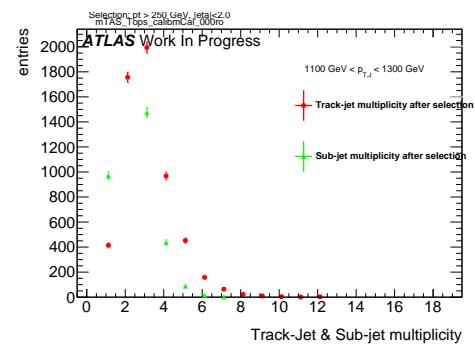


Figure 123: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

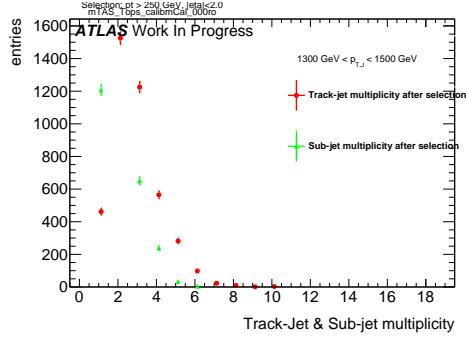


Figure 124: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

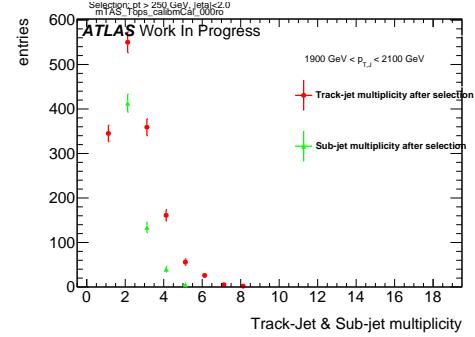


Figure 127: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

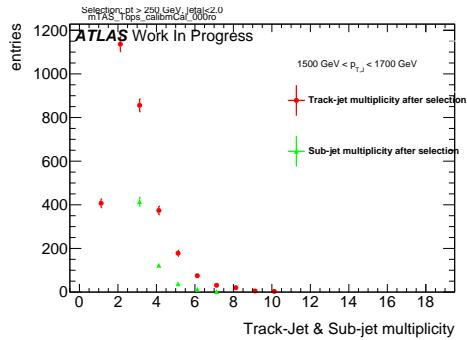


Figure 125: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

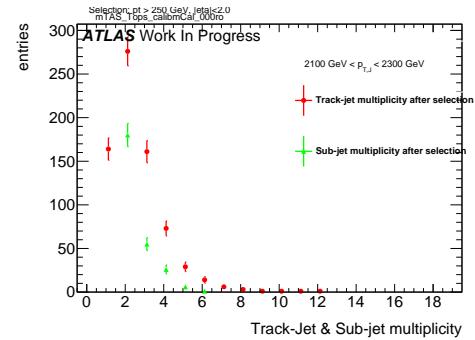


Figure 128: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

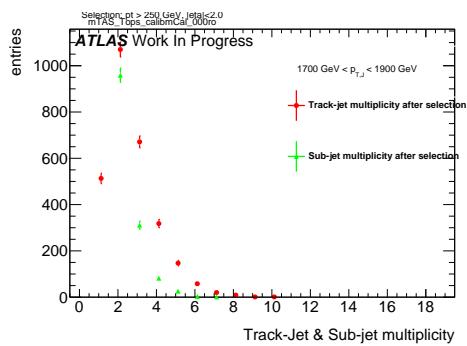


Figure 126: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

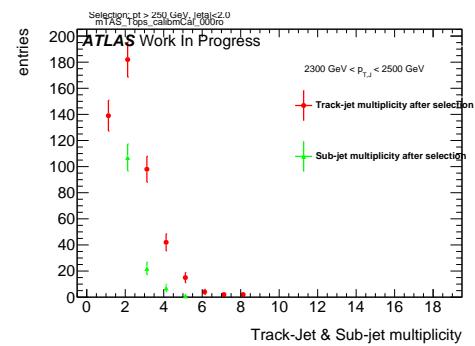


Figure 129: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

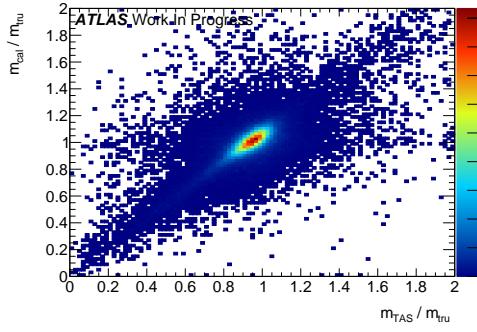
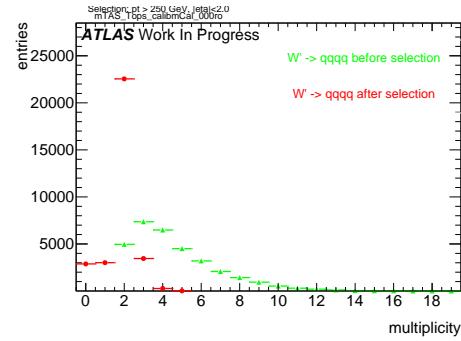
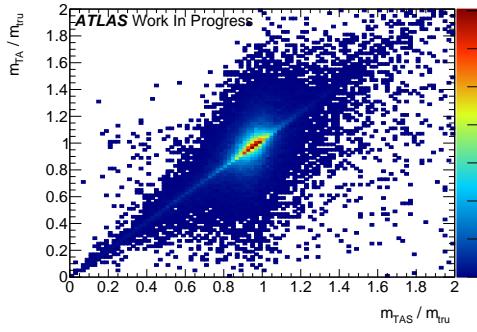
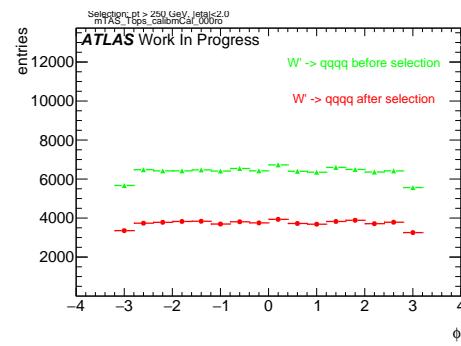
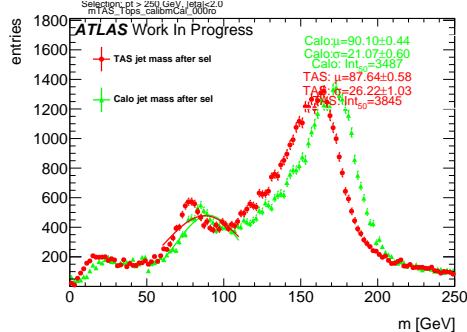
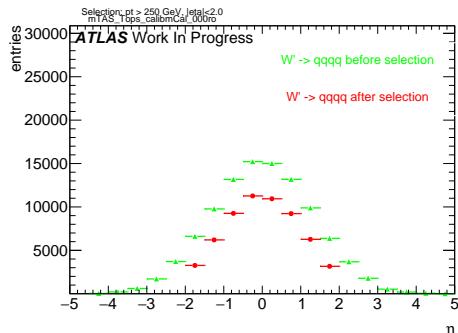
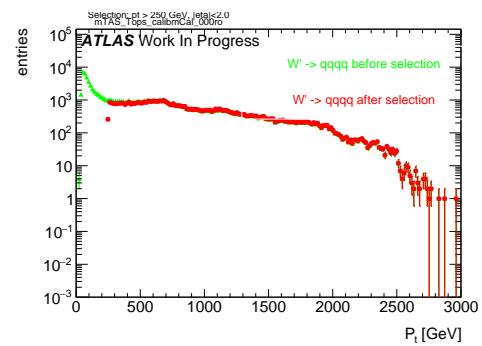
Figure 130: Scatter plot m^{TAS} versus m^{calo} responses

Figure 134: large-R jet Multiplicity, before and after selection

Figure 131: Scatter plot m^{TAS} versus m^A responsesFigure 135: ϕ distribution of the large-R jet, before and after selectionFigure 132: m^{TAS} distribution in all the p_T binsFigure 133: η distribution of the large-R jet, before and after selectionFigure 136: p_T distribution of the large-R jet, before and after selection

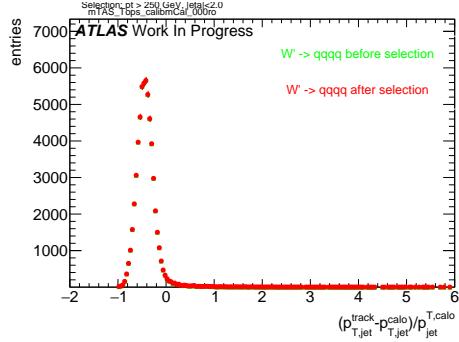


Figure 137: p_T resolution: $\frac{p_{T,jet}^{track} - p_{T,jet}^{calo}}{p_{T,jet}^{calo}}$, before and after selection

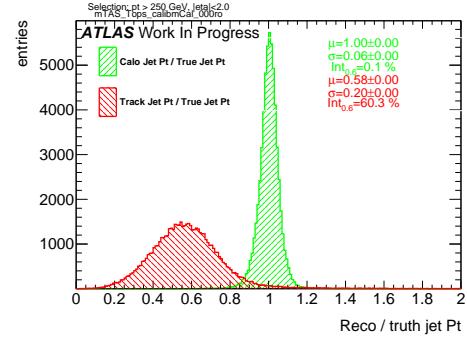


Figure 140: Transverse momentum response $p_T^{\text{Reco}} / p_T^{\text{Truth}}$ for calorimeter and tracks

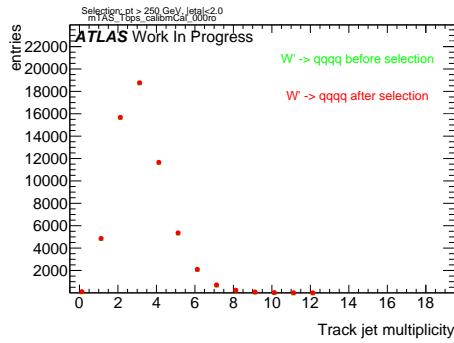


Figure 138: Multiplicity of track-jets R=0.2 per large-R jet

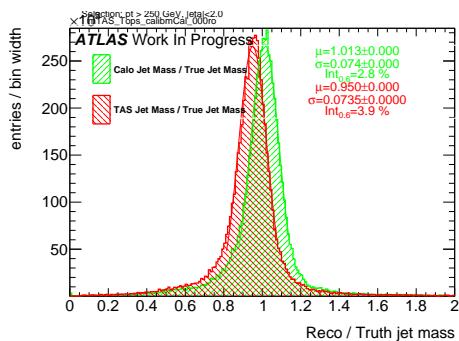


Figure 139: Response $m^{\text{Reco}} / m^{\text{Truth}}$ for all the p_T bins

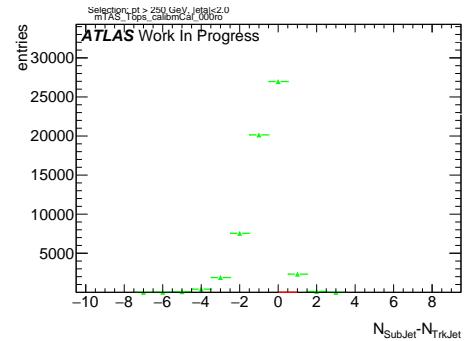


Figure 141: sub-jet - track-jet Multiplicity

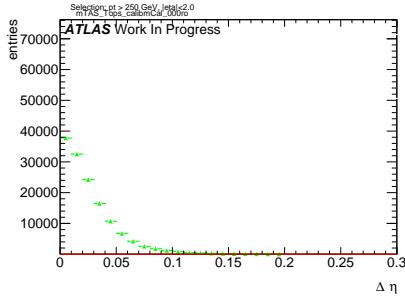


Figure 142: $|\eta_{\text{sub-jet}} - \eta_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

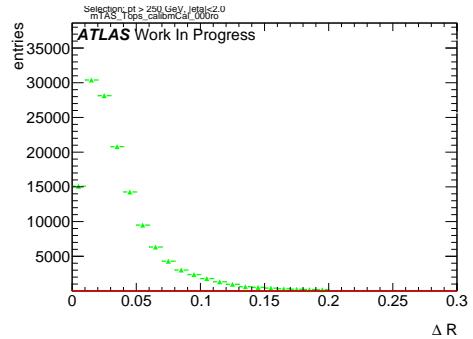


Figure 146: $|R_{\text{sub-jet}} - R_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

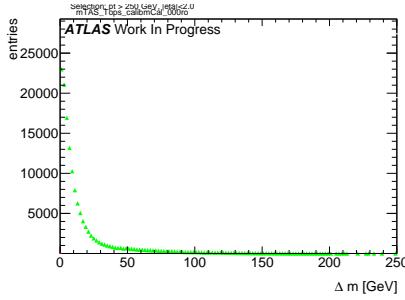


Figure 143: $|m_{\text{sub-jet}} - m_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

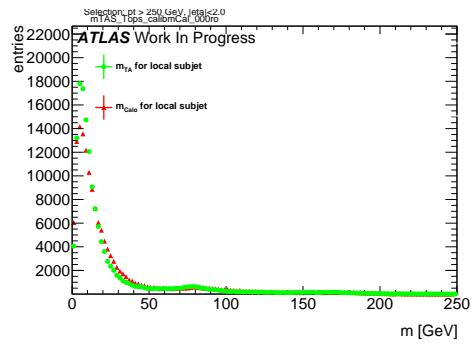


Figure 147: Mass distribution of the sub-jet, calorimeter and track-assisted

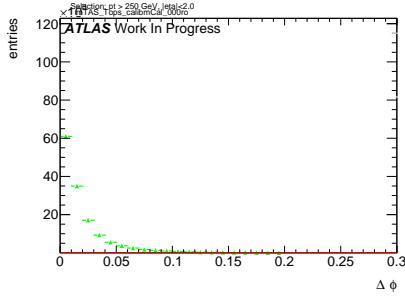


Figure 144: $|\phi_{\text{sub-jet}} - \phi_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

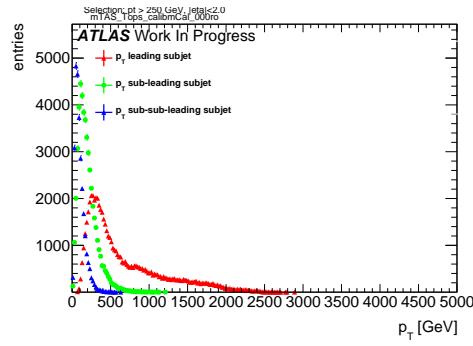


Figure 148: p_T distribution for leading, sub-leading and sub-sub-leading sub-jets

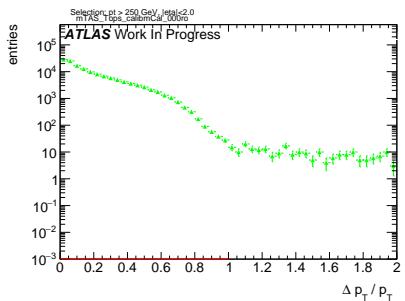


Figure 145: $|p_{T,\text{sub-jet}} - p_{T,\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

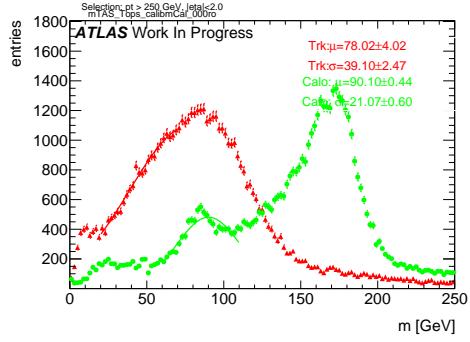


Figure 149: Mass distribution for calorimeter and tracks associated to the large-R jet

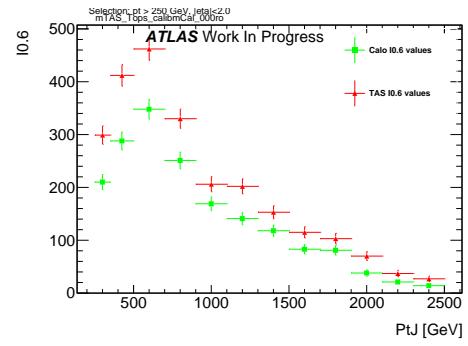


Figure 152: Left integral, $\int_0^{0.6}$ of the mass response, vs bin of p_T^J

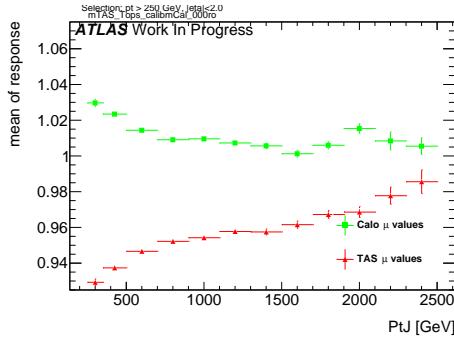


Figure 150: μ from fit of the mass Response vs bin of p_T^J

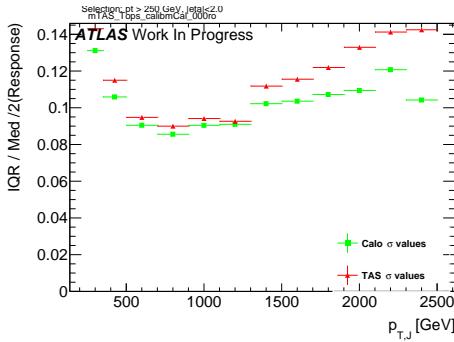


Figure 151: σ from fit of the mass Response vs bin of p_T^J

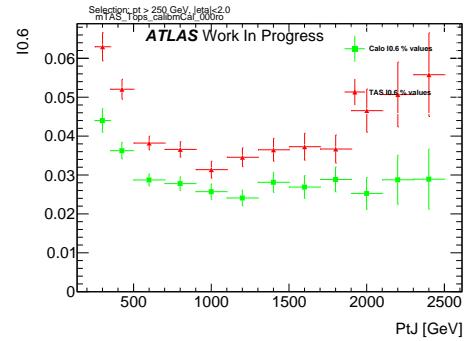
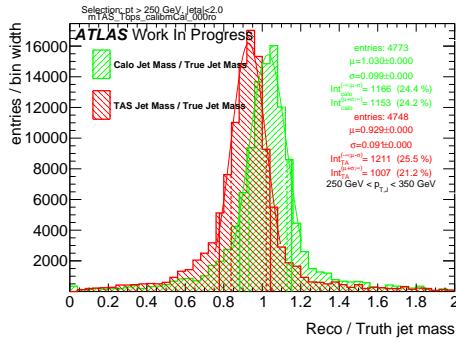
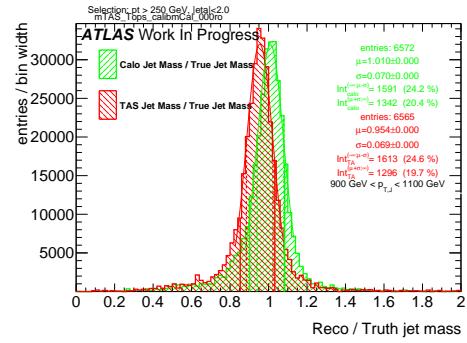
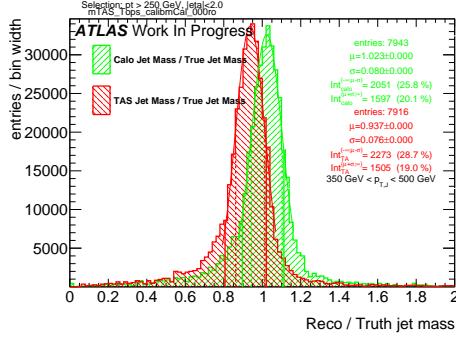
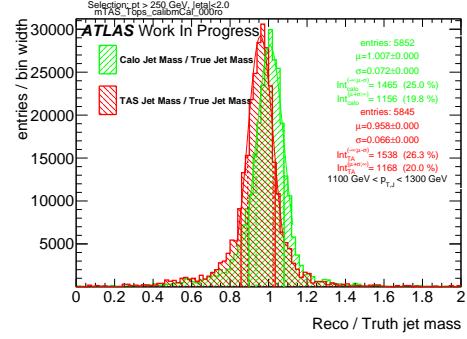
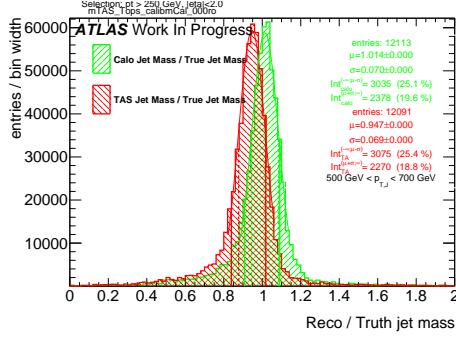
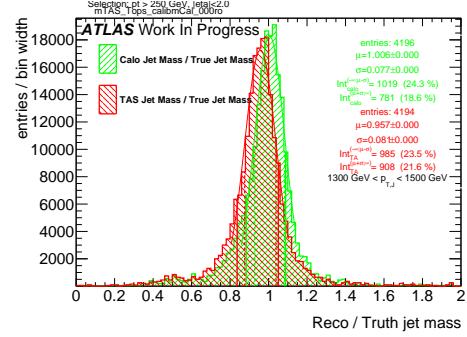
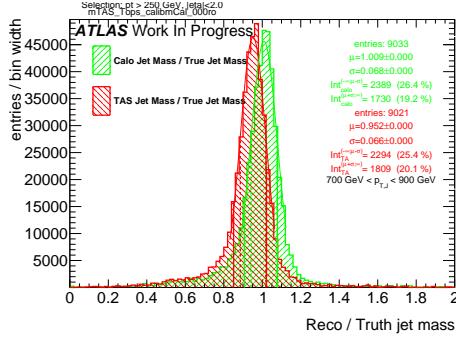
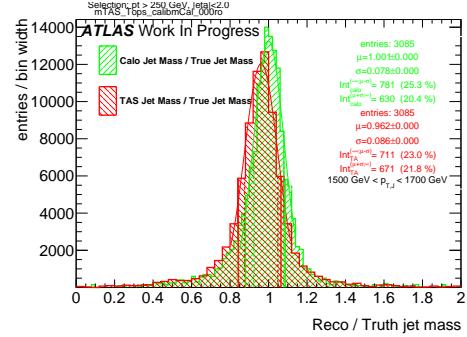
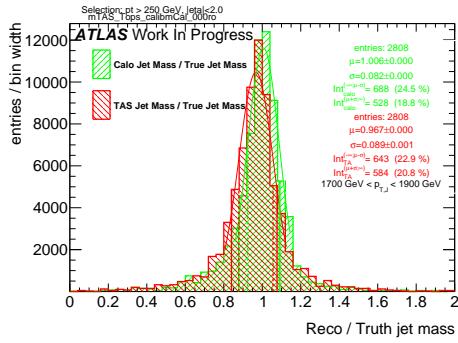
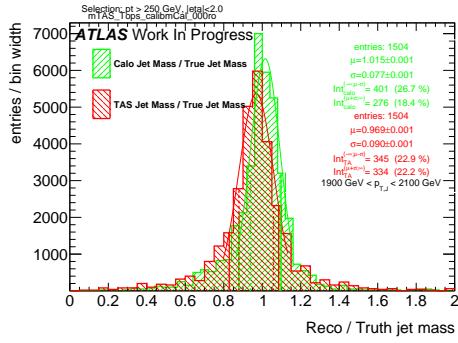
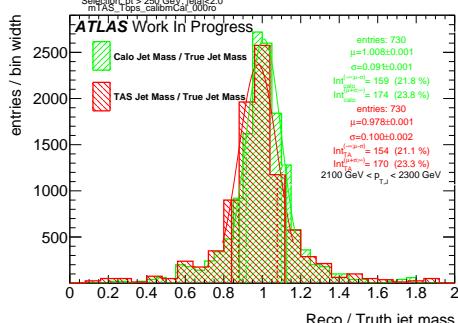
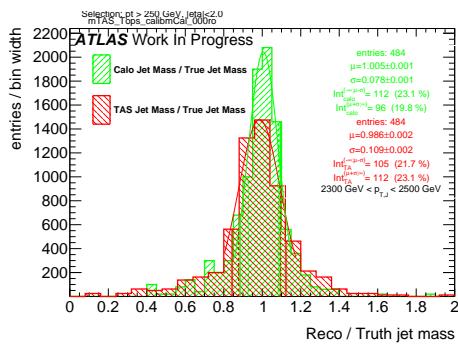


Figure 153: Left integral normalized, $\int_0^{0.6}$ of the mass response, vs bin of p_T^J

Figure 154: Response in bin of p_T^J (indicated on plot)Figure 158: Response in bin of p_T^J (indicated on plot)Figure 155: Response in bin of p_T^J (indicated on plot)Figure 159: Response in bin of p_T^J (indicated on plot)Figure 156: Response in bin of p_T^J (indicated on plot)Figure 160: Response in bin of p_T^J (indicated on plot)Figure 157: Response in bin of p_T^J (indicated on plot)Figure 161: Response in bin of p_T^J (indicated on plot)

Figure 162: Response in bin of p_T^J (indicated on plot)Figure 163: Response in bin of p_T^J (indicated on plot)Figure 164: Response in bin of p_T^J (indicated on plot)Figure 165: Response in bin of p_T^J (indicated on plot)

532 **C m^{TAS} distributions, boosted higgs**

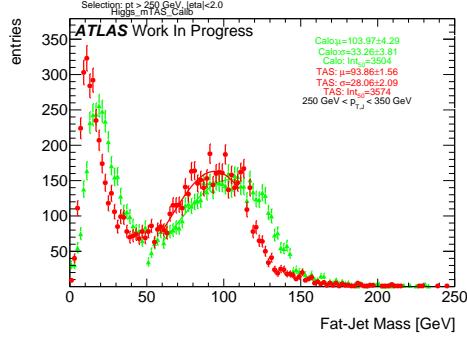


Figure 166: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

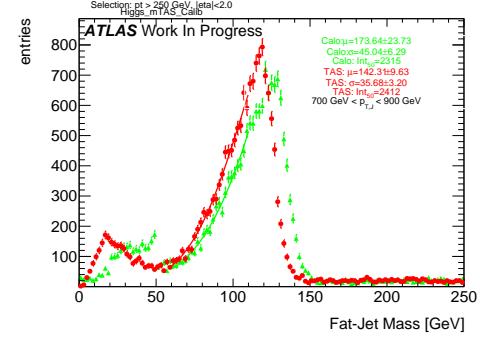


Figure 169: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

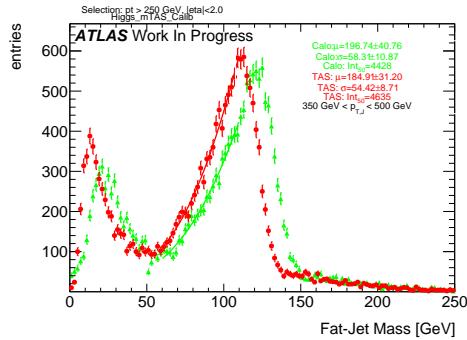


Figure 167: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

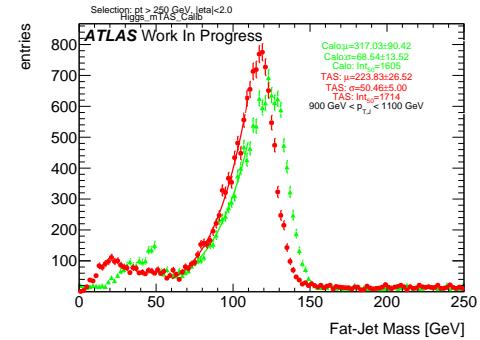


Figure 170: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

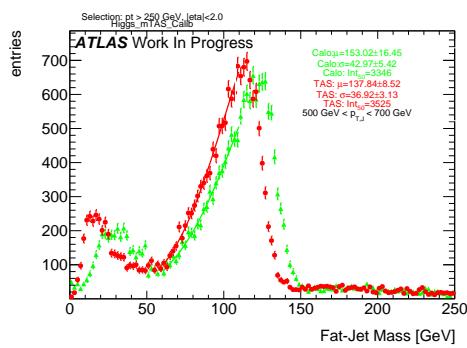


Figure 168: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

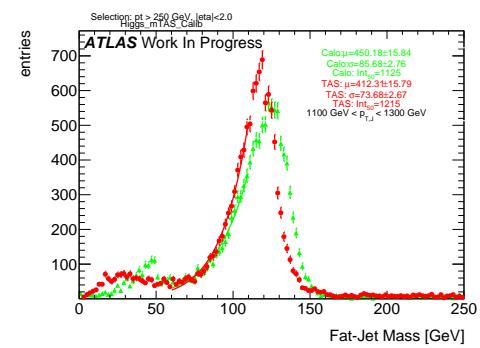


Figure 171: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

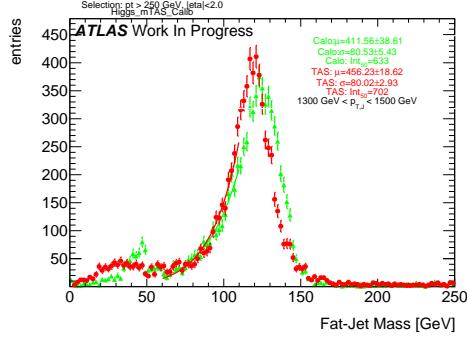


Figure 172: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

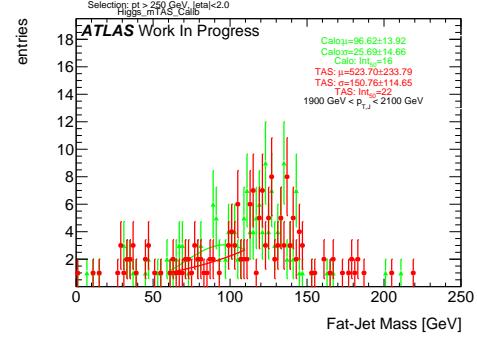


Figure 175: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

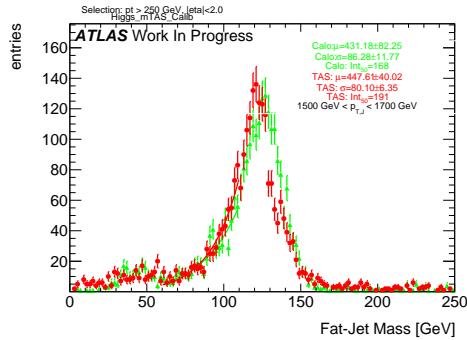


Figure 173: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

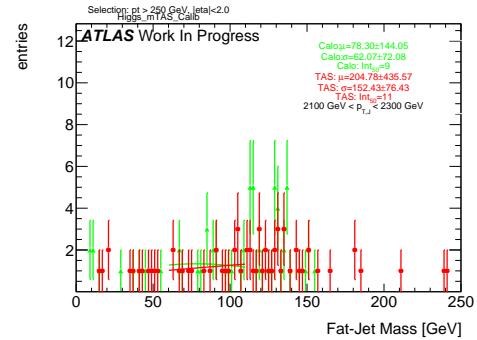


Figure 176: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

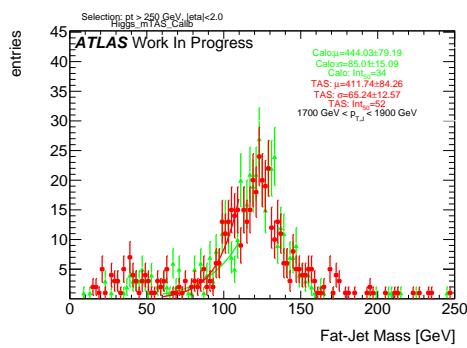


Figure 174: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

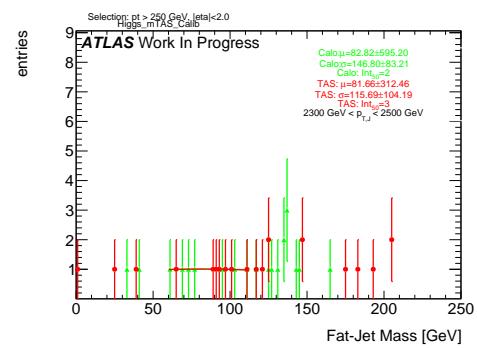


Figure 177: m^{TAS} and m^{calo} for p_T^J bin (indicated on plot)

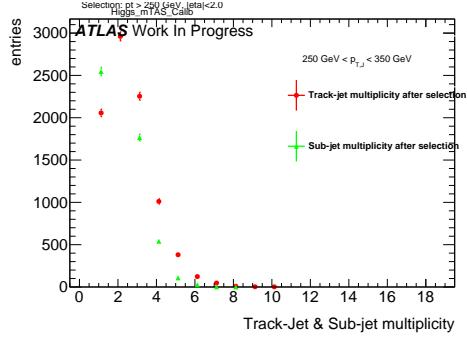


Figure 178: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

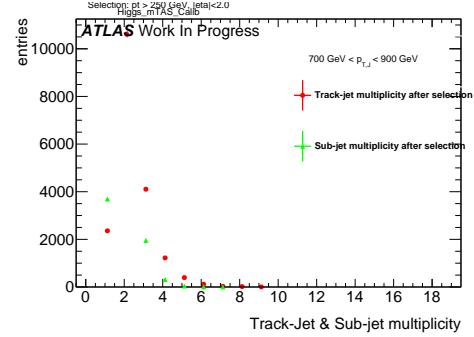


Figure 181: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

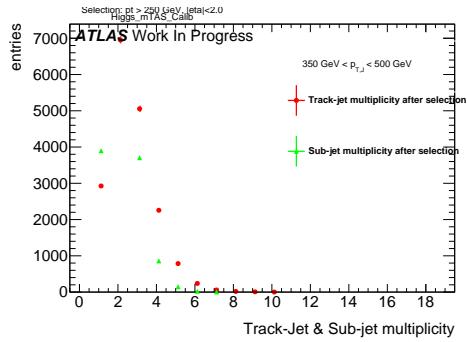


Figure 179: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

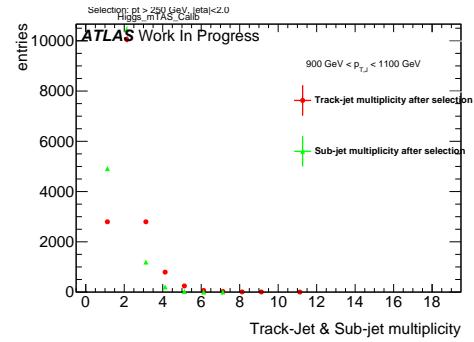


Figure 182: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

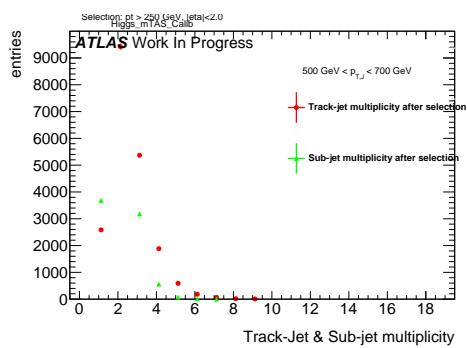


Figure 180: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

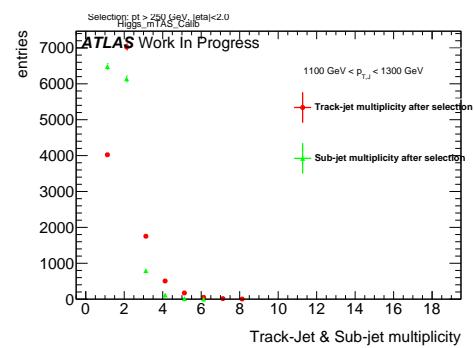


Figure 183: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

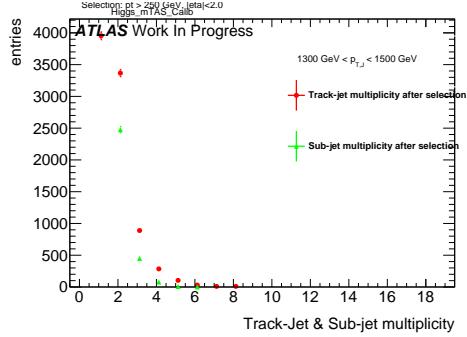


Figure 184: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

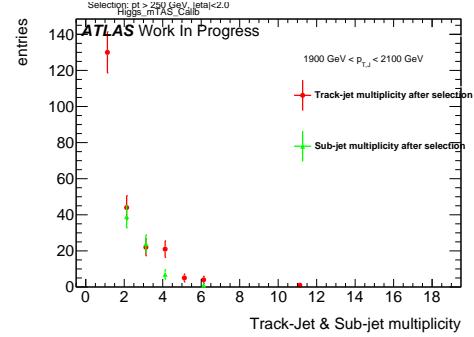


Figure 187: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

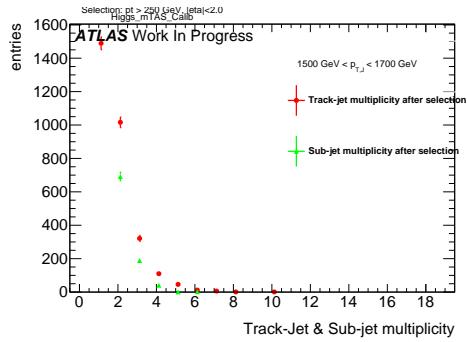


Figure 185: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

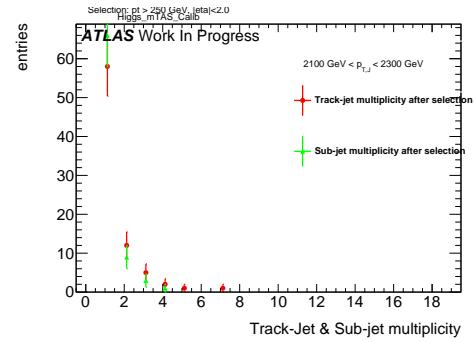


Figure 188: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

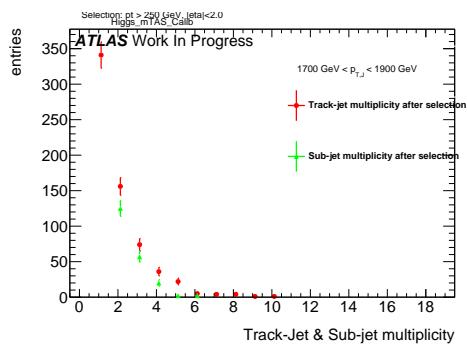


Figure 186: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

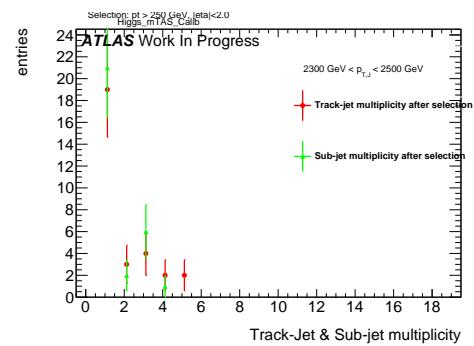


Figure 189: Track-jet R=0.2 and sub-jet multiplicity for p_T^J bin (indicated on plot)

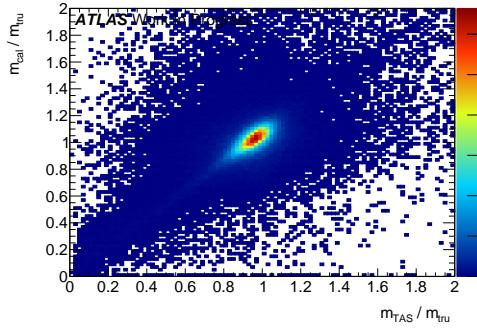
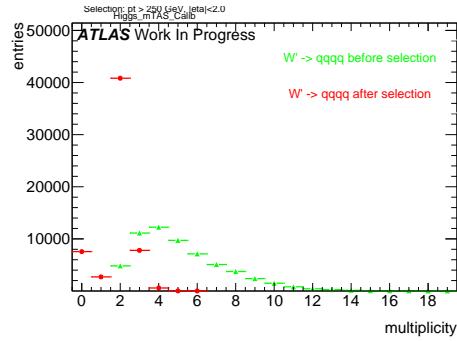
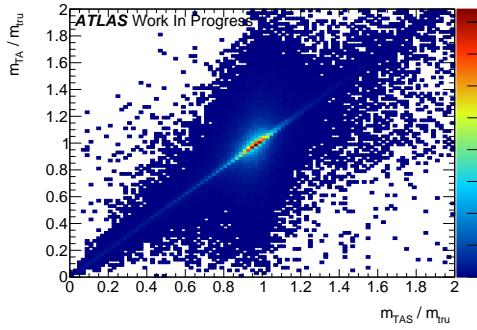
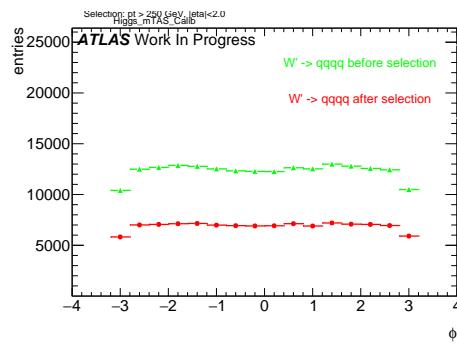
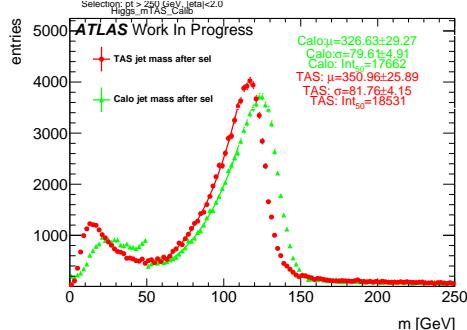
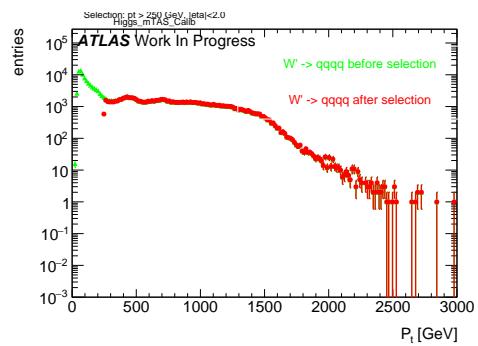
Figure 190: Scatter plot m^{TAS} versus m^{calo} responses

Figure 194: large-R jet Multiplicity, before and after selection

Figure 191: Scatter plot m^{TAS} versus m^T responsesFigure 195: ϕ distribution of the large-R jet, before and after selectionFigure 192: m^{TAS} distribution in all the p_T binsFigure 196: p_T distribution of the large-R jet, before and after selectionFigure 193: η distribution of the large-R jet, before and after selection

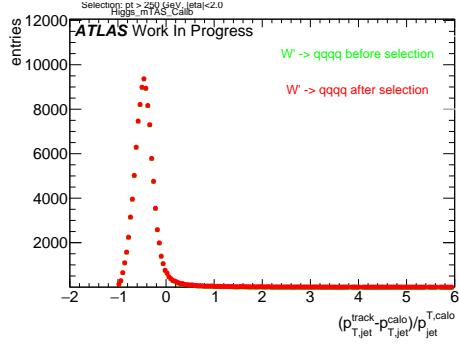


Figure 197: p_T resolution: $\frac{p_{T,jet}^{track} - p_{T,jet}^{fat}}{p_{T,jet}^{fat}}$, before and after selection

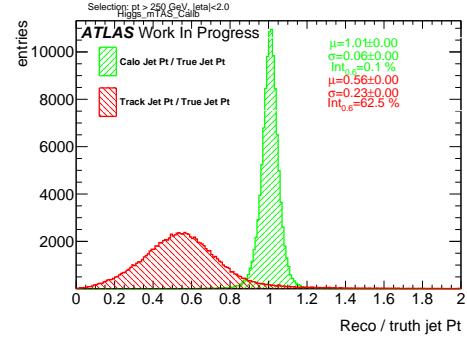


Figure 200: Transverse momentum response p_T^{Reco} / p_T^{Truth} for calorimeter and tracks

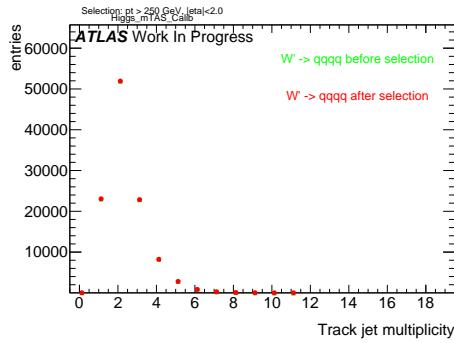


Figure 198: Multiplicity of track-jets R=0.2 per large-R jet

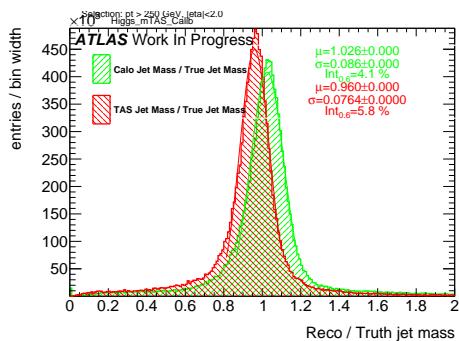


Figure 199: Response m^{Reco} / m^{Truth} for all the p_T bins

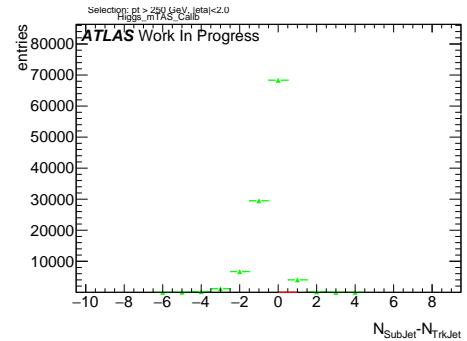


Figure 201: sub-jet - track-jet Multiplicity

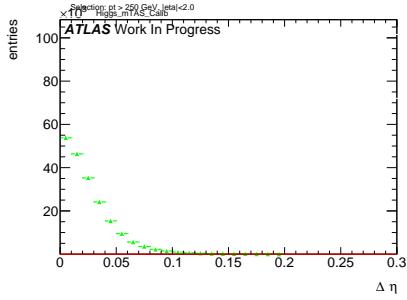


Figure 202: $|\eta_{\text{sub-jet}} - \eta_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

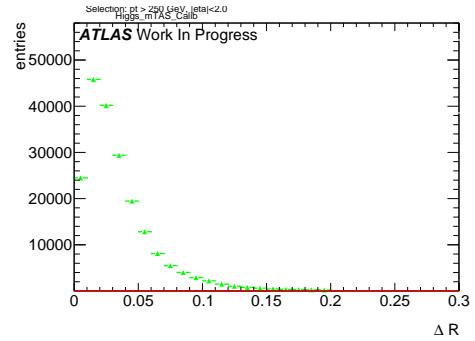


Figure 206: $|R_{\text{sub-jet}} - R_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

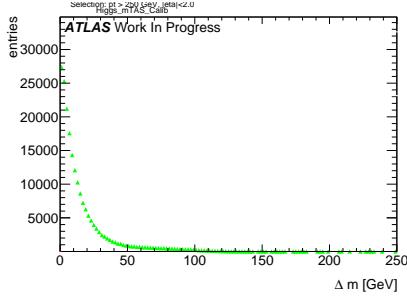


Figure 203: $|m_{\text{sub-jet}} - m_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

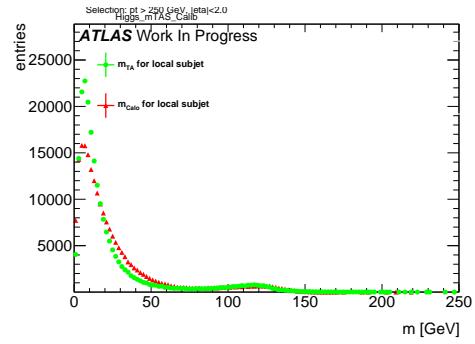


Figure 207: Mass distribution of the sub-jet, calorimeter and track-assisted

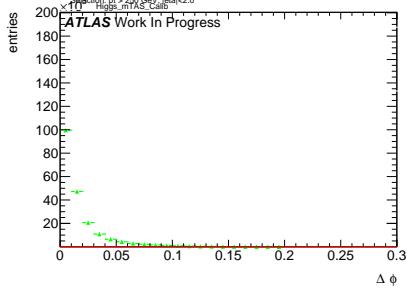


Figure 204: $|\phi_{\text{sub-jet}} - \phi_{\text{track-jet}}|$ distribution, where sub-jet and track-jet are the closest

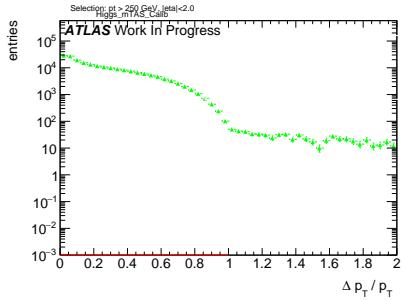


Figure 205: $|p_{\text{T,sub-jet}} - p_{\text{T,track-jet}}|$ distribution, where sub-jet and track-jet are the closest

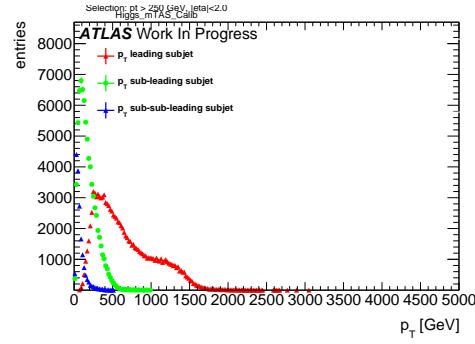


Figure 208: p_{T} distribution for leading, sub-leading and sub-sub-leading sub-jets

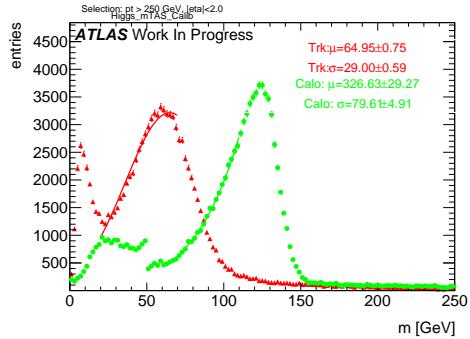


Figure 209: Mass distribution for calorimeter and tracks associated to the large-R jet

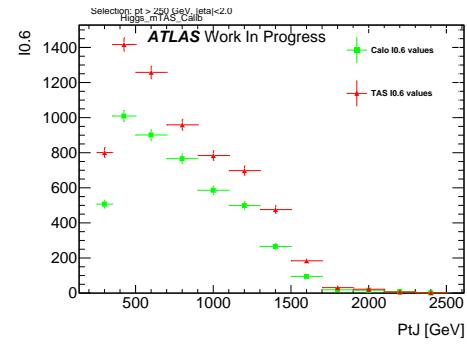


Figure 212: Left integral, $\int_0^{0.6}$ of the mass response, vs bin of p_T^J

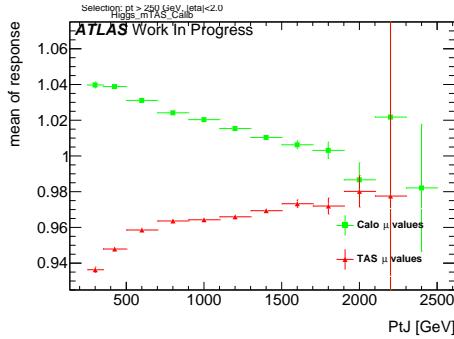


Figure 210: μ from fit of the mass Response vs bin of p_T^J

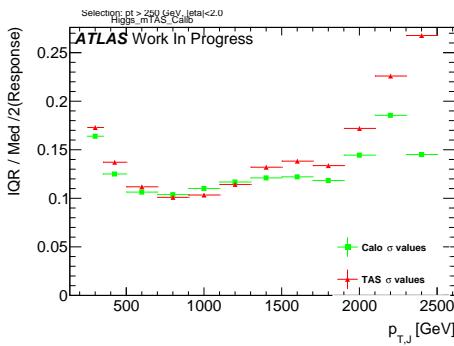


Figure 211: σ from fit of the mass Response vs bin of p_T^J

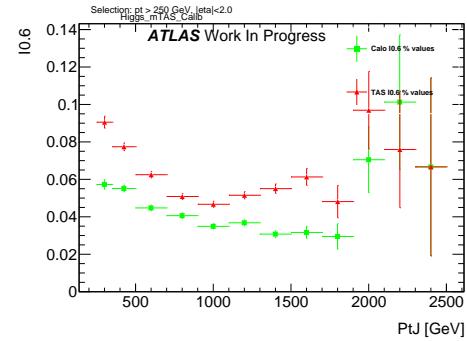
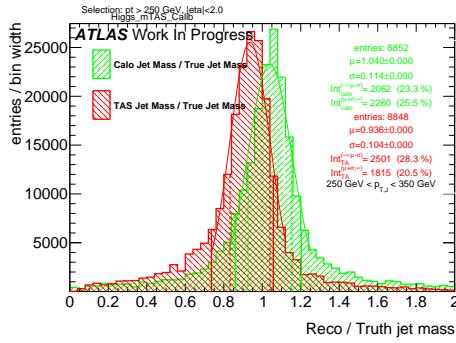
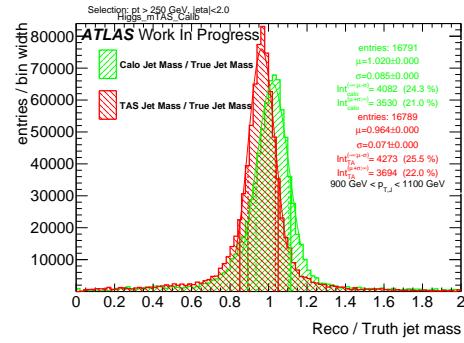
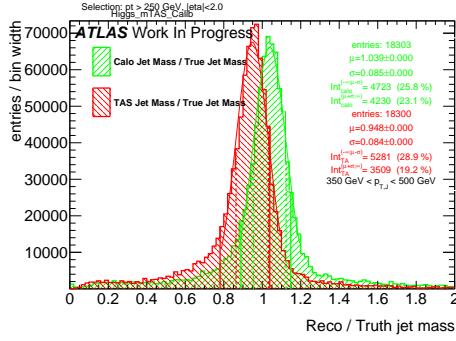
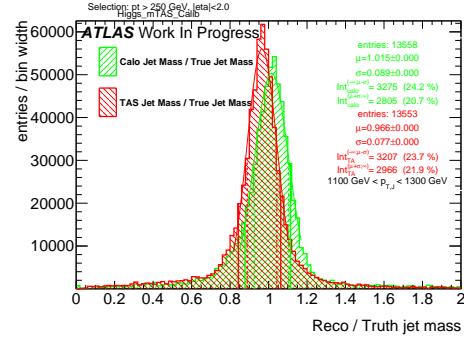
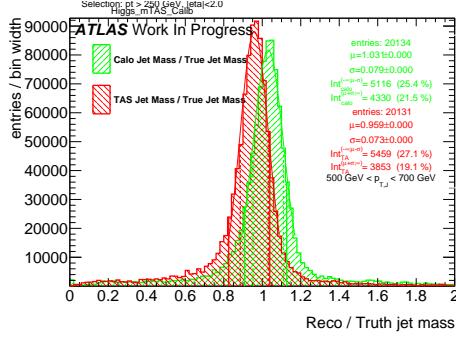
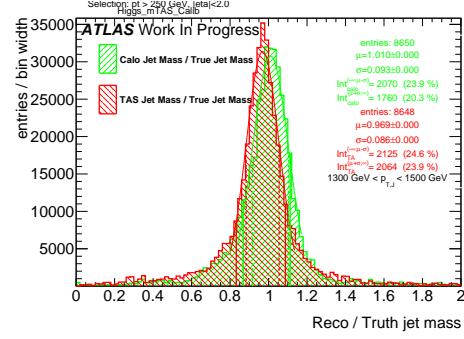
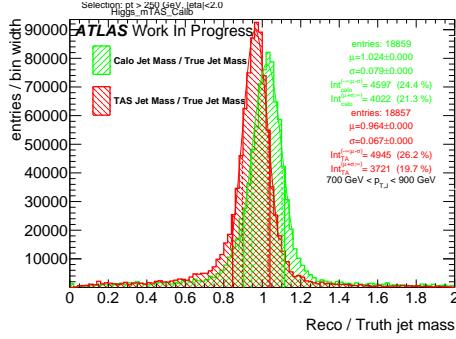
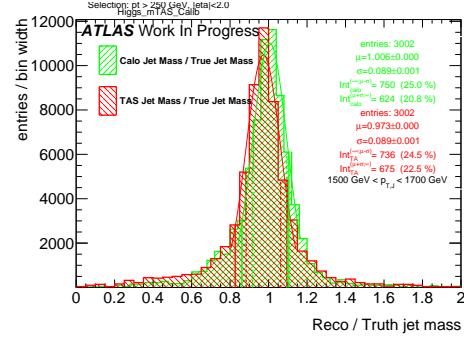


Figure 213: Left integral normalized, $\int_0^{0.6}$ of the mass response, vs bin of p_T^J

Figure 214: Response in bin of p_T^J (indicated on plot)Figure 218: Response in bin of p_T^J (indicated on plot)Figure 215: Response in bin of p_T^J (indicated on plot)Figure 219: Response in bin of p_T^J (indicated on plot)Figure 216: Response in bin of p_T^J (indicated on plot)Figure 220: Response in bin of p_T^J (indicated on plot)Figure 217: Response in bin of p_T^J (indicated on plot)Figure 221: Response in bin of p_T^J (indicated on plot)

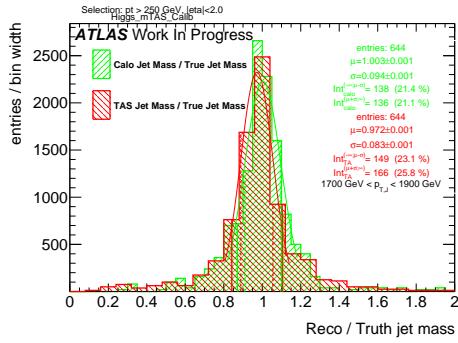


Figure 222: Response in bin of p_T^J (indicated on plot)

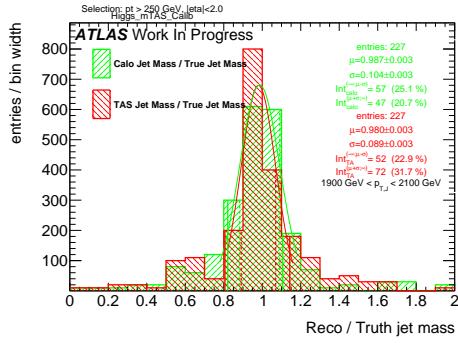


Figure 223: Response in bin of p_T^J (indicated on plot)

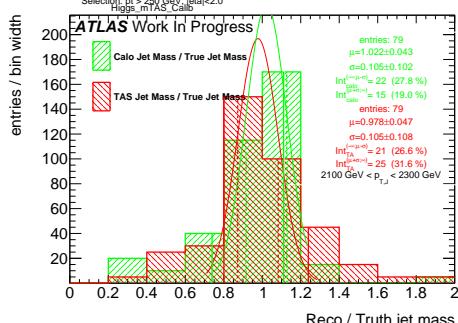


Figure 224: Response in bin of p_T^J (indicated on plot)

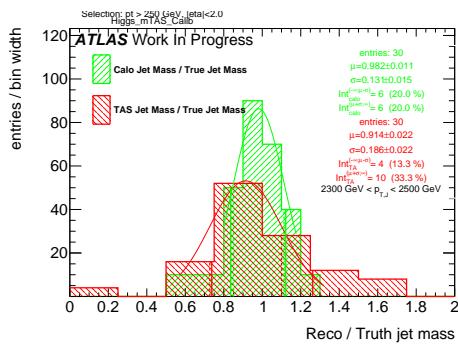
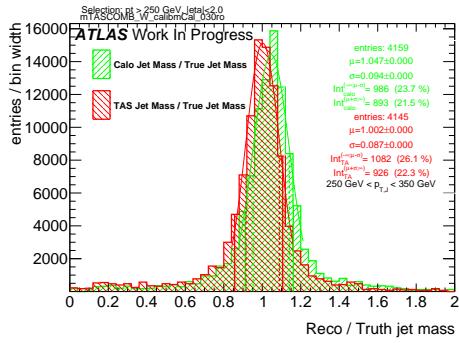
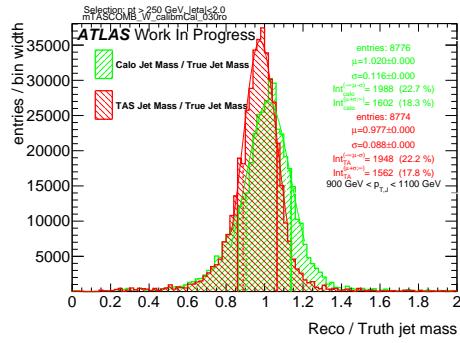
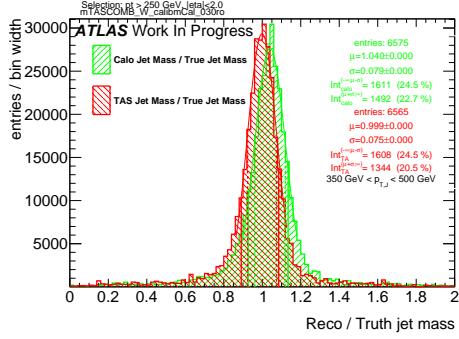
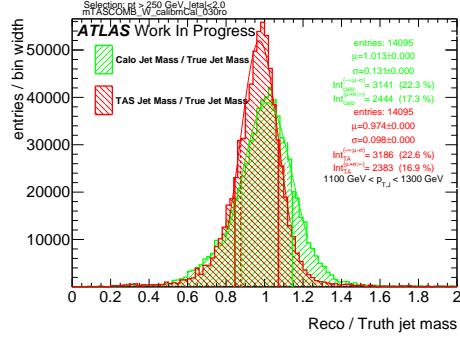
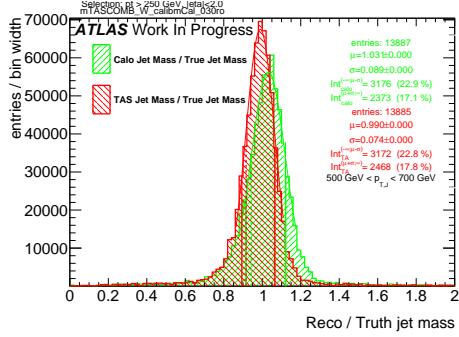
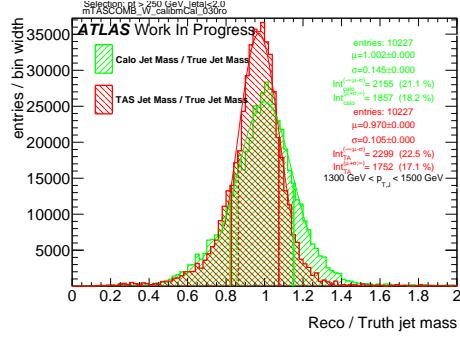
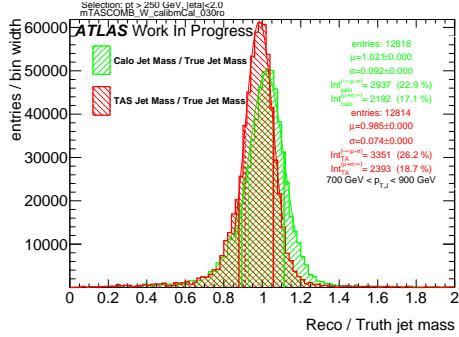
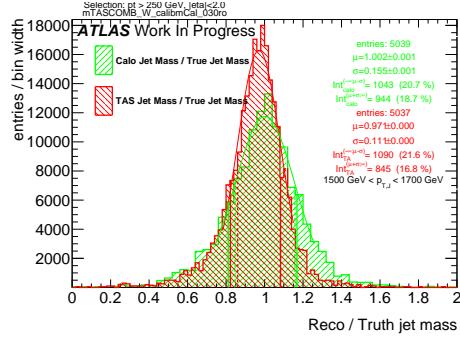
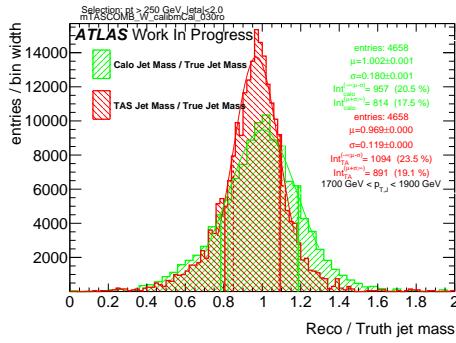
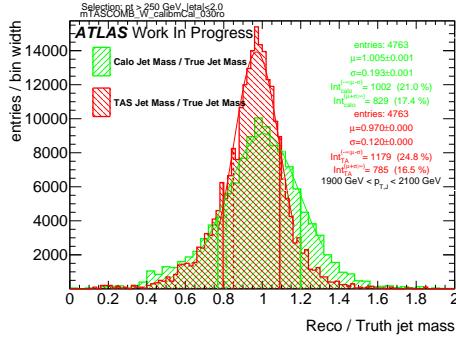
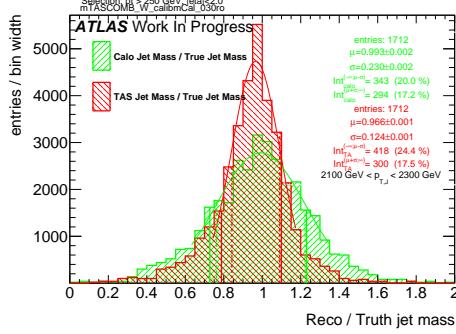
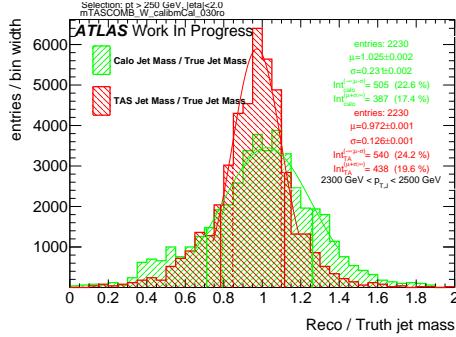


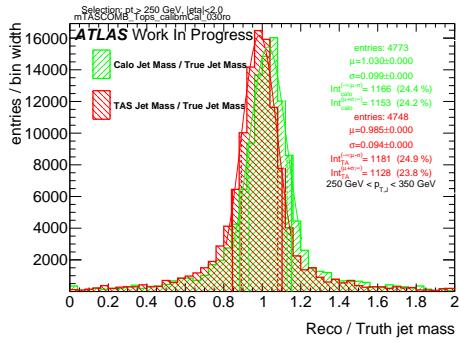
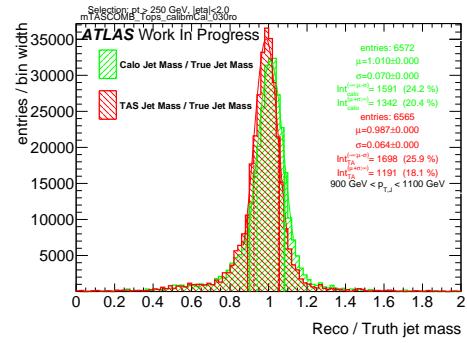
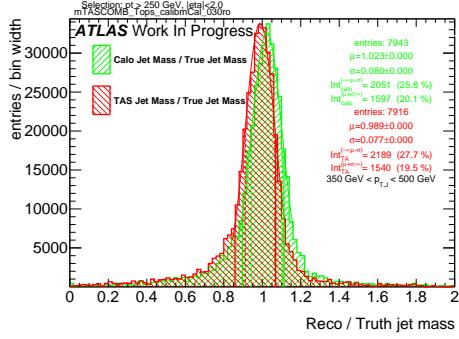
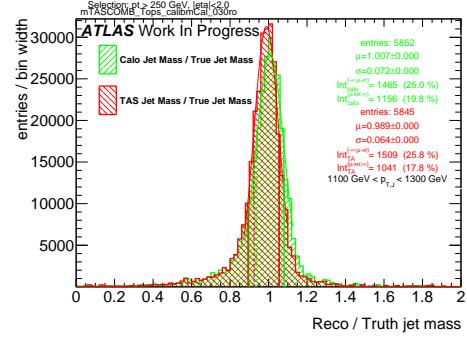
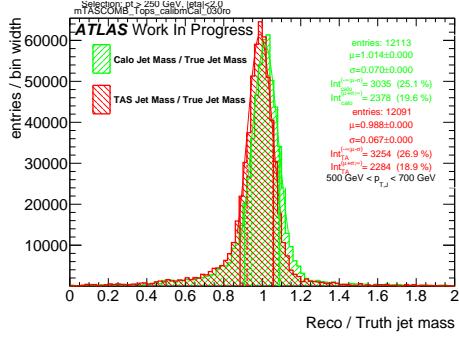
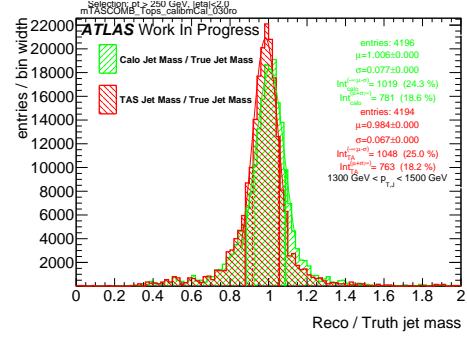
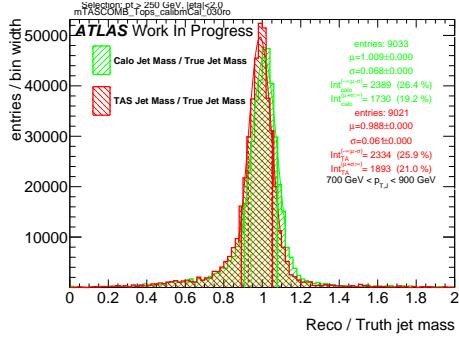
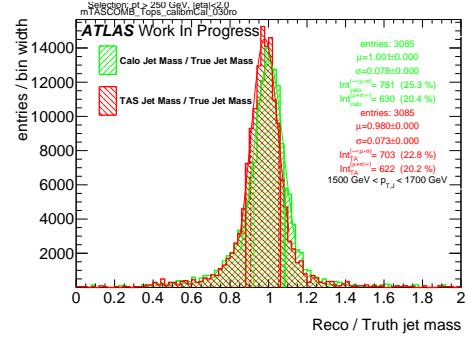
Figure 225: Response in bin of p_T^J (indicated on plot)

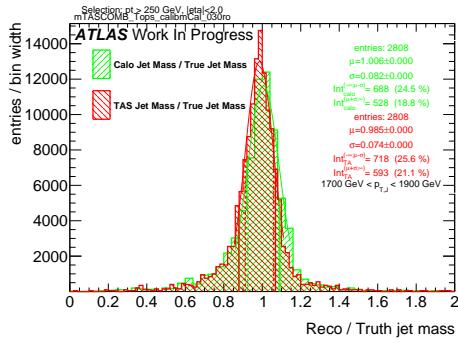
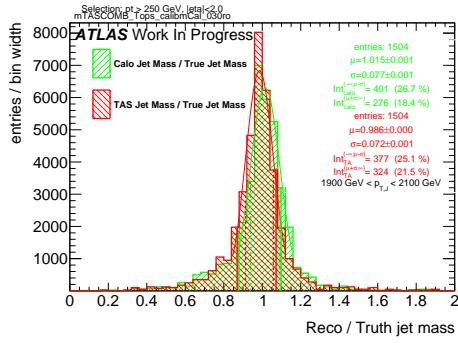
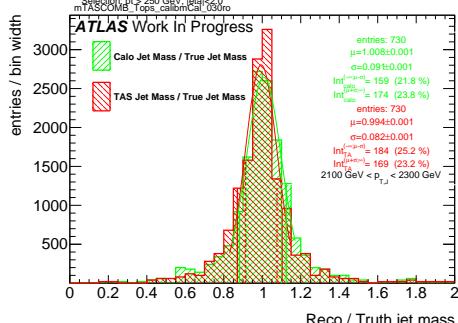
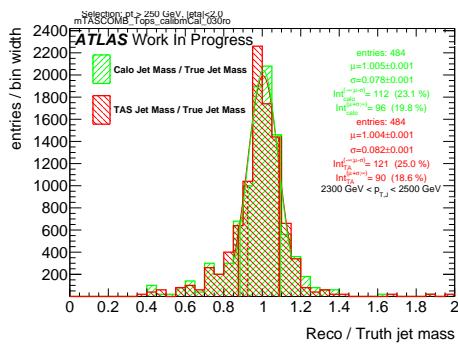
533 **D m_{TAS}^{comb} response distributions, boosted W/Z**

Figure 226: Response in bin of p_T^J (indicated on plot)Figure 230: Response in bin of p_T^J (indicated on plot)Figure 227: Response in bin of p_T^J (indicated on plot)Figure 231: Response in bin of p_T^J (indicated on plot)Figure 228: Response in bin of p_T^J (indicated on plot)Figure 232: Response in bin of p_T^J (indicated on plot)Figure 229: Response in bin of p_T^J (indicated on plot)Figure 233: Response in bin of p_T^J (indicated on plot)

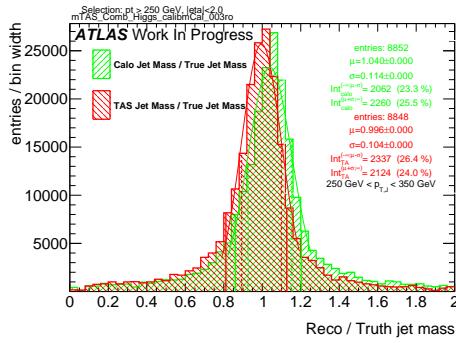
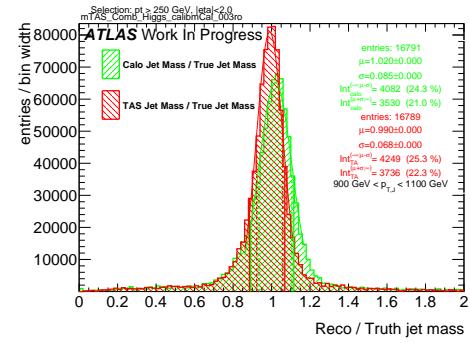
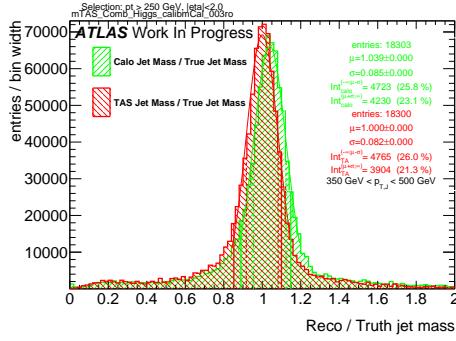
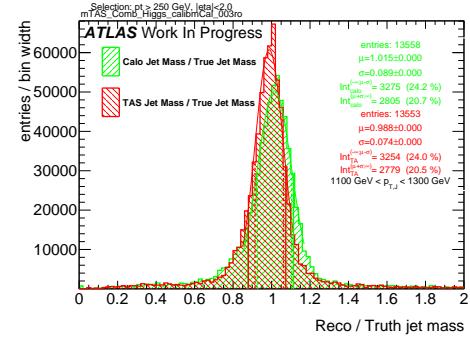
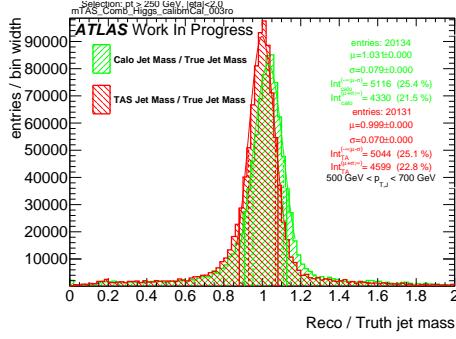
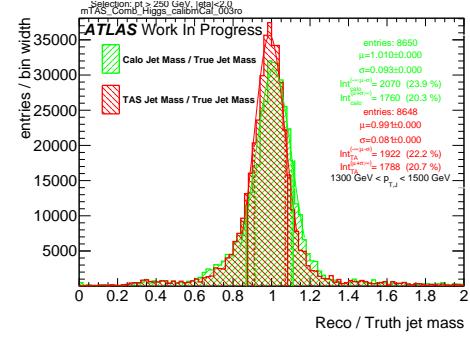
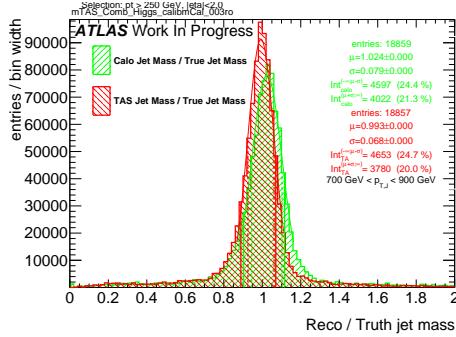
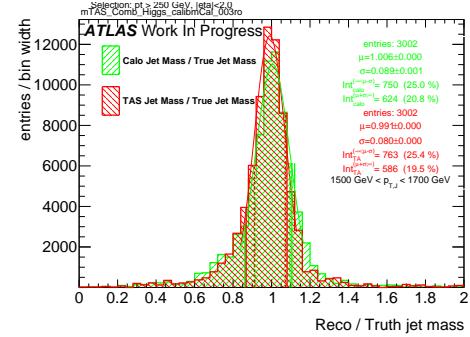
Figure 234: Response in bin of p_T^J (indicated on plot)Figure 235: Response in bin of p_T^J (indicated on plot)Figure 236: Response in bin of p_T^J (indicated on plot)Figure 237: Response in bin of p_T^J (indicated on plot)

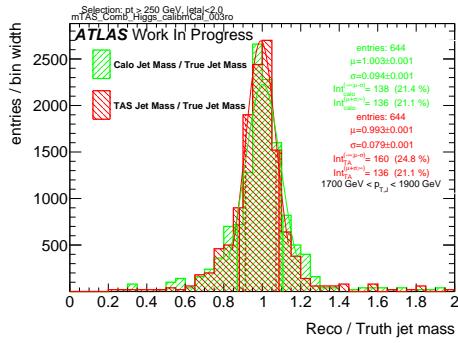
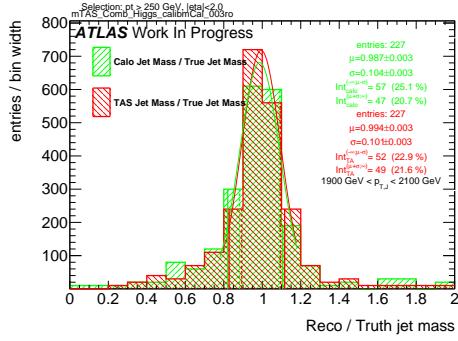
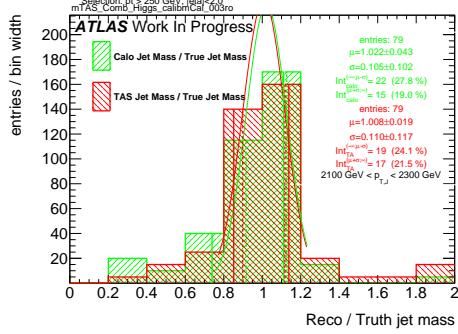
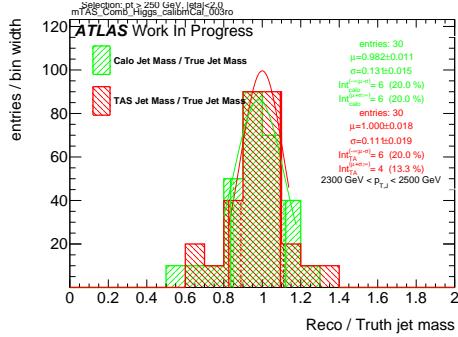
⁵³⁴ E m_{TAS}^{comb} response distributions, boosted tops

Figure 238: Response in bin of p_T^J (indicated on plot)Figure 242: Response in bin of p_T^J (indicated on plot)Figure 239: Response in bin of p_T^J (indicated on plot)Figure 243: Response in bin of p_T^J (indicated on plot)Figure 240: Response in bin of p_T^J (indicated on plot)Figure 244: Response in bin of p_T^J (indicated on plot)Figure 241: Response in bin of p_T^J (indicated on plot)Figure 245: Response in bin of p_T^J (indicated on plot)

Figure 246: Response in bin of p_T^J (indicated on plot)Figure 247: Response in bin of p_T^J (indicated on plot)Figure 248: Response in bin of p_T^J (indicated on plot)Figure 249: Response in bin of p_T^J (indicated on plot)

535 F m_{TAS}^{comb} response distributions, Higgs

Figure 250: Response in bin of p_T^J (indicated on plot)Figure 254: Response in bin of p_T^J (indicated on plot)Figure 251: Response in bin of p_T^J (indicated on plot)Figure 255: Response in bin of p_T^J (indicated on plot)Figure 252: Response in bin of p_T^J (indicated on plot)Figure 256: Response in bin of p_T^J (indicated on plot)Figure 253: Response in bin of p_T^J (indicated on plot)Figure 257: Response in bin of p_T^J (indicated on plot)

Figure 258: Response in bin of p_T^J (indicated on plot)Figure 259: Response in bin of p_T^J (indicated on plot)Figure 260: Response in bin of p_T^J (indicated on plot)Figure 261: Response in bin of p_T^J (indicated on plot)

536 **References**

- 537 [1] Particle Data Group Collaboration, K. Olive *et al.*, *Review of Particle Physics*, Chinese Physics C 38
538 (2014) 090001.
- 539 [2] Paul Langacker, *The Physics of Heavy Z Gauge Bosons*. In: Rev. Mod. Phys. 81 (2009), pp. 1199–1228.
540 doi: 10.1103/RevModPhys.81.1199. arXiv: 0801.1345 [hep-ph].
- 541 [3] Martin Schmaltz and Christian Spethmann, *Two Simple W' Models for the Early LHC*. In: JHEP 07
542 (2011), p. 046. doi: 10.1007/JHEP07(2011)046. arXiv: 1011. 5918 [hep-ph].
- 543 [4] Georges Aad *et al.* *Combination of searches for WW, WZ, and ZZ resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*. In: Phys. Lett. B755 (2016), pp. 285–305. doi:
544 10.1016/j.physletb.2016.02.015. arXiv: 1512.05099 [hep-ex].
- 546 [5] The CMS Collaboration, *Search for massive resonances decaying into pairs of boosted bosons in semi-leptonic final states at $\sqrt{s} = 8$ TeV*, CERN-PH-EP/2013-037 2014/09/03 arXiv:1405.3447v2
547 [hep-ex].
- 549 [6] C. Patrignani *et al.*, Particle Data Group, Chin. Phys. C, 40, 100001 (2016).
550 <http://pdg.lbl.gov/2015/reviews/rpp2015-rev-wprime-searches.pdf>
- 551 [7] C. Patrignani *et al.*, Particle Data Group, Chin. Phys. C, 40, 100001 (2016).
552 <http://pdg.lbl.gov/2015/reviews/rpp2015-rev-zprime-searches.pdf>
- 553 [8] L. Randall and R. Sundrum, *A Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett.
554 83 (1999) 3370, arXiv: hep-ph/9905221.
- 555 [9] H. Davoudiasl, J.L. Hewett and T.G. Rizzo, *Warped Phenomenology*, Stanford Linear Accelerator
556 Center Stanford CA 94309, USA, arXiv:hep-ph/9909255v1 6 Sep 1999.
- 557 [10] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1 & 2*
558 , Tech. Rep. ATLAS TDR 14, CERN/LHCC 99-14, Geneva, 1999.
- 559 [11] L. Evans and P. Bryant, *LHC Machine*, JINST 3 (2008) S08001.
- 560 [12] C. Patrignani *et al.*, Particle Data Group, Chin. Phys. C, 40, 100001 (2016).
561 <http://pdg.lbl.gov/2010/reviews/rpp2010-rev-qcd.pdf>
- 562 [13] The CMS Collaboration, *Measurement of the inclusive 3-jet production differential cross section in proton-proton collisions at 7 TeV and determination of the strong coupling constant in the TeV range*,
563 CERN-PH-EP/2013-037 2015/05/04 arXiv:hep-ph/14121633.
- 565 [14] The ATLAS Collaboration, *Improved luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C (2013) 73: 2518. doi:10.1140/epjc/s10052-013-2518-3.
- 567 [15] W. Bartel *et al.* [JADE Collaboration], *Observation of Planar Three Jet Events in e + e - Annihilation and Evidence for Gluon Bremsstrahlung*, Phys. Lett. B 91, 142 (1980).
- 569 [16] Ahmed Ali (DESY), Gustav Kramer (Univ. Hamburg), *Jets and QCD: A Historical Review of the Discovery of the Quark and Gluon Jets and its Impact on QCD*, arXiv:1012.2288v2 [hep-ph].
- 571 [17] Matteo Cacciari, Gavin P. Salam, Gregory Soyez, *The anti- k_t jet clustering algorithm*,
572 arXiv:0802.1189v2 [hep-ph].

- 573 [18] F. Halzen, A.D. Martin and D.M. Scott, Phys. Rev. D25 (1982) 754, Phys. Lett. B112 (1982) 160.
- 574 [19] Torbjorn Sjostrand, *Monte Carlo Generators*, hep-ph/0611247, CERN-LCGAPP-2006-06.
- 575 [20] W. Lampl *et al.*, *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-
576 PUB-2008-002.
- 577 [21] The ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at $\sqrt{s} = 7$ TeV recorded*
578 *in 2010 with the ATLAS detector*, CERN-PH-EP-2012-191 arXiv:1210.6210v1 [hep-ex].
- 579 [22] GEANT4 Collaboration (S. Agostinelli *et al.*), *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth.
580 A506 (2003) 250-303 DOI: 10.1016/S0168-9002(03)01368-8 SLAC-PUB-9350, FERMILAB-PUB-
581 03-339.
- 582 [23] Ariel Schwartzman, *Jet energy calibration at the LHC*, SLAC National Accelerator Laboratory, 2575
583 Sand Hill Road Menlo Park, California 94025, USA, arXiv:1509.05459v1 [hep-ex] 17 Sep 2015.
- 584 [24] The ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*,
585 ATLAS-CONF-2016-035 19 July 2016.
- 586 [25] The ATLAS Collaboration, *Boosted Higgs ($\rightarrow b\bar{b}$) Boson Identification with the ATLAS Detector at*
587 *$\sqrt{s} = 13$ TeV*, ATLAS-CONF-2016-039 30th July 2016.
- 588 [26] The ATLAS Collaboration, *Search for high-mass diboson resonances with boson-tagged jets in*
589 *proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, arXiv:1506.00962v3 [hep-ex].
- 590 [27] The CMS Collaboration, *Combination of searches for WW, WZ, ZZ, WH, and ZH resonances at $\sqrt{s}=$*
591 *8 and 13 TeV*, CMS PAS B2G-16-007.
- 592 [28] The ATLAS Collaboration, *Search for heavy particles decaying to pairs of highly-boosted top quarks*
593 *using lepton-plus-jets events in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
594 ATLAS-CONF-2016-014.
- 595 [29] The ATLAS Collaboration, *Search for resonances with boson-tagged jets in 15.5 fb^{-1} of p-p*
596 *collisions at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, ATLAS-CONF-2016-055.
- 597 [30] The ATLAS Collaboration, *Search for pair production of Higgs bosons in the $b\bar{b}\bar{b}\bar{b}$ final state using*
598 *proton proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, ATLAS-CONF-2016-049.
- 599 [31] The ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in*
600 *proton-proton collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*, CERN-PH-EP-2013-069
601 arXiv:1306.4945 [hep-ex].
- 602 [32] Oleg Zenin, on behalf of the ATLAS Collaboration, *Soft QCD and underlying event measurements*
603 *at ATLAS*, Nuclear and Particle Physics Proceedings Volumes 273275, April-June 2016, Pages 2053-
604 2058 37th International Conference on High Energy Physics (ICHEP).
- 605 [33] Page title: JetSubstructureECFA2014 AtlasPublic TWiki Website name: Twiki.cern.ch
606 URL:<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetSubstructureECFA2014> Access date: 27th
607 October 2016.
- 608 [34] E. Lehmann and G. Casella, *Theory of Point Estimation*, Springer Verlag, 1998, isbn: 0387985026.

609 **List of contributions**

610

611 **Auxiliary material**