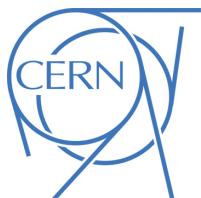




## ATLAS Note

KIP-2016



Draft version 0.2

1

# 2 Jet Observables using Subjet-assisted Tracks

3 Oleg Brandt<sup>a</sup>, Sascha Dreyer<sup>a</sup>, Fabrizio Napolitano<sup>a</sup>, Merve Sahinsoy<sup>a</sup>,  
4 Benjamin P. Nachman<sup>b</sup>

5 <sup>a</sup>*Heidelberg University*

6 <sup>b</sup>*Lawrence Berkeley National Laboratory*

7 29th December 2016

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

# 17 **Contents**

18	<b>1 Introduction</b>	4
19	<b>2 MC Samples and event weighting</b>	4
20	<b>3 Object Definition</b>	5
21	3.1 Standard Large-Radius jet	5
22	3.1.1 Grooming and Selection	5
23	3.1.2 Calorimeter Mass	6
24	3.1.3 Track Mass	6
25	3.1.4 Track-Assisted Mass ( $m^{TA}$ )	6
26	3.2 The Track-Assisted Sub-jet Mass ( $m^{TAS}$ )	7
27	3.2.1 Observable Definition: Inputs	7
28	3.2.2 Observable Definition: Procedure	8
29	3.3 The Combined Mass	10
30	3.3.1 Combination $m^{TA} - m^{calo}$	10
31	3.3.2 Combination $m^{TAS} - m^{calo}$	10
32	3.4 Jet Substructure Observables with (assisted) Tracks	11
33	3.4.1 Prior Mass-Cut	12
34	3.4.2 Energy Correlation Functions	13
35	3.4.3 n-Subjettiness	13
36	3.4.4 TAS Procedure for Jet Substructure Observables	15
37	<b>4 Figures of Merit for Performance Studies</b>	15
38	4.1 For jet mass	15
39	4.2 Receiver Operator Characteristics	16
40	<b>5 Performance of Track-assisted subjet mass</b>	18
41	5.1 Performance in $W \rightarrow q'\bar{q}$ Decays	18
42	5.2 Performance in $h \rightarrow b\bar{b}$ Decays	18
43	5.3 Performance in $t \rightarrow q'\bar{q}b$ Decays	18
44	5.4 Performance in QCD Multijet Events	18
45	5.5 Performance in Massive $\tilde{W} \rightarrow q'\bar{q}$ Decays with $m_{\tilde{W}} = m_t$	19
46	5.6 Stability of Mean of Response and Left-Hand-Side Integral	20
47	5.7 Potential Improvements from Sub-jet Calibration	21
48	5.7.1 Simple Sub-jet Calibration	21
49	5.8 Limitation of $m^{TAS}$ from tracking	22
50	5.9 Performance with Alternate Inputs to the $m^{TAS}$	24
51	<b>6 Performance of Combined Calorimeter and Track-Assisted Sub-Jet Mass</b>	26
52	6.1 Performance in $W \rightarrow q'\bar{q}$ Decays	26
53	6.2 Performance in $h \rightarrow b\bar{b}$ Decays	27
54	6.3 Performance in $t \rightarrow q'\bar{q}b$ Decays	27
55	<b>7 Performance of Jet Substructure Observables with (assisted) Tracks</b>	28
56	7.1 Effect of Track Selection	28
57	7.2 Performance with default $\beta$	29

---

58	7.2.1	Performance for $W$ Tagging	29
59	7.2.2	Tracks and TAS at very high $p_T$	29
60	7.2.3	Performance for Higgs Tagging	31
61	7.2.4	Performance for Top Tagging	32
62	<b>8</b>	<b>Performance of Jet Substructure Observables with (assisted) Tracks and optimised <math>\beta</math></b>	<b>32</b>
63	8.0.1	Optimization for $W$ Tagging	33
64	8.0.2	Optimization for Higgs Tagging	34
65	8.0.3	Optimization for Top Tagging	37
66	8.1	Summary of the Results	37
67	8.1.1	$W$ boson Tagging	39
68	8.1.2	Higgs Boson Tagging	40
69	8.1.3	Top Quark Tagging	41
70	<b>9</b>	<b>Uncertainties on observables with sub-jet-assisted tracks</b>	<b>42</b>
71	<b>10</b>	<b>Conclusions &amp; Outlook</b>	<b>43</b>
72	<b>Appendix</b>		<b>44</b>
73	<b>A</b>	<b>Monte Carlo Samples</b>	<b>44</b>
74	<b>B</b>	<b>Trimming</b>	<b>44</b>
75	<b>C</b>	<b>Tracks details</b>	<b>45</b>
76	<b>D</b>	<b>Alternative Performance Figure of Merit (FoM)</b>	<b>46</b>
77	D.1	Gaussian Fit	46
78	<b>E</b>	<b><math>m^{TAS}</math> distributions, boosted <math>W/Z</math></b>	<b>54</b>
79	<b>F</b>	<b><math>m^{TAS}</math> distributions, boosted tops</b>	<b>66</b>
80	<b>G</b>	<b><math>m^{TAS}</math> distributions, boosted higgs</b>	<b>77</b>
81	<b>H</b>	<b><math>m_{TAS}^{comb}</math> response distributions, boosted <math>W/Z</math></b>	<b>88</b>
82	<b>I</b>	<b><math>m_{TAS}^{comb}</math> response distributions, boosted tops</b>	<b>91</b>
83	<b>J</b>	<b><math>m_{TAS}^{comb}</math> response distributions, Higgs</b>	<b>94</b>
84	<b>K</b>	<b><math>p_T</math> reweighting</b>	<b>97</b>
85	<b>L</b>	<b><math>p_T</math> Dependence of Substructure Observables</b>	<b>97</b>
86	<b>M</b>	<b>Results of <math>\beta</math> Optimisation</b>	<b>98</b>
87	M.1	Performance for $W$ tagging	98
88	M.2	Performance for Higgs tagging	100

---

89	M.3 Performance for Top tagging	102
90	<b>N Signal and Background Distributions</b>	<b>104</b>
91	N.1 W Distributions	104
92	N.2 Higgs Distributions	124
93	N.3 Top Distributions	144
94	<b>Auxiliary material</b>	<b>155</b>

## 95 1 Introduction

96 Jets are collimated streams of particles resulting from quarks and gluons fragmentation and hadronization.  
 97 The distribution of energy inside a jet contains information about the initiating particle. When a massive  
 98 particle such as a top quark, Higgs boson or W/Z bosons is produced with significant Lorentz boost and  
 99 decays into quarks, the entire hadronic decay may be captured inside a single jet. The mass of such jets  
 100 (jet mass) is one of the most powerful tools for distinguishing massive particle decays from the continuum  
 101 multijet background; the ratios of Energy Correlation Functions C2 and D2 (see Reference [bib:ECF])  
 102 and n-Subjettiness  $\tau_{21}$  and  $\tau_{32}$  (see Reference [bib:nsub]) provide an ad-hoc tool pupusely developed for  
 103 the multijet background and constitue a fundamental part of many for boson taggers. This note documents  
 104 the so-called subjet-assisted techniques with the ATLAS detector. The track-assisted subjet mass  $m^{TAS}$   
 105 definition is presented and confronted with the standard development in ATLAS,  $m^{comb}$  and  $m^{TA}$ . Energy  
 106 Correlation Functions and n-Subjettiness with the modified subjet-assisted technique are presented and  
 107 confronted with the standard one in ATLAS. The note ends with conclusions for the jet observables using  
 108 subjet-assisted tracks.

## 109 2 MC Samples and event weighting

110 The samples used are divided into two main groups: SM background and beyond SM signal. The  
 111 SM background includes the QCD multijet samples, produced with a falling  $p_T$  spectrum. The beyond  
 112 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  
 113 ( $t \rightarrow W(\rightarrow q\bar{q}')b$ )) and RS-Graviton  $\rightarrow hh \rightarrow b\bar{b}b\bar{b}$ , i.e. final states have only jets in all the samples.  
 114 The details of the samples are given in Table 6; the masses considered span from 0.5 to 5 TeV to improve  
 115 and diversify the kinematic space covered.

116 The substructure observables are compared via their performance in QCD rejection. While the  $p_T$   
 117 distribution of the multi-jet sample falls exponentially, the  $p_T$  of the signal samples features characteristic  
 118 peaks related to the resonance masses. To avoid bias in the comparison, the signal sample is given weights  
 119 such that the truth  $p_T$  distribution of the leading jet matches the one of the background sample (see Figure  
 120 268 in the Appendix). Furthermore, the spectrum is split into six different  $p_T$  regions to study the behavior  
 121 with rising energy.

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
$W' \rightarrow \tilde{W}\tilde{W}$ 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.

## 122 3 Object Definition

123 This section gives an overview of the objects used for the observables based on subjet-assisted tracks,  
 124 which are the large-radius jet mass, the Energy Correlation Functions and the n-Subjettiness, as they are  
 125 used within ATLAS.

### 126 3.1 Standard Large-Radius jet

127 Large-radius jet, or large- $R$  jets are jets constructed with a radius parameter of the re-clustering algorithm  
 128 of 1.0 for those built using the anti- $k_t$  algorithm and 1.2 for the C/A algorithm. Since the active area of  
 129 this jets is typically six times bigger than their counterparts of radius 0.4 which is the usual choice of jet  
 130 radii within ATLAS, the necessity of further techniques is required to have control over the effect of soft  
 131 radiation contamination from Pile-Up (PU) and Underlying Event (UE).

#### 132 3.1.1 Grooming and Selection

133 **Grooming** In order to use large- $R$  jets, it is necessary to gain additional information on the interior of  
 134 these objects, i.e. using techniques that exploit its substructure allowing a jet-by-jet discrimination of the  
 135 energy deposit most likely coming from the hard-scattering to other soft radiation.

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

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

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

144 The grooming algorithms presented here are the most important ones in ATLAS: the *Trimming*; other  
 145 used as well, the *Split-Filtering* and the *Pruning* can be found in [substructure1]. Details on Trimming,  
 146 the most used within ATLAS and in this note, are given in the Appendix.

<sup>147</sup> **Selection** The selection applied is typical for many Beyond the Standard Model searches:  $p_T > 250$  GeV  
<sup>148</sup> and  $|\eta| < 2.0$  for the large-radius jet. No other requirements were made for the purpose of the performance  
<sup>149</sup> studies here shown if not stated differently, e.g. the mass cut selection.

### <sup>150</sup> 3.1.2 Calorimeter Mass

<sup>151</sup> Once the collection of topo-clusters from the large- $R$  jet is groomed, i.e. cleaned from PU contamination  
<sup>152</sup> through the trimming technique, it is possible to use them for the measure of physical related properties  
<sup>153</sup> such as the jet mass, since the possible sources of soft radiation 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,  $E_i$ ,  $\eta$  and  $\phi$ , the mass can be simply calculated from 4-vector properties:

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

<sup>154</sup> where  $J$  labels the Large- $R$  jet and assuming the topo-clusters as massless.

### <sup>155</sup> 3.1.3 Track Mass

<sup>156</sup> There are significant advantages and few disadvantages of the use of tracks for large-radius jet mass  
<sup>157</sup> reconstruction, inherited both from the detector experimental properties and from the underlying physical  
<sup>158</sup> processes. Main advantages are: performance of angular separation and the association of the tracks to  
<sup>159</sup> the primary vertex for rejection of soft radiation background. Tracks can additionally required to be well  
<sup>160</sup> reconstructed from the detector and they are classified in LOOSE, MEDIUM and TIGHT for increasing  
<sup>161</sup> quality criteria. The mass  $m^{track}$  is then calculated summing up the 4-momenta of those tracks which  
<sup>162</sup> passed the selection and are ghost associated to the groomed jet.

<sup>163</sup> The important disadvantage comes from the complete blindness of the tracker system to the electrically  
<sup>164</sup> neutral component (mostly  $\pi^0$ ) of the jet. As seen in Figure 1, the track mass (red distribution) is not only  
<sup>165</sup> shifted towards lower values than the calorimeter mass (green distribution), but its width also degrades.

<sup>166</sup> Tracks could be used either for independent mass reconstruction or, most importantly, as an additional  
<sup>167</sup> information to the calorimeter measurement.

### <sup>168</sup> 3.1.4 Track-Assisted Mass ( $m^{TA}$ )

<sup>169</sup> The track-assisted mass,  $m^{TA}$ , was one of the first attempts to combine the information form the tracker  
<sup>170</sup> system and from the calorimeter. It is defined as  $m^{TA} = \frac{p_T^{calo}}{p_T^{track}} \times m^{track}$ , where the  $p_T^{track}$  and the  $m^{track}$   
<sup>171</sup> are calculated from the tracks which are associated to the large-radius jet, adding up their 4-momenta (hence  
<sup>172</sup> exploiting the superior angular resolution of the tracker system); the  $p_T^{calo}$  is the transverse momentum  
<sup>173</sup> as measured from the calorimeter system. The ratio  $p_T^{calo}/p_T^{track}$  restores the fraction of the missing  
<sup>174</sup> neutral component in the  $m^{track}$ . The  $m^{TA}$  has a better performance on the reconstruction of boosted  
<sup>175</sup> objects such as  $W/Z$  in the extreme kinematic regime ( $\sim 1$  TeV) and above in the transverse momentum

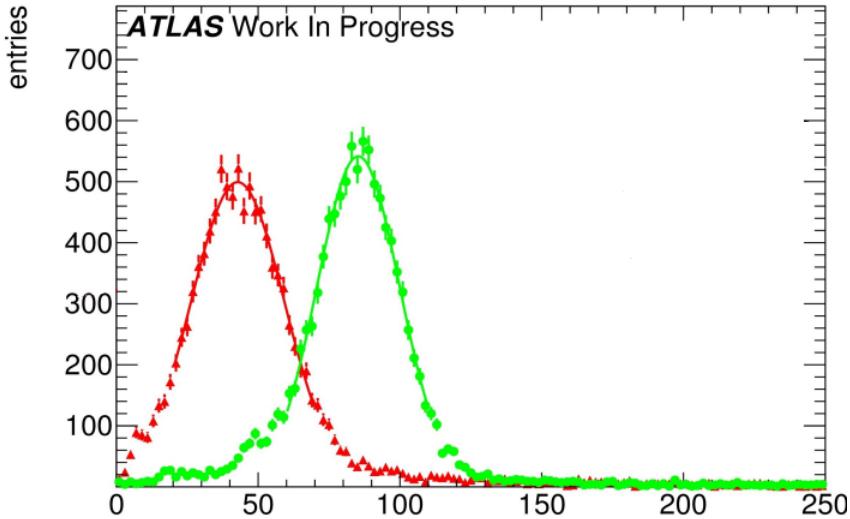


Figure 1: Mass distribution boosted  $W/Z$ : in green the  $m^{calo}$  and in red the  $m^{track}$ .

of the decaying electroweak object. Another advantage of this observable shows up as it comes to the systematic uncertainties: in particular jet mass scale and jet mass resolution uncertainty on  $m^{TA}$  can be estimated by propagating the track reconstruction uncertainties and calorimeter-jet  $p_T$  uncertainties through the definition of the variable given above. The tracking uncertainties are smaller for  $m^{TA}$  rather than  $m^{calo}$  because a larger extent of the uncertainty cancels in the ratio  $m^{track}/p_T^{track}$ . Apart all of this advantages, the track-assisted mass shows its limits when it comes to intermediate transverse momentum regimes and below ( $p_T < 1$  TeV) in  $W/Z$  and for Higgs and top quarks throughout the whole kinematic space. Full description of this variable is given in the ATLAS CONF Note [art35].

### 3.2 The Track-Assisted Sub-jet Mass ( $m^{TAS}$ )

In this section the main outcome of the optimization of the large-radius jet mass reconstruction is presented: the *track-assisted sub-jet mass* ( $m^{TAS}$ ). The main idea takes inspiration from the track-assisted mass: if one can use tracks to exploit the better angular resolution and correct the missing neutral component jet-by-jet, there is an additional information that can be used. The neutral fraction, in fact, varies stochastically not only per-jet basis, but even per-sub-jet basis, since each quark follows a different parton showering and hadronization process. Correcting the missed neutral component per-sub-jet, it should perform better already at an intuitive level, as it accesses information from jet substructure.

#### 3.2.1 Observable Definition: Inputs

There are two inputs to the  $m^{TAS}$ : tracks and sub-jets. The definition of the standard inputs are given here; alternative approaches are given in subsection 5.9.

**Tracks** Only the tracks that satisfy the quality criteria and primary vertex association, described in the appendix 3.1.3, are used. The tracks are additionally required to be ghost associated to the sub-jets of the

197 groomed jet; namely only the sub-jets which survived the trimming procedure and are described in the  
 198 next subsection. Ghost association provides a clear correspondence of tracks to the sub-jets set and was  
 199 therefore chosen and preferred to other kind of assignments.

200 **Sub-jets** The choice of sub-jets must follow a simple requirement: of course we want to take those which  
 201 most likely come from the hard-scattering. This means that the choice of taking them after grooming is  
 202 strongly favored.

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

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

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

### 208 3.2.2 Observable Definition: Procedure

209 There are two ways of sub-jet assisting the tracks for the calculation of the  $m^{TAS}$ : assisting track-jets  
 210 changing the mass or assisting single tracks changing the transverse momentum. The first approach was  
 211 the first one also historically, adopted because of higher versatility and feasibility of implementation. The  
 212 second approach was also found to be equivalent to the first [presentation].

213 For the substructure variable, however, the first approach cannot be used because of simple cancellation  
 214 in the computation of the variable.

215 To generalize the scheme adopted for both, tracks should be assisted singularly. In this note will be shown  
 216 the  $m^{TAS}$  obtained assisting track-jets, since the differences are negligible.

217 **Assisting Track-Jets** Having tracks and sub-jets now well defined, we can describe the recipe to produce  
 218 the  $m^{TAS}$ . For brevity we will call the sub-jets SJ in the formulae below.

219 As said, the tracks are the ones ghost-associated to the sub-jets; however, tracks which fall inside the area  
 220 of the large- $R$  jet, but not inside the sub-jets area, are still much probably coming from the hard-scattering.  
 221 They are then associated again to the closest sub-jets via  $\Delta R$  association.

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

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

226 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  
 227 level, i.e.

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

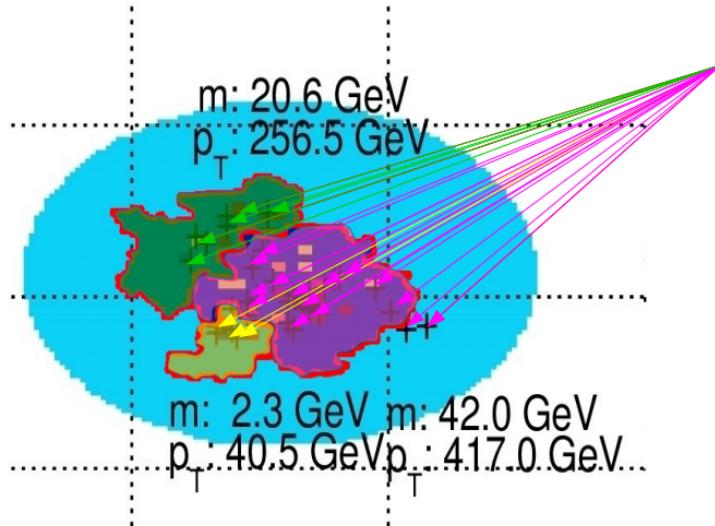


Figure 2: Pictorial event display showing the  $\eta \phi$  region of a large- $R$  anti- $k_t$  trimmed jet, (in blue the catchment area of the anti- $k_t$ ) showing the different  $k_t$  sub-jets: they are highlighted in green, fuchsia and yellow. The associated track-jets (here indicated as arrows pointing the calorimeter area) are colored with the same color of the correspondent sub-jet. Some tracks associated with  $\Delta R$  procedure can be seen in the fuchsia sub-jet. The transverse momenta and mass values are also shown for the sub-jets.

Where the summation symbol between quotation mark symbolize that the sum must be intended at 4-vector level: since now we are working inside the sub-jets, in fact, we need to change the sub-jet's 4-vector itself and not only the mass. If we call  $p_\mu^{TJ}$  the Lorentz vector of the track-jet,

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

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

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

**Assisting Single Tracks** This correction is now applied on single track rather than the whole track-jet and on the transverse momentum, not the mass. The TAS correction reads:

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

235 The corection factor  $\frac{p_T^{SJ}}{p_T^{TJ}}$  refers to the  $p_T$  of the sub-jet in which the track is associated and the  $p_T$  of the  
 236 track-jet associated to it.

237 As before, these four-momenta are then summed together to give this alternative definition:

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

238 where now the sum refers from the first to the M-th tracks associated to the large- $R$  jet.

239 An important remark is that, in the case of a large- $R$  jet with only one sub-jet, the  $m^{TAS}$  has exactly  
 240 the same definition of the  $m^{TA}$ . This implies, since the angular separation of the decay product scales  
 241 inversely with  $p_T$ , that the performance should approach the one of the  $m^{TA}$  at very high transverse  
 242 momenta. However, the space for improvement is precisely in the low-intermediate  $p_T$  regime.

### 243 3.3 The Combined Mass

244 Since the calorimeter large- $R$  jet mass is not explicitly used in the track-assisted (sub-jet) mass, it may be  
 245 possible to improve the performance creating a new observable which combines both mass definitions.  
 246 This is discussed in great details in the BOOST 2016 Conf Note [[art35](#)].

#### 247 3.3.1 Combination $m^{TA} - m^{calo}$

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

$$\begin{aligned} m^{comb} &= a \times m^{calo} + b \times m^{TA}, \\ a &= \frac{\sigma_{calo}^{-2}}{\sigma_{calo}^{-2} + \sigma_{TA}^{-2}} & b &= \frac{\sigma_{TA}^{-2}}{\sigma_{calo}^{-2} + \sigma_{TA}^{-2}} \end{aligned} \quad (1)$$

249 where  $\sigma_{calo}$  and  $\sigma_{TA}$  are the  $m^{calo}$ 's and  $m^{TA}$ 's resolution functions. The  $m^{comb}$  then is the  $m^{TA} - m^{calo}$   
 250 combination. The weights are here and also afterwards computed from the mass response distribution;  
 251 the sigma parameter corresponds to the width of the Gaussian distribution, which is estimated using the  
 252 InterQuantile range.

#### 253 3.3.2 Combination $m^{TAS} - m^{calo}$

254 There is a main difference between the  $m^{TAS}$  and  $m^{TA}$  when it comes to combination: since the  $m^{TAS}$  is  
 255 using sub-jet level information but  $m^{TA}$  not, the correlation with the  $m^{calo}$  is expected to be higher. This  
 256 can be seen e.g. in the plots in Figure 3 (additional plots shown in Figure 43 in Appendix), where the  
 257 correlation is not only higher for the simple  $W/Z$  and Higgs jets, but above 50% for tops. The assumption

of independent variables here falls, since the observable are only approximately Gaussian. The Ansatz is to take into account the correlation via the formula:

$$m_{TAS}^{comb} = w \times m^{calo} + (1 - w) \times m^{TAS},$$

$$w = \frac{\sigma_{TAS}^2 - \rho\sigma_{calo}\sigma_{TAS}}{\sigma_{calo}^2 + \sigma_{TAS}^2 - 2\rho\sigma_{calo}\sigma_{TAS}} \quad (2)$$

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

$$m_{TAS}^{comb} = a \times m^{calo} + b \times m^{TAS},$$

$$a = \frac{\sigma_{TAS}^2 - \rho\sigma_{calo}\sigma_{TAS}}{\sigma_{calo}^2 + \sigma_{TAS}^2 - 2\rho\sigma_{TAS}\sigma_{calo}} \quad b = \frac{\sigma_{calo}^2 - \rho\sigma_{calo}\sigma_{TAS}}{\sigma_{calo}^2 + \sigma_{TAS}^2 - 2\rho\sigma_{TAS}\sigma_{calo}} \quad (3)$$

which reduces to equation (1) after simple algebra for the case when  $\rho = 0$ . Of course, this value can be set to the value of the specific sample considered, or to an average of 0.3 if one wants to give a definition generally valid for all the cases considered; in this case, the performance would be slightly sub-optimal.

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

1. For the given sample, the  $m^{TAS}$  and  $m^{calo}$  are calculated;
2. The mass responses are also produced for the given ranges of  $p_T$ ;
3. For each of these responses, the value of the  $\frac{68\% \text{ IQnR}}{2}$  (identified as the  $\sigma$  in Eq 2) as defined previously is calculated and stored;
4. The average correlation factor of 0.3 (an average value for the samples considered) is assumed;
5. With the formula 2,  $m_{TAS}^{comb}$  is calculated using the  $m^{TAS}$ ,  $m^{calo}$  and the values stored in step 3.

In this note, the IQnR weights are produced for each sample specifically. In order to give a sample-independent definition of the  $m_{TAS}^{comb}$ , following also the procedure adopted for the  $m^{comb}$ , these weights could be taken from a QCD multijet sample and applied indiscriminately to the particular case. Here of course the performance would be again sub-optimal, being the variable not developed in an ad-hoc way for each signal sample, but from QCD multijet only.

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

### 3.4 Jet Substructure Observables with (assisted) Tracks

Jet substructure observables are used to distinguish between different jet topologies, corresponding to signal and background events.  $W$  and Higgs jets originate from two quarks, expected is a structure of two angular separated, hard energy depositions (two-prong). Top jets feature three-prong structure from the  $b$  quark and the hadronically decaying  $W$  boson if all decay products are caught in one large radius jet (typically  $p_T \gtrsim 2m_{Top}$ ). The background are QCD jets which are characterized by a single hard substructure with diffuse, soft wide-angle radiation. These observables are calculated from the jet

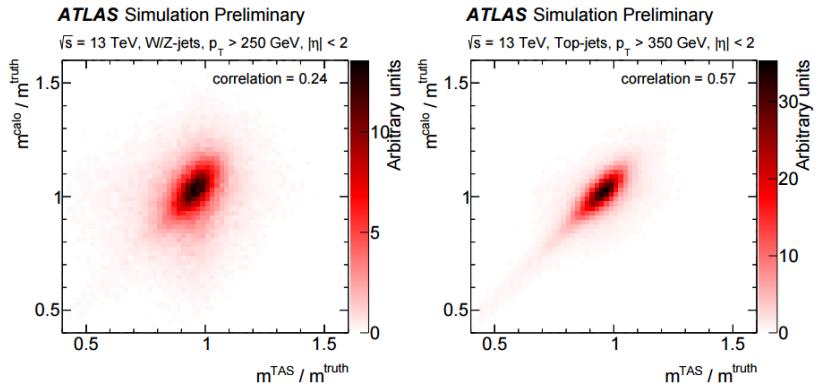


Figure 3: The calorimeter based jet mass mass response versus the track-assisted sub-jet mass response, on the left for boosted  $W/Z$  on the right for boosted tops.

$p_T$ [GeV]	W boson		Higgs boson		Top quark	
	Mass [GeV]	$\frac{1}{\epsilon_{b,gr}}$	Mass [GeV]	$\frac{1}{\epsilon_{b,gr}}$	Mass [GeV]	$\frac{1}{\epsilon_{b,gr}}$
250 - 500	63 - 85	10.8	56 - 167	3.8	77 - 191	6.3
500 - 800	72 - 92	13.6	92 - 150	7.3	117 - 205	6.9
800 - 1200	76 - 104	9.6	98 - 143	9.5	122 - 218	6.5
1200 - 1600	77 - 107	7.3	103 - 149	9.0	122 - 227	6.3
1600 - 2000	79 - 115	5.6	91 - 170	4.4	121 - 235	5.6
> 2000	80 - 126	4.2	/	/	123 - 251	4.8

Table 2: Studied  $p_T$  regions and corresponding calculated 68% mass intervals along with the background rejections from the mass cut for  $W$  boson, Higgs boson and Top quark jets.

287 constituents, which are calorimeter Topo-clusters in the default case or tracks and subjet-assisted tracks  
288 as studied here.

### 289 3.4.1 Prior Mass-Cut

290 Tagging variables are usually used after applying a mass-cut around the interval that contains 68% of the  
291 signal events. Therefore, a cut is applied on the calibrated mass of the large-R calorimeter jet which is  
292 calculated to cover the smallest interval around the peak mass that contains 68% of the signal events. In  
293 the Higgs tagging case, there is not enough statistics to derive a conclusive result for  $p_T > 2000 \text{ GeV}$ .  
294 Hence this study is restricted to the five lower  $p_T$  bins.

295 Prior to tagging with the n-Subjettiness or C2/D2 variables, a cut on the calibrated calorimeter jet mass is  
296 applied, given that the mass is the main discriminant in QCD jet rejection. This cut is defined to choose  
297 the smallest interval around the peak mass containing 68% of the signal. However, the reconstructed  
298 mass depends on the  $p_T$  region, therefore a different cut was calculated for every region to meet the  
299 requirements.

300 **3.4.2 Energy Correlation Functions**

301 The C2 and D2 observables were found to be very powerful to distinguish between one- and two-  
 302 prong like jets, see e.g. Reference [bib:power\_counting]. They are defined as ratios of the so called  
 303 *Energy Correlation Functions* ECF( $N, \beta$ ) which explore the substructure of a jet using a sum over the  
 304 constituents. These variables are IRC-safe ( $\beta \geq 0$ ) and theoretically well understood, see Reference  
 305 [bib:analytic\_ECF]. D2 was found to perform slightly better for tagging  $W$  jets as C2 in Reference  
 306 [bib:w\_tagging], most notably due to a more  $p_T$  robust cut value and somewhat higher background  
 307 rejections.

$$\begin{aligned} C2(\beta) &= \frac{\text{ECF}(3, \beta) \cdot \text{ECF}(1, \beta)}{\text{ECF}(2, \beta)^2} \\ D2(\beta) &= \frac{\text{ECF}(3, \beta) \cdot \text{ECF}(1, \beta)^3}{\text{ECF}(2, \beta)^3} \end{aligned} \quad (4)$$

308 The constituents are correlated via a product of their  $p_T$ , weighted by an angular factor. This factor is  
 309 calculated from the pairwise angular distances of the considered constituents. An additional exponent  
 310  $\beta$  scales this factor. The default value for  $\beta$  is 1, corresponding to angular and momentum parts being  
 311 weighted equally.

$$\begin{aligned} \text{ECF}(1) &= \sum_{\text{constituents}} p_T \\ \text{ECF}(2, \beta) &= \sum_{i=1}^n \sum_{j=i+1}^n p_{T,i} p_{T,j} \Delta R_{ij}^\beta \\ \text{ECF}(3, \beta) &= \sum_{i=1}^n \sum_{j=i+1}^n \sum_{k=j+1}^n p_{T,i} p_{T,j} p_{T,k} (\Delta R_{ij} \Delta R_{ik} \Delta R_{jk})^\beta \end{aligned} \quad (5)$$

312 ECF(2,  $\beta$ ) uses pairwise correlation and is sensitive to two-prong structures. ECF(3,  $\beta$ ) relies on triple-  
 313 wise correlations to identify three-prong structures. ECF(1,  $\beta$ ) corresponds to the  $p_T$  of the whole jet and  
 314 serves as normalization to minimize the dependence on the energy scale.

315 For collinear or soft configurations of  $N$  constituents, ECF( $N, \beta$ ) acquires very small values. Given  
 316 ECF(2,  $\beta$ ), only well angular separated and non-soft constituent pairs act as non-negligible sum terms.  
 317 This directly leads to the picture of two hard substructures inside the jet. A similar conclusion can be  
 318 done for ECF(3,  $\beta$ ) and three hard substructures. A jet with  $N$  or more hard substructures features a higher  
 319 ECF( $N, \beta$ ) value as a jet consisting of fewer than  $N$  prongs. Consequently, the ratios C2 and D2 can be  
 320 used to separate these topologies.

321 A  $W$  jet features a small ECF(3,  $\beta$ ) but a high ECF(2,  $\beta$ ) value resulting in small C2/D2, corresponding  
 322 to a high agreement with the two-prong hypothesis. QCD jets feature a very small ECF(3,  $\beta$ ) and a small  
 323 ECF(2,  $\beta$ ) value. Considering the power of ECF(2,  $\beta$ ), this results in a higher C2/D2 value as for a  $W$  jet.  
 324 A  $W$  tagging example for C2 and D2 with clusters as input is given in Figure 4.

325 **3.4.3 n-Subjettiness**

326 Similar to C2/D2, ratios of *n-Subjettiness* variables  $\tau_N$  can be used to distinguish between different jet  
 327 topologies. Often used for Top tagging is the observable:

$$\tau_{32} = \frac{\tau_3}{\tau_2} \quad (6)$$

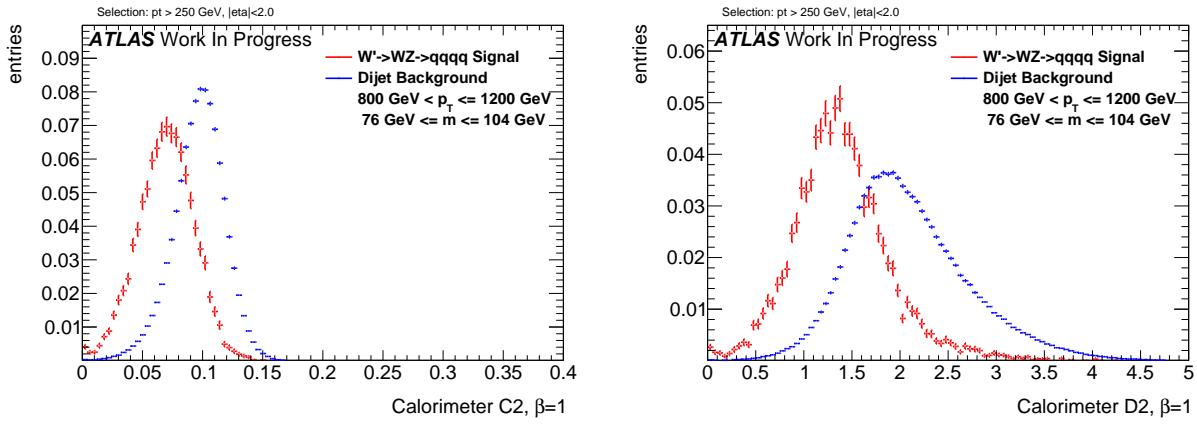


Figure 4: Exemplary C2 (left) and D2 (right) distributions for  $W$  signal jets and QCD background jets, calculated with clusters. In both cases the background tends to higher values.

328 N-Subjettiness  $\tau_N$  quantifies the level of agreement between a given large-R jet and a certain number  
 329  $N$  of sub-jet axes. The two mainly used sub-jet axes definitions are  $k_T$ -axes and the  $k_T$ -WTA (Winner  
 330 Takes All) definition. The jet is reclustered with an exclusive  $k_T$ -algorithm that runs the recombination  
 331 just until  $N$  sub-jets are clustered. The  $k_T$ -axes are defined by the four-momenta of the  $k_T$ -sub-jets, WTA  
 332 correspond to the four-momentum of the hardest constituent in each  $k_T$ -sub-jet. Used in this study is the  
 333  $k_T$ -WTA axis definition. Calculated via a sum over the jets constituents (calorimeter clusters as default),  
 334 the variable is defines as follows.

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k})^\beta \quad (7)$$

335 The constituents  $p_T$  is multiplied by the angular distance to the nearest sub-jet axis. The overall value is  
 336 normalised with a sum over the constituents  $p_T$  times the characteristic radius parameter  $R$  of the large jet.  
 337

$$d_0 = \sum_k p_{T,k} R_0 \quad (8)$$

338 Again, the angular part can be scaled relative to the  $p_T$  factor via the exponent  $\beta$ . N-Subjettiness is an  
 339 IRC-safe variable for values of  $\beta \geq 0$ .

340 Small values of  $\tau_N$  correspond to a jet with all constituents more or less aligned or near to the given  $N$   
 341 sub-jet axes. Hence, the jet is compatible with the assumption to be composed of  $N$  or fewer sub-jets.  
 342 A higher value indicates a consistency with more than  $N$  sub-jets as a non-negligible part is located  
 343 apart of the  $N$  sub-jet axes. Consequently, Top jets are likely to feature a small  $\tau_3$  and a high  $\tau_2$  value.  
 344 QCD jets with their one-prong structure result in a high  $\tau_3$  and a small  $\tau_2$  value. While  $\tau_2$  and  $\tau_3$  alone  
 345 provide only slightly separation, their ratio has proven to be very powerful. A Top tagging example with  
 346  $\tau_{32}$  calculated with clusters in given in Figure 5. The  $\tau_{21}$  ratio can as well be used to tag  $W$  or Higgs  
 347 jets. Nevertheless, C2 and D2 were found to yield higher background rejections for  $W$  tagging, see e.g.  
 348 Reference [bib:w\_tagging].

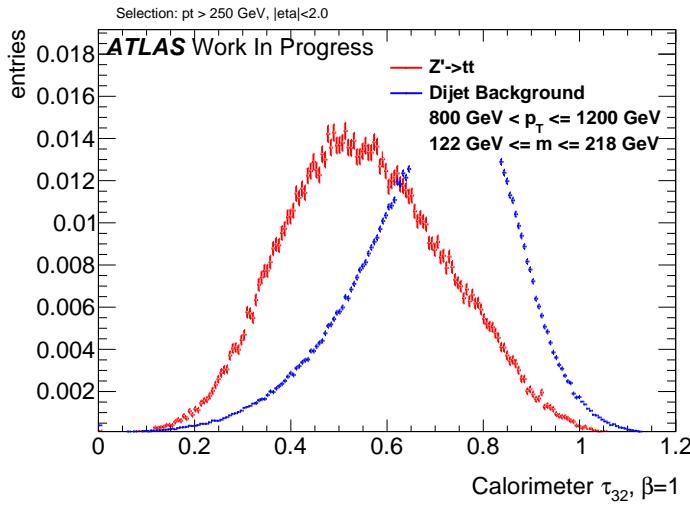


Figure 5: Exemplary  $\tau_{32}$  distributions for Top signal jets and QCD background jets, calculated with clusters. The background tends to higher values.

#### 349 3.4.4 TAS Procedure for Jet Substructure Observables

350 The concept of track assisting with the  $p_T$  ratio of the whole jet is without effect for the studied substructure  
 351 variables. This can be understood from the definitions of the weighted  $p_T$  sums. If corrected with only  
 352 one ratio, all tracks are scaled by the same factor  $c$ , which then can be put in front of the sum and cancels  
 353 as soon as the ratios  $\tau_{21}$  and  $\tau_{32}$ , respectively C2 and D2 are formed.

$$\begin{aligned}\tau_N &= \frac{1}{d_0} \sum_k p_{T,k} c \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k})^\beta \\ &= \frac{c}{d_0} \sum_k p_{T,k} \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k})^\beta\end{aligned}\quad (9)$$

354 Track assisting with ghost association to subjets (TAS), see Section ?? for  $m^{TAS}$  works with different  
 355 scaling factors depending on the corresponding sub-jet  $c_k$ , which also affect ratios:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} c_k \min(\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k})^\beta \quad (10)$$

356 Therefore, used for the substructure observables is the TAS method assisting single tracks as introduced  
 357 in Section 3.2.2. Studied are tracks and assisted tracks (TAS) as well as calorimeter clusters as input for  
 358 these observables. The used selection for tracks is the same as for TAS, only the part of  $p_T$  scaling is  
 359 omitted.

## 360 4 Figures of Merit for Performance Studies

### 361 4.1 For jet mass

362 The general idea of Figure of Merit (FoM) is given in the Appendix; here the InterQuantile range  
 363 is described since used in this note and identical to the one used in the conference BOOST 2016.

- 364 The InterQuantile range (IQnR) is here defined as it corresponds to a sigma of a “perfect” Gaussian  
 365 distribution:  $q84\% - q16\%$  where  $q84\%$  is the 84<sup>th</sup> percentile and  $q16\%$  is the 16<sup>th</sup>, not to be confused  
 366 with the InterQuartile Range (IQR) which is the  $q75\% - q25\%$  and does not correspond to the sigma.  
 367 The final descriptor is then divided by the Median ( $\frac{1}{2} \times 68\% \text{ IQnR}/\text{median}$ ). It provides stability and high  
 368 sensitivity to left-hand-side and right-hand-side tails.
- 369 The IQnR is then applied to the response distribution Figure of Merit: given the reconstructed mass  
 370 (calorimeter, track etc.) one can compare it to its *truth* mass ( $m^{truth}$ ), computed from the particle at MC  
 371 level before the interaction with the detector:

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

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

373 In Figure 6 a mass response for a single range of transverse momentum is shown, for the calorimeter mass.  
 374 On the plot the contours of a standard deviation and of  $q16\%$  and  $q84\%$  are drawn with dashed and solid  
 375 lines, respectively, showing the difference induced by the tail. This sort of plot is the key when looking  
 376 quantitatively to the observable performance and can be found in the Appendix for each of the process  
 377 studied in every  $p_T$  range considered.

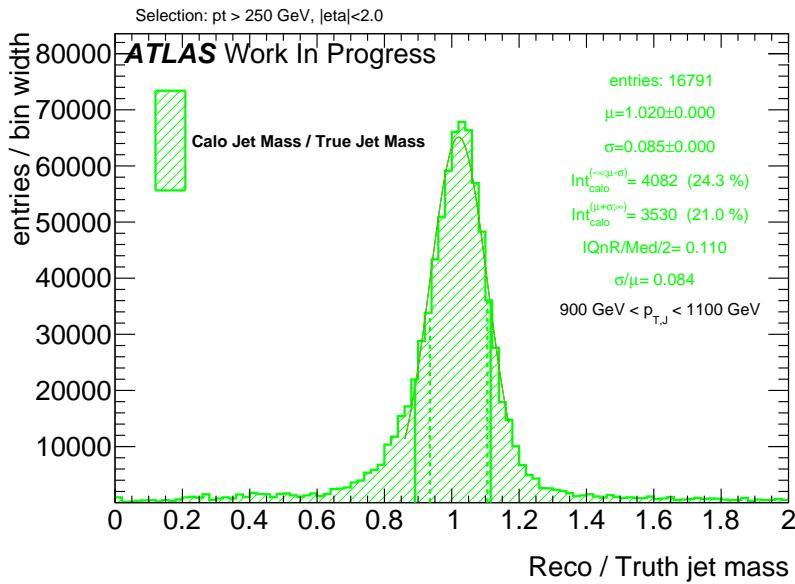


Figure 6: Calorimeter mass response plot for boosted  $W/Z$ . One the plot, right, are shown: the number of entries, the mean and the width of the fit to the Gaussian core, the integral from 0 to  $\mu - \sigma$  and the one from  $\mu + \sigma$  to  $+\infty$ , the values  $\frac{1}{2} \times 68\% \text{ IQnR}/\text{median}$  and  $\sigma/\mu$ . On the distribution the dashed vertical lines represent the points  $\mu - \sigma$  and  $\mu + \sigma$  and the solid lines represent the  $q16\%$  and  $q84\%$ . These lines also explicitly show the asymmetry between the left-hand-side flank, in general more pronounced, and the right-hand-side one

## 378 4.2 Receiver Operator Characteristics

- 379 The separation power of discrimination variables can be studied quite intuitively by comparing the signal  
 380 and background distributions of a certain variable. Another used figure of merit for the performance,

especially for comparisons of different variables, is to use *Receiver Operator Characteristics* (ROC) which show the achieved background rejection for different values of signal efficiency (signal fraction left after performing a cut). Each point is calculated from the underlying signal and background distributions by integrating the background distribution from zero <sup>1</sup> to the point where the desired signal fraction is achieved. The fraction of background events contained in this region are kept when cutting at this signal efficiency, hence the inverse of this fraction,  $\frac{1}{\epsilon_{background}}$  is an estimate for the background rejection. The lower the fraction of background events in the region, the better is the achieved exclusion. Accordingly, a good discrimination variable is represented by a ROC with preferably high values of background rejection up to high signal efficiencies.

---

<sup>1</sup> If the signal distribution lies at lower values as the background.

## 390 5 Performance of Track-assisted subjet mass

391 The track-assisted subjet mass takes inspiration from the simpler development which is already imple-  
 392 mented within ATLAS, the track-assisted mass which is described briefly below for completeness.

### 393 5.1 Performance in $W \rightarrow q' \bar{q}$ Decays

394 The boosted  $W/Z$  was the first one looked at, and with which the  $m^{TAS}$  was designed. The  $m^{calo}$  shows a  
 395 fast deterioration of the performance at high  $p_T$ , and, as shown in the previous section, the  $m^{TA}$  prevents  
 396 this deterioration but suffers at low transverse momenta ( $p_T < 1$  TeV). The  $m^{TAS}$  has a similar behavior in  
 397 the extreme transverse momentum regime as the  $m^{TA}$ , since the sub-jet multiplicity peaks at one, where  
 398 there are no differences between the two observables. In the low- $p_T$  regime, on the contrary, it exploits  
 399 the difference in charged to neutral ratio for each sub-jet, achieving a better performance. This is shown  
 400 in Figure 7(a) as a function of  $p_T$ : below  $\sim 1$  TeV achieves lower values of the IQnR converging from  
 401 below to the  $m^{TA}$  as the number of sub-jets decreases to one.

### 402 5.2 Performance in $h \rightarrow b\bar{b}$ Decays

403 In the Randall-Sundrum graviton to di-Higgs to four b-quark, the performance is again problematic for the  
 404  $m^{TA}$  with respect to  $m^{calo}$ , which is far beyond the latter, while the performance of the  $m^{TAS}$  is partially  
 405 similar to the boosted top-quark sample, but degrades much more in the extreme  $p_T$  regime, following the  
 406  $m^{TA}$ . Shown in Figure 7(b).

### 407 5.3 Performance in $t \rightarrow q' \bar{q} b$ Decays

408 The boosted tops are shown on Figure 7(c); the  $m^{TAS}$  is comparable yet slightly worse than the  $m^{calo}$   
 409 in the low-middle  $p_T$  regime, while degrades at higher  $p_T$  approaching the  $m^{TA}$ , which is far beyond the  
 410 track-assisted sub-jet mass in performance. As already noted, the worse performance can be ascribed both  
 411 to the higher top-quark mass, and to its different and more complex decay topology.

### 412 5.4 Performance in QCD Multijet Events

413 The behavior of the QCD multijet sample is similar to the boosted  $W/Z$  sample, where the  $m^{TA}$  exhibits  
 414 a crossing point in the middle-low regime  $p_T \simeq 900$  GeV and proceeds with a better performance at high  
 415 transverse momenta. Again the  $m^{TAS}$  follows this similarity showing no crossing point and an optimal  
 416 overall behavior, both with respect to calorimeter- and track-assisted-based mass definition. On Figure  
 417 8.

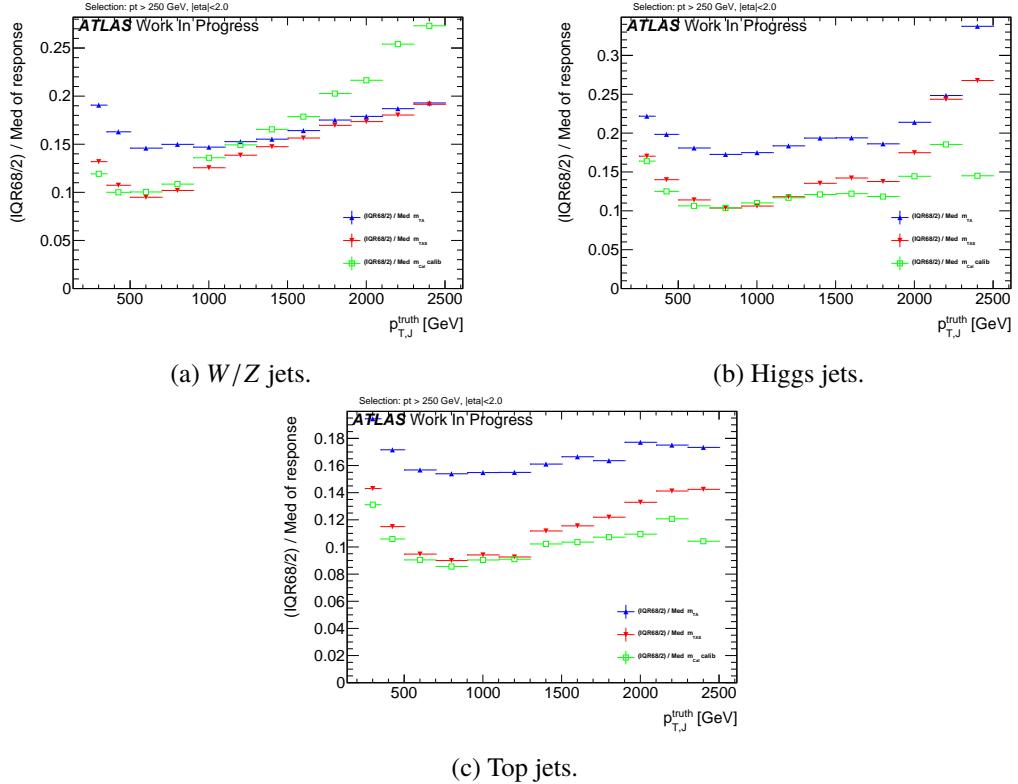


Figure 7: Performance of the  $m^{TAS}$  versus the  $m^{calo}$  and  $m^{TA}$  for  $W/Z$ , top left, where  $m^{TA}$  is not better than  $m^{calo}$  in the low  $p_T$  range but is outperformed by the  $m^{TAS}$ ; Higgs decay, where  $m^{calo}$  is everywhere better than  $m^{TA}$ , yet comparable with  $m^{TAS}$  and top decays where the more complex topology makes critical the high  $p_T$  regime

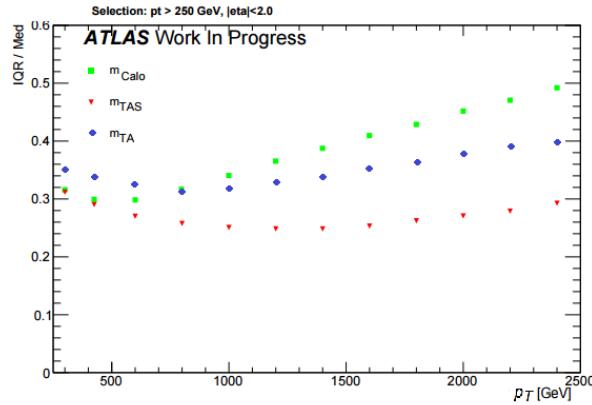


Figure 8: Performance of the  $m^{TAS}$  versus the  $m^{calo}$  and  $m^{TA}$  for the QCD multijet shows a much better behavior of the track-assisted sub-jet mass. Here shown 50% IQnR/median and not the  $\frac{1}{2} \times 68\%$  IQnR/median.

## 418 5.5 Performance in Massive $\tilde{W} \rightarrow q' \bar{q}$ Decays with $m_{\tilde{W}} = m_t$

419 The massive  $W$  sample is a special sample which was used to understand the behavior of the boosted tops,  
420 whether its worse resolution was coming from the higher mass of the top quark or from the more complex  
421 decay topology (three-pronged instead of two-pronged decay and  $b$ -quark presence). The sample is almost

422 identical to the boosted  $W/Z$  one ( $W' \rightarrow WZ$ ) but in this case the SM electroweak boson have the mass  
 423 of the top quark  $m_W = m_t$ . In fact, from the rule  $\Delta R = 2m/p_T$ , a bigger separation is expected between  
 424 quarks from the hadronic decay. The comparison with  $m^{calo}$  is shown in Figure 9, together with the  
 425 boosted top-quark for completeness. As seen here, the performance of the latter is clearly worse than the  
 426 former, the trend is yet very similar. This difference is interpreted in terms of different and more complex  
 427 topology and hence higher sub-jet multiplicity: in the three sub-jet structure, resolving accurately the  
 428 components is more challenging.

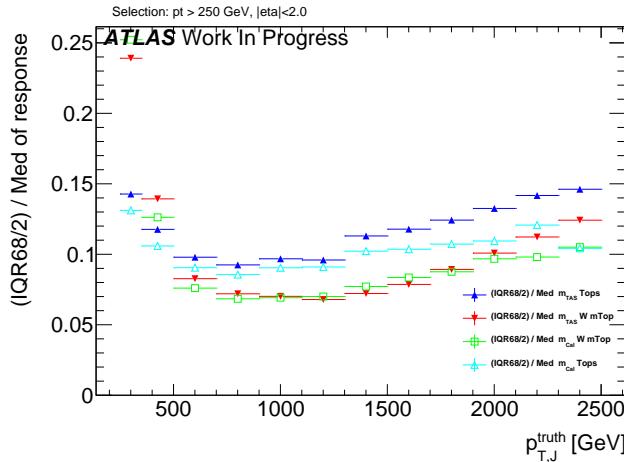


Figure 9: Performance of the  $m^{TAS}$  versus the  $m^{calo}$  for the massive  $W/Z$  (in red and green); shown on the same plot also the boosted top sample (in blue and light blue).

## 429 5.6 Stability of Mean of Response and Left-Hand-Side Integral

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

435 The mean of the response distribution is shown for boosted  $W/Z$  decays in Figure 10, left; as seen here,  
 436 despite the mean being constantly below unity, its behavior is much more flat and independent of  $p_T$ ,  
 437 especially in the low-intermediate regime. This is surprising since the  $m^{calo}$  is already shown after the  
 438 calibration step, which is not taken instead for the  $m^{TAS}$ . Conversely the left-hand-side tail of the mass  
 439 response which is shown in the same figure, right, shows a more enhanced behavior than the  $m^{calo}$ , but  
 440 still never reaches the 10%. Of course an enhancement of the tail causes a loss of gaussianity and a  
 441 number of jets which are reconstructed with a lower mass than they should, but it is still comparable with  
 442 the calorimeter mass.

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

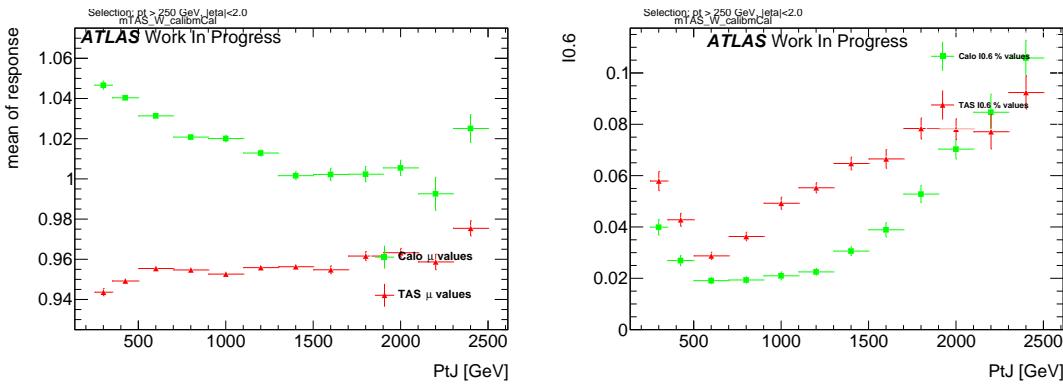


Figure 10: Stability quantifiers which were checked for the  $m^{TAS}$ : mean on the left and normalized left-hand side integral of the mass response distribution on the right. The mean is calculated from a Gaussian fit and the integral goes from 0 to 0.6.

## 445 5.7 Potential Improvements from Sub-jet Calibration

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

449 The *perfect calibration* refers to the procedure of using  $m^{TAS}$  with truth-level information for calorimeter  
450 and tracker system, i.e. looking at the best possible scenario with an ideal detector. The performance is of  
451 course expected to be optimal, because of the use of the truth-level. This step was necessary as feasibility  
452 study, to understand whether ulterior efforts in this direction were meaningful. Truth-level tracks are the  
453 particles in the jet which have an electric charge and are stable, truth-level sub-jets are all the particles,  
454 charged and not, which are ghost associated to the calorimeter sub-jets. There are few possibilities in  
455 doing so, here some nomenclature for this study will be introduced:

- 456 •  $m^{TAS}$  using truth-level sub-jets and tracks; normal tracks (with all detector effects) are used to assist  
457 the truth-level sub-jets;
- 458 •  $m^{TAS}$  using truth-level tracks and truth-level sub-jets; the truth-level tracks are used to assist the  
459 truth-level sub-jets;
- 460 •  $m^{calo}$  truth, calculated using only the truth sub-jets.

### 461 5.7.1 Simple Sub-jet Calibration

462 The perfect calibration using truth level sub-jets and tracks is shown in Figure 11 in blue dots; since the  
463 performance exhibits room for big improvement below  $\sim 1$  TeV and moderate to small improvement above  
464 this value, the second step of a simple calibration was tried.

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

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

474 This procedure was called *poor man's calibration* or PM calibration or *simple calibration*. A check on  
 475 the  $p_T$  response before and after calibration together with the mean of the entire Large- $R$  jet response is  
 476 shown in Figure 44 and 45 in Appendix.

477 The results are on Figure 11; there are only marginal improvements in few ranges of low transverse  
 478 momentum where the scale factors are further away from unity, and the overall observable is not performing  
 479 better than the standard  $m^{TAS}$ . This is interpreted both in terms of a missing calibration as a function of  
 480 the  $\eta$  variables (having hence a befit from the crack region) and because the correction done on average  
 481 does not provide the sufficient handle in a jet-by-jet basis, especially when all the sub-jets are rescaled by  
 482 similar factors (which translates into a similarity of  $p_T$ s of the sub-jets, often the case for e.g. boosted  
 483  $W/Z$ , less for boosted tops entirely contained in the large- $R$  jet).

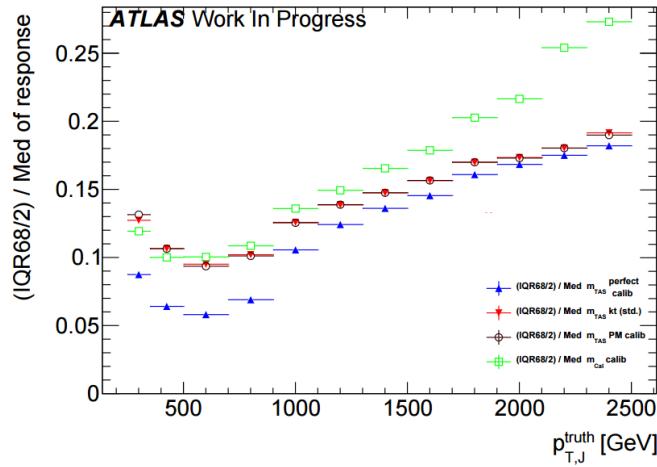


Figure 11: Performance of the poor man's calibration. The improvement is marginal throughout the entire transverse momentum space.

## 484 5.8 Limitation of $m^{TAS}$ from tracking

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

489 The preliminary investigation in this direction was then the study on the track mass resolution: a response  
 490 of the mass of the tracks associated to the jet ( $m^{track}$ ), was constructed, using the truth-level tracks.

- 491 The result is shown on Figure 12: for the samples considered, it shows a linear degradation of the mass of  
 492 the tracks associated to the jet ( $m^{track}$ ), both for massive and SM  $W/Z$ .

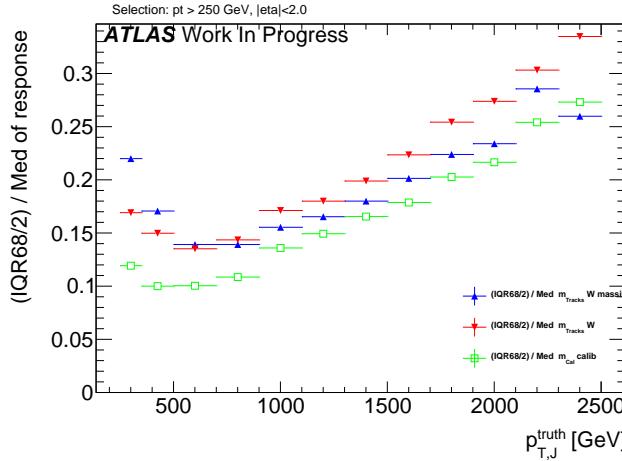


Figure 12: The performance of the track mass ( $m^{track}$ ) in blue and red for massive  $W$  sample and boosted  $W/Z$  respectively; for reference in green the calorimeter mass of the large- $R$  jet.

- 493 The hypothesis of the degradation of the  $m^{TAS}$  driven by the tracks is also supported by the Figure 46  
 494 in Appendix, where the truth-level tracks are used instead of real tracks to compute the variable; it can  
 495 be seen the flat behavior at high  $p_T$ , hence ascribing the worsening of the resolution to tracks at higher  
 496 transverse momenta.

497 A complete breakdown of the variable in terms of truth-level particles is given in Figure 13(a), where  
 498 all the different components are separated. In particular the black dots show the  $m^{TAS}$  using truth-level  
 499 sub-jets but real tracks for the track assistance procedure. Even combining this truth-level information, in  
 500 fact, it shows a large worsening of the performance (truth-level sub-jets only are shown as blue dots). On  
 501 the other side using again truth-level tracks for the track assistance procedure of the truth-level sub-jet,  
 502 shows a recovery of the loss in performance.

503 Additional studies on the limitation of the  $m^{TAS}$  based on MC studies without detector interactions are also  
 504 presented. In particular, the truth study presented for boosted  $W/Z$  decay in were extended for boosted  
 505 top quark decays.

506 As seen on Figure 13(b), the breakdown of the  $m^{TAS}$  shows that, in particular for the high transverse  
 507 momenta regimes, the tracks are subjected to fast degradation which makes their combination with the  
 508 calorimeter mass not anymore an advantage.

509 This is a limitation which was expected and understood from the detector performance point of view,  
 510 and here shows the impossibility, with the variables which are presented here  $m^{TA}$  and  $m^{TAS}$  to reach a  
 511 competitive standpoint with the  $m^{calo}$  in the extreme kinematic regime for the top quark decay.

512 In black, in fact, the performance of the  $m^{TAS}$  variable using tracks with detector effect and sub-jets  
 513 without those effects, shows this intrinsic limit which takes place already at 1.5 TeV.

514 The crossing point is, as already pointed out for the boosted tops, present because of the optimal perform-  
 515 ance of the calorimeter system caused by the higher mass of the top quark, and partially also because of  
 516 its more complex decay structure and difficulty to be resolved in sub-jets.

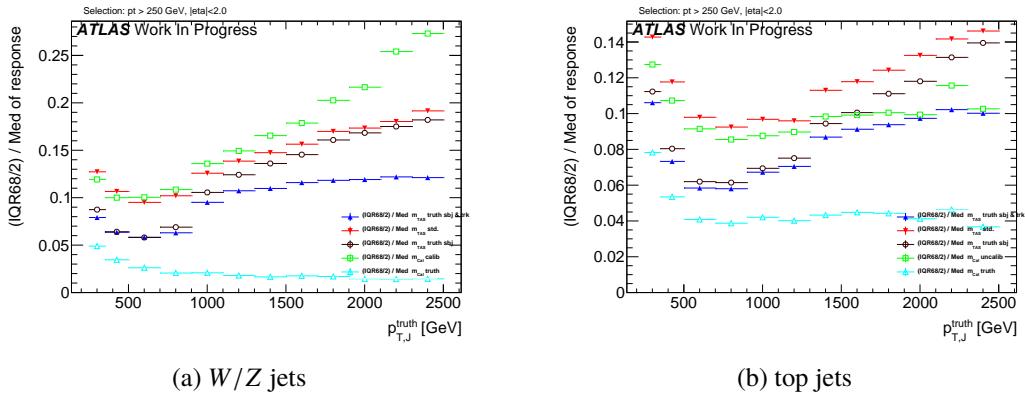


Figure 13: Breakdown of the  $m^{TAS}$  in its component using truth-level information for boosted  $W/Z$  decays, on the left. In blue the  $m^{TAS}$  using truth-level sub-jets and truth level tracks, in black  $m^{TAS}$  using truth level sub-jets but real tracks and in light blue for reference the mass of the truth level particles associated to the sub-jets. As usual, in red and green the standard  $m^{TAS}$  and the  $m^{calo}$ . On the right the same for top jets.

## 5.9 Performance with Alternate Inputs to the $m^{TAS}$

There are quite a few ways to modify the track-assisted sub-jet mass; however, all the alternative approaches showed worse performance, and they are mentioned here for completeness only. The per-track four momentum correction scheme which is used for the ECF and the n-Subjettiness and also explored with the  $m^{TAS}$  with no significant difference was described in 3.2.2.

The other alternatives considered were:

- for the tracks:
  - use of tracks not as input directly, but only taking those belonging to anti- $k_t$  reclustered track-jet with radius of 0.3 or 0.2;
  - tighter or looser quality conditions were explored;
  - tighter or looser primary vertex association requirement were explored.
- for the sub-jets:
  - the trimming procedure was modified: various radii  $R_{\text{sub}}$  of the sub-jets were tested;
  - the sub-jets were reclustered using not only the standard  $k_t$ , but also anti- $k_t$  and C/A.
- for the procedure: different 4-momentum correction scheme was also studied in more details, see 3.2.2.

The different reclustering algorithm choice has a deep impact and was studied in details, since it changes the topo-cluster added to the sub-jets and the tracks associated to them. The situation is depicted in the event-display in Figure 14; the display on the left shows the standard choice of  $k_t$ , the one on the right shows the modified approach anti- $k_t$ .

In Figure 15(a) 15(c) 15(b) the performance for boosted  $W/Z$ , tops and Higgs are shown, respectively. It can be seen that the  $k_t$  algorithm provides the best observable definition, in all the samples considered. However, the anti- $k_t$  algorithm provides similar performances; this was an important check as the jet

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

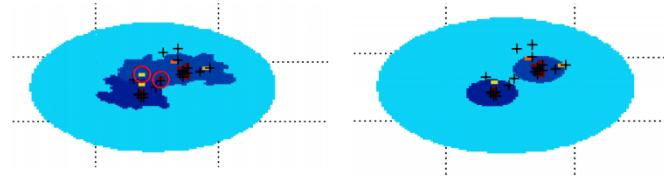


Figure 14: An example of event-display shows the differences in the reclustering algorithm used for the sub-jets: on the right  $k_t$ , and on the left anti- $k_t$ . Highlighted some constituents trimmed away with the second choice.

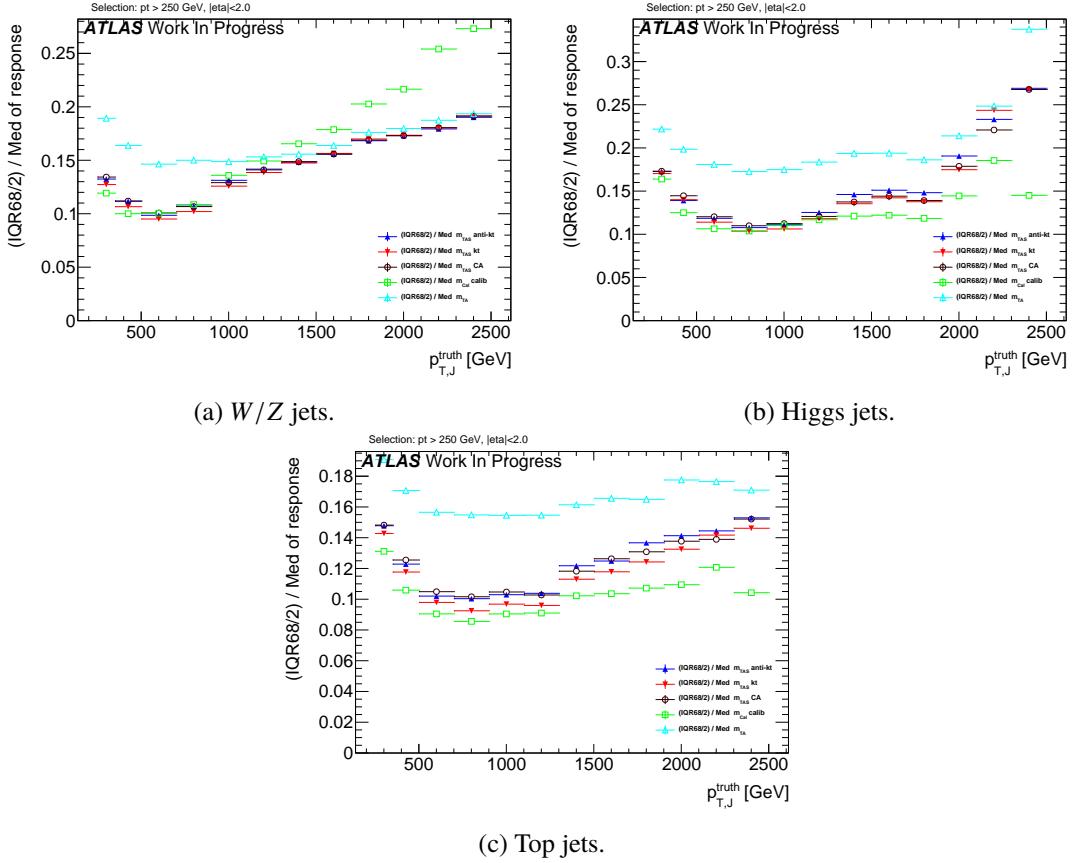


Figure 15: Performance of  $m^{TAS}$  with different reclustering algorithms for the sub-jets: anti- $k_t$ ,  $k_t$  and C/A and for  $W/Z$  jets, top left, Higgs jets, top right and top jets, bottom. In all the cases shown, the  $k_t$  is producing the better results, but all the three have a very similar performance.

## 543 6 Performance of Combined Calorimeter and Track-Assisted Sub-Jet 544 Mass

545 This section presents the achievement of the variable obtained combining the  $m^{TAS}$  and the  $m^{calo}$ , the  
546  $m_{TAS}^{comb}$  with respect to the combination of the  $m^{TA}$  and the  $m^{calo}$ , the  $m^{comb}$ . Both these variables were  
547 defined in 3.3

### 548 6.1 Performance in $W \rightarrow q' \bar{q}$ Decays

549 On the boosted  $W/Z$ s sample, the  $m_{TAS}^{comb}$  outperforms all the other definitions throughout all the transverse  
550 momentum space; on Figure 16(a) they are shown for reference together with the  $m^{TAS}$ . It can be noted  
551 here that the track-assisted sub-jet mass, although being sub-optimal, has comparable performance, yet  
552 presenting fewer complications due to the combination procedure.

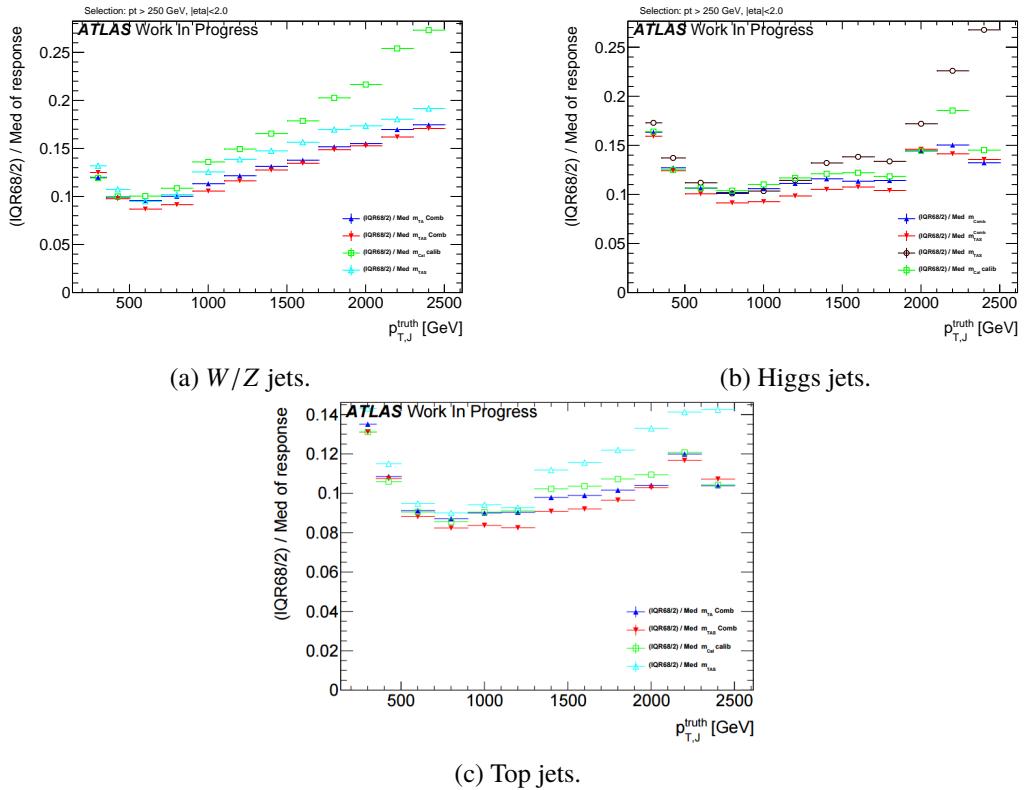


Figure 16: Performance of  $m^{comb}$  and  $m_{TAS}^{comb}$  for different samples: the  $W/Z$  jets, top left, the Higgs jets, top right and the top jets, bottom. The  $m_{TAS}^{comb}$  outperforms the other definitions throughout the whole spectrum of transverse momentum. The  $m^{TAS}$ , although being sub-optimal follows with similar performance the  $m^{comb}$ . The Higgs and top jets presents the same properties as shown before, and the combined mass reflects these properties.

553 6.2 Performance in  $h \rightarrow b\bar{b}$  Decays

Again, for the Higgs decay there are similarities as for the top sample; on Figure 16(b) the two definitions of the combined mass, together with the simpler  $m^{TAS}$ . Although this variable is slightly sub-optimal yet still comparable in the low to intermediate range in transverse momenta, where the tracks are driving a decrease in performance for the high to very-high  $p_T$ . The  $m_{TAS}^{comb}$  uses this advantage to achieve optimal behavior in the entire transverse momentum spectrum, outperforming both  $m^{calo}$  and  $m^{comb}$  almost everywhere.

## 560 6.3 Performance in $t \rightarrow q' \bar{q} b$ Decays

The boosted top sample remains the most challenging one also with the combined mass; as seen on Figure 16(c), the  $m^{comb}$  performs quite similarly to the calorimeter based mass definition, yet behaving considerably better than the  $m^{TAS}$  especially at high transverse momentum. The  $m_{TAS}^{comb}$ , however, outperforms all the other definitions, and shows its optimal observable strength at intermediate  $p_T$  i.e. in the range  $0.8 < p_T < 1.6$  TeV.

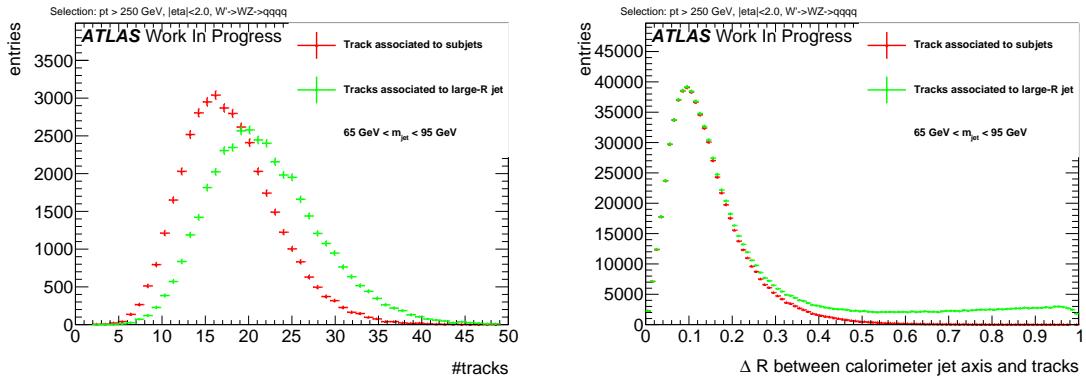


Figure 17: The number of tracks ghost associated to the large-R jet and to the sub-jets (left) and angular distance of associated tracks to the large-R calorimeter jet axis (right). Signal events were not reweighted at this step.

## 566 7 Performance of Jet Substructure Observables with (assisted) Tracks

567 This section shows the performance of (assisted) tracks compared to clusters as input for the substructure  
 568 observables C2/D2 and  $\tau_{21}$  for  $W$  and Higgs tagging as well as  $\tau_{32}$  for Top tagging. For now, the angular  
 569 weighting parameter  $\beta$  is set to the default value of 1.

### 570 7.1 Effect of Track Selection

571 Compared are tracks ghost associated to the un-groomed large-R jet and the collection used for  $m^{TAS}$ , see  
 572 Section 3.2; ghost association to  $k_T$ -subjets and  $\Delta R$  matching of tracks close to sub-jets.

573 The distributions showing the number of tracks associated to a calorimeter jet, see the left side of Figure  
 574 17, indicate, that on average around four tracks less are associated to the sub-jets compared to the un-  
 575 groomed jet. The right side of Figure 17 shows the angular distance  $\Delta R$  between the single tracks and  
 576 the axis of the large-R calorimeter jet. Both distributions are aligned in the lower  $\Delta R$  region while the  
 577 histogram representing the tracks associated to the un-groomed jet shows an enhancement towards larger  
 578  $\Delta R$ . Accordingly, the additional tracks feature an angular separation from the jet axis of more than 0.3.  
 579 Given the required primary vertex association, it is unlikely that these tracks originate from pile-up.  
 580 Instead, the origin might be found in final- or initial state radiation.

581 Figure 18 shows the signal distributions of the C2/D2, and  $\tau_{21}$ , calculated with both selections of tracks  
 582 for  $W$  boson jets. The large  $\Delta R$  to the jet axis of the differing tracks push the substructure variables to  
 583 higher, more background like values. The broader distributions are a result of the variating nature of these  
 584 tracks. C2 and D2 are more sensitive to tracks with a large  $\Delta R$  to the jet axis, because the angular distance  
 585 between all pairs and triples of tracks is considered, among tracks on possibly opposite sides of the large-R  
 586 jet. The distances to  $k_T$ -WTA axes used by  $\tau_{21}$  are more robust against these scenarios. For comparison,  
 587 the signal and background distributions for the variables calculated with calorimeter clusters are shown as  
 588 well. We expect that variables calculated with (assisted) tracks do not perform worse than cluster based  
 589 pendants. In contrast to the jet mass, ratios of  $ECF(N,\beta)$  and  $\tau_N$  are quite energy scale independent. They  
 590 are found to not be as sensitive to the missing neutral fraction with un-assisted tracks, justifying the study  
 591 of normal tracks next to TAS as input.

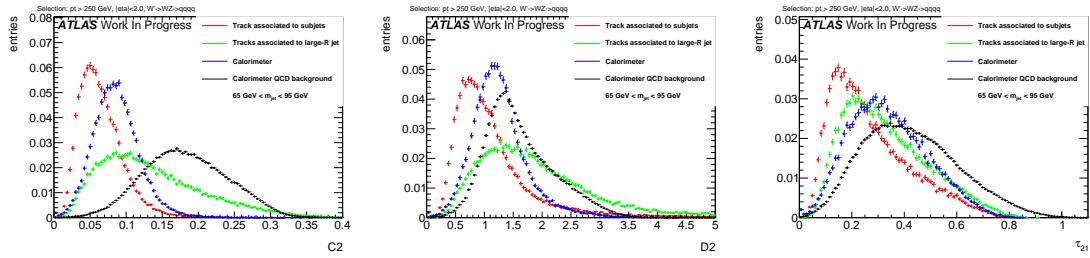


Figure 18: Substructure variables (left)  $C_2$ , (right)  $D_2$  and (below)  $\tau_{21}$  calculated with calorimeter clusters as well as tracks associated to sub-jets and to the large-R jet. Signal events were not reweighted at this step.

## 592 7.2 Performance with default $\beta$

593 The stated signal efficiencies are calculated after the mass cut plus tagging with n-Subjettiness or  $C_2/D_2$ .  
 594 Therefore, the endpoint of the ROCs lies at 68% signal efficiency, the fraction kept after the mass cut.  
 595 Consequently, it is required to achieve a tagging only signal efficiency of  $\frac{0.5}{0.68} \sim 0.74$  for a signal efficiency  
 596 of 0.5 after mass cut and tagging. Similarly, the stated and compared background rejections result from the  
 597 multiplication of both, thus representing the QCD rejection of the combined mass- and tagging variable-  
 598 cut. The complete set of signal and background distributions for different inputs can be found in the  
 599 Appendix N.

### 600 7.2.1 Performance for $W$ Tagging

601 Shown in Figure 19 are exemplary signal and background distributions in intermediate  $p_T$  bins compared  
 602 for different inputs. This shows throughout narrower signal distributions rising slightly sharper for tracks  
 603 and assisted tracks compared to calorimeter clusters which can be explained by the high angular resolution.  
 604 The right handed tails of the signal distributions are similar to the calorimeter variables. Similarly, the  
 605 background distributions shift as well, but not as distinct as seen for the signal. The  $p_T$  dependence of the  
 606 substructure variables can be found in Appendix L.

607 The ROCs in Figure 20, 21 and 22 show the actual achieved background rejection at different  $p_T$  values.  
 608 For lower  $p_T$  values, TAS perform comparably to calorimeter clusters. Tracks without assisting achieve  
 609 a considerably lower background rejection with  $D_2$  and  $\tau_{21}$  for lower energies. Tracks and TAS perform  
 610 equally well at high energies for  $D_2$  and  $\tau_{21}$  and for  $C_2$  over the whole studied range. At higher boosts,  
 611 the angular resolution of the tracks becomes more and more relevant as the separation between jet  
 612 constituents shrinks. Consequently, tracks and TAS start to outperform calorimeter based variables and  
 613 become increasingly effective with rising energy.

### 614 7.2.2 Tracks and TAS at very high $p_T$

615 The  $C_2$  variable was found to perform equally well with tracks and TAS as input. This variable seems to  
 616 be relative insensitive to the track assisting and tracks alone already perform well.  $D_2$  and  $\tau_{21}$  in contrast,  
 617 feature a visibly worse separation with tracks than with assisted tracks. In these cases, the scale difference  
 618 due to the missing neutral fraction seems to have a greater influence.

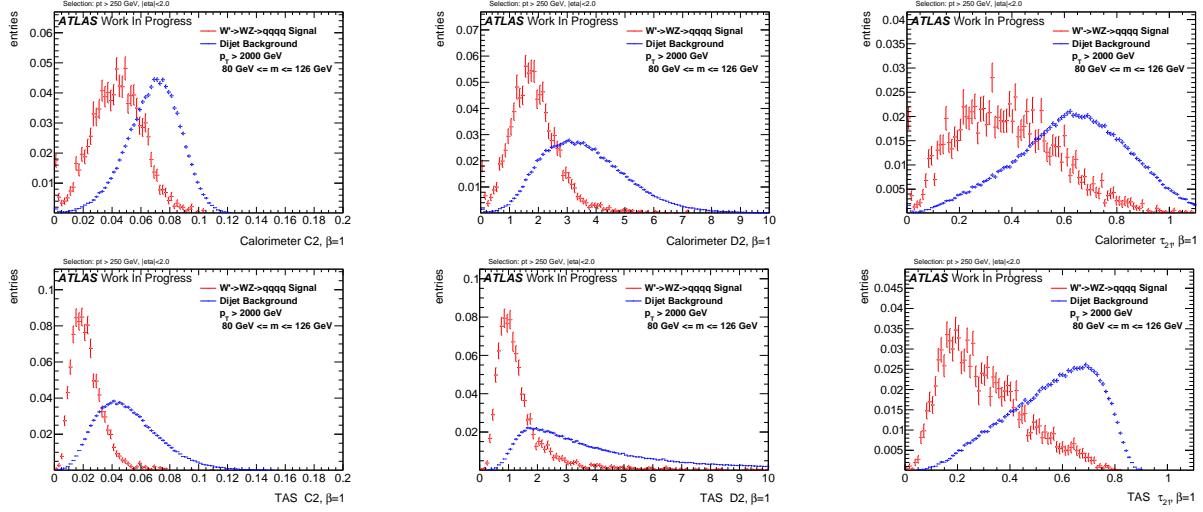


Figure 19:  $W$  boson signal and QCD background distributions for calorimeter (left) and TAS (right) at  $\beta = 1$  with C2 (top) and D2 (middle) and  $\tau_{21}$  (bottom) for  $>2000$  GeV

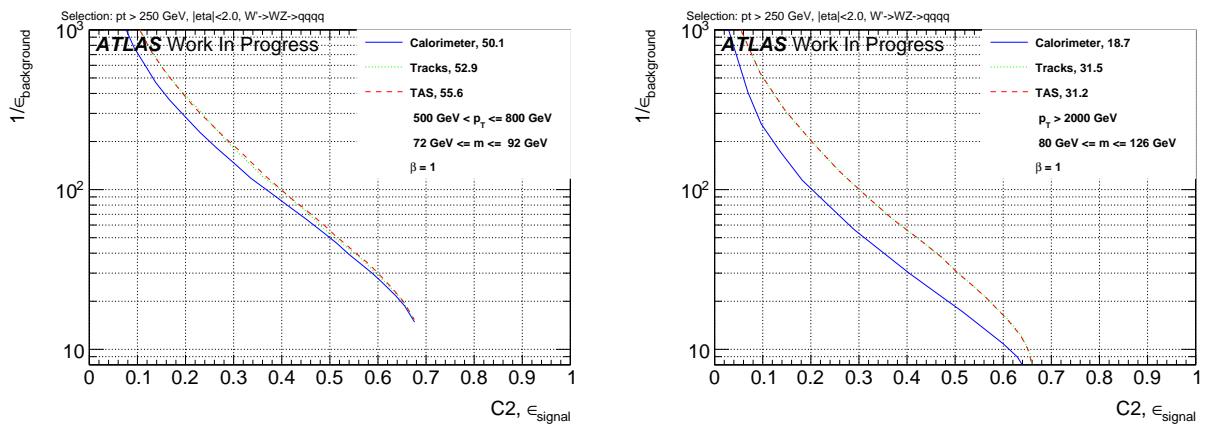


Figure 20: ROCs showing QCD rejection against  $W$  boson efficiency for tracks, TAS and colorimeter C2 at  $\beta = 1$  for 500-800 GeV (left) and  $>2000$  GeV (right). The numbers in the legend indicate the achieved background rejection at 50% signal efficiency.

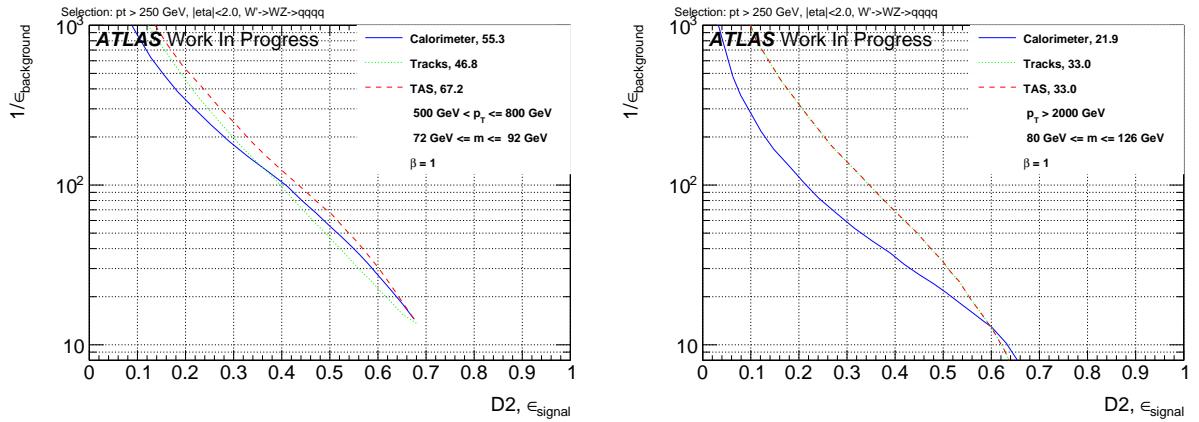


Figure 21: ROCs showing QCD rejection against  $W$  boson efficiency for tracks, TAS and calorimeter D2 at  $\beta = 1$  for 500-800 GeV (left) and  $>2000$  GeV (right). The numbers in the legend indicate the achieved background rejection at 50% signal efficiency.

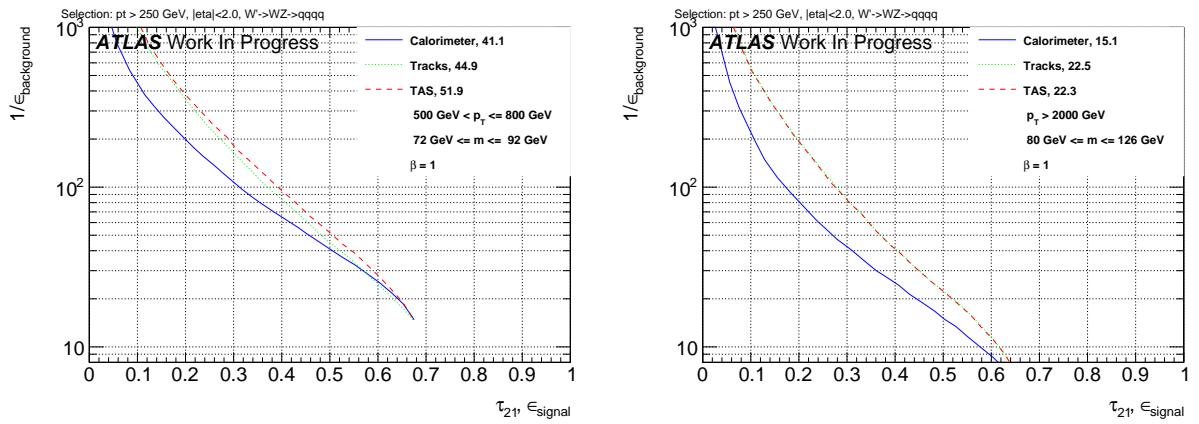


Figure 22: ROCs showing QCD rejection against  $W$  boson efficiency for tracks, TAS and colorimeter  $\tau_{21}$  at  $\beta = 1$  for 500-800 GeV (left) and  $>2000$  GeV (right). The numbers in the legend indicate the achieved background rejection at 50% signal efficiency.

619 For very high  $p_T$  values however, it is often the case that the large-R calorimeter jet features only one  
 620  $R = 0.2$  sub-jet after trimming due to the now small separation of constituents. A single sub-jet results in  
 621 the TAS procedure to fall back to TA. As stated in Section 3.4, TA has no impact on the ratios. Therefore,  
 622 C2/D2 and  $\tau_{21}$  perform equally well when calculated with tracks or TAS for events with only one sub-jet  
 623 and thereby the difference between both decreases for very high energies.

### 624 7.2.3 Performance for Higgs Tagging

625 The Higgs boson is heavier than the  $W$  or  $Z$  boson, resulting in a higher angular separation of the jet  
 626 constituents considering the rule of thumb  $\delta R \sim \frac{2m}{p_T}$  for decay products. As a result, angular resolution  
 627 effects won't have the same impact as for the  $W$  boson. This can be verified by the performance of  
 628 track-based variables in the ROCs found in Figure 23.

629 For Higgs boson tagging and an angular weight of  $\beta = 1$ , found were no distinct improvements with TAS  
 630 or tracks compared to calorimeter clusters. The C2 variable performs better with calorimeter clusters,  
 631 D2 yields an equal QCD discrimination with TAS and calorimeter clusters. The n-Subjettiness ratio  $\tau_{21}$

632 benefits from TAS in some  $p_T$  regions, while the calorimeter pendant performs better in the other regions. Furthermore, tracks and TAS perform comparable over the whole studied  $p_T$  range.

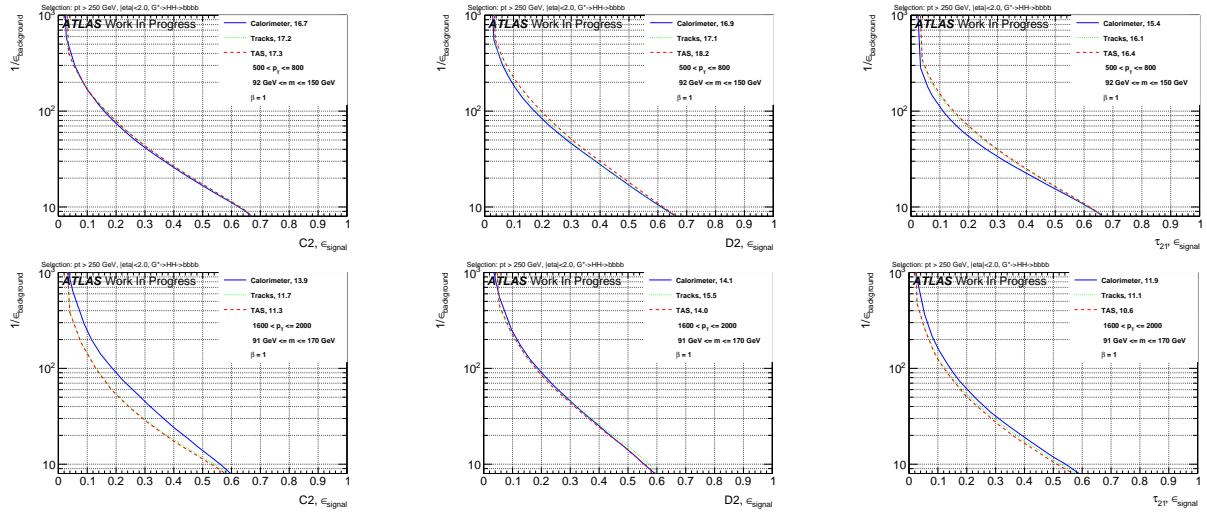


Figure 23: ROCs showing QCD rejection against Higgs boson efficiency for tracks, TAS and colorimeter. C2 (left), D2 (middle) and  $\tau_{21}$  (right) at  $\beta = 1$ . Shown is the energy range between 500-800 GeV (top) and 1600-2000 GeV (bottom). The numbers in the legend second  $p_T$  bin (left) and highest bin (right). The numbers in the legend indicate the achieved background rejection at 50% signal efficiency.

633

#### 634 7.2.4 Performance for Top Tagging

635 The top quark features a characteristic three body decay and a very high mass around 173 GeV. Studied  
636 here is the n-Subjettiness ratio  $\tau_{32}$  to distinguish the three prong like top quark jets and QCD background  
637 jets.

638 The ROCs in Figure 24 show the accompanying improvements in the separation power of  $\tau_{32}$  possible  
639 with TAS. Tagging tops quark events with  $\tau_{32}$  is found to greatly benefit from the excellent angular  
640 resolution of tracks. This is especially the case for high  $p_T$  where the limitation of the calorimeter cell  
641 size clearly diminishes the possible identification of three distinct substructures inside a large-R jet. The  
642 enhancements are not as articulated for the low  $p_T$  regions, nevertheless TAS  $\tau_{32}$  performs here at least  
643 equally well as calorimeter  $\tau_{32}$ . Furthermore, tracks are observed to perform slightly worse in comparison  
644 with TAS for the lower  $p_T$  regions, but match the TAS performance for very large boosts as expected.

## 645 8 Performance of Jet Substructure Observables with (assisted) Tracks and 646 optimised $\beta$

647 The observed background rejection of variables calculated with TAS and tracks is at least as high as of  
648 calorimeter cluster based variables due to the high angular resolution of tracks. Therefore, studied are the  
649 effects of a higher weighting of the angular part of the substructure variables. For completeness, considered

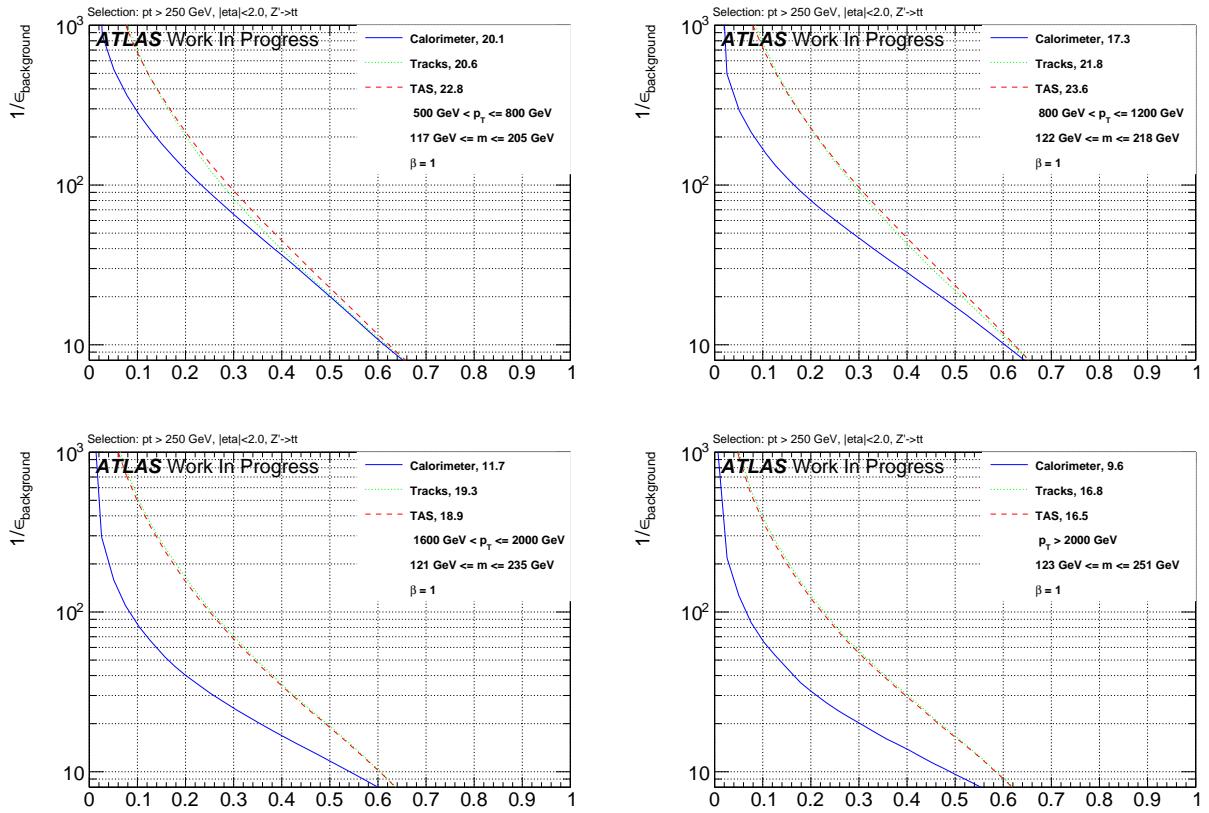


Figure 24: ROCs showing QCD rejection against top quark efficiency for tracks, TAS and calorimeter  $\tau_{32}$  at  $\beta = 1$ ,  $p_T$  ordering from upper left to lower right.

as well as a lower weighting. Previous studies of default calorimeter variables for  $W$  boson tagging, see e.g. Reference [bib:w\_tagging], found  $\beta = 1$  to maximize the separation power of calorimeter variables. A scan over the values  $\beta = 0.5, 1, 1.7, 2, 3$  is performed in order to identify the best variables for the specific scenarios of tagging  $W$  boson, Higgs boson or top quark jets. The background rejections, achieved at the 50 % working point after mass cut and tagging are summarized in tables. The corresponding ROCs can be found at the end of this section 8.1. Pseudo-experiments were used to propagate the uncertainties on the signal and background distributions due to the finite size of the MC samples to the background rejections.

### 8.0.1 Optimization for $W$ Tagging

The results of the optimization for  $W$  jets can be found in Appendix M.1. As expected, tracks and TAS perform visibly worse with a low angular weighting. For higher values of  $\beta$ , tracks and TAS gain in separation power, verifying the significance of the angular part for track based variables. Nevertheless, the separation is observed to degrade for angular weightings too high compared to the  $p_T$  part, here  $\beta = 3$ .

A  $\beta$  of around 2 maximizes the separation power of tracks and TAS. The advantages of  $\beta = 2$  compared to  $\beta = 1$  are found at higher  $p_T$  values, minor losses are visible in the lowest energy regions. A slightly lower

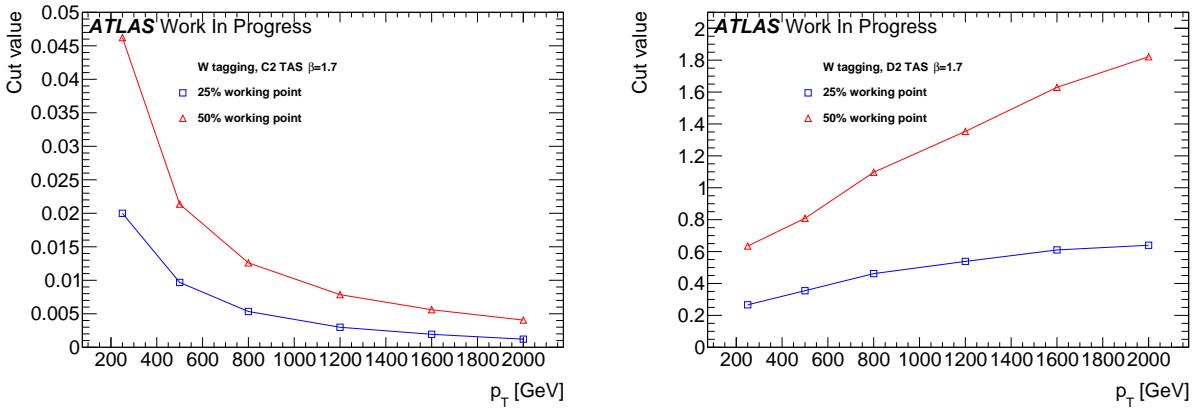


Figure 25: Cut values for  $C2_{TAS}^{(\beta=1.7)}$  (left) and  $D2_{TAS}^{(\beta=1.7)}$  (right) to achieve 50 % and 25 %  $W$  boson efficiency.

value of  $\beta = 1.7$  was able to retain the great background rejection of a large angular weighting at high  $p_T$  while still performing well at lower energies. Variables calculated with clusters are not as sensitive to a variation of the angular weighting.

$C2$  is found to be relatively insensitive to the track assisting, whereas separation with  $\tau_{21}$  and  $D2$  (for lower energies) is visibly poorer with tracks compared to TAS. Starting around 1200 GeV, tracks perform comparably and for even higher energies equally well as TAS due to the rising amount of large-R jets with only one sub-jet. Comparing variables independently for the different inputs,  $\tau_{21}$  performs worse than  $C2$  and  $D2$ . However, e.g.  $\tau_{21}$ , TAS can outperform  $C2_{calo}$  and  $D2_{calo}^{(β=1)}$ .

The variables achieving the highest background rejections for  $W$  boson tagging are  $D2_{TAS}^{(\beta=1.7)}$  and  $C2_{TAS}^{(\beta=1.7)}$ , depending on the energy.  $D2$  cut values were shown to be more  $p_T$  robust,  $C2$  could also be used with tracks instead of TAS, which saves the assistance. For TAS variables, uncertainties on the sub-jets, whose derivation is being worked at, need to be propagated while tracks feature already well-known uncertainties.

Shown in Figure 25 are the cut values for 50 % and 25 % signal efficiency for  $D2_{TAS}^{(\beta=1.7)}$  and  $C2_{TAS}^{(\beta=1.7)}$ . As for the default cluster variables, the  $D2_{TAS}^{(\beta=1.7)}$  cut is more  $p_T$  robust than the cut on  $C2_{TAS}^{(\beta=1.7)}$ .

Table 3 lists the background rejections for  $D2_{TAS}^{(\beta=1.7)}$ ,  $C2_{TAS}^{(\beta=1.7)}$  and the currently used  $D2_{calo}^{(β=1)}$  along with the corresponding improvements. For lower energies,  $D2_{TAS}^{(\beta=1.7)}$  is the best choice. For very high boosts of the  $W$  boson,  $C2_{TAS}^{(\beta=1.7)}$  performs superior, especially for 25 %  $\epsilon_{signal}$ , where the background rejection with  $C2_{TAS}^{(\beta=1.7)}$  is around 3.5 times as large as the QCD rejection with  $D2_{calo}^{(β=1)}$ . These enormous improvements at lower  $\epsilon_{signal}$  are due to the signal distributions for TAS and tracks rising much steeper than for clusters. The tail to higher, background like values in contrast, is more comparable, leading to an alignment of the background rejection for very large  $\epsilon_{signal}$ .

### 8.0.2 Optimization for Higgs Tagging

The results of the optimization for Higgs jets can be found in Appendix M.2. The study of  $\beta = 1$  in the Higgs boson case, see section 7.2.3, showed no improvements in the rejection of QCD events due to tracks

<b>50% <math>\epsilon_{signal}</math></b>	<b>W tagging</b>		
$p_T$ [GeV]	D2 <sub>calo</sub> <sup>(<math>\beta=1</math>)</sup>	D2 <sub>TAS</sub> <sup>(<math>\beta=1.7</math>)</sup>	C2 <sub>TAS</sub> <sup>(<math>\beta=1.7</math>)</sup>
250 - 500	$35.0 \pm 2.0$	$35.4 \pm 2.3 (+1 \pm 9\%)$	$28.9 \pm 1.5 (-17 \pm 6\%)$
500 - 800	$55.3 \pm 2.6$	$67.6 \pm 3.2 (+22 \pm 8\%)$	$58.6 \pm 2.6 (+6 \pm 7\%)$
800 - 1200	$41.1 \pm 2.0$	$54.9 \pm 2.4 (+34 \pm 9\%)$	$54.6 \pm 2.8 (+33 \pm 9\%)$
1200 - 1600	$38.1 \pm 1.9$	$50.8 \pm 1.8 (+33 \pm 8\%)$	$53.8 \pm 2.7 (+41 \pm 10\%)$
1600 - 2000	$25.4 \pm 1.3$	$37.8 \pm 2.0 (+49 \pm 11\%)$	$50.9 \pm 4.3 (+100 \pm 20\%)$
> 2000	$21.9 \pm 1.7$	$36.3 \pm 2.0 (+66 \pm 16\%)$	$46.1 \pm 4.7 (+111 \pm 27\%)$

<b>25% <math>\epsilon_{signal}</math></b>	<b>W tagging</b>		
$p_T$ [GeV]	D2 <sub>calo</sub> <sup>(<math>\beta=1</math>)</sup>	D2 <sub>TAS</sub> <sup>(<math>\beta=1.7</math>)</sup>	C2 <sub>TAS</sub> <sup>(<math>\beta=1.7</math>)</sup>
250 - 500	$139.6 \pm 9.8$	$146.0 \pm 12.4 (+5 \pm 12\%)$	$108.2 \pm 7.5 (-22 \pm 8\%)$
500 - 800	$243.7 \pm 13.2$	$360.1 \pm 21.1 (+48 \pm 12\%)$	$298.4 \pm 15.9 (+22 \pm 9\%)$
800 - 1200	$181.0 \pm 8.8$	$308.5 \pm 19.3 (+70 \pm 14\%)$	$313.2 \pm 24.4 (+78 \pm 16\%)$
1200 - 1600	$156.9 \pm 8.3$	$295.4 \pm 17.8 (+88 \pm 15\%)$	$354.6 \pm 25.6 (+126 \pm 20\%)$
1600 - 2000	$84.6 \pm 5.7$	$219.6 \pm 10.9 (+160 \pm 22\%)$	$320.5 \pm 31.4 (+279 \pm 45\%)$
> 2000	$78.9 \pm 7.6$	$233.5 \pm 14.7 (+196 \pm 34\%)$	$288.4 \pm 33.3 (+266 \pm 55\%)$

Table 3: Listing of the background rejections after the jet mass cut and tagging at 50% and 25%  $W$  boson efficiency for the identified best variables D2<sub>TAS</sub><sup>( $\beta=1.7$ )</sup> & C2<sub>TAS</sub><sup>( $\beta=1.7$ )</sup> together with the improvements over the standard choice D2<sub>calo</sub><sup>( $\beta=1$ )</sup>. Highlighted in red is the best variable in each studied energy range.

and TAS as input. As for the  $W$  boson, the performance of tracks and TAS diminishes considerably with an angular weighting of  $\beta = 0.5$ .

No improvement of  $\tau_{21}$  is observed with tracks or TAS, clusters perform equally well for lower  $p_T$  and slightly better at high energies. Again, the QCD rejection achieved with  $\tau_{21}$  is exceeded by C2 and D2. The discrimination with clusters profits from a slightly higher angular weighting, although the gain is not as significant as for tracks and TAS. This consistently shows the lower sensitivity to a variation of the angular weight. The small gain is connected to the higher separation of the Higgs decay products compared to the  $W$  boson case.

For boosted Higgs tagging, D2 outperforms C2 over the whole studied energy range. Values of  $\beta = 1.7 \& 2$  yield the highest background rejection for track and TAS based D2. D2<sub>TAS</sub><sup>( $\beta=1.7, 2$ )</sup> and D2<sub>track</sub><sup>( $\beta=1.7, 2$ )</sup> perform superior to D2<sub>calo</sub> at high boosts, due to the low angular separation of constituents, and equally well at lower energies.

The differences between  $\beta = 1.7$  and  $\beta = 2$  are inconclusive with minor advantages at high and slight inferiorities at low  $p_T$  for  $\beta = 2$ . Tracks perform slightly worse than TAS for lower energies but similarly better in the two highest studied  $p_T$  regions. Chosen for further examination are D2<sub>TAS</sub><sup>( $\beta=1.7$ )</sup> and D2<sub>track</sub><sup>( $\beta=1.7$ )</sup>.

Shown in Figure 26 are the cut values for 50 % and 25 % signal efficiency for D2<sub>TAS</sub><sup>( $\beta=1.7$ )</sup> and D2<sub>track</sub><sup>( $\beta=1.7$ )</sup>. The cut value shows a slight upward trend for rising  $p_T$ . Moreover, cut values for the first bin are higher as for the second, in contrast to the overall upward trend of D2. This is the result of the low boost in the lowest  $p_T$  region resulting in a left shoulder of the mass distributions representing large-R jets containing only part of the Higgs boson decay. These jets feature one-prong structure and result in background-like

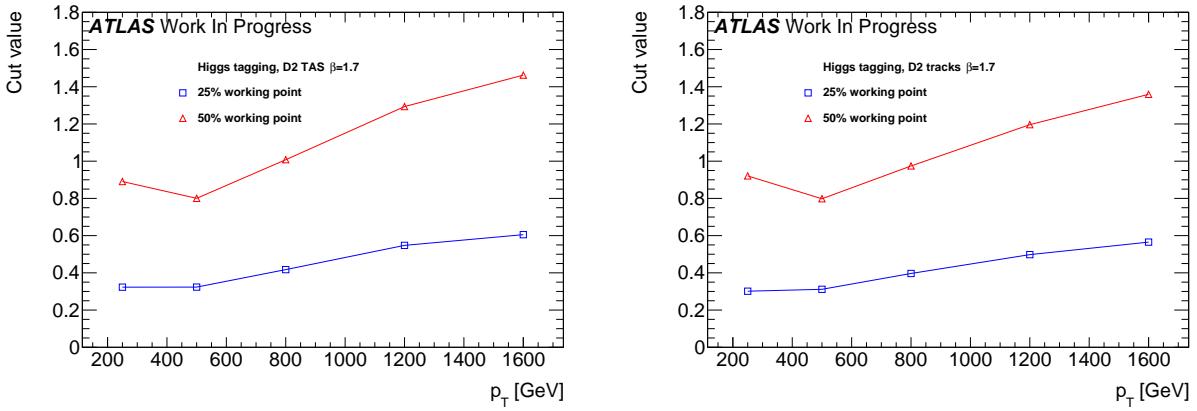


Figure 26: Cut values for  $D2_{\text{TAS}}^{(\beta=1.7)}$  (left) and  $D2_{\text{track}}^{(\beta=1.7)}$  (right) to achieve 50% and 25% Higgs boson efficiency.

50% $\epsilon_{\text{signal}}$		Higgs tagging		
$p_T$ [GeV]	$D2_{\text{calo}}^{(\beta=1)}$	$D2_{\text{TAS}}^{(\beta=1.7)}$	$D2_{\text{track}}^{(\beta=1.7)}$	
250 - 500	$8.4 \pm 0.2$	$8.5 \pm 0.2 (+1 \pm 4\%)$	$8.3 \pm 0.2 (-1 \pm 3\%)$	
500 - 800	$17.7 \pm 0.4$	$18.7 \pm 0.4 (+6 \pm 3\%)$	$17.9 \pm 0.4 (+1 \pm 3\%)$	
800 - 1200	$26.3 \pm 0.6$	$28.5 \pm 0.7 (+8 \pm 4\%)$	$28.0 \pm 0.7 (+6 \pm 4\%)$	
1200 - 1600	$27.0 \pm 0.8$	$28.9 \pm 0.7 (+7 \pm 4\%)$	$30.0 \pm 0.8 (+11 \pm 4\%)$	
1600 - 2000	$14.9 \pm 0.8$	$17.7 \pm 0.8 (+19 \pm 8\%)$	$18.5 \pm 0.8 (+24 \pm 9\%)$	

25% $\epsilon_{\text{signal}}$		Higgs tagging		
$p_T$ [GeV]	$D2_{\text{calo}}^{(\beta=1)}$	$D2_{\text{TAS}}^{(\beta=1.7)}$	$D2_{\text{track}}^{(\beta=1.7)}$	
250 - 500	$25.1 \pm 0.6$	$28.9 \pm 0.7 (+15 \pm 4\%)$	$30.5 \pm 0.8 (+22 \pm 4\%)$	
500 - 800	$54.1 \pm 1.4$	$69.6 \pm 1.9 (+29 \pm 5\%)$	$64.9 \pm 1.8 (+20 \pm 5\%)$	
800 - 1200	$90.8 \pm 2.5$	$121.3 \pm 3.4 (+34 \pm 5\%)$	$117.9 \pm 3.2 (+30 \pm 5\%)$	
1200 - 1600	$97.6 \pm 3.1$	$117.7 \pm 3.8 (+21 \pm 5\%)$	$122.4 \pm 4.2 (+25 \pm 6\%)$	
1600 - 2000	$54.6 \pm 3.5$	$74.0 \pm 5.7 (+36 \pm 14\%)$	$75.0 \pm 5.1 (+37 \pm 13\%)$	

Table 4: Listing of the background rejections after the jet mass cut and tagging at 50% and 25% Higgs signal efficiency for the identified best variables  $D2_{\text{TAS}}^{(\beta=1.7)}$  &  $D2_{\text{track}}^{(\beta=1.7)}$  together with the improvements over the best variable with clusters,  $D2_{\text{calo}}^{(\beta=1)}$ . Highlighted in red is the best variable in each studied energy range.

712 D2 values. The TAS D2 cut is marginally higher than the corresponding track D2 cut since the assisted  
 713 tracks have a higher  $p_T$  and the D2 cut features a rising tendency with  $p_T$ .

714 Listed in Table 4 are the background rejections for  $D2_{\text{TAS}}^{(\beta=1.7)}$ ,  $D2_{\text{track}}^{(\beta=1.7)}$ , and for the best calorimeter  
 715 variable, which is  $D2_{\text{calo}}^{(\beta=1)}$ , with the corresponding improvements due to the use of TAS respectively  
 716 tracks instead of clusters. At very high energies, the angle between the  $b\bar{b}$  pair is small despite the high  
 717 Higgs boson mass and the effect of the calorimeter cell size becomes significant.

### 718 8.0.3 Optimization for Top Tagging

719 The results of the optimization for Top jets can be found in Appendix M.3. Studied was  $\tau_{32}$  with values  
 720 of  $\beta \geq 1$ , since the  $W$  boson and Higgs boson parts affirmed the expected lower performance of track  
 721 and TAS based variables with an angular weighting of  $\beta \leq 1$ . The calorimeter  $\tau_{32}$  variable profits from a  
 722 higher angular weighting up to around  $\beta = 2$ , but degrades in performance for  $\beta = 3$ . Since the involved  
 723 three prong structure of the top quark decay requires a good angular separation of the jet constituents to  
 724 be resolved, tracks and TAS perform superior to clusters. A higher angular weighting does not improve  
 725 the separation power of track and TAS variables,  $\beta = 2$  already diminishes the performance. The best  
 726 discrimination is achieved with TAS and  $\beta = 1, 1.7$ . The marginal differences between both values of  
 727  $\beta$  depend on the considered  $p_T$  region. Track  $\tau_{32, \text{track}}$  achieves lower separation as  $\tau_{32, \text{TAS}}$ , except for  
 728 regions with very high boosts, but as well outperforms the cluster variable.

729 Shown in Figure 27 are the cut values for 50 % and 25 % signal efficiency for  $\tau_{32, \text{TAS}}^{(\beta=1.7)}$  and  $\tau_{32, \text{TAS}}^{(\beta=1.7)}$ . The  
 730 crack between the first and second  $p_T$  bin is more evident since the top quark with its much higher mass is  
 731 here very unlikely to be reconstructed into a single large-R jet, resulting in background like signal events.  
 732 Furthermore,  $\tau_{32} (\beta = 1.7)$  needs to be cut at lower values as  $\tau_{32} (\beta = 1)$  to achieve a certain signal  
 733 efficiency. This is the result of the higher angular weighting that shifts the overall distributions to lower  
 734 values, because the angular distance between two constituents inside a (highly) boosted large-R jet is in  
 735 the majority of cases lower than one. Thus, the angular part of  $\tau_{32}$  decreases with  $\beta > 1$ . The TAS  $\tau_{32}$  cut  
 736 value is observed to be robust against variations of  $p_T$ , in accordance to the results of the  $p_T$  correlation  
 plots, see 271.

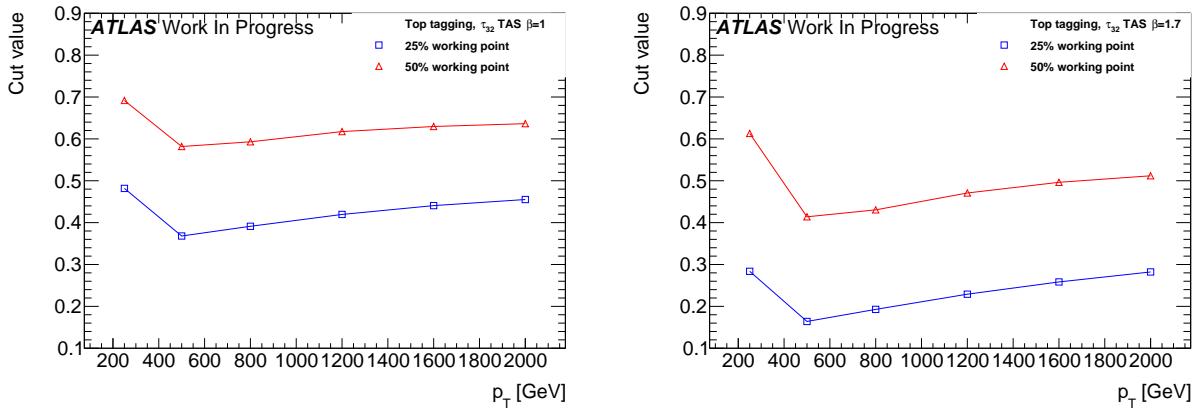


Figure 27: Cut values for  $\tau_{32, \text{TAS}}^{(\beta=1)}$  (left) and  $\tau_{32, \text{TAS}}^{(\beta=1.7)}$  (right) to achieve 50% and 25% Top quark efficiency

737  
 738 Listed in Table 5 are the background rejections for  $\tau_{32, \text{TAS}}^{(\beta=1)}$ ,  $\tau_{32, \text{TAS}}^{(\beta=1.7)}$  and the best cluster based variable,  
 739  $\tau_{32, \text{calo}}^{(\beta=2)}$ . The differences between both values of  $\beta$  with TAS are marginal, as well for lower signal  
 740 efficiencies. Improvements due to the use of TAS instead of clusters are possible for Top quark tagging  
 741 over the whole studied  $p_T$  range.

## 742 8.1 Summary of the Results

743 The following ROCs show the observables with the highest QCD rejection with the cluster-based default  
 744 variables for  $W$  8.1.1, Higgs 8.1.2 and Top 8.1.3 tagging for all studied energy ranges. (Assisted) tracks

<b>50% <math>\epsilon_{signal}</math></b>	<b>Top Tagging</b>		
$p_T$ [GeV]	D2 <sub>calo</sub> <sup>(<math>\beta=2</math>)</sup>	$\tau_{32, TAS}^{(\beta=1)}$	$\tau_{32, TAS}^{(\beta=1.7)}$
250 - 500	$9.5 \pm 0.2$	$10.7 \pm 0.2 (+13 \pm 3 \%)$	$10.1 \pm 0.2 (+6 \pm 3 \%)$
500 - 800	$22.4 \pm 0.6$	$22.8 \pm 0.6 (+2 \pm 4 \%)$	$22.8 \pm 0.6 (+2 \pm 4 \%)$
800 - 1200	$20.6 \pm 0.5$	$23.6 \pm 0.6 (+15 \pm 4 \%)$	$24.1 \pm 0.6 (+17 \pm 4 \%)$
1200 - 1600	$16.6 \pm 0.4$	$22.0 \pm 0.6 (+33 \pm 5 \%)$	$22.3 \pm 0.6 (+34 \pm 5 \%)$
1600 - 2000	$13.3 \pm 0.4$	$18.9 \pm 0.6 (+42 \pm 6 \%)$	$18.8 \pm 0.6 (+41 \pm 6 \%)$
> 2000	$10.9 \pm 0.4$	$16.5 \pm 0.7 (+51 \pm 8 \%)$	$15.7 \pm 0.7 (+44 \pm 8 \%)$

<b>25% <math>\epsilon_{signal}</math></b>	<b>Top Tagging</b>		
$p_T$ [GeV]	D2 <sub>calo</sub> <sup>(<math>\beta=2</math>)</sup>	$\tau_{32, TAS}^{(\beta=1)}$	$\tau_{32, TAS}^{(\beta=1.7)}$
250 - 500	$33.7 \pm 1.0$	$37.6 \pm 1.4 (+12 \pm 5 \%)$	$36.7 \pm 1.2 (+9 \pm 5 \%)$
500 - 800	$114.7 \pm 3.3$	$138.0 \pm 4.3 (+20 \pm 5 \%)$	$139.1 \pm 4.2 (+21 \pm 5 \%)$
800 - 1200	$97.0 \pm 2.7$	$144.6 \pm 4.9 (+49 \pm 7 \%)$	$149.6 \pm 5.2 (+54 \pm 7 \%)$
1200 - 1600	$68.6 \pm 2.1$	$133.2 \pm 4.6 (+94 \pm 9 \%)$	$134.7 \pm 5.1 (+96 \pm 10 \%)$
1600 - 2000	$47.5 \pm 1.6$	$100.3 \pm 4.2 (+111 \pm 11 \%)$	$99.9 \pm 4.4 (+110 \pm 12 \%)$
> 2000	$36.3 \pm 1.6$	$80.2 \pm 5.0 (+121 \pm 17 \%)$	$75.5 \pm 4.9 (+108 \pm 16 \%)$

Table 5: Listing of the background rejections after the jet mass cut and tagging at 50% and 25% top signal efficiency for the identified best variables  $\tau_{32, TAS}^{(\beta=1,1.7)}$  together with the improvements over the best variable with clusters, D2<sub>calo</sub><sup>( $\beta=2$ )</sup>. Highlighted in red is the best variable in each studied energy range.

745 outperform clusters and input for substructure observables in almost every studied case due to their  
 746 angular resolution. For tagging only slightly boosted Higgs bosons, they perform as good as clusters.  
 747 These improvements are rising with energy since the jet constituents angular separation shrinks.

### 748 8.1.1 W boson Tagging

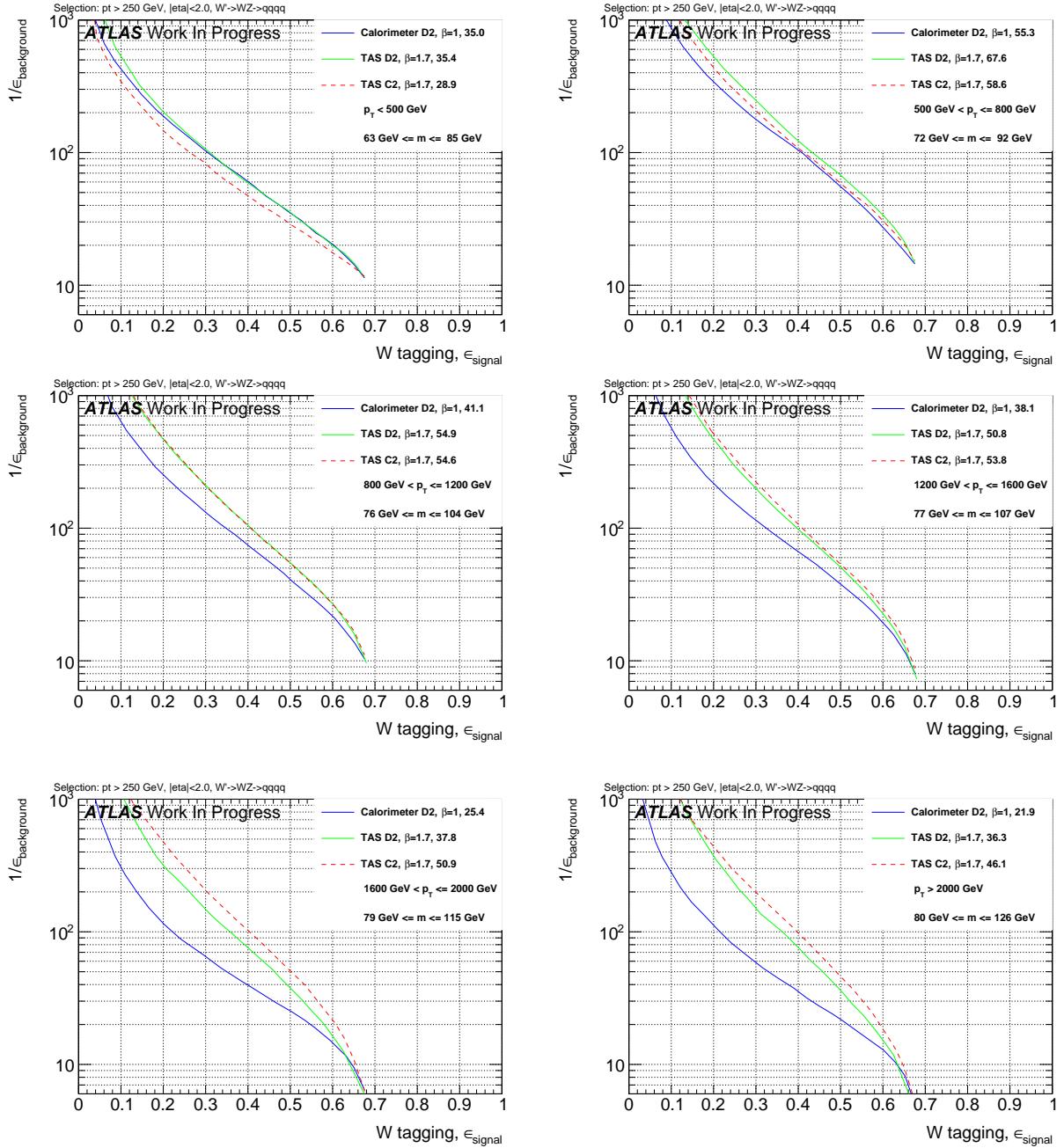


Figure 28: ROCs showing QCD rejection against  $W$  jet efficiency for  $D2_{\text{TAS}}^{(\beta=1.7)}$  &  $C2_{\text{TAS}}^{(\beta=1.7)}$  compared to  $D2_{\text{calo}}^{(\beta=1)}$ .

### 749 8.1.2 Higgs Boson Tagging

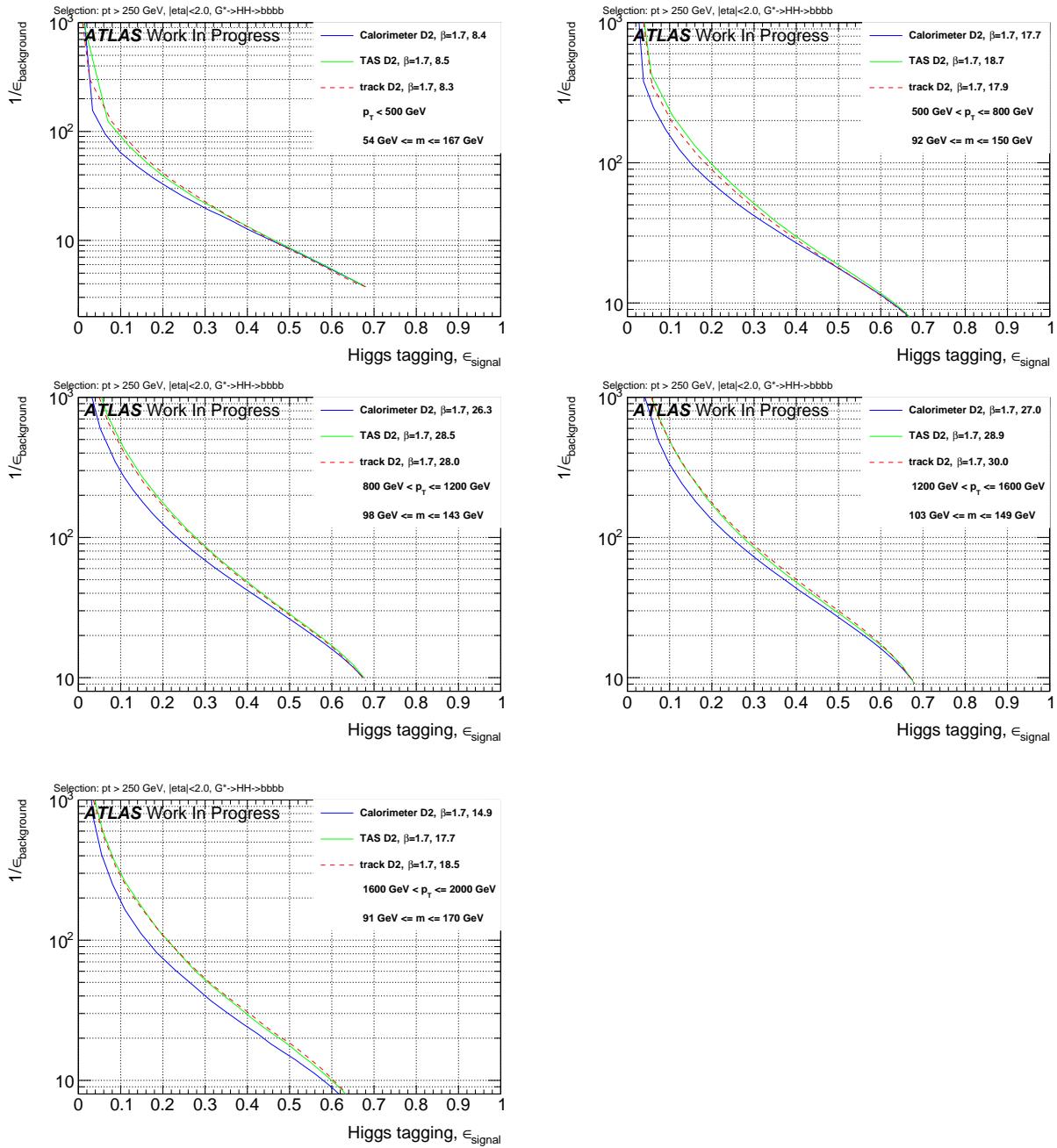


Figure 29: ROCs showing QCD rejection against Higgs jet efficiency for  $D2_{\text{TAS}}^{(\beta=1.7)}$  &  $D2_{\text{track}}^{(\beta=1.7)}$  compared to  $D2_{\text{calo}}^{(\beta=1)}$ .

### 750 8.1.3 Top Quark Tagging

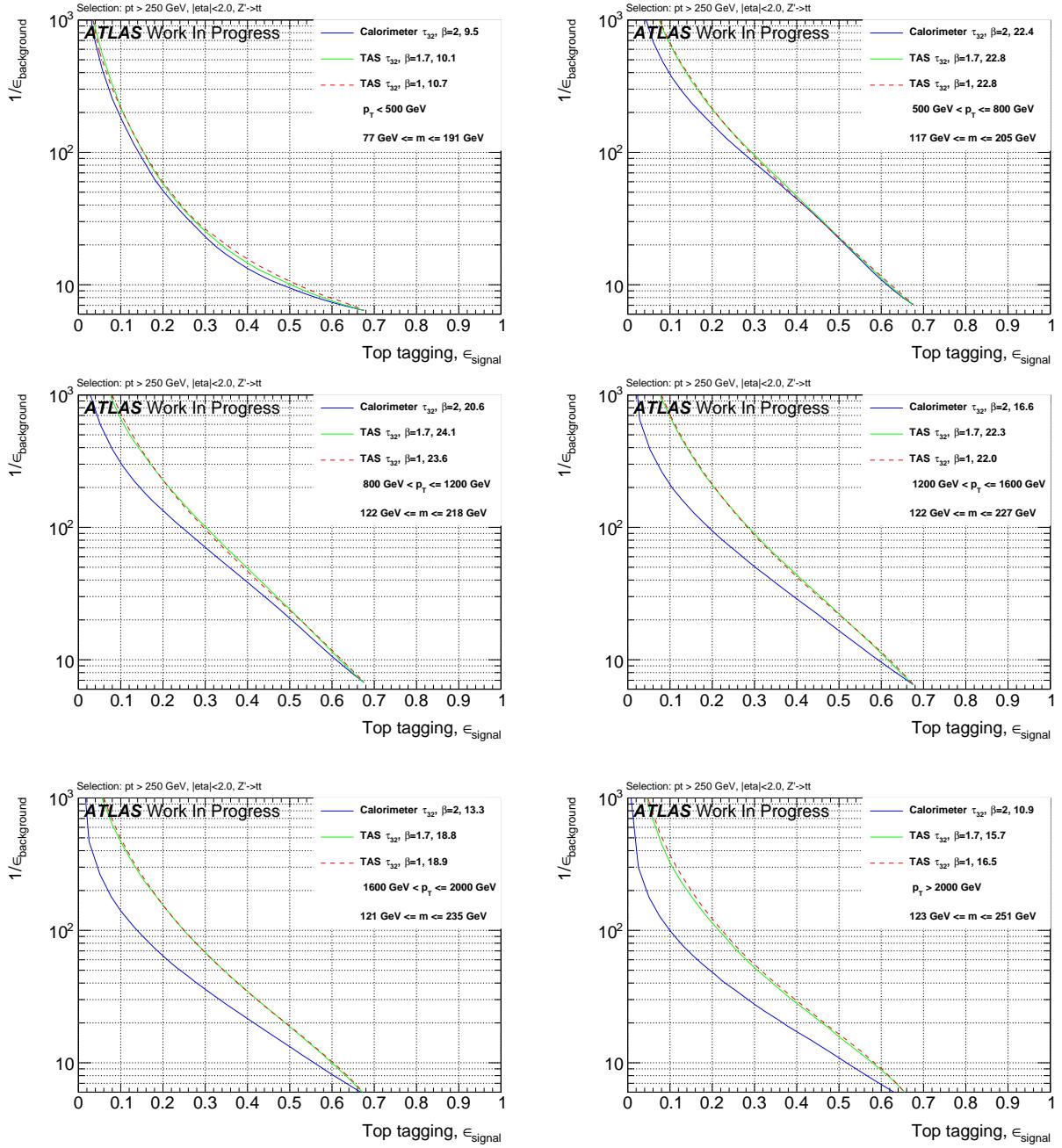


Figure 30: ROCs showing QCD rejection against Top jet efficiency for  $\tau_{32, \text{TAS}}^{(\beta=1)}$  &  $\tau_{32, \text{TAS}}^{(\beta=1.7)}$  compared to  $\tau_{32, \text{TAS}}^{(\beta=2)}$

## 751 9 Uncertainties on observables with sub-jet-assisted tracks

752 This chapter gives a brief overview of the uncertainties on the track-assisted (sub-jet) mass variable.  
 753 For  $m^{TA}$  the uncertainties are smaller than calorimeter-based jet mass variables because of the way it  
 754 is constructed,  $m^{TA} = m^{trk} \times p_T^{calo}/p_T^{trk}$ : the ratio  $m^{trk}/p_T^{trk}$  causes a cancellation of the tracking  
 755 uncertainties to a large extent, which are smaller than  $m^{calo}$ . The remaining term  $p_T^{calo}$  is the additional  
 756 one where uncertainties on this variable need to be evaluated with special care.

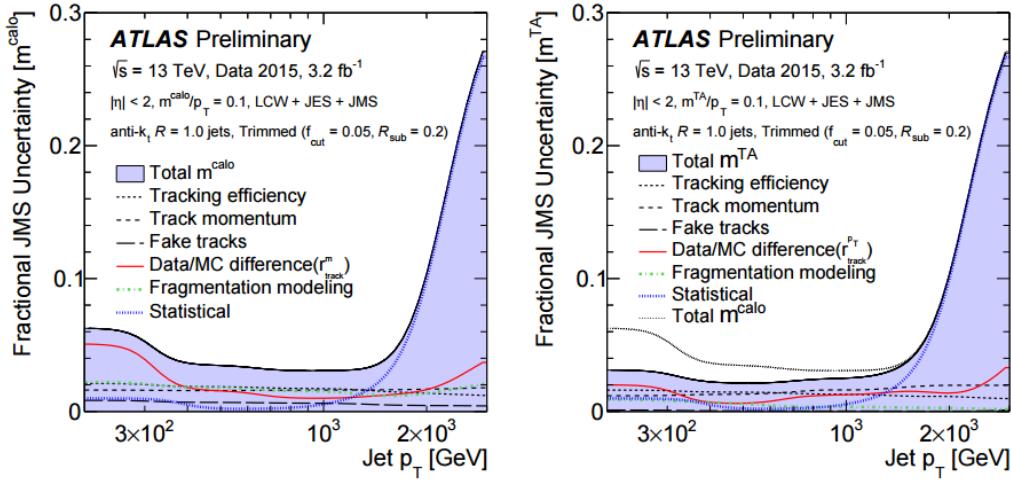


Figure 31: Comparison of the uncertainties for  $m^{calo}$ , on the left, and  $m^{TA}$ , on the right the rise on the high jet  $p_T$  is due to statistics. From the [art35].

757 For what concerns the  $m^{TAS}$ , the tracking uncertainties are expected to be identical to the  $m^{TA}$ , because  
 758 of the identical use of tracks in both variables, as also discussed in the BOOST Conference Note [art35].  
 759 The only significant difference in this regard of  $m^{TA}$  with respect to  $m^{TAS}$  is the  $p_T^{jet}$  instead of the  $p_T^{subjet}$ :  
 760 the uncertainties in the first one are calculated in-situ using  $p_T$  balance methods, and they are generally  
 761 well-behaved; for the second one, the uncertainties are also expected to lay in the same order of magnitude.  
 762 In Figure 31 the comparison of the uncertainties for the  $m^{calo}$  on the left and  $m^{TA}$  on the right shows the  
 763 smaller fractional JMS uncertainties in the use of the track-assisted method.  
 764 The path to bring  $m^{TAS}$  ready to usage needs of course to evaluate the uncertainties on the  $p_T^{subjet}$  which  
 765 can be provided with the R-Scan procedure which includes the anti- $k_t$  sub-jets of radius of 0.2; those were  
 766 already showed to have a similar performance of the  $k_t$  sub-jets used as standard.  
 767 The  $m^{TAS}$  moreover is expected to have little to none benefit from the calibration procedure which could  
 768 be provided from the R-Scan as well, as already shown and discussed in the previous chapter.

## 769 10 Conclusions & Outlook

770 The  $m^{TAS}$  variable was developed for the large- $R$  jet mass; it combines the information of the tracker- and  
 771 calorimeter-system to achieve an higher precision in the jet mass reconstruction, correcting the missed  
 772 neutral fraction which is absent in the tracker but not in the calorimeter. With respect to the  $m^{TA}$ , it  
 773 applies this correction at sub-jet by sub-jet level and not at jet by jet level, therefore providing a more  
 774 accurate reconstruction. It was shown in Monte Carlo simulation to be a very good observable confronting  
 775 quantitatively with the other definitions which are either standard or in preparation,  $m^{calo}$ ,  $m^{TA}$  and  $m^{comb}$ .  
 776 The optimal configuration of  $m^{TAS}$  is shown and confronted with different approaches, in particular in  
 777 terms of different trimming procedure of the large- $R$  jet to be used as an input. All the components of  
 778 the observable have been studied with the use of truth Monte Carlo information without detector effect, in  
 779 order to evaluate quantitatively its limits and strengths; the track  $p_T$  measure degradation was found to be  
 780 the cause of the variable decreasing performance at higher transverse momenta.

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

787 The performance of the Energy Correlation Functions and n-Subjettiness observables was studied with  
 788 tracks and subjet-assisted tracks (TAS) compared to the default calculation with calorimeter clusters. The  
 789 assisting procedure for these observables is the same single-track scaling method used for  $m^{TAS}$ . It was  
 790 shown, that the better angular resolution of tracks can help to reduce the QCD background in boosted  
 791  $W$ , Higgs and Top tagging cases, especially for very high energies. Optimal results can be achieved by  
 792 tuning the angular weighting to slightly higher values. In our study, a value of  $\beta = 1.7$  yield the highest  
 793 background rejections in most of the cases.

794 Subjet-assisted tracks performed superior for lower and intermediate boosts. At very high energies, the  
 795 assisting was shown to loose effect on the studied observables and tracks and TAS equally outperform the  
 796 cluster based variables in terms of QCD rejection.

797 For the three cases of  $W$ , Higgs and Top tagging, the optimal variables were identified and the  $p_T$  behaviour  
 798 of the cut-value for two different working points were studied.

799 For the very conclusion, the variables constructed in the work of this study,  $m^{TAS}$  and  $m^{comb}_{TAS}$ , as well as the  
 800 TAS and track based  $C2/D2/\tau_{21}/\tau_{32}$ , exhibit a better performance as their counterparts,  $m^{TA}$  and  $m^{comb}$ ,  
 801 respectively JSS observables calculated from topo-clusters. The introduced observables are now ready to  
 802 use or in preparation within the ATLAS collaboration, and share the same advantages -and disadvantages.  
 803 Further steps are necessary to get this observables to usage: calibration (mass) and uncertainties (mass  
 804 and further substructure observables).

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
$W' \rightarrow \tilde{W}\tilde{W}$ with $m_{\tilde{W}} = m_t$	Pythia 8	NNPDF23LO	A14	1.5, 2.5, 3, 4, 5 TeV

Table 6: 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.

## 805 Appendix

### 806 A Monte Carlo Samples

807 The samples used are divided into two main groups: SM background and beyond SM signal. The  
 808 SM background includes the QCD multijet samples, produced with a falling  $p_T$  spectrum. The beyond  
 809 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  
 810 ( $t \rightarrow W(\rightarrow q\bar{q}')b$ )) and RS-Graviton  $\rightarrow hh \rightarrow b\bar{b}b\bar{b}$ , i.e. final states have only jets in all the samples.  
 811 The details of the samples are given in Table 6; the masses considered span from 0.5 to 5 TeV to improve  
 812 and diversify the kinematic space covered.

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

### 820 B Trimming

821 The trimming algorithm is the most important in ATLAS and the one mainly used in this note. It takes  
 822 advantage of the fact that contamination from soft radiation has a much lower  $p_T$  with respect to the  
 823 hard-scattering component. Therefore uses a transverse momentum ratio to distinguish among those. The  
 824 algorithm works on a two-dimensional parameter space:  $R_{sub}$  and  $f_{cut}$ . The steps are as follows:

- 825 •  $k_t$  algorithm (but of course other choices are also possible) is used to create sub-jets with a smaller  
 826 radius  $R_{sub}$ , aiming at separating the soft radiation from the hard one in different sub-jets. Typical  
 827 choices are 0.2 and 0.3 (0.2 is used as standard);
- 828 • for each sub-jet, the ratio  $f_{cut}$  between its  $p_T$  and the parent jet  $p_T^{jet}$  is calculated: if then this ratio  
 829 is below a certain value, the sub-jet is removed. Standard choice is  $\frac{p_T}{p_T^{jet}} > f_{cut} = 0.05$ ;

- 830 • the sub-jets which survived this procedure are the only one which compose the trimmed jet.

831 The trimming procedure is also explained in Figure 32, an example of performance in simulation with  
 832 standard parameters is shown in Appendix (Figure 34).

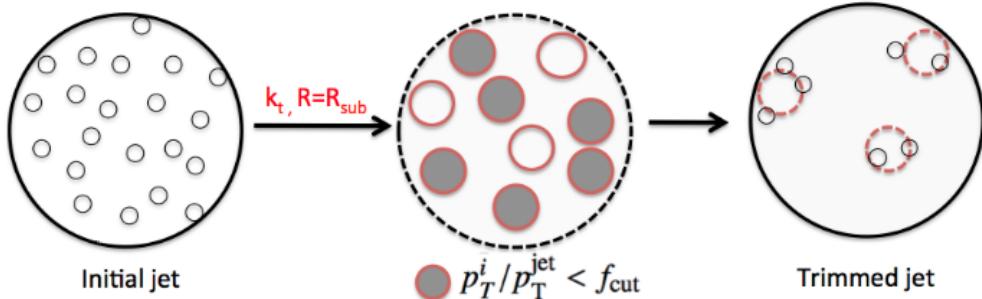


Figure 32: Schematic of the trimming algorithm.

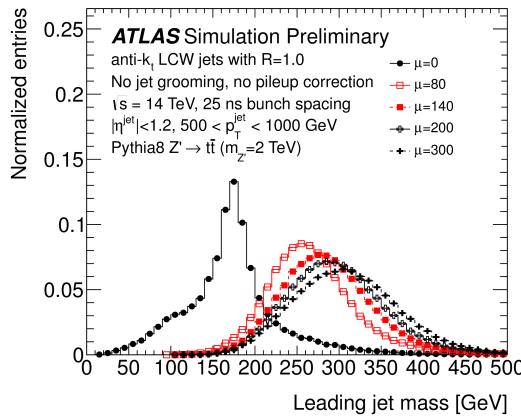


Figure 33: Effect of pile-up contamination in large- $R$  jets: here shown different PU conditions parametrized by  $\langle \mu \rangle$ . From [highlumi].

## 833 C Tracks details

834 The requirements applied on the track used in the work presented in this note are given here:

- 835 •  $p_T^{track} > 400$  MeV;
- 836 •  $|\eta| < 2.5$ ;
- 837 • Maximum 7 hits in the Pixel and STC sub-detectors;
- 838 • Maximum 1 Pixel hole;
- 839 • Maximum 2 silicon holes;
- 840 • Less than 3 shared modules;

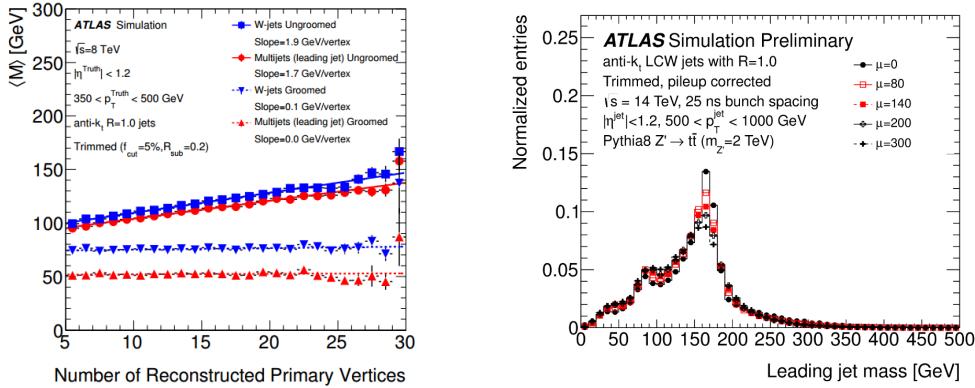


Figure 34: Left: mass reconstructed as a function of the number of primary vertices (parameterizing PU) for different samples; after trimming procedure the mass is pretty much independent of PU for all the samples. Right: mass distributions for different PU conditions: after trimming the reconstruction is not degraded as much as Figure 33.

- 841 • Maximum 2 mm of displacement along beam axis ( $z_0$ ) from the primary vertex;  
 842 • Maximum 2.5 mm of distance in x-y plane from the primary vertex and point of closest approach  
 843 ( $d_0$ ).

## 844 D Alternative Performance Figure of Merit (FoM)

845 A concrete, quantitative feature has to be defined in order to understand which observable is “better”, in  
 846 the sense that we would prefer one or the other according to this criterion. This is often referred to as  
 847 *Figure of Merit* or simply FoM.

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

### 853 D.1 Gaussian Fit

854 The important feature to keep in mind, in fact, is the underlying physics which brings us to calculate the  
 855 mass of a jet. In figure 35 this is made clear: if the width of the invariant mass distribution of the jet is  
 856 smaller (highlighted), it allows a bigger background rejection, here shown as the QCD dijet, for the same  
 857 signal efficiency, by means of a simple mass requirement.

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

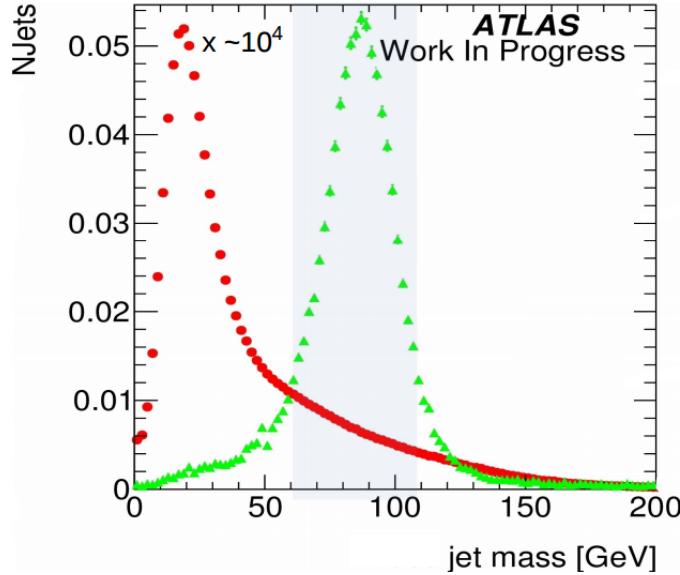


Figure 35: Mass distributions: in red the QCD dijet background rescaled, in green the  $W/Z$  from the  $W'$  sample. Highlighted the width of the 68% of the  $W/Z$  distribution.

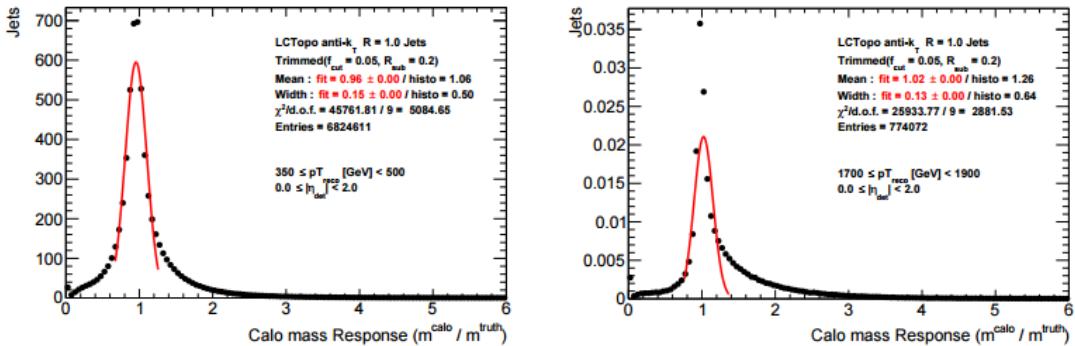


Figure 36: Mass Response distributions for the QCD multijet for various  $p_T$  ranges: on the right the failure of the Gaussian fit shows the limitation of this approach to serve as the Figure of Merit. On the plot the fit parameters and transverse momentum ranges.

The situation is depicted e.g. in Figure 36, 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 consider 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 is presented in the body of this note.

Jet Mass Observable Distribution Kinematic distribution for all the samples,  $p_T \eta$  and  $\phi$  is shown.

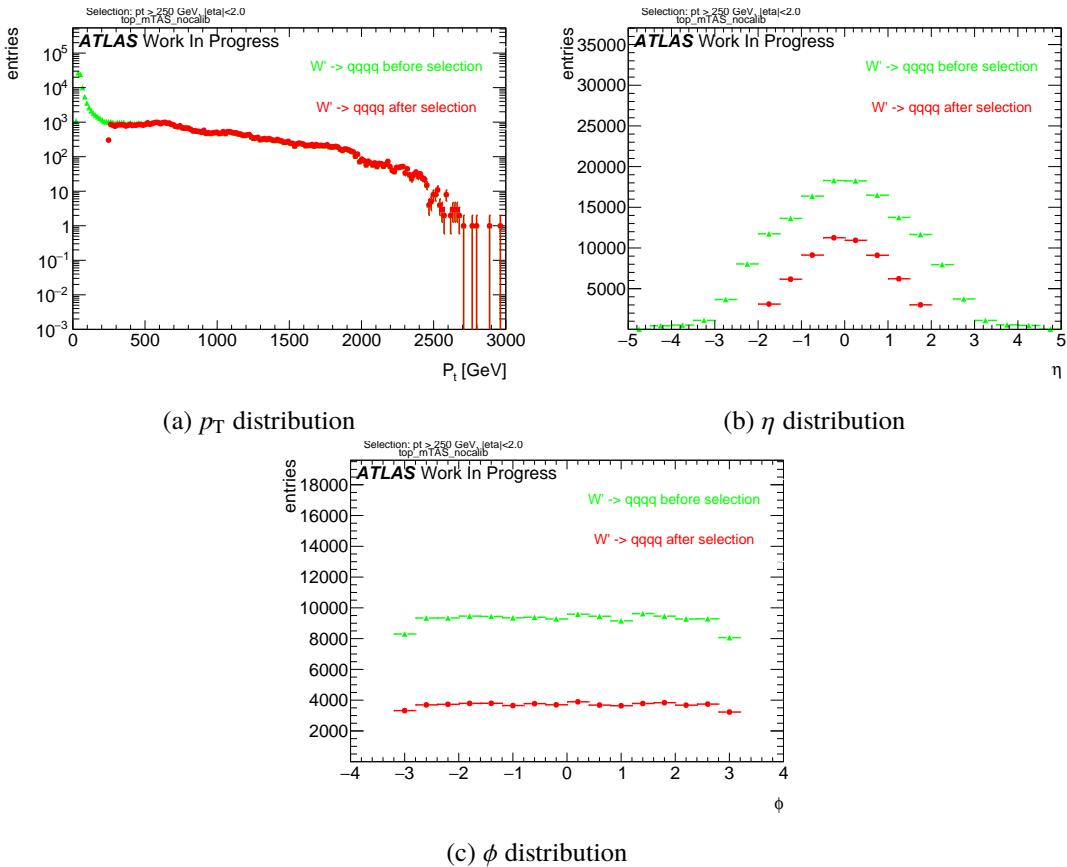


Figure 37: Boosted tops kinematic distribution.

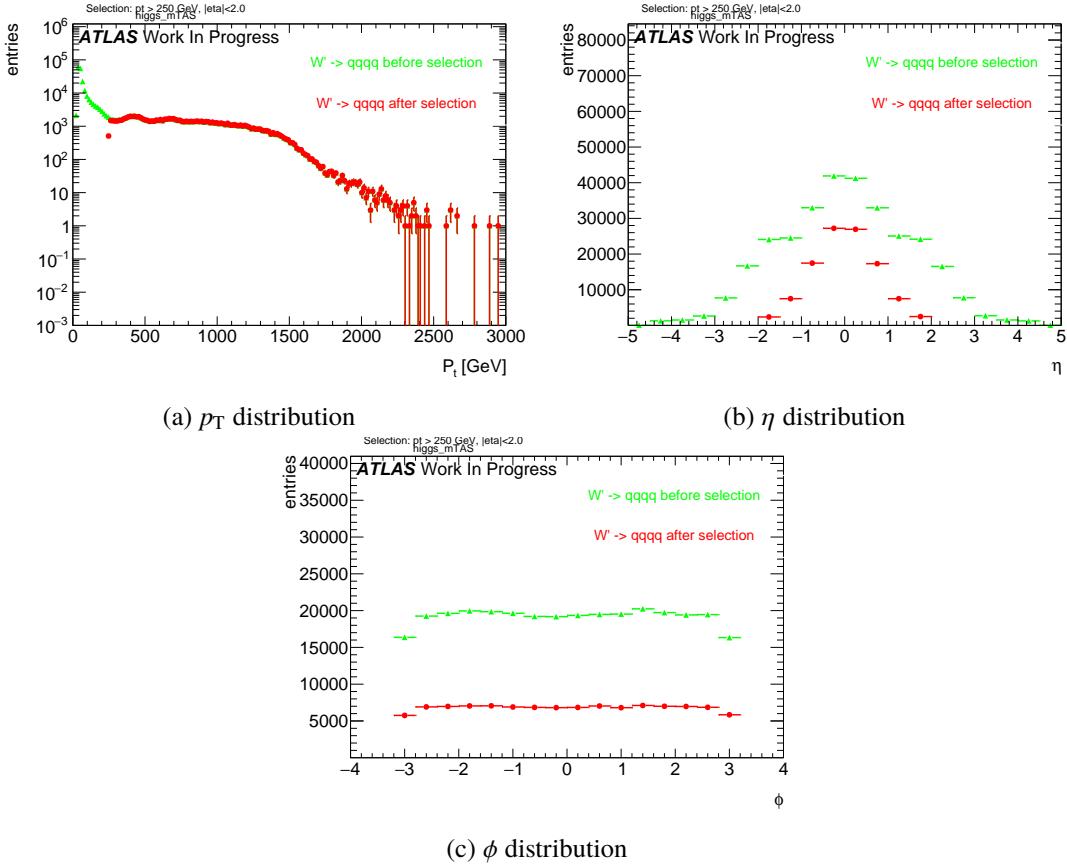


Figure 38: RS-Graviton kinematic distribution.

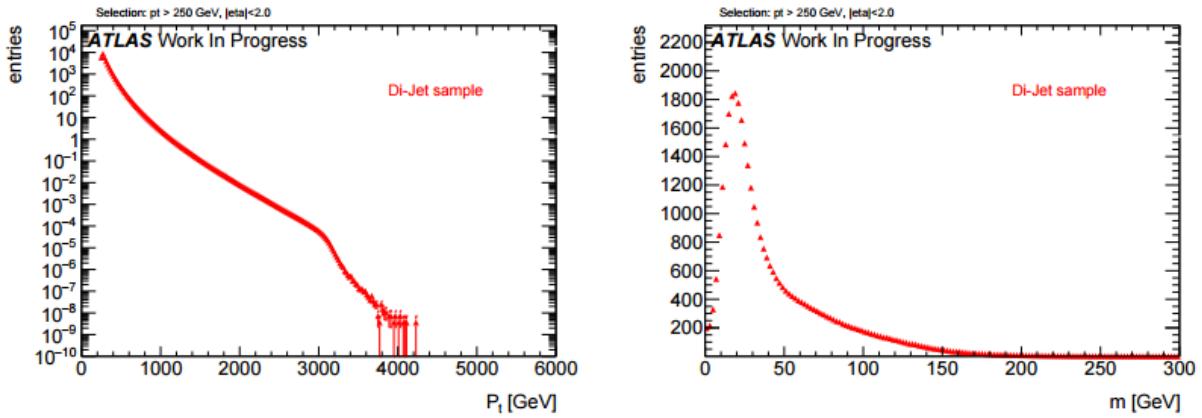


Figure 39: QCD dijet transverse momentum and mass distributions.

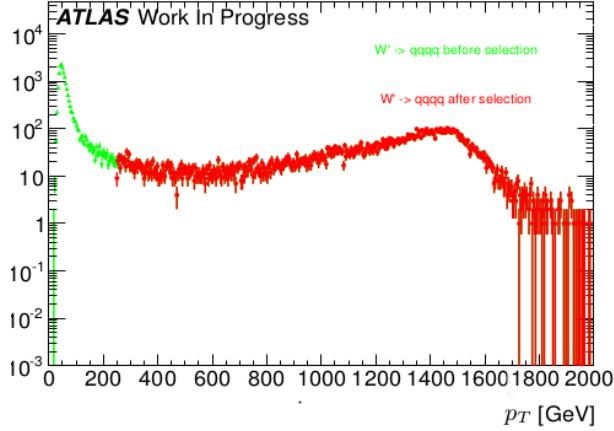


Figure 40: The  $p_T$  distribution of a 3 TeV resonance from the hadronically decaying  $W$  or  $Z$ , in logarithmic plot. As can be seen, the jacobian peak is around  $p_T \simeq m_{W'}/2 \simeq 1.5$  TeV.

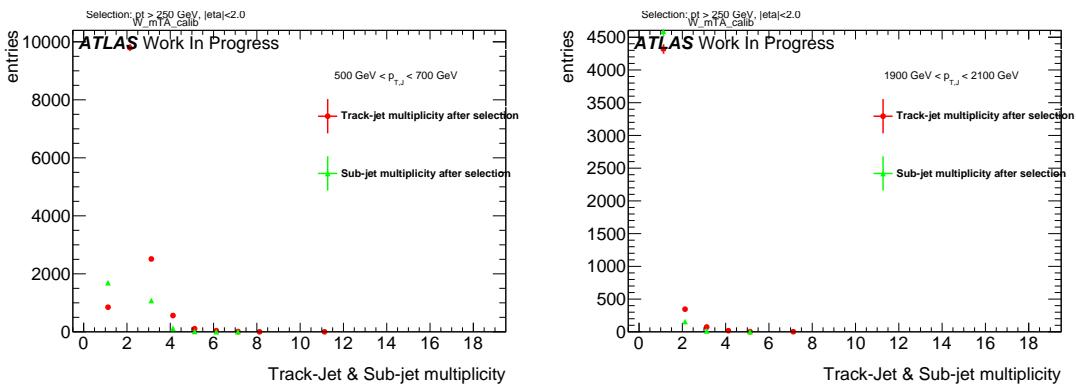


Figure 41: Sub-jet and Track-jet (jets created having tracks as input) multiplicity, for selected bins of transverse momentum.

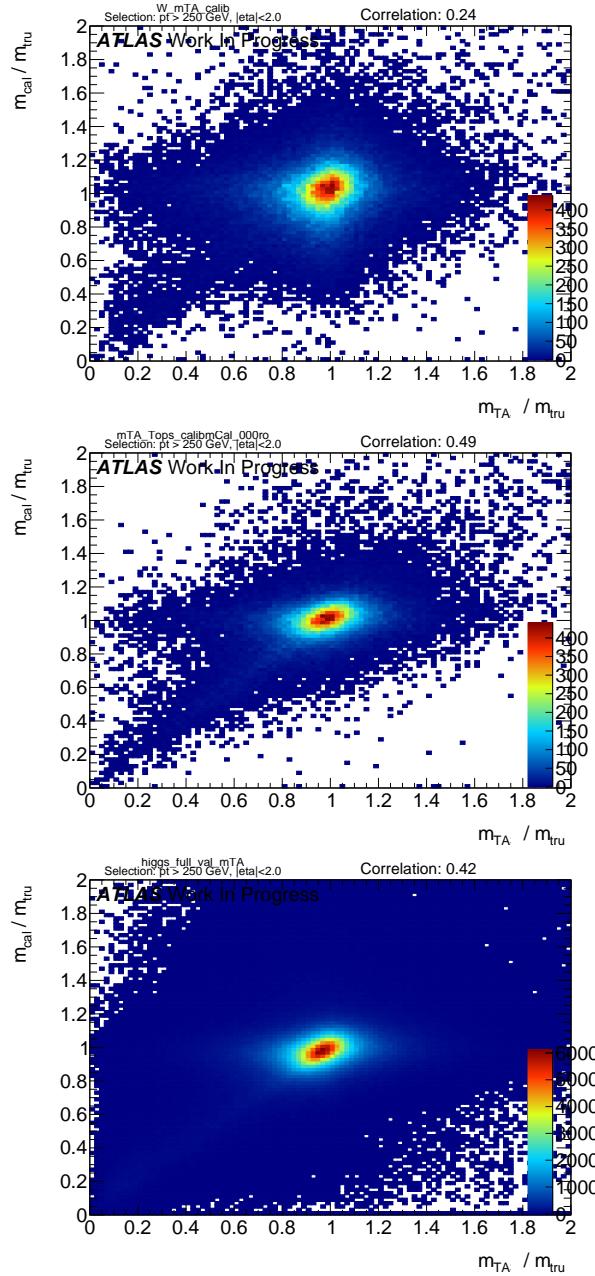


Figure 42: Calorimeter based jet mass response vs the track-assisted mass response for the three signal samples. Correlation coefficient is indicated on the top right.

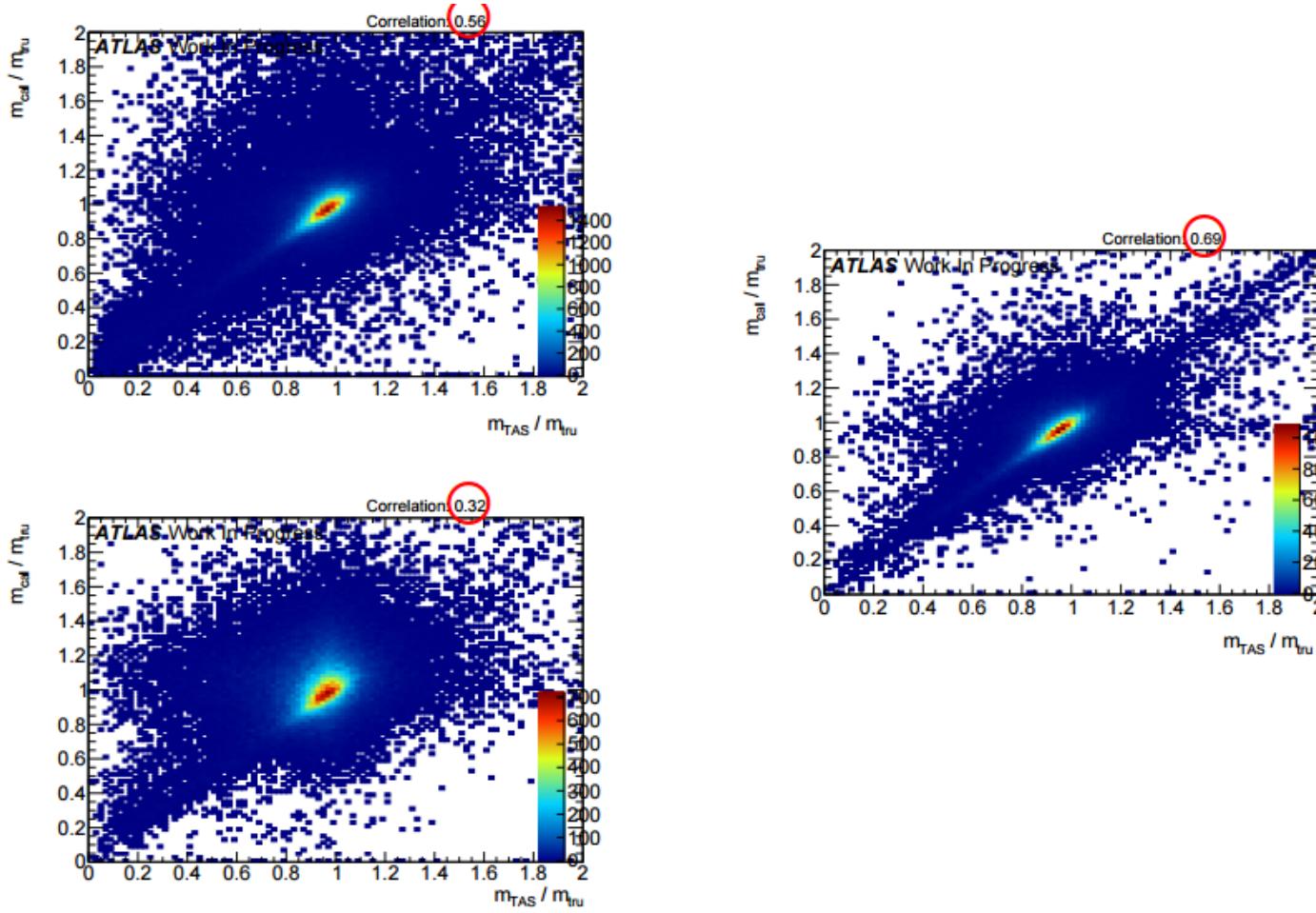


Figure 43: Calorimeter based jet mass response vs the track-assisted sub-jet mass response for the three signal samples. Correlation coefficient is indicated on the top right and highlighted. On the left, top, the higgs sample, bottom, the  $W/Z$ ; on the right the top-quark sample.

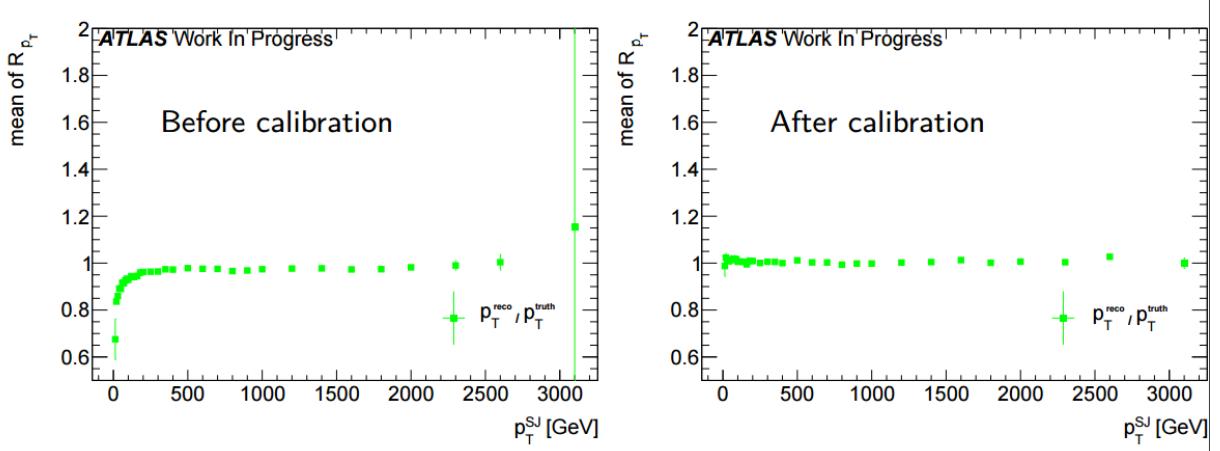


Figure 44: Poor's man calibration effect on mean of transverse momentum's response of the sub-jet, before, left, and after, right, the procedure.

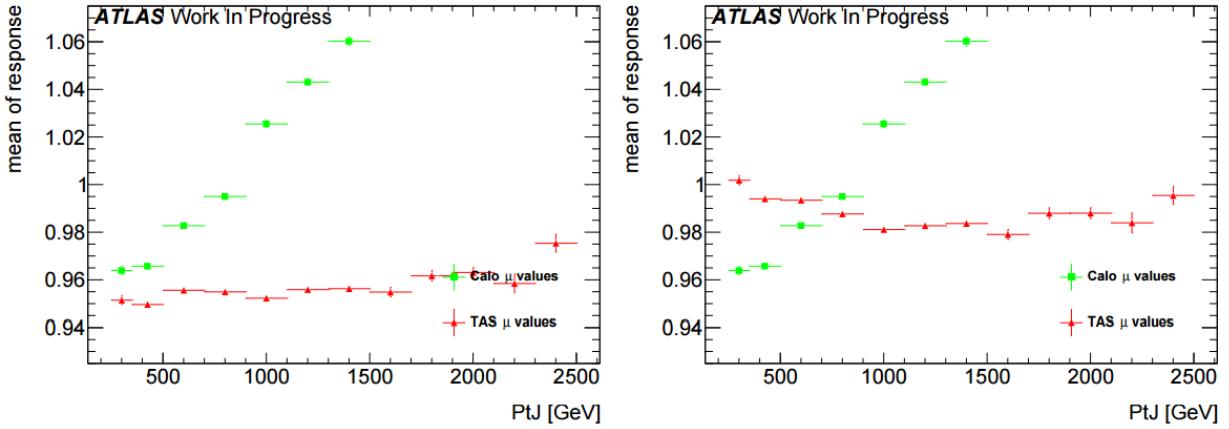


Figure 45: Poor's man calibration effect on the mean of the mass response of the large-R jet, before, left, and after, right, the procedure.

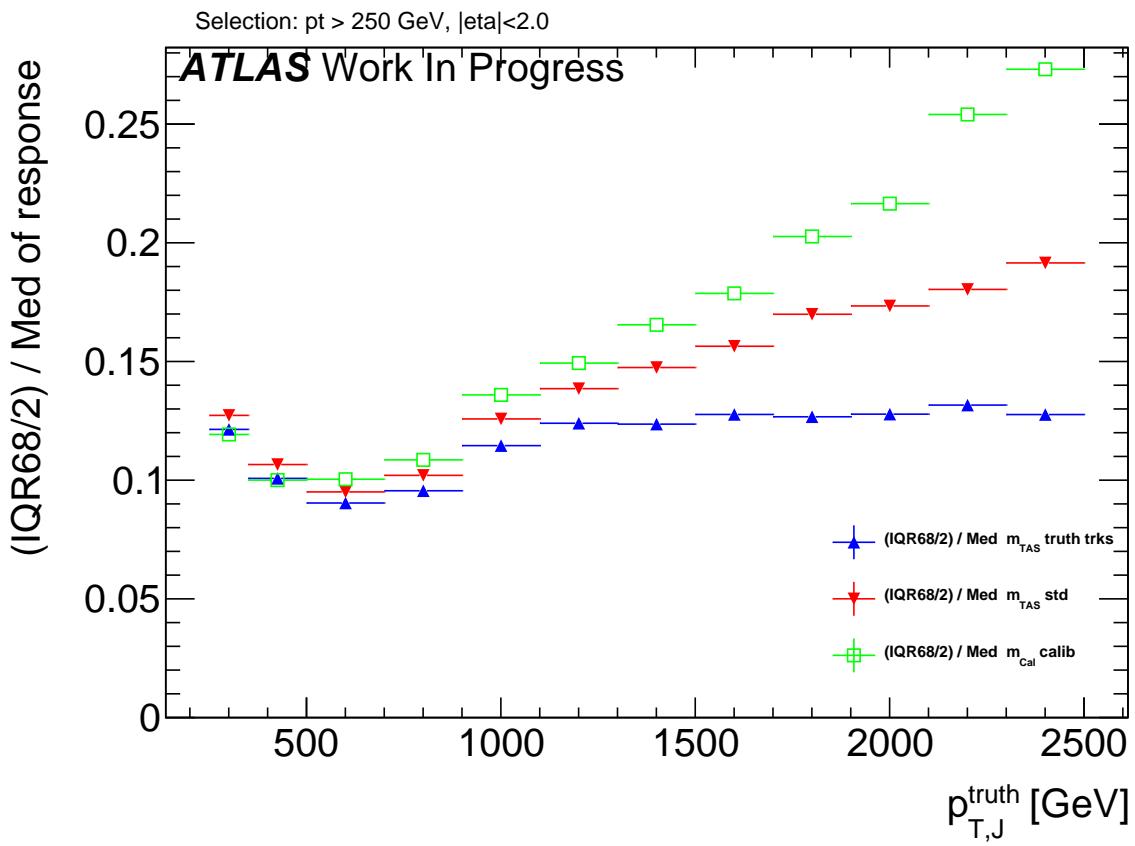
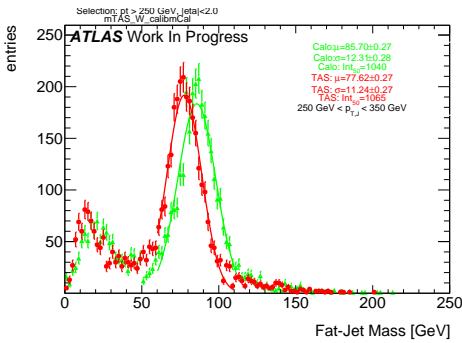
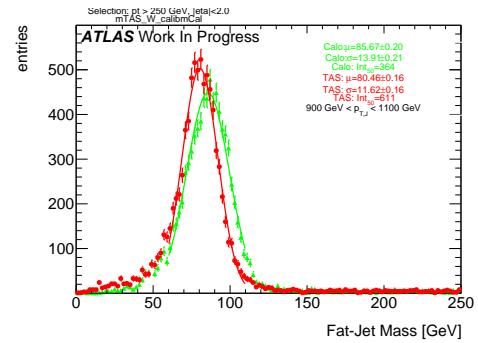
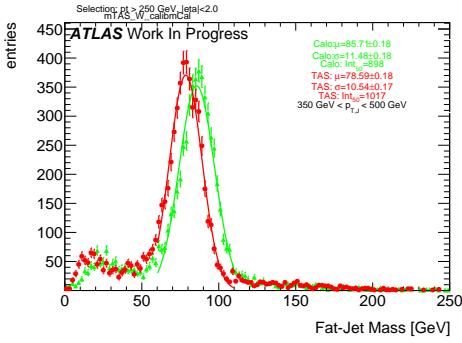
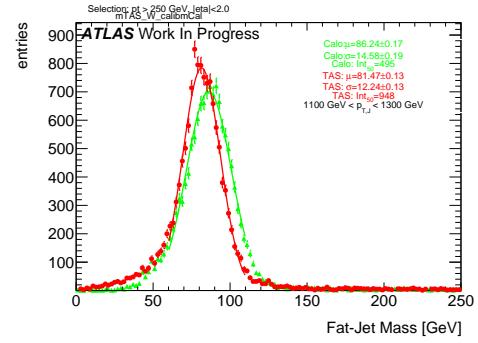
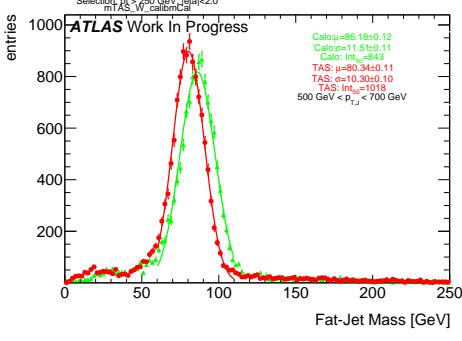
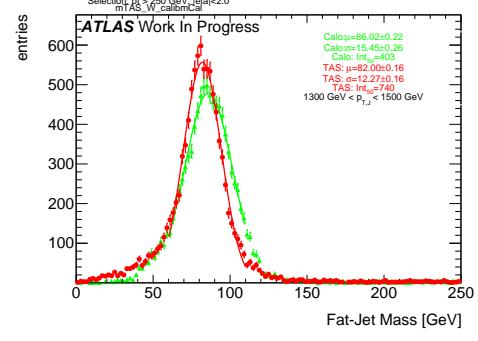
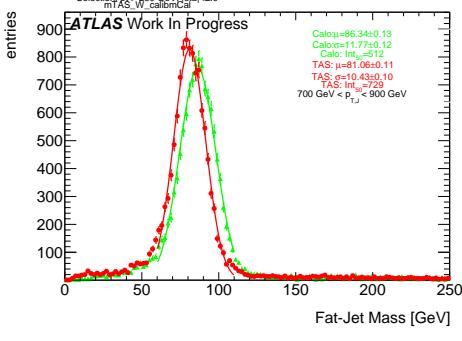
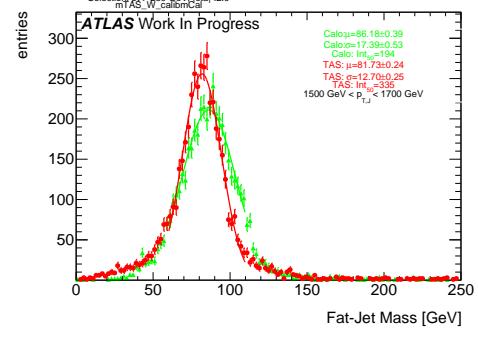


Figure 46: Comparison of the  $m^{\text{TAS}}$  and the same variable using truth-level information for the tracks.

<sub>871</sub> E  **$m^{TAS}$  distributions, boosted W/Z**

Figure 47:  $m^{TAS}$  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 48:  $m^{TAS}$  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 49:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)Figure 53:  $m^{TAS}$  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 54:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

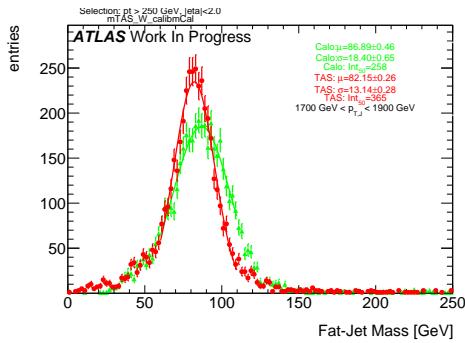


Figure 55:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

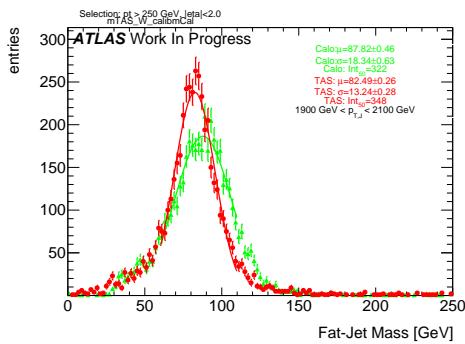


Figure 56:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

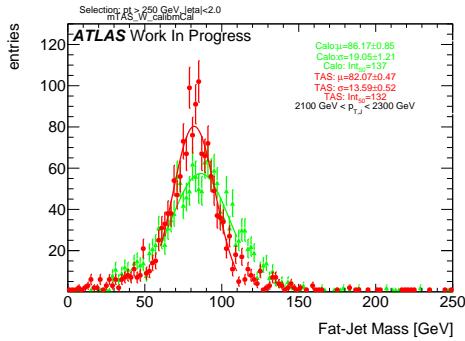


Figure 57:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

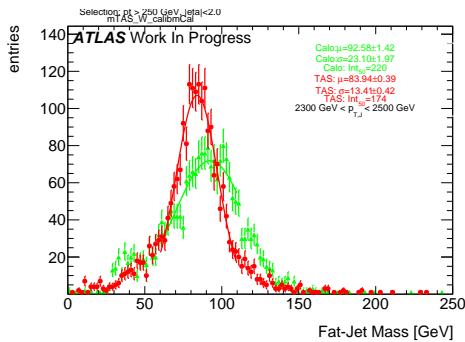


Figure 58:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

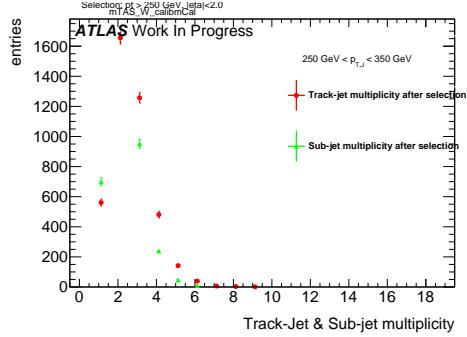


Figure 59: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

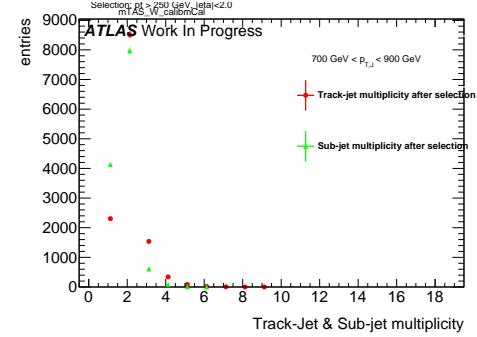


Figure 62: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

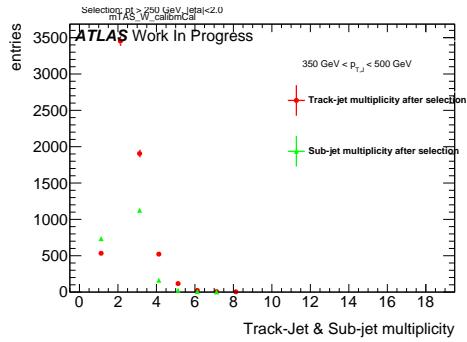


Figure 60: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

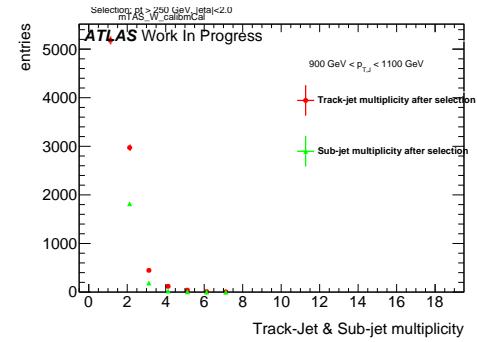


Figure 63: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

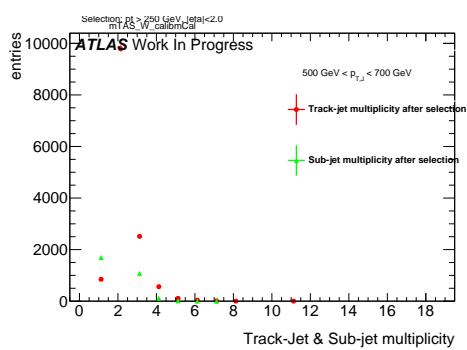


Figure 61: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

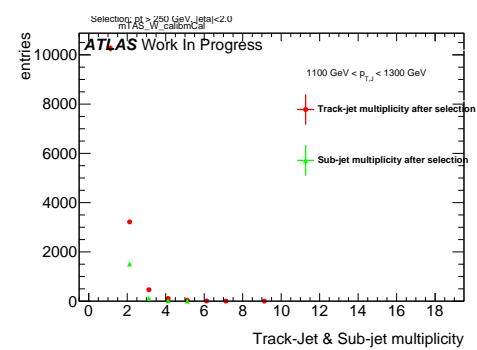


Figure 64: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

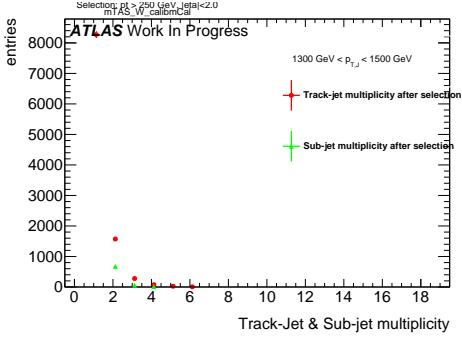


Figure 65: Track-jet  $R=0.2$  and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

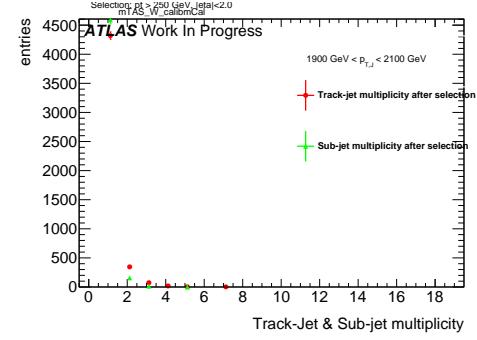


Figure 68: Track-jet  $R=0.2$  and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

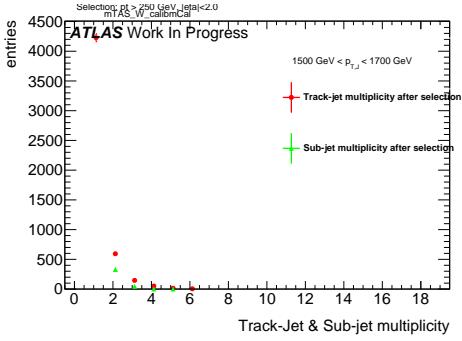


Figure 66: Track-jet  $R=0.2$  and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

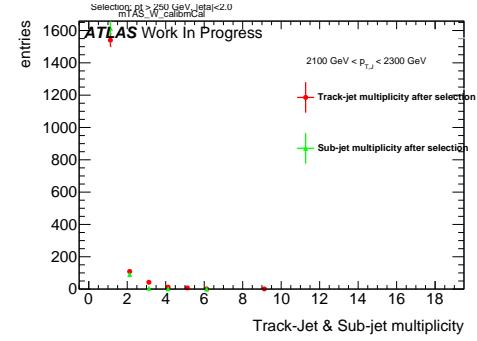


Figure 69: Track-jet  $R=0.2$  and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

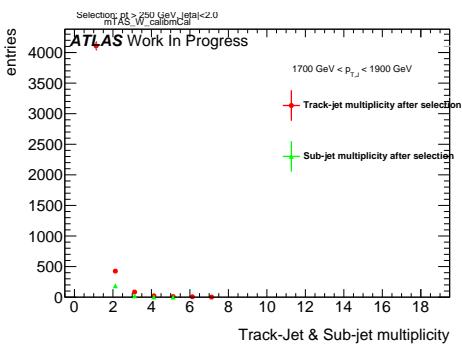


Figure 67: Track-jet  $R=0.2$  and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

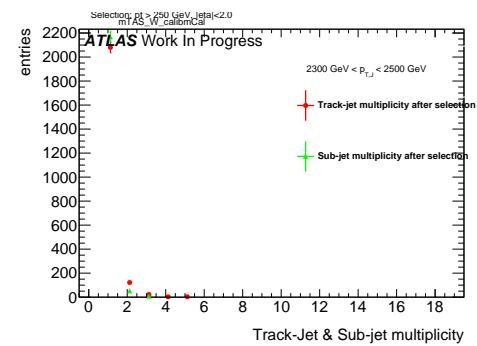


Figure 70: Track-jet  $R=0.2$  and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

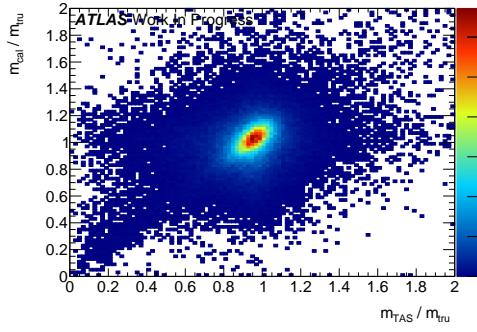
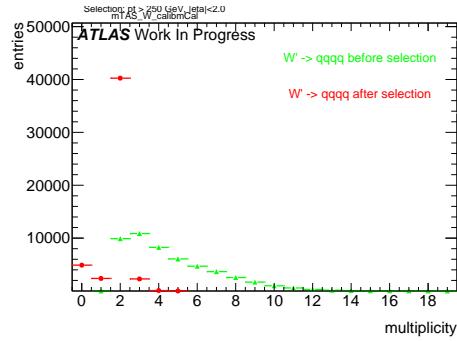
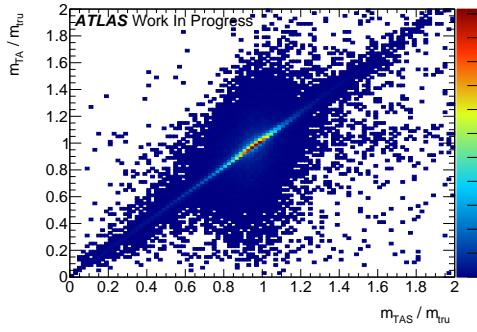
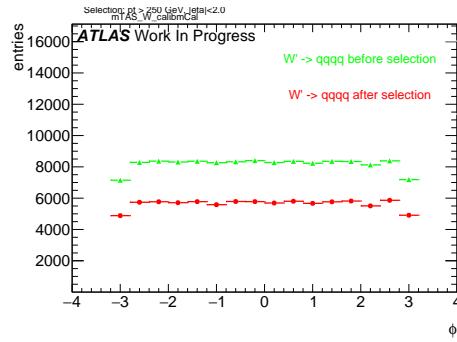
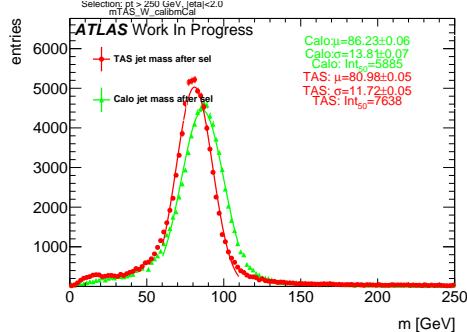
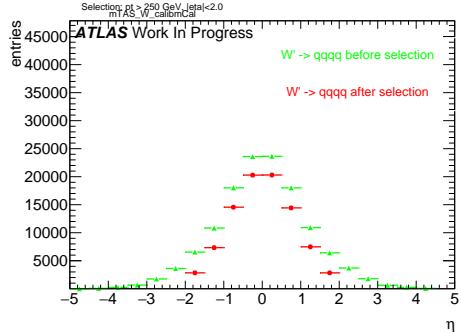
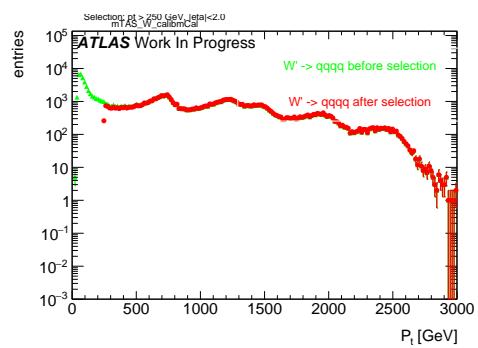
Figure 71: Scatter plot  $m^{TAS}$  versus  $m^{calo}$  responses

Figure 75: large-R jet Multiplicity, before and after selection

Figure 72: Scatter plot  $m^{TAS}$  versus  $m^TA$  responsesFigure 76:  $\phi$  distribution of the large-R jet, before and after selectionFigure 73:  $m^{TAS}$  distribution in all the  $p_T$  binsFigure 74:  $\eta$  distribution of the large-R jet, before and after selectionFigure 77:  $p_T$  distribution of the large-R jet, before and after selection

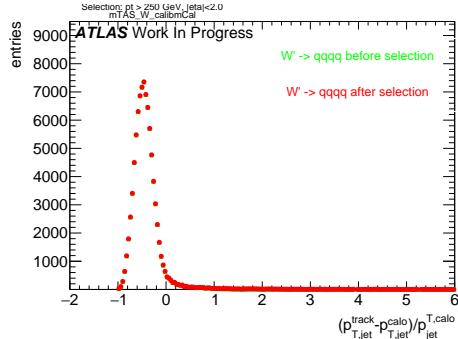


Figure 78:  $p_T$  resolution:  $\frac{p_{T,jet}^{track} - p_{T,jet}^{calo}}{p_{T,jet}^{calo}}$ , before and after selection

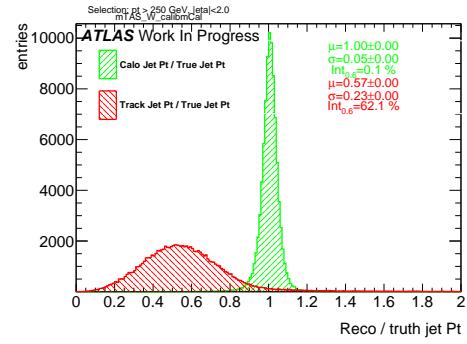


Figure 81: Transverse momentum response  $p_T^{Reco} / p_T^{Truth}$  for calorimeter and tracks

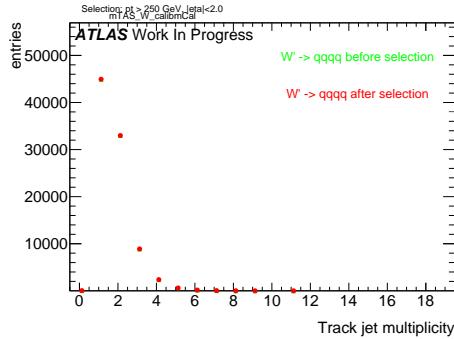


Figure 79: Multiplicity of track-jets R=0.2 per large-R jet

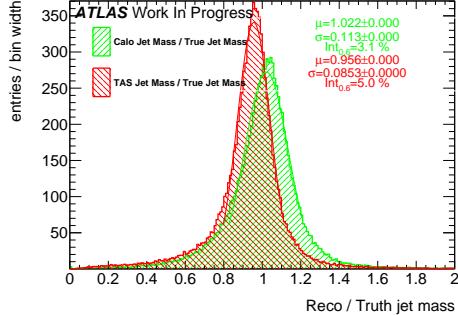


Figure 80: Response  $m^{Reco} / m^{Truth}$  for all the  $p_T$  bins

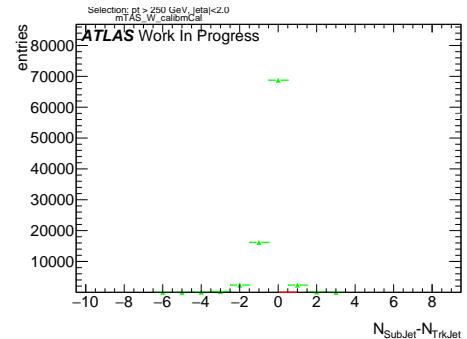


Figure 82: sub-jet - track-jet Multiplicity

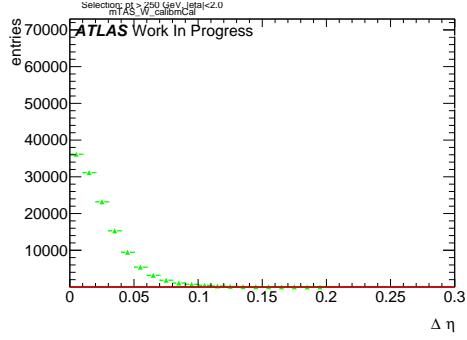


Figure 83:  $|\eta_{\text{sub-jet}} - \eta_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

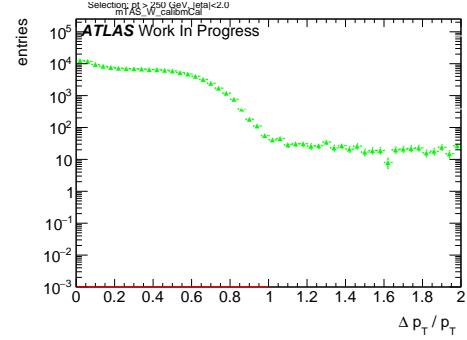


Figure 86:  $|p_{T,\text{sub-jet}} - p_{T,\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

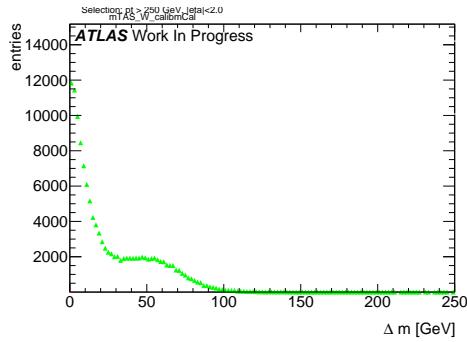


Figure 84:  $|m_{\text{sub-jet}} - m_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

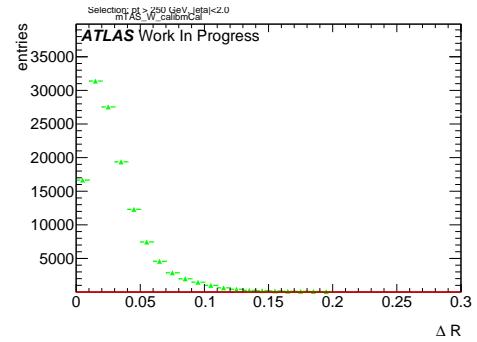


Figure 87:  $|R_{\text{sub-jet}} - R_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

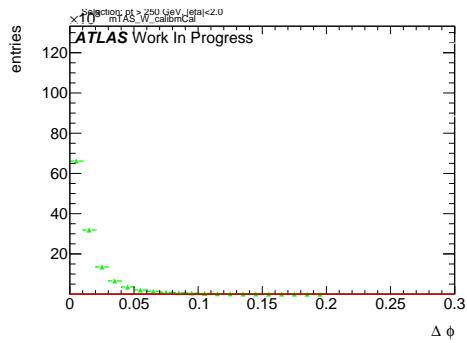


Figure 85:  $|\phi_{\text{sub-jet}} - \phi_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

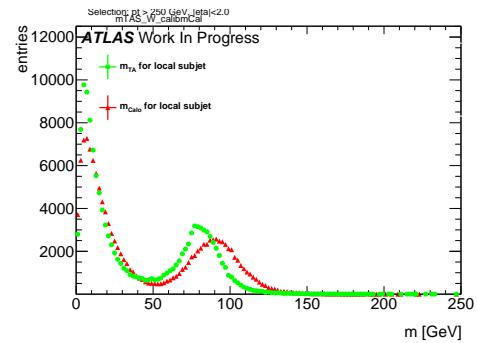


Figure 88: Mass distribution of the sub-jet, calorimeter and track-assisted

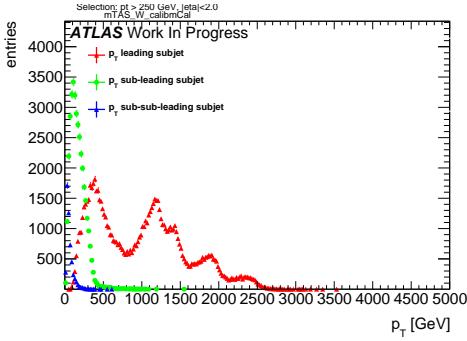


Figure 89:  $p_T$  distribution for leading, sub-leading and sub-sub-leading sub-jets

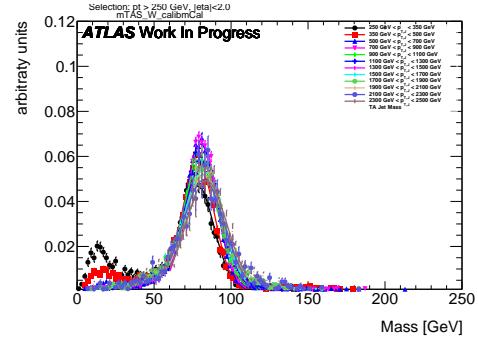


Figure 92:  $m^{TAS}$  for  $p_T^J$  bin, superimposed

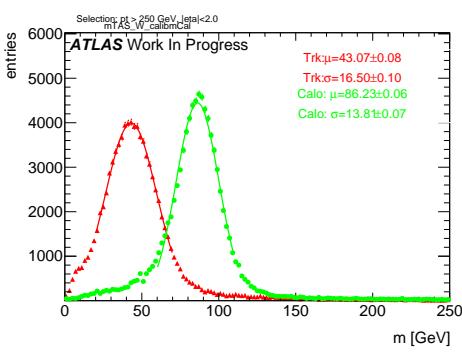


Figure 90: Mass distribution for calorimeter and tracks associated to the large-R jet

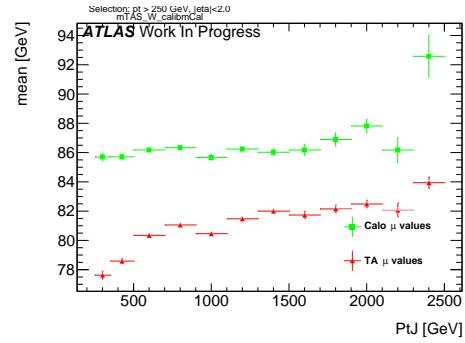


Figure 93:  $\mu$  from fit of the mass distribution vs bin of  $p_T^J$

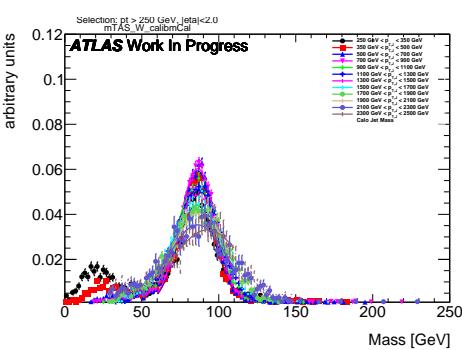


Figure 91:  $m^{calo}$  for  $p_T^J$  bin, superimposed

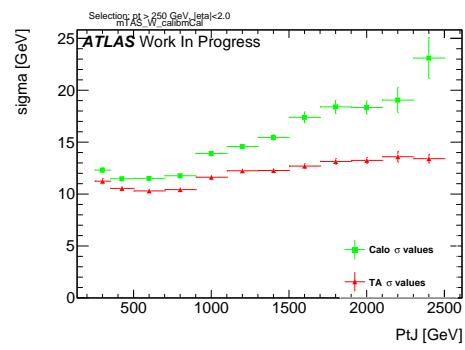


Figure 94:  $\sigma$  from fit of the mass distribution vs bin of  $p_T^J$

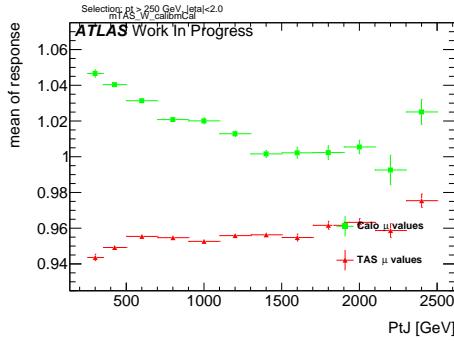


Figure 95:  $\mu$  from fit of the mass Response vs bin of  $p_T^J$

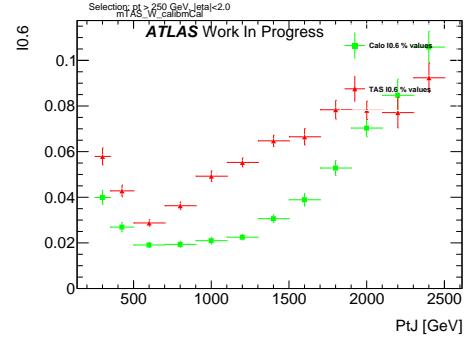


Figure 98: Left integral normalized,  $\int_0^{0.6}$  of the mass response, vs bin of  $p_T^J$

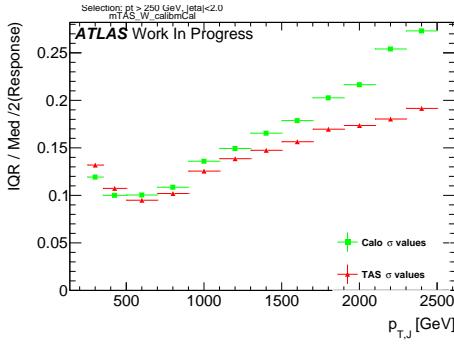


Figure 96:  $\sigma$  from fit of the mass Response vs bin of  $p_T^J$

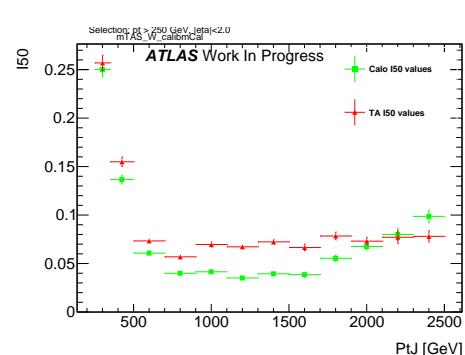


Figure 99:  $\int_0^{50GeV}$  from fit of the mass distribution vs bin of  $p_T^J$  (normalized)

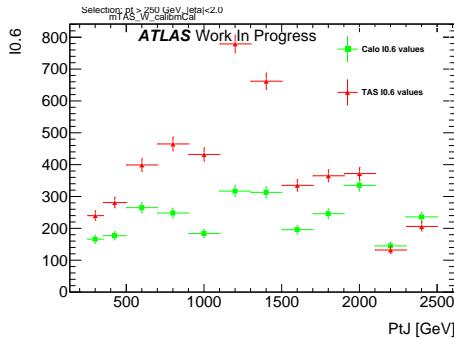


Figure 97: Left integral,  $\int_0^{0.6}$  of the mass response, vs bin of  $p_T^J$

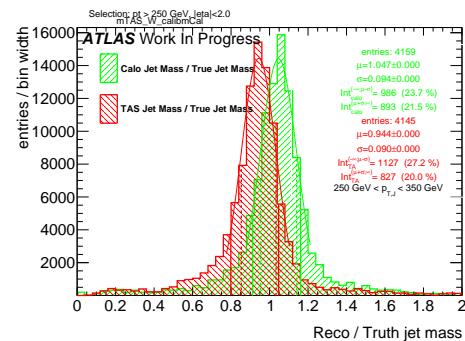
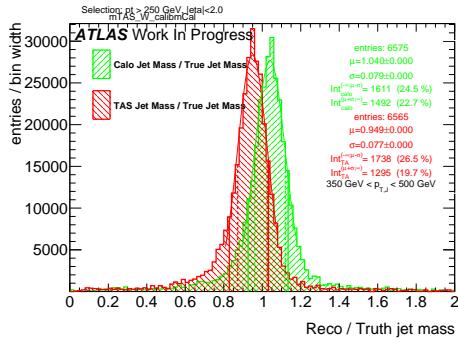
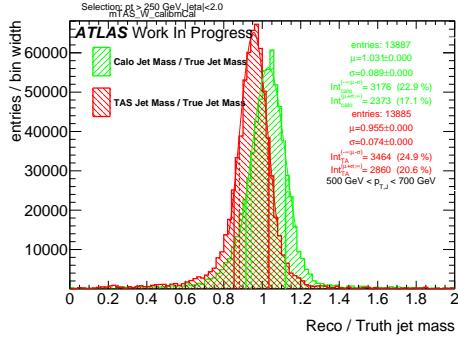
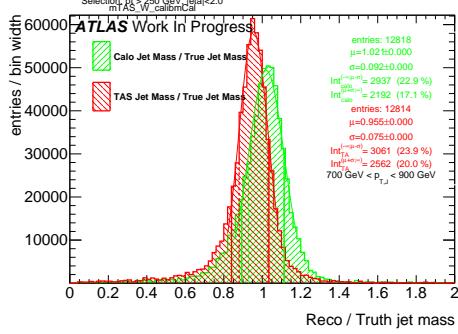
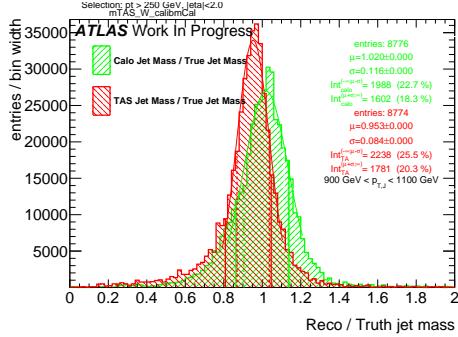
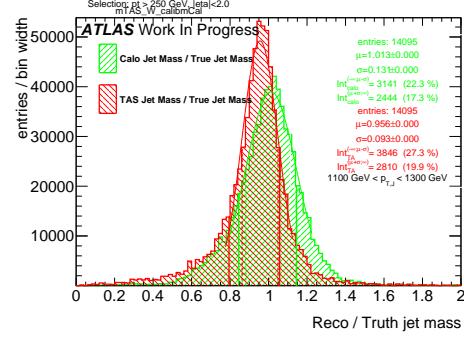
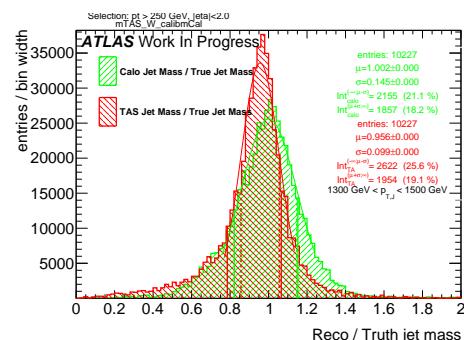
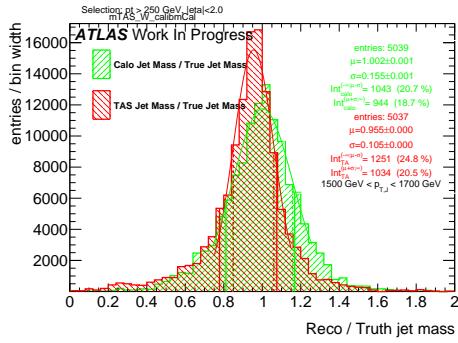
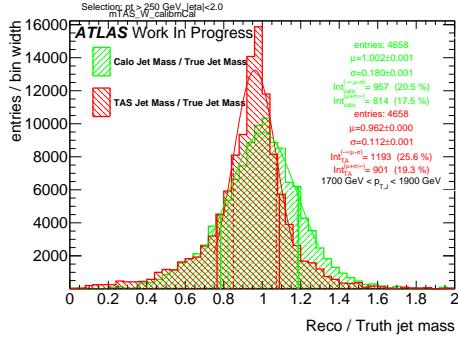
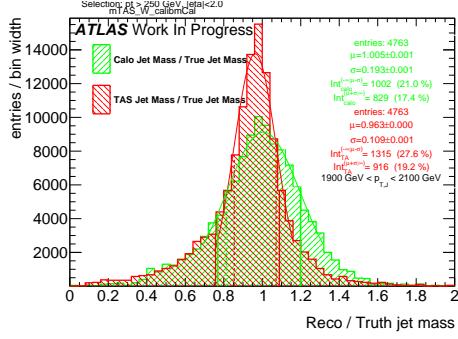
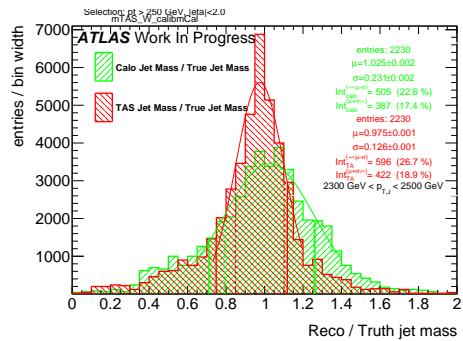
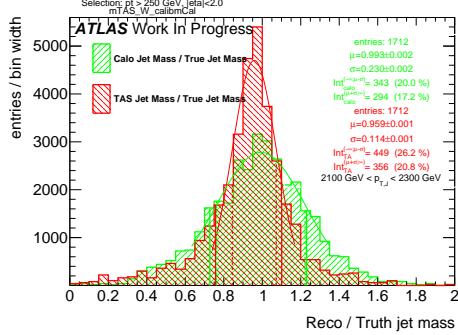


Figure 100: Response in bin of  $p_T^J$  (indicated on plot)

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)Figure 106: Response in bin of  $p_T^J$  (indicated on plot)

Figure 107: Response in bin of  $p_T^J$  (indicated on plot)Figure 108: Response in bin of  $p_T^J$  (indicated on plot)Figure 109: Response in bin of  $p_T^J$  (indicated on plot)Figure 111: Response in bin of  $p_T^J$  (indicated on plot)Figure 110: Response in bin of  $p_T^J$  (indicated on plot)

872 **F  $m^{TAS}$  distributions, boosted tops**

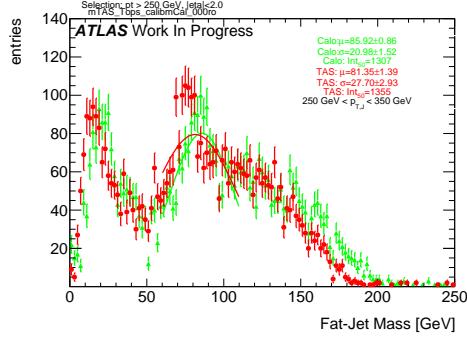


Figure 112:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

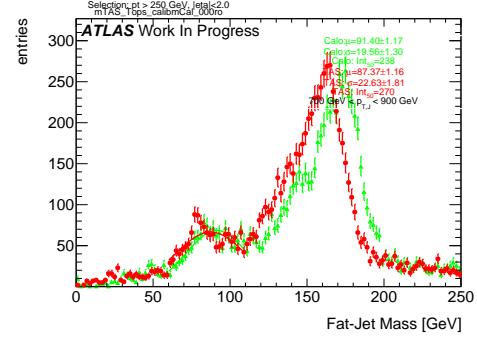


Figure 115:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

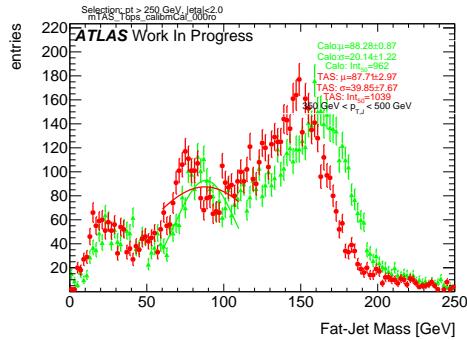


Figure 113:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

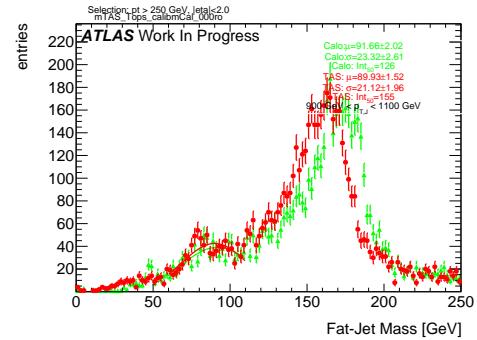


Figure 116:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

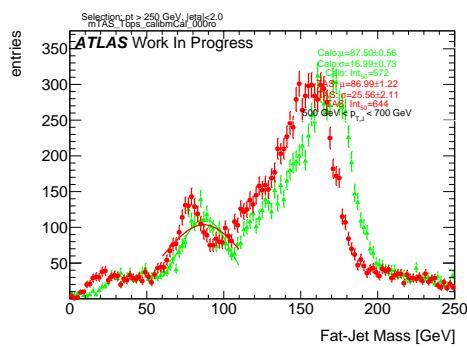


Figure 114:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

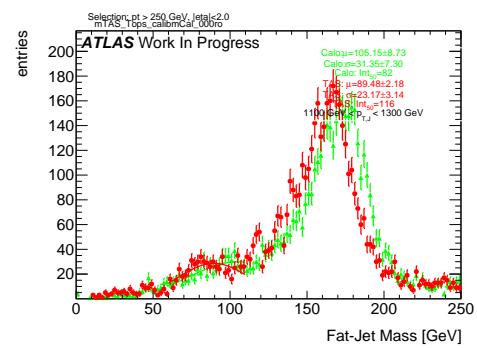


Figure 117:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

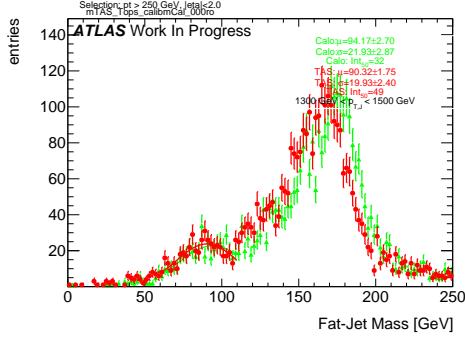


Figure 118:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

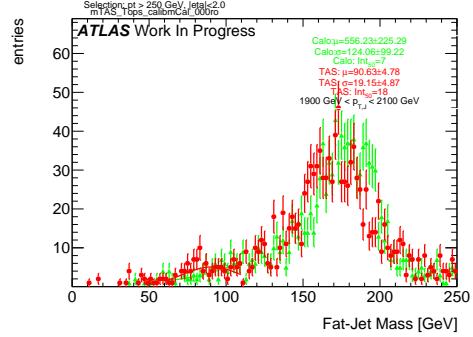


Figure 121:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

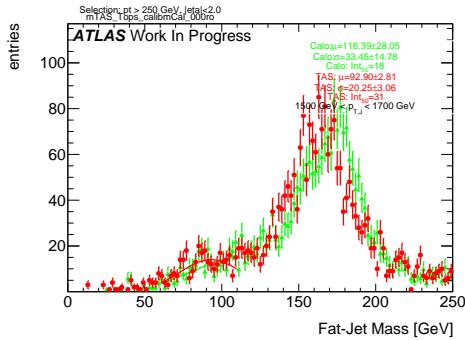


Figure 119:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

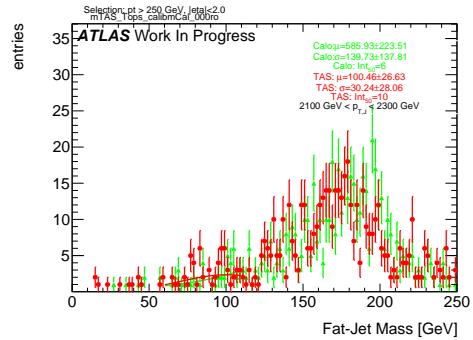


Figure 122:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

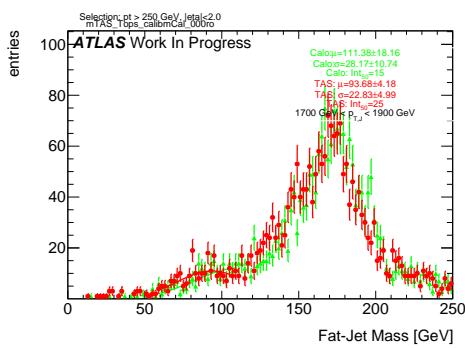


Figure 120:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

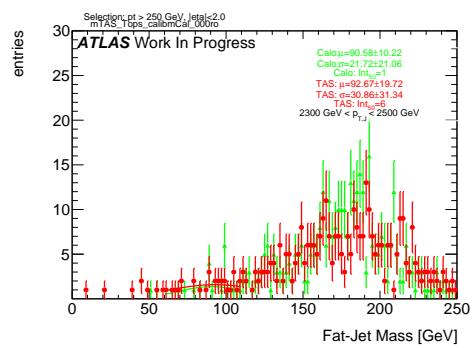


Figure 123:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

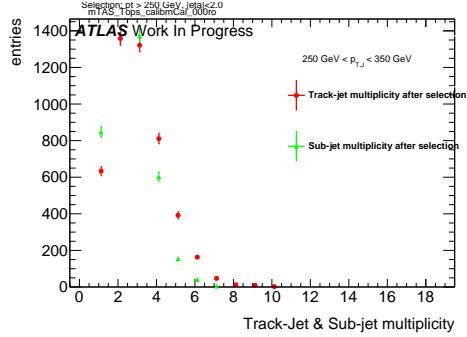


Figure 124: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

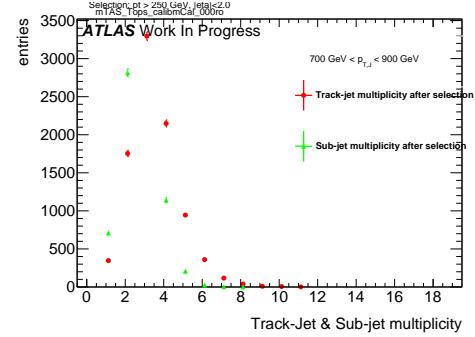


Figure 127: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

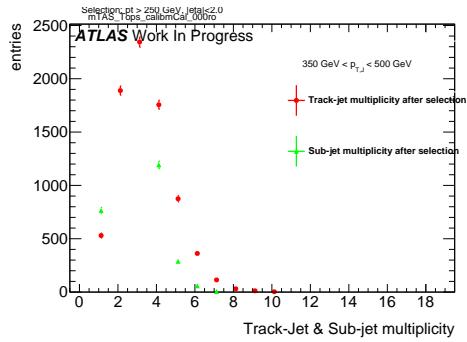


Figure 125: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

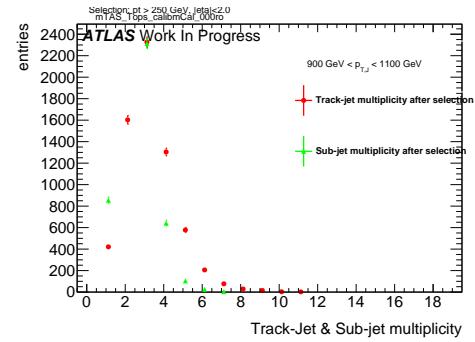


Figure 128: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

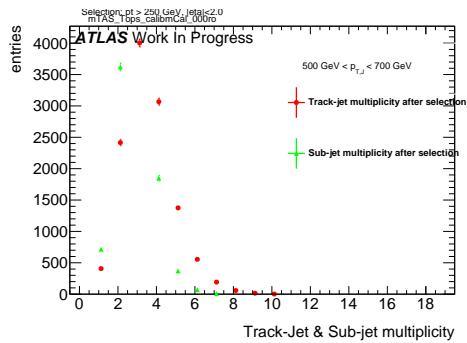


Figure 126: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

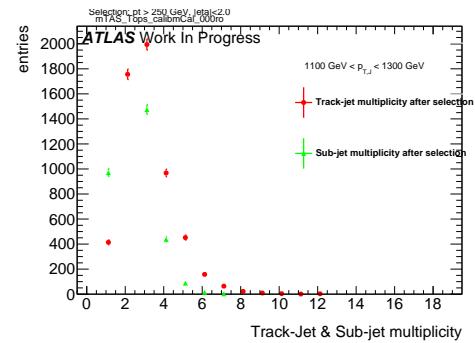


Figure 129: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

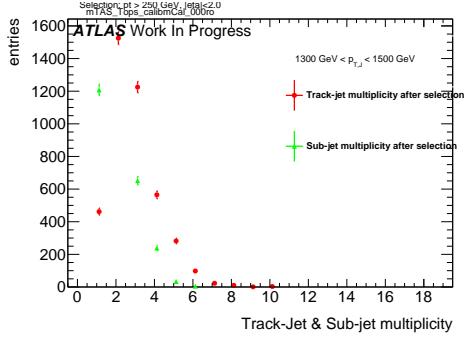


Figure 130: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

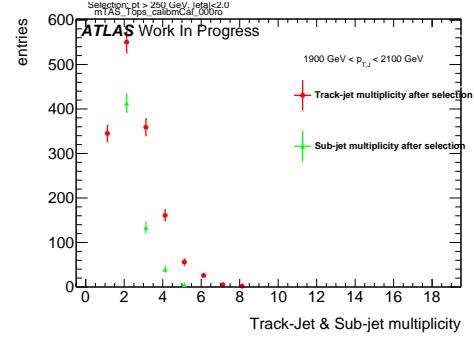


Figure 133: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

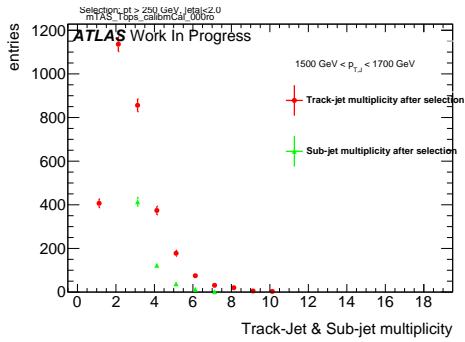


Figure 131: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

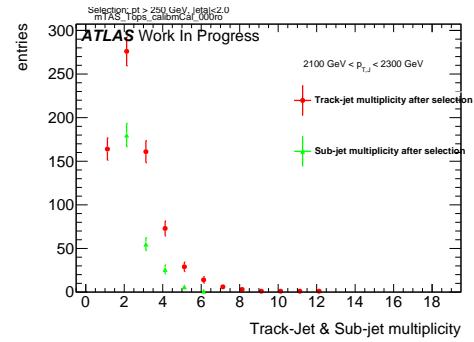


Figure 134: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

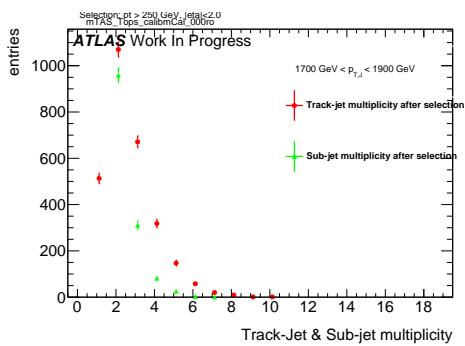


Figure 132: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

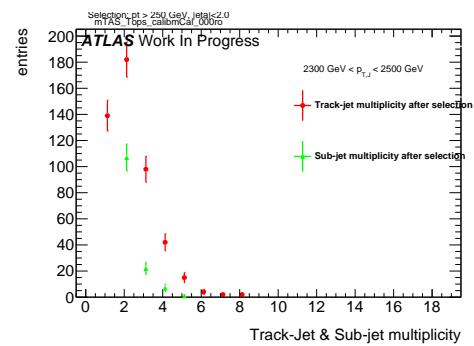


Figure 135: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

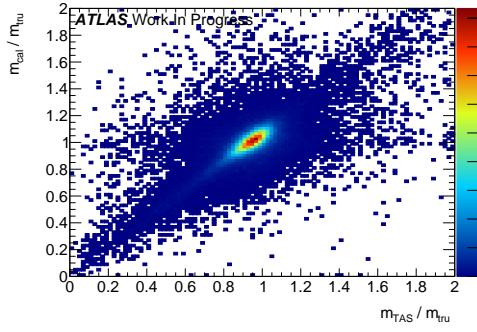
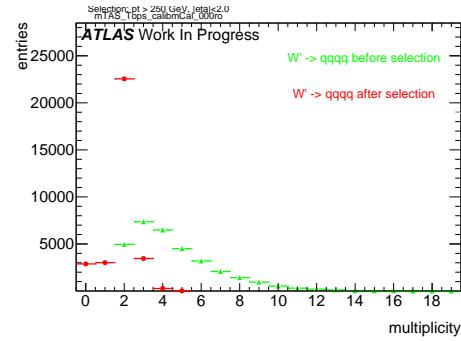
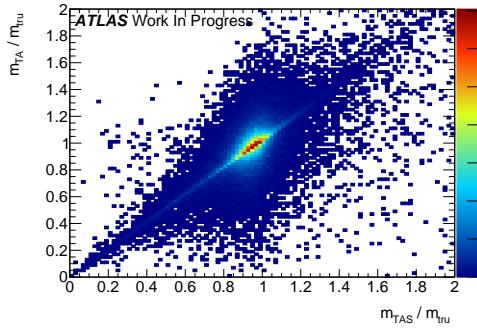
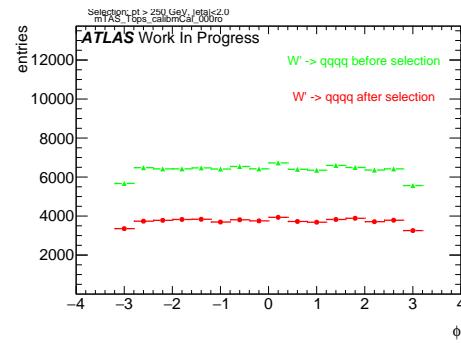
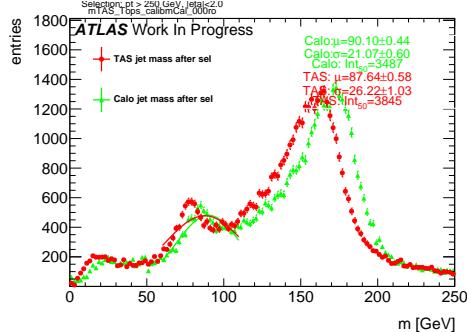
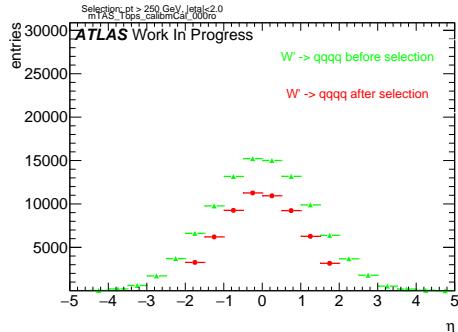
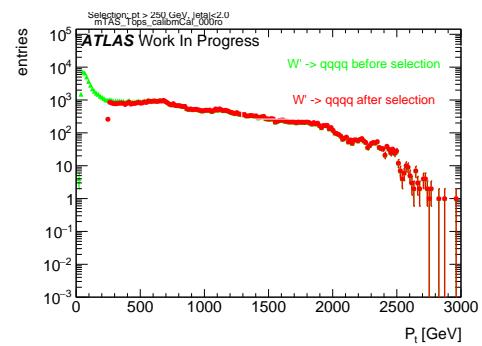
Figure 136: Scatter plot  $m^{TAS}$  versus  $m^{calo}$  responses

Figure 140: large-R jet Multiplicity, before and after selection

Figure 137: Scatter plot  $m^{TAS}$  versus  $m^A$  responsesFigure 141:  $\phi$  distribution of the large-R jet, before and after selectionFigure 138:  $m^{TAS}$  distribution in all the  $p_T$  binsFigure 139:  $\eta$  distribution of the large-R jet, before and after selectionFigure 142:  $p_T$  distribution of the large-R jet, before and after selection

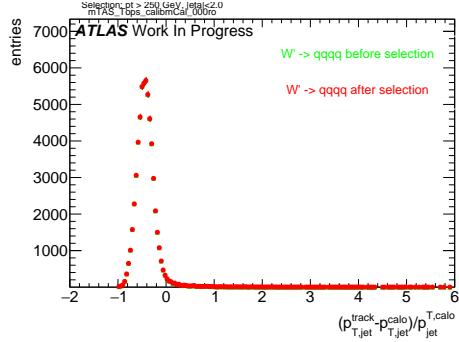


Figure 143:  $p_T$  resolution:  $\frac{p_{T,jet}^{track} - p_{T,jet}^{fat}}{p_{T,jet}^{fat}}$ , before and after selection

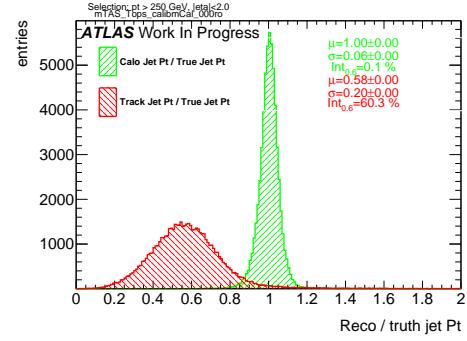


Figure 146: Transverse momentum response  $p_T^{Reco}/p_T^{Truth}$  for calorimeter and tracks

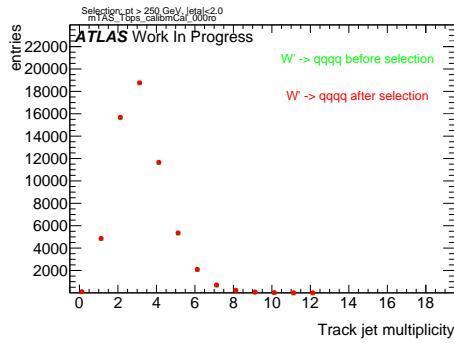


Figure 144: Multiplicity of track-jets R=0.2 per large-R jet

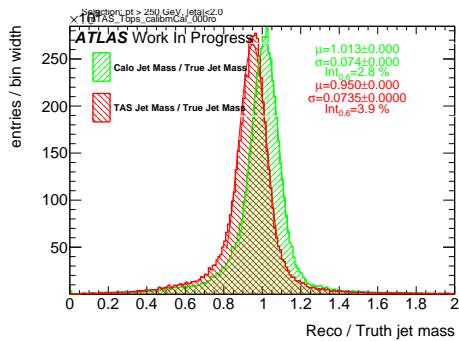


Figure 145: Response  $m^{Reco}/m^{Truth}$  for all the  $p_T$  bins

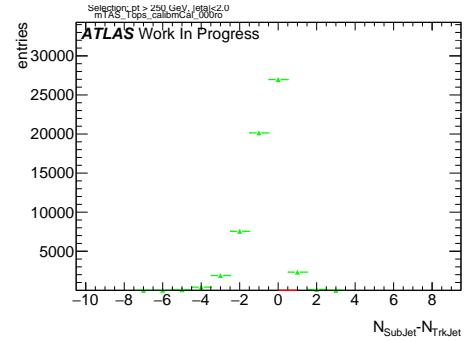


Figure 147: sub-jet - track-jet Multiplicity

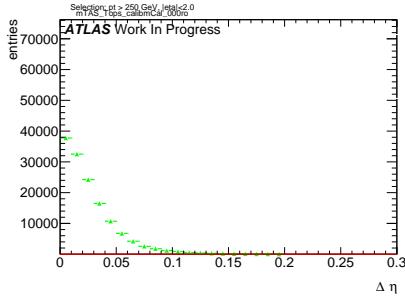


Figure 148:  $|\eta_{\text{sub-jet}} - \eta_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

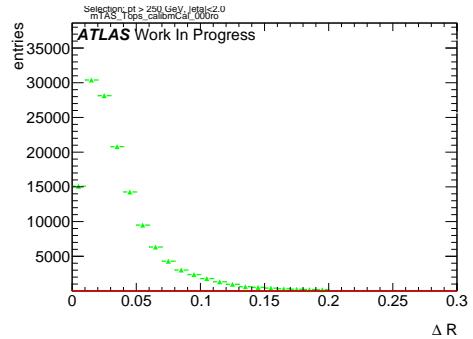


Figure 152:  $|R_{\text{sub-jet}} - R_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

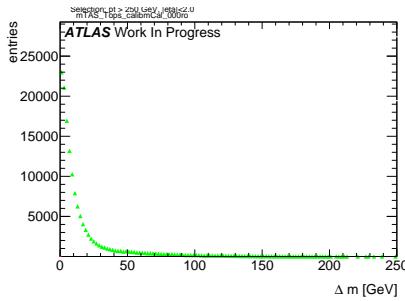


Figure 149:  $|m_{\text{sub-jet}} - m_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

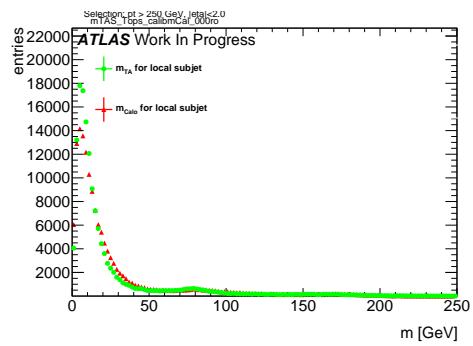


Figure 153: Mass distribution of the sub-jet, calorimeter and track-assisted

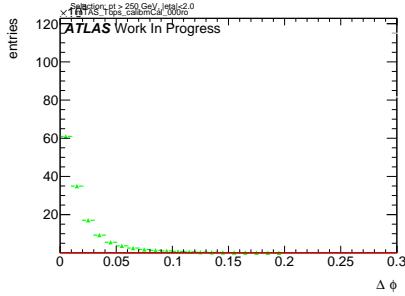


Figure 150:  $|\phi_{\text{sub-jet}} - \phi_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

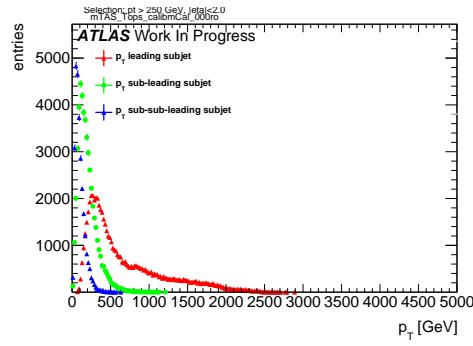


Figure 154:  $p_T$  distribution for leading, sub-leading and sub-sub-leading sub-jets

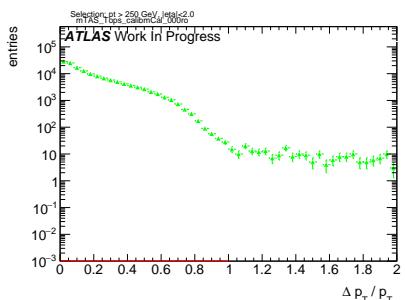


Figure 151:  $|p_{T,\text{sub-jet}} - p_{T,\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

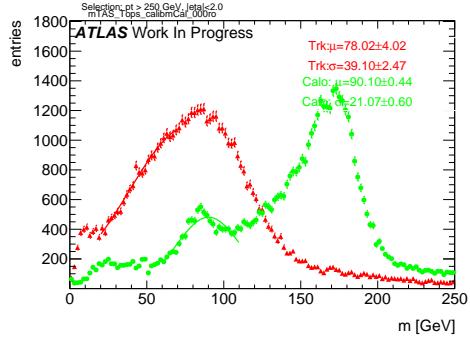


Figure 155: Mass distribution for calorimeter and tracks associated to the large-R jet

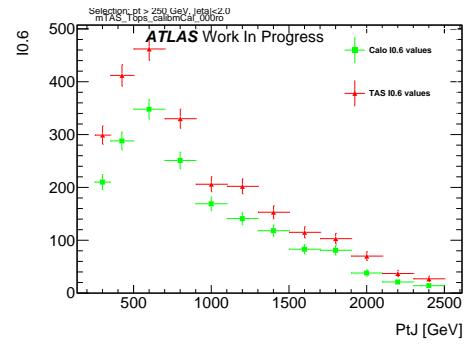


Figure 158: Left integral,  $\int_0^{0.6}$  of the mass response, vs bin of  $p_T^J$

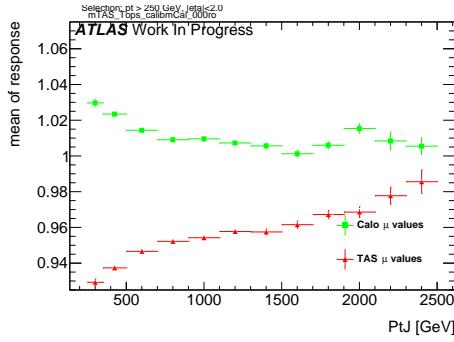


Figure 156:  $\mu$  from fit of the mass Response vs bin of  $p_T^J$

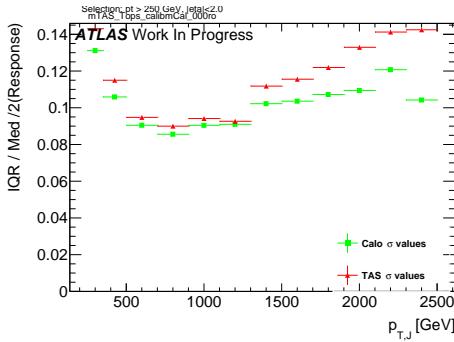


Figure 157:  $\sigma$  from fit of the mass Response vs bin of  $p_T^J$

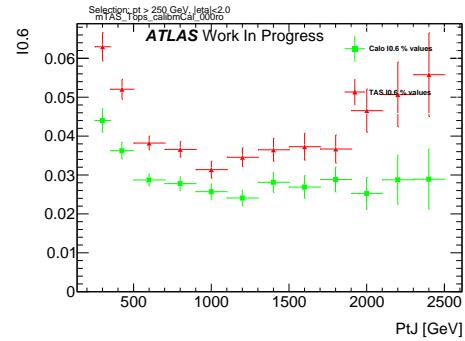
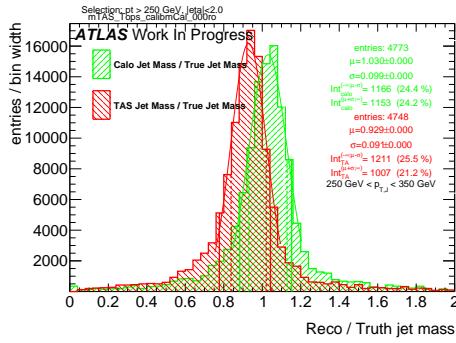
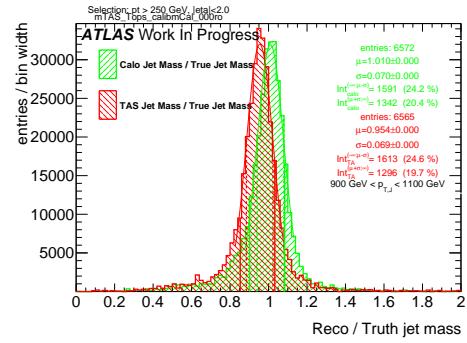
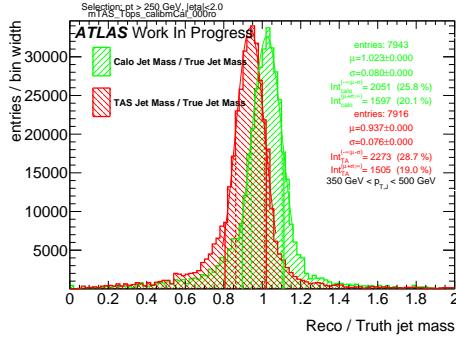
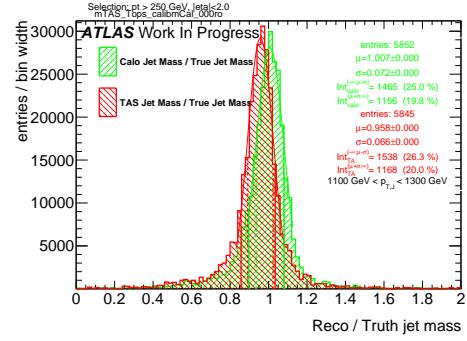
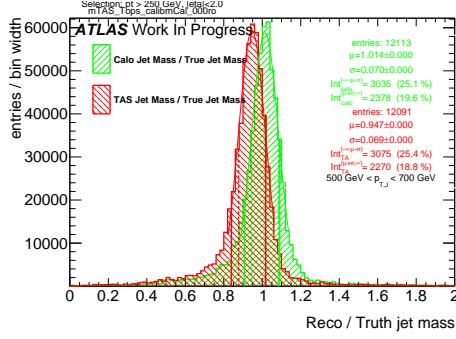
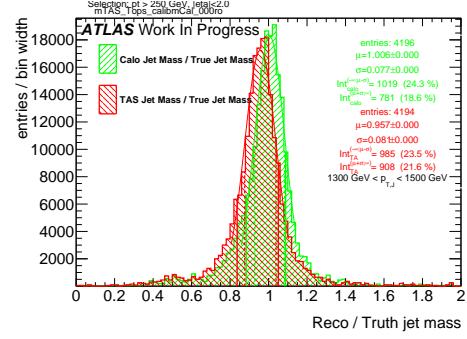
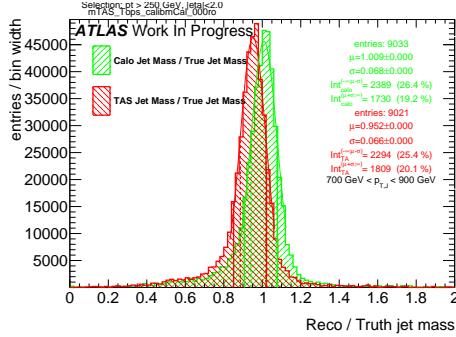
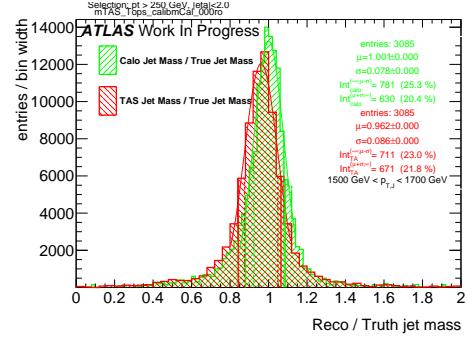
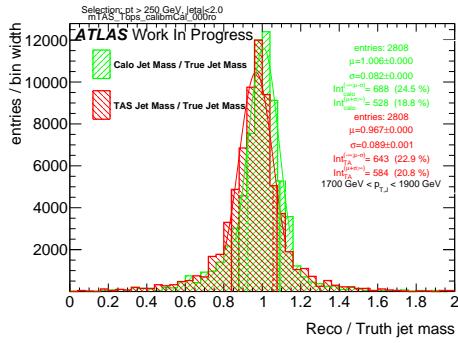
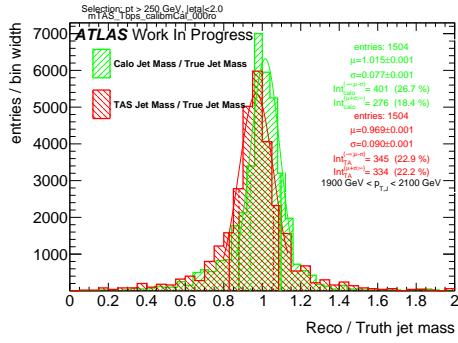
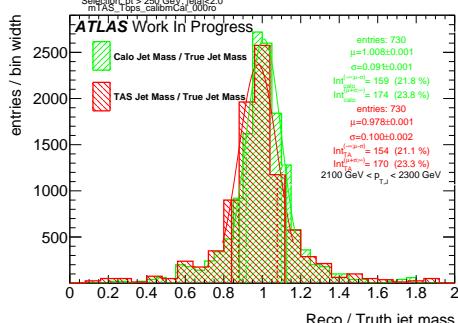
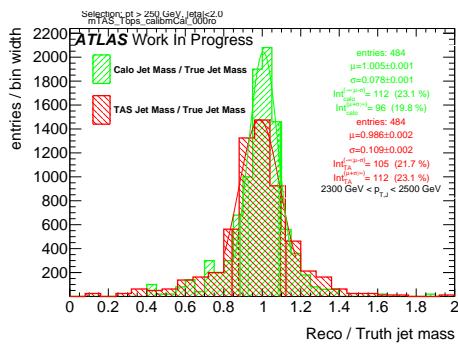


Figure 159: Left integral normalized,  $\int_0^{0.6}$  of the mass response, vs bin of  $p_T^J$

Figure 160: Response in bin of  $p_T^J$  (indicated on plot)Figure 164: Response in bin of  $p_T^J$  (indicated on plot)Figure 161: Response in bin of  $p_T^J$  (indicated on plot)Figure 165: Response in bin of  $p_T^J$  (indicated on plot)Figure 162: Response in bin of  $p_T^J$  (indicated on plot)Figure 166: Response in bin of  $p_T^J$  (indicated on plot)Figure 163: Response in bin of  $p_T^J$  (indicated on plot)Figure 167: Response in bin of  $p_T^J$  (indicated on plot)

Figure 168: Response in bin of  $p_T^J$  (indicated on plot)Figure 169: Response in bin of  $p_T^J$  (indicated on plot)Figure 170: Response in bin of  $p_T^J$  (indicated on plot)Figure 171: Response in bin of  $p_T^J$  (indicated on plot)

873 **G  $m^{TAS}$  distributions, boosted higgs**

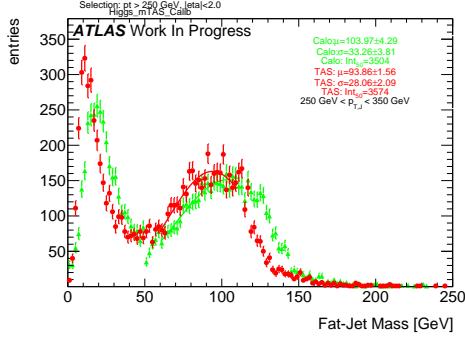


Figure 172:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

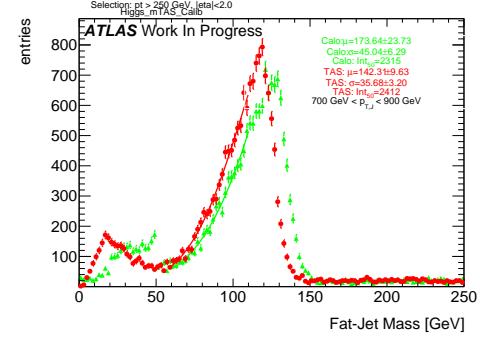


Figure 175:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

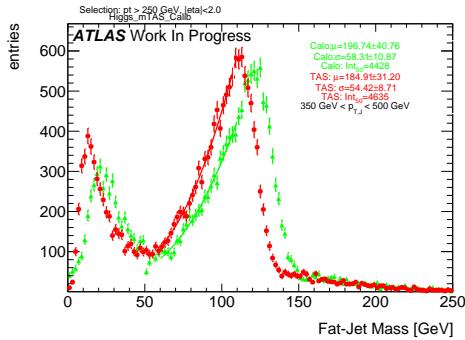


Figure 173:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

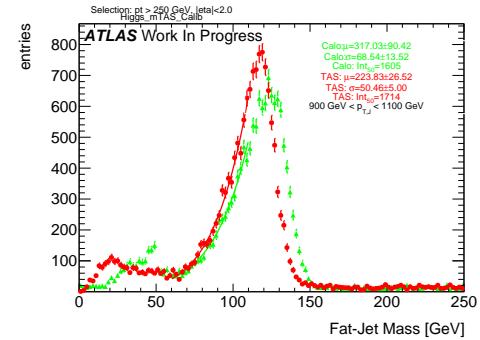


Figure 176:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

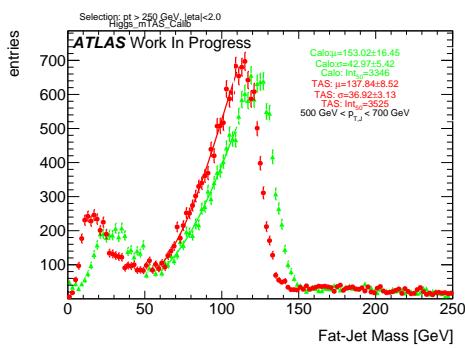


Figure 174:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

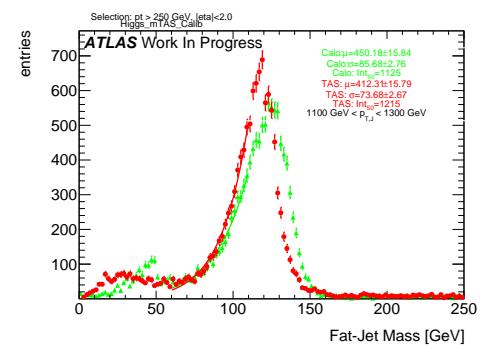


Figure 177:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

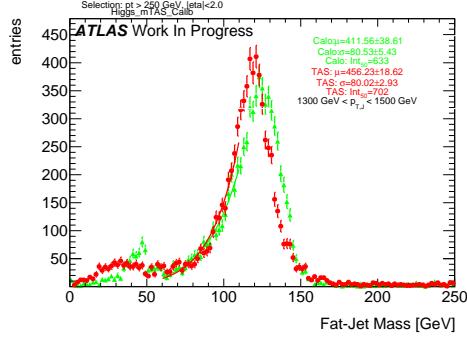


Figure 178:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

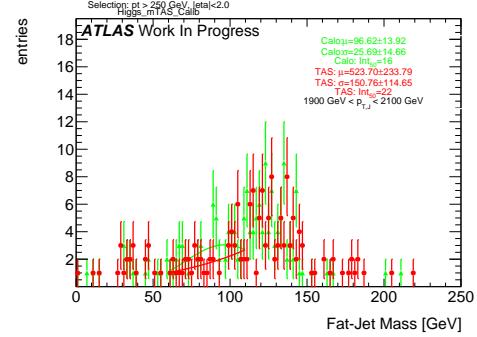


Figure 181:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

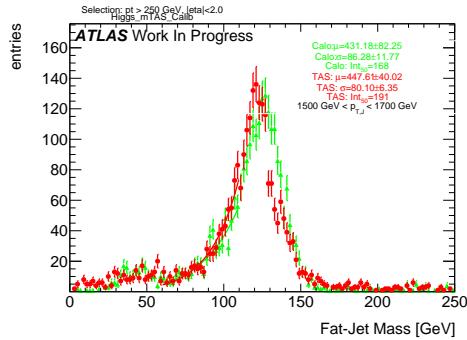


Figure 179:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

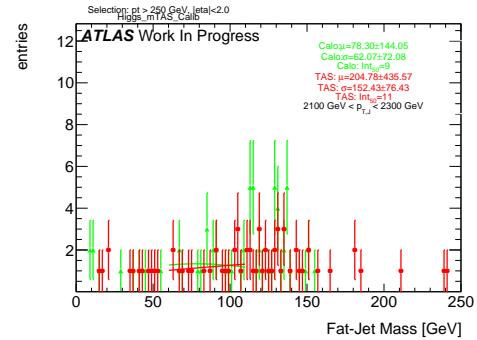


Figure 182:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

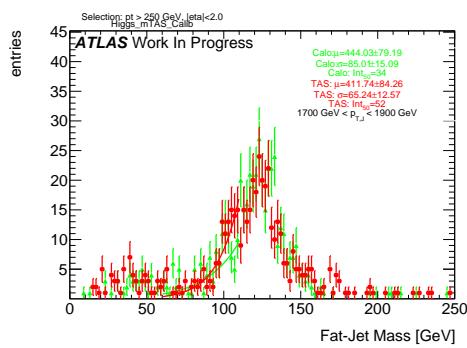


Figure 180:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

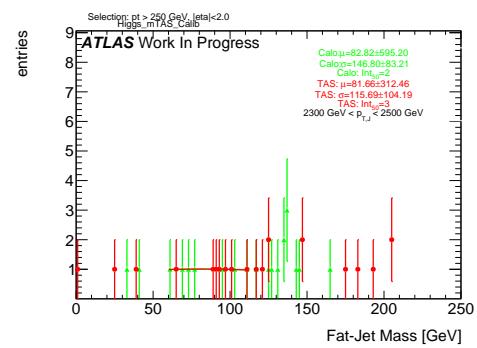


Figure 183:  $m^{TAS}$  and  $m^{calo}$  for  $p_T^J$  bin (indicated on plot)

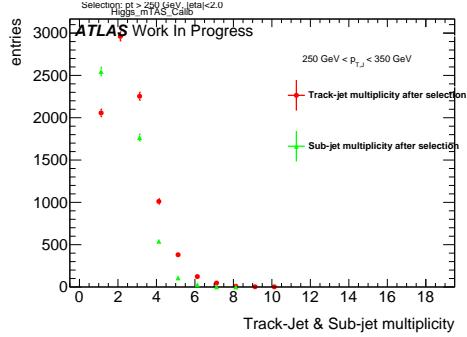


Figure 184: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

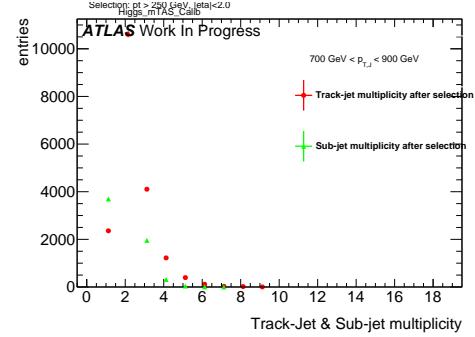


Figure 187: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

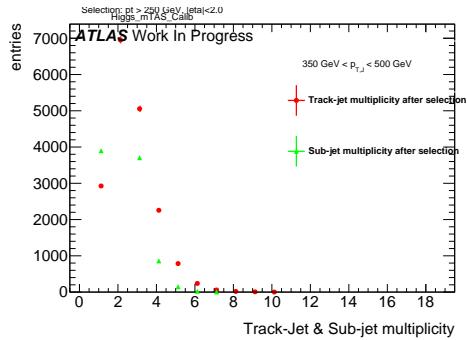


Figure 185: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

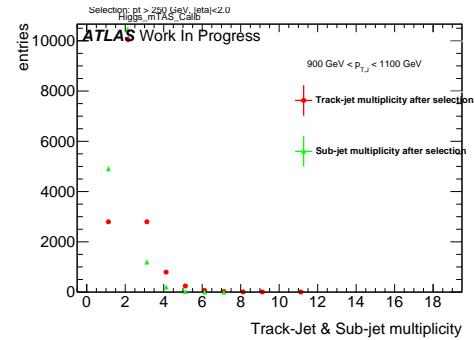


Figure 188: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

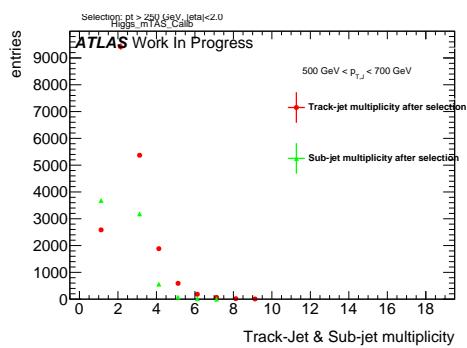


Figure 186: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

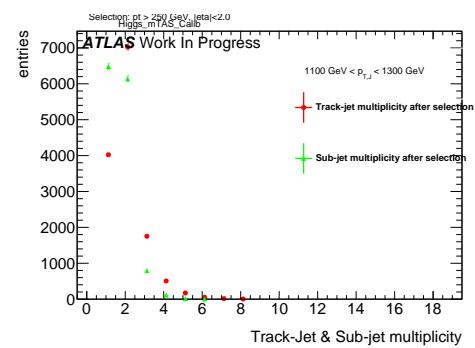


Figure 189: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

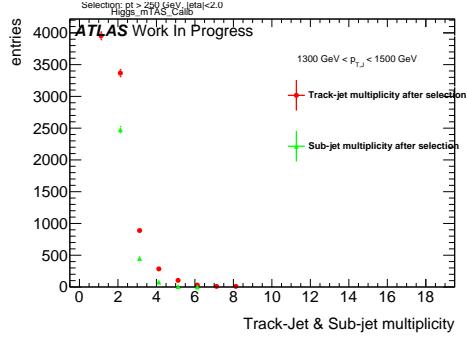


Figure 190: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

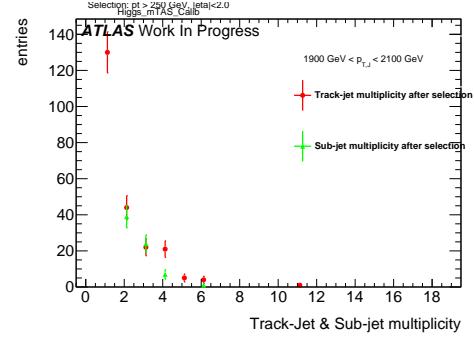


Figure 193: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

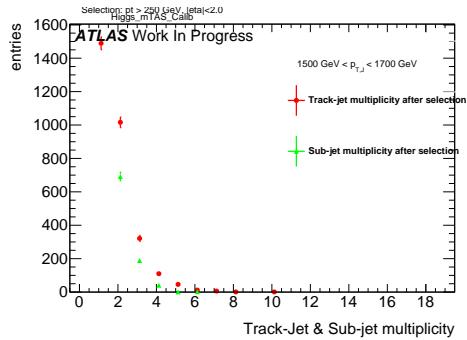


Figure 191: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

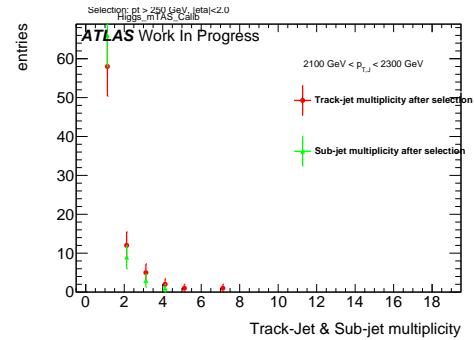


Figure 194: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

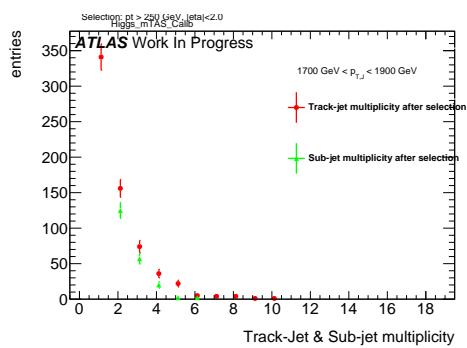


Figure 192: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

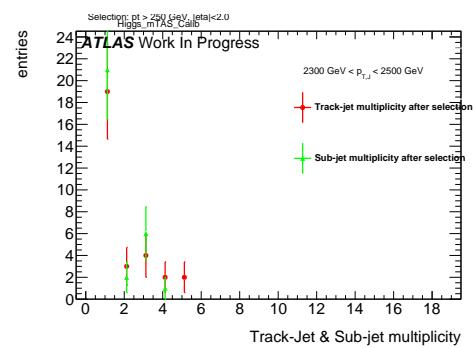


Figure 195: Track-jet R=0.2 and sub-jet multiplicity for  $p_T^J$  bin (indicated on plot)

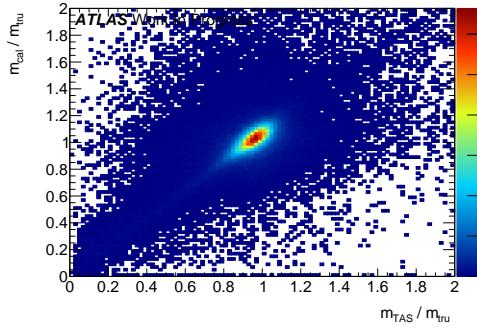
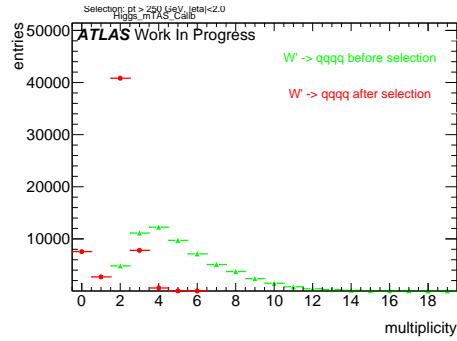
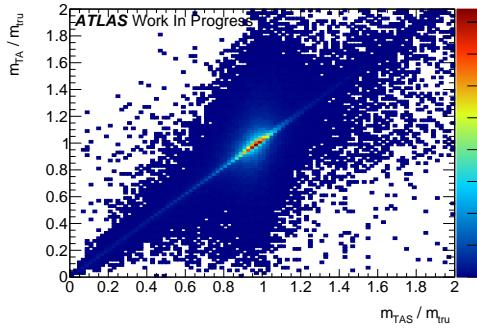
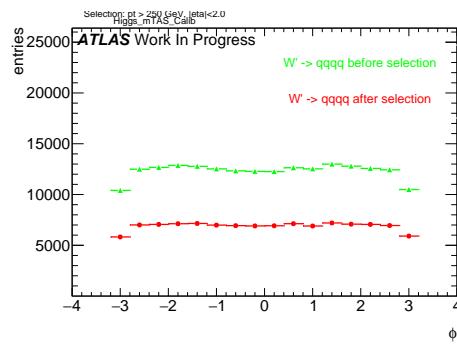
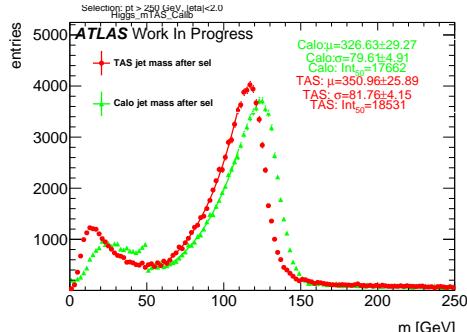
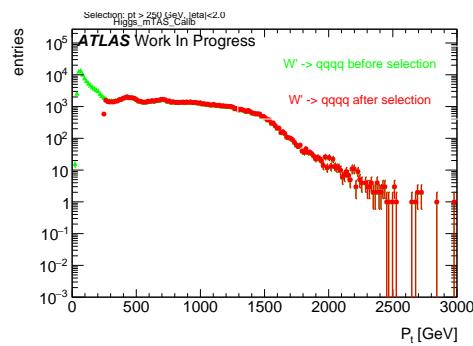
Figure 196: Scatter plot  $m^{TAS}$  versus  $m^{calo}$  responses

Figure 200: large-R jet Multiplicity, before and after selection

Figure 197: Scatter plot  $m^{TAS}$  versus  $m^T$  responsesFigure 201:  $\phi$  distribution of the large-R jet, before and after selectionFigure 198:  $m^{TAS}$  distribution in all the  $p_T$  binsFigure 202:  $p_T$  distribution of the large-R jet, before and after selectionFigure 199:  $\eta$  distribution of the large-R jet, before and after selection

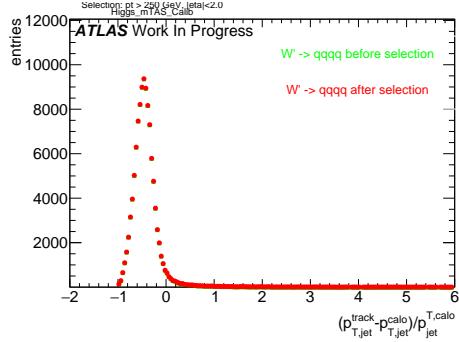


Figure 203:  $p_T$  resolution:  $\frac{p_{T,jet}^{track} - p_{T,jet}^{fat}}{p_{T,jet}^{fat}}$ , before and after selection

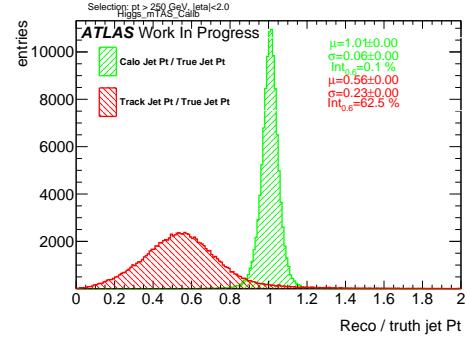


Figure 206: Transverse momentum response  $p_T^{Reco} / p_T^{Truth}$  for calorimeter and tracks

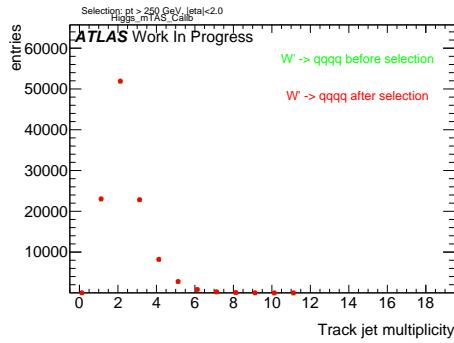


Figure 204: Multiplicity of track-jets R=0.2 per large-R jet

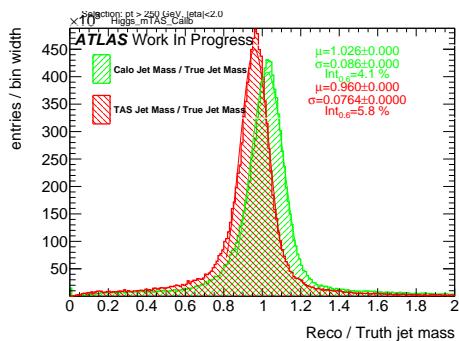


Figure 205: Response  $m^{Reco} / m^{Truth}$  for all the  $p_T$  bins

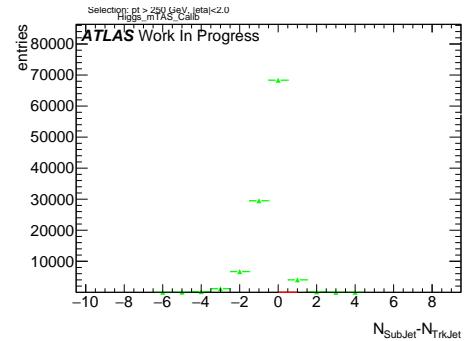


Figure 207: sub-jet - track-jet Multiplicity

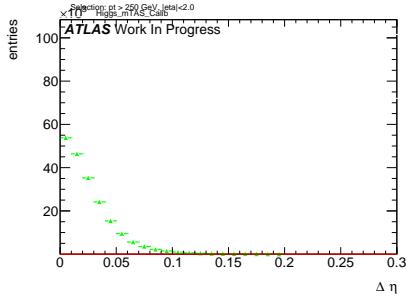


Figure 208:  $|\eta_{\text{sub-jet}} - \eta_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

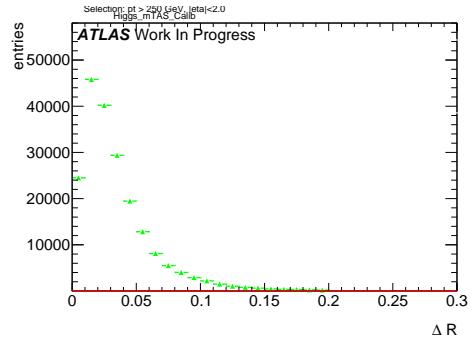


Figure 212:  $|R_{\text{sub-jet}} - R_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

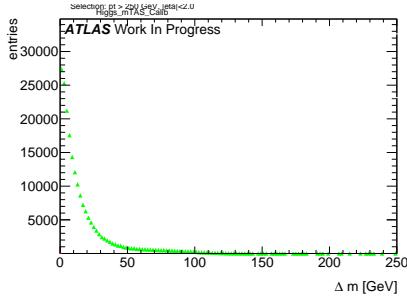


Figure 209:  $|m_{\text{sub-jet}} - m_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

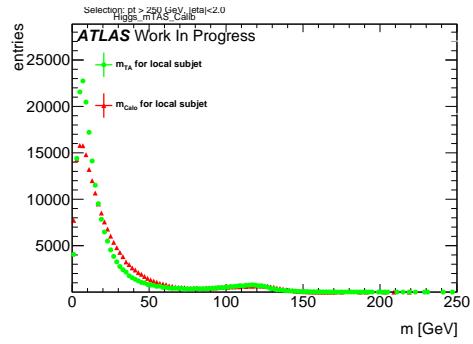


Figure 213: Mass distribution of the sub-jet, calorimeter and track-assisted

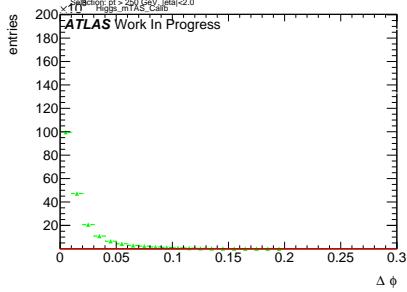


Figure 210:  $|\phi_{\text{sub-jet}} - \phi_{\text{track-jet}}|$  distribution, where sub-jet and track-jet are the closest

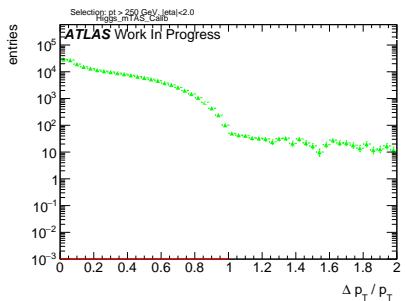


Figure 211:  $|p_{\text{T,sub-jet}} - p_{\text{T,track-jet}}|$  distribution, where sub-jet and track-jet are the closest

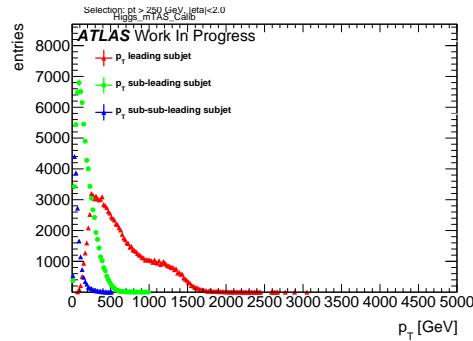


Figure 214:  $p_{\text{T}}$  distribution for leading, sub-leading and sub-sub-leading sub-jets

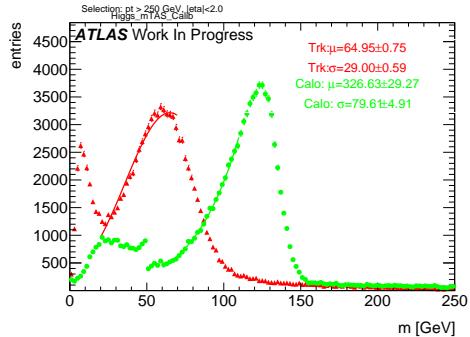


Figure 215: Mass distribution for calorimeter and tracks associated to the large-R jet

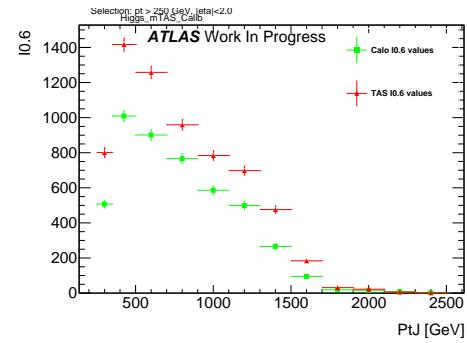


Figure 218: Left integral,  $\int_0^{0.6}$  of the mass response, vs bin of  $p_T^J$

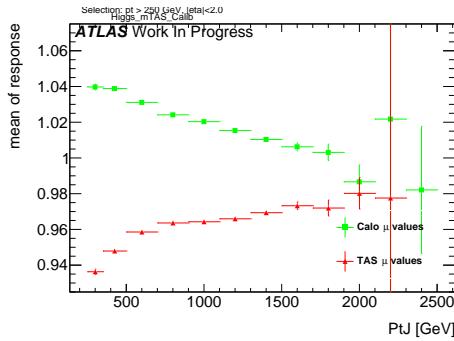


Figure 216:  $\mu$  from fit of the mass Response vs bin of  $p_T^J$

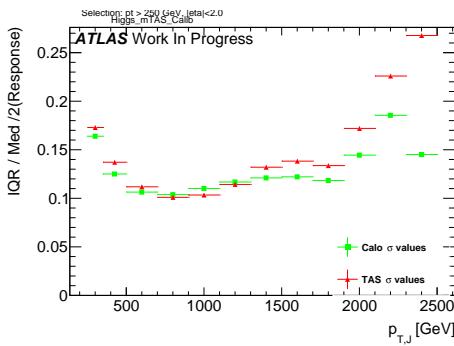


Figure 217:  $\sigma$  from fit of the mass Response vs bin of  $p_T^J$

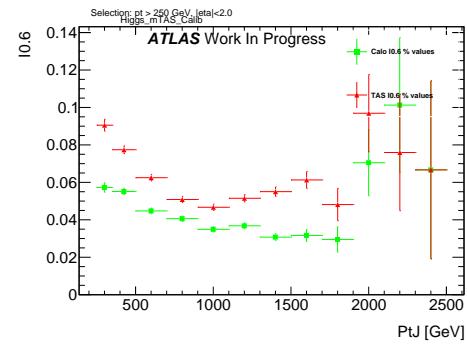
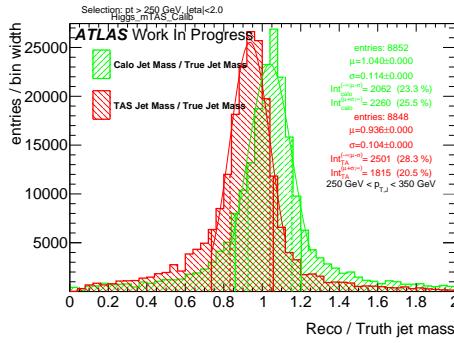
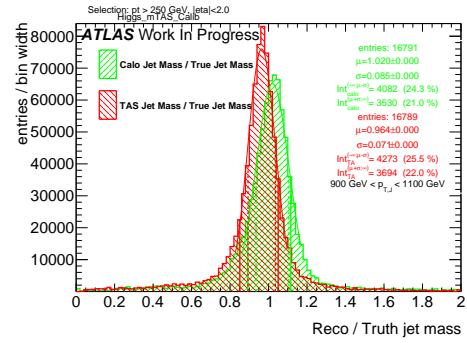
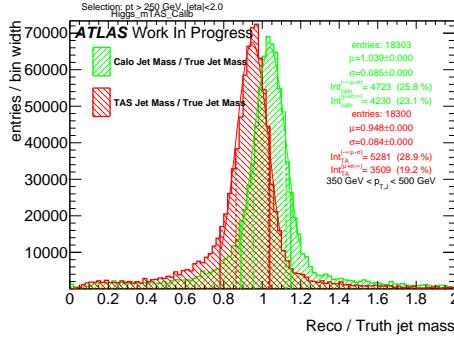
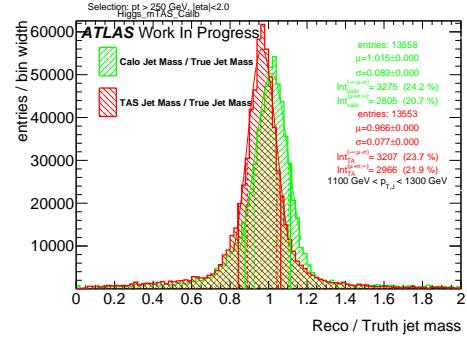
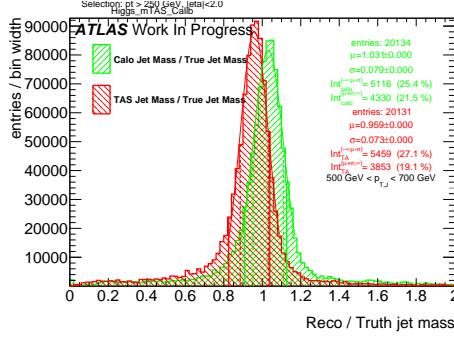
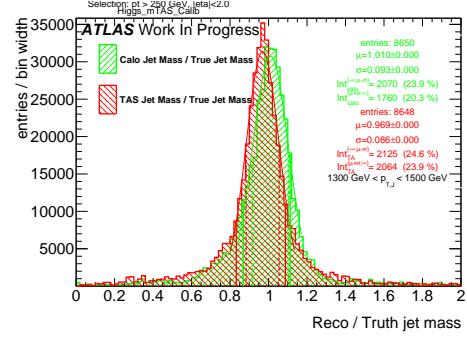
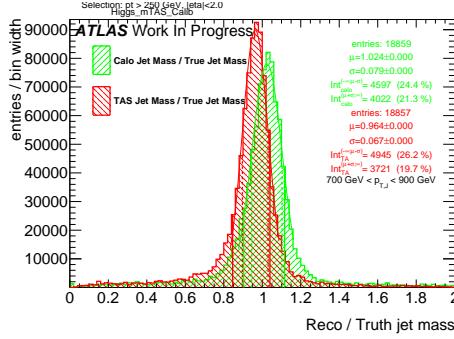
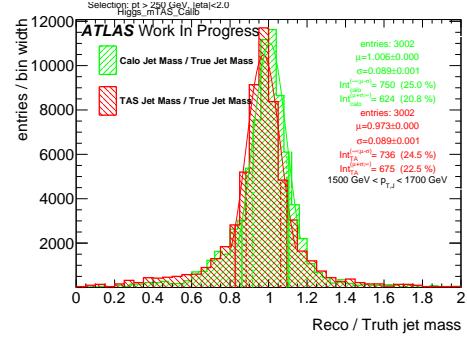


Figure 219: Left integral normalized,  $\int_0^{0.6}$  of the mass response, vs bin of  $p_T^J$

Figure 220: Response in bin of  $p_T^J$  (indicated on plot)Figure 224: Response in bin of  $p_T^J$  (indicated on plot)Figure 221: Response in bin of  $p_T^J$  (indicated on plot)Figure 225: Response in bin of  $p_T^J$  (indicated on plot)Figure 222: Response in bin of  $p_T^J$  (indicated on plot)Figure 226: Response in bin of  $p_T^J$  (indicated on plot)Figure 223: Response in bin of  $p_T^J$  (indicated on plot)Figure 227: Response in bin of  $p_T^J$  (indicated on plot)

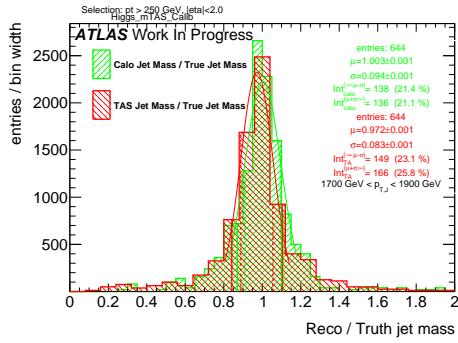


Figure 228: Response in bin of  $p_T^J$  (indicated on plot)

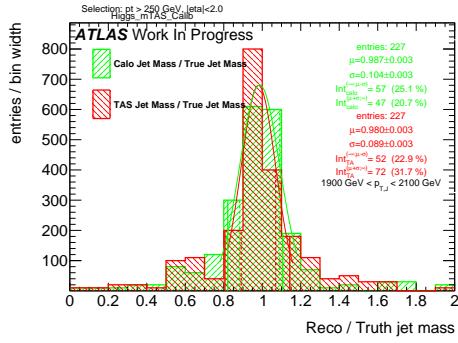


Figure 229: Response in bin of  $p_T^J$  (indicated on plot)

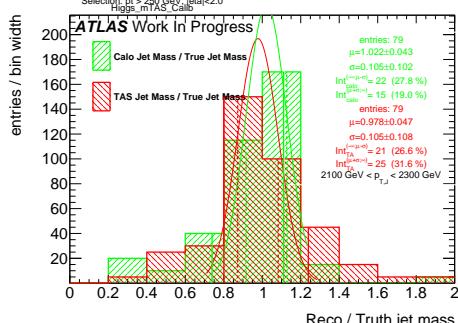


Figure 230: Response in bin of  $p_T^J$  (indicated on plot)

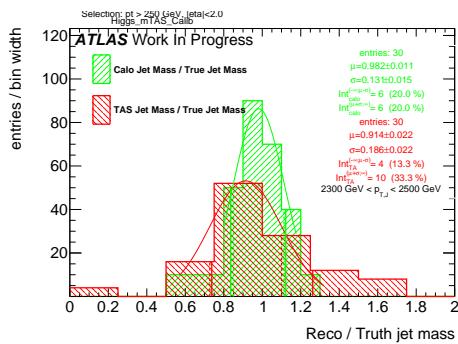
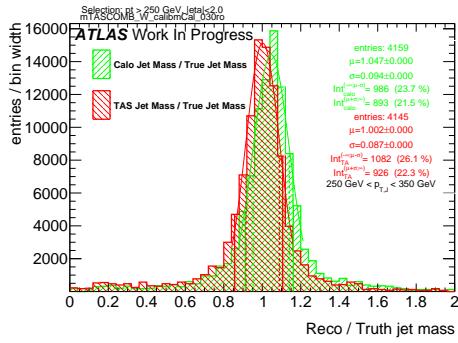
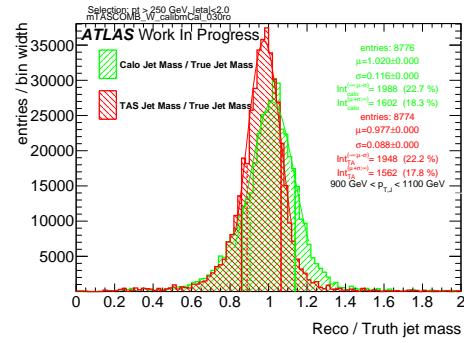
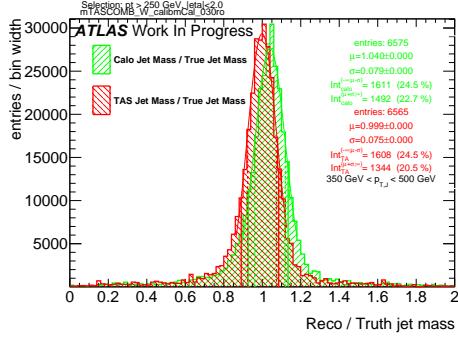
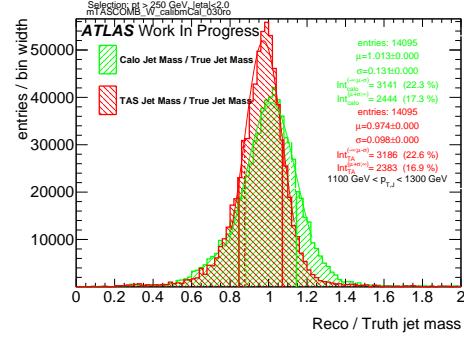
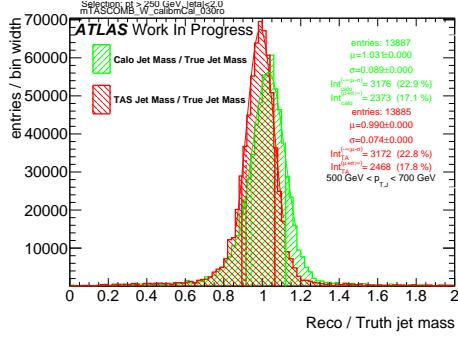
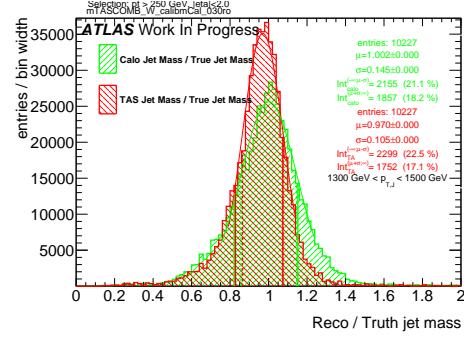
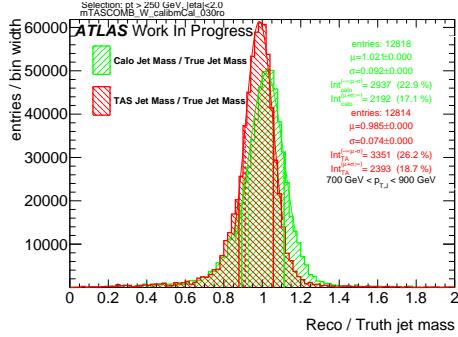
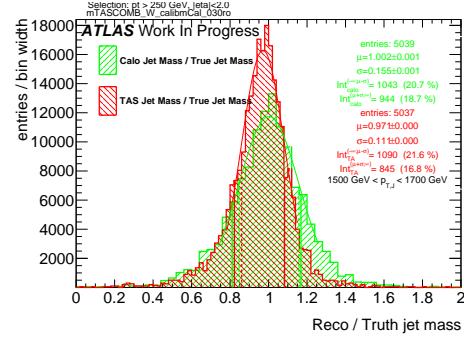
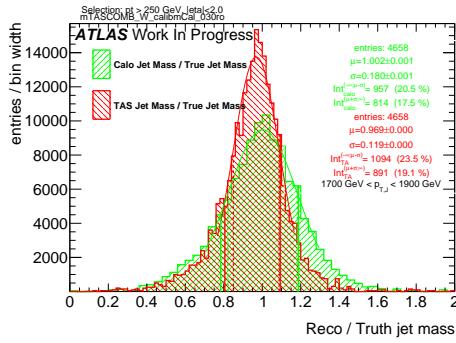
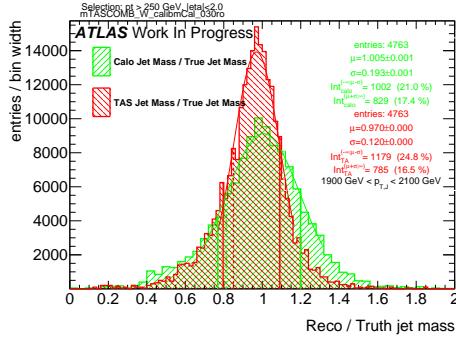
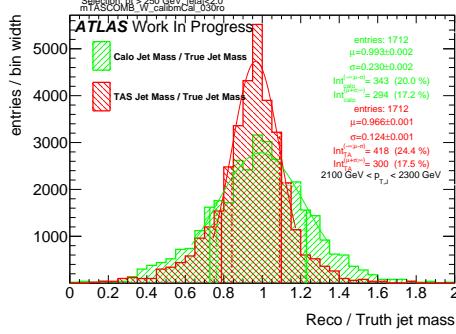
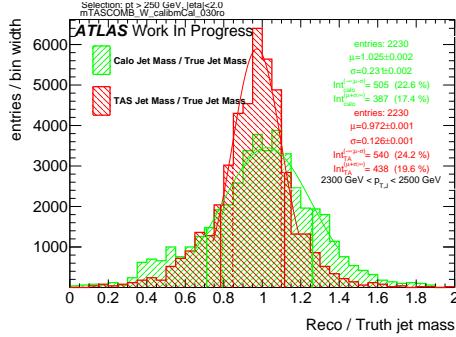


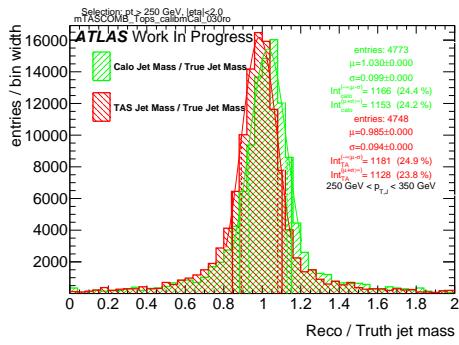
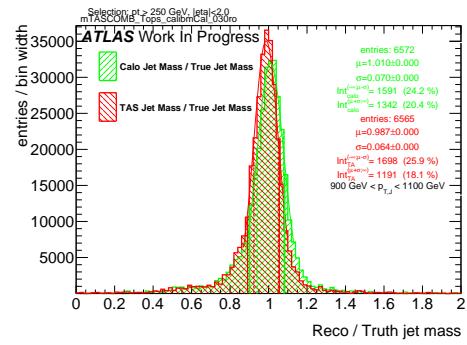
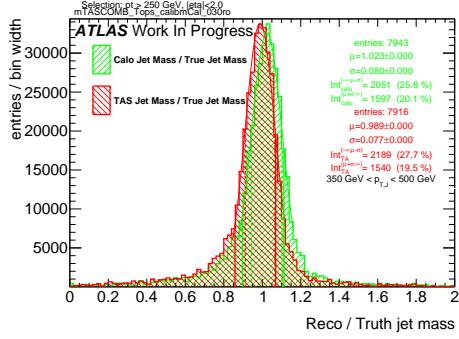
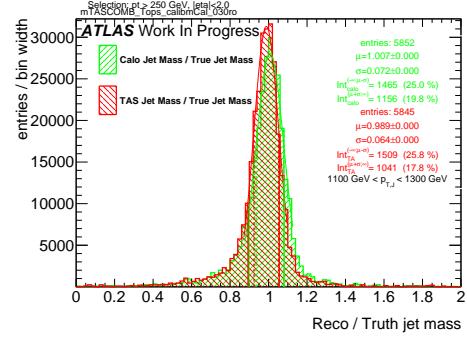
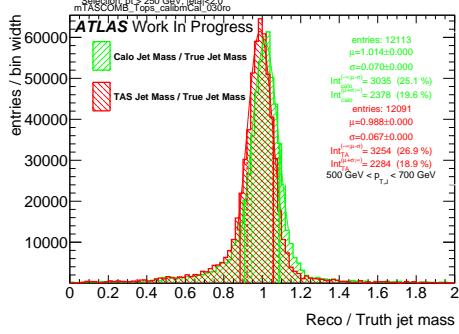
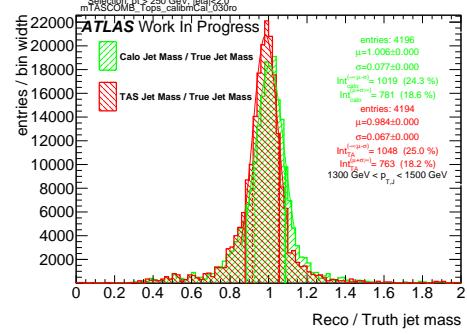
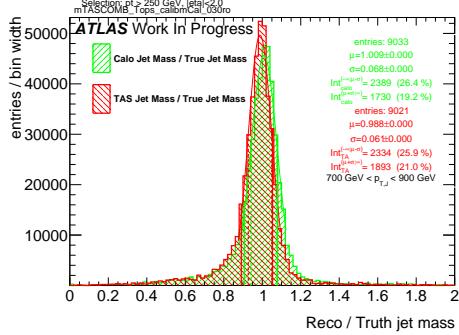
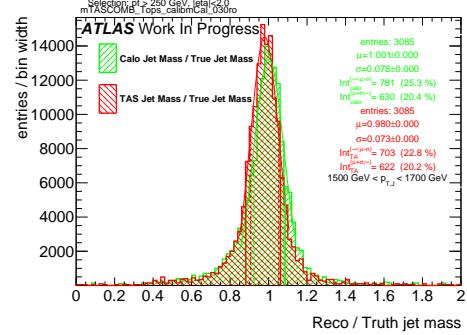
Figure 231: Response in bin of  $p_T^J$  (indicated on plot)

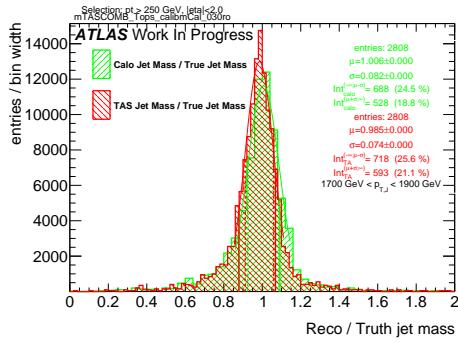
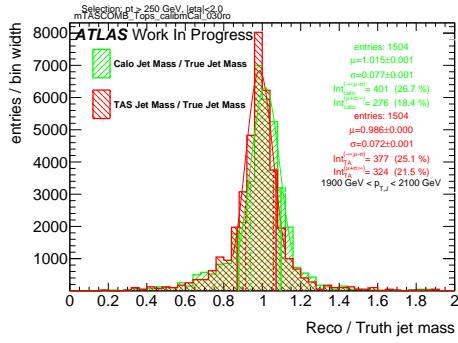
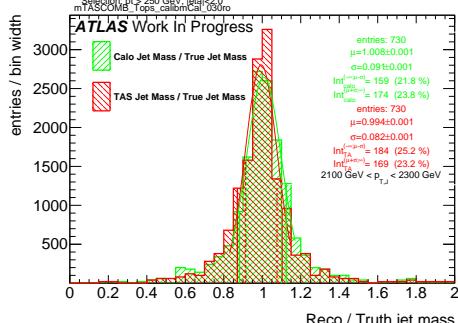
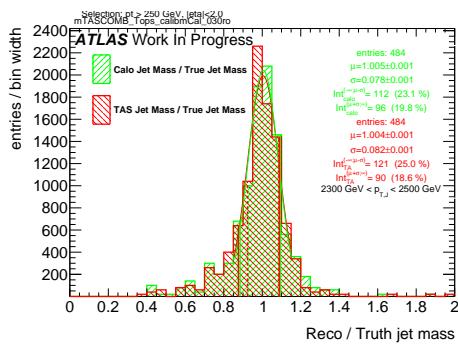
<sup>874</sup> H  $m_{TAS}^{comb}$  response distributions, boosted W/Z

Figure 232: Response in bin of  $p_T^J$  (indicated on plot)Figure 236: Response in bin of  $p_T^J$  (indicated on plot)Figure 233: Response in bin of  $p_T^J$  (indicated on plot)Figure 237: Response in bin of  $p_T^J$  (indicated on plot)Figure 234: Response in bin of  $p_T^J$  (indicated on plot)Figure 238: Response in bin of  $p_T^J$  (indicated on plot)Figure 235: Response in bin of  $p_T^J$  (indicated on plot)Figure 239: Response in bin of  $p_T^J$  (indicated on plot)

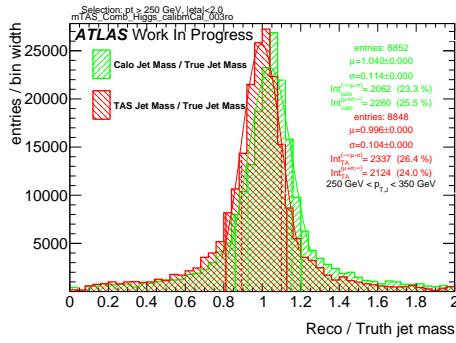
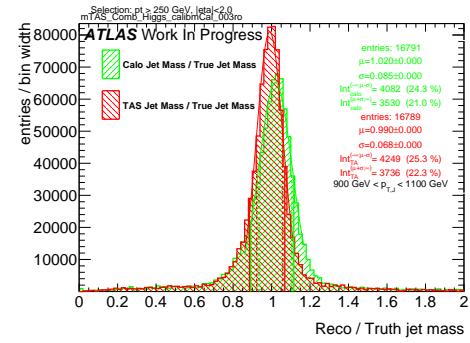
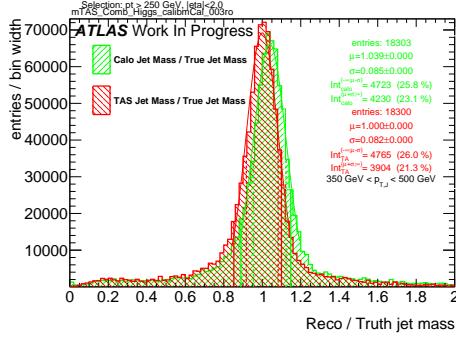
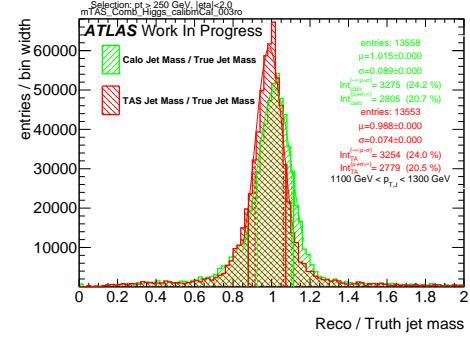
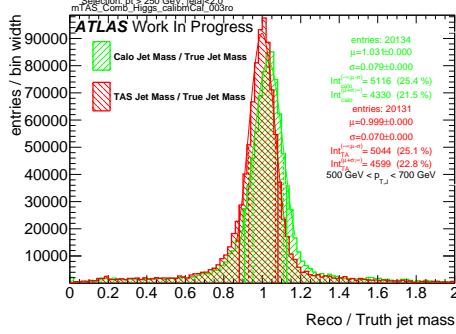
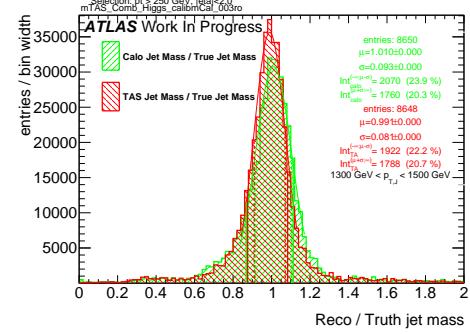
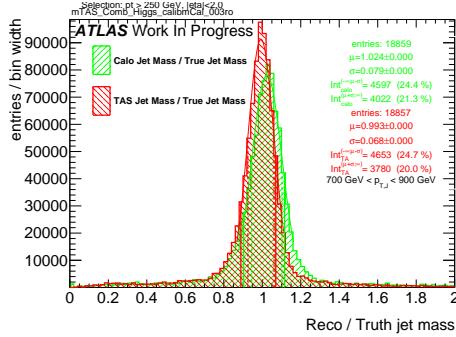
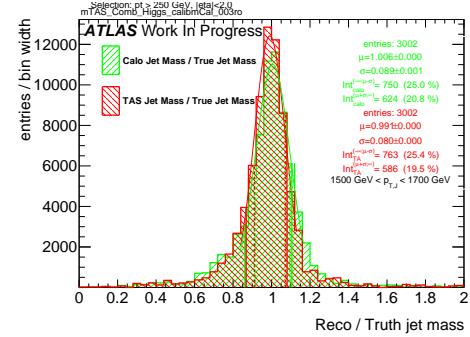
Figure 240: Response in bin of  $p_T^J$  (indicated on plot)Figure 241: Response in bin of  $p_T^J$  (indicated on plot)Figure 242: Response in bin of  $p_T^J$  (indicated on plot)Figure 243: Response in bin of  $p_T^J$  (indicated on plot)

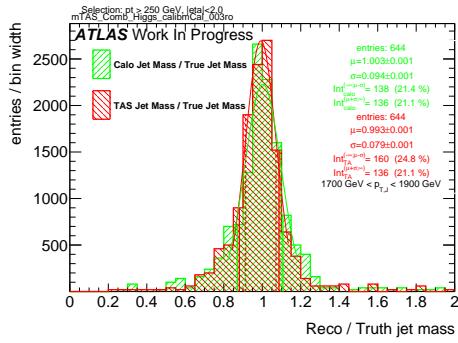
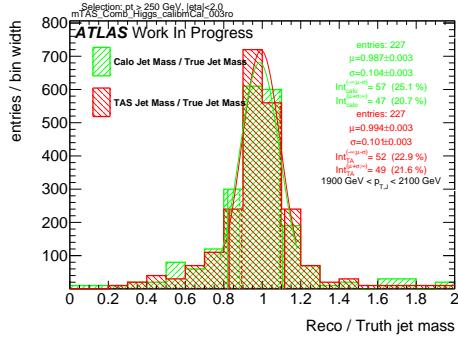
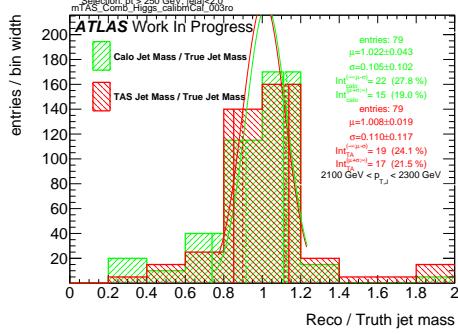
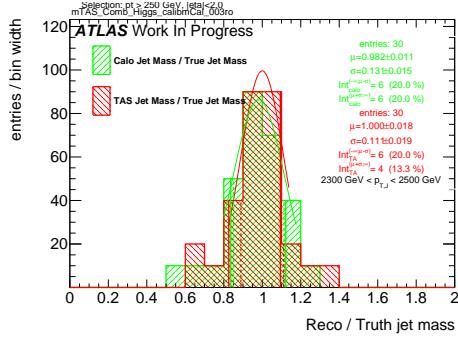
<sup>875</sup> I  $m_{TAS}^{comb}$  response distributions, boosted tops

Figure 244: Response in bin of  $p_T^J$  (indicated on plot)Figure 248: Response in bin of  $p_T^J$  (indicated on plot)Figure 245: Response in bin of  $p_T^J$  (indicated on plot)Figure 249: Response in bin of  $p_T^J$  (indicated on plot)Figure 246: Response in bin of  $p_T^J$  (indicated on plot)Figure 250: Response in bin of  $p_T^J$  (indicated on plot)Figure 247: Response in bin of  $p_T^J$  (indicated on plot)Figure 251: Response in bin of  $p_T^J$  (indicated on plot)

Figure 252: Response in bin of  $p_T^J$  (indicated on plot)Figure 253: Response in bin of  $p_T^J$  (indicated on plot)Figure 254: Response in bin of  $p_T^J$  (indicated on plot)Figure 255: Response in bin of  $p_T^J$  (indicated on plot)

<sup>876</sup> **J  $m_{TAS}^{comb}$  response distributions, Higgs**

Figure 256: Response in bin of  $p_T^J$  (indicated on plot)Figure 260: Response in bin of  $p_T^J$  (indicated on plot)Figure 257: Response in bin of  $p_T^J$  (indicated on plot)Figure 261: Response in bin of  $p_T^J$  (indicated on plot)Figure 258: Response in bin of  $p_T^J$  (indicated on plot)Figure 262: Response in bin of  $p_T^J$  (indicated on plot)Figure 259: Response in bin of  $p_T^J$  (indicated on plot)Figure 263: Response in bin of  $p_T^J$  (indicated on plot)

Figure 264: Response in bin of  $p_T^J$  (indicated on plot)Figure 265: Response in bin of  $p_T^J$  (indicated on plot)Figure 266: Response in bin of  $p_T^J$  (indicated on plot)Figure 267: Response in bin of  $p_T^J$  (indicated on plot)

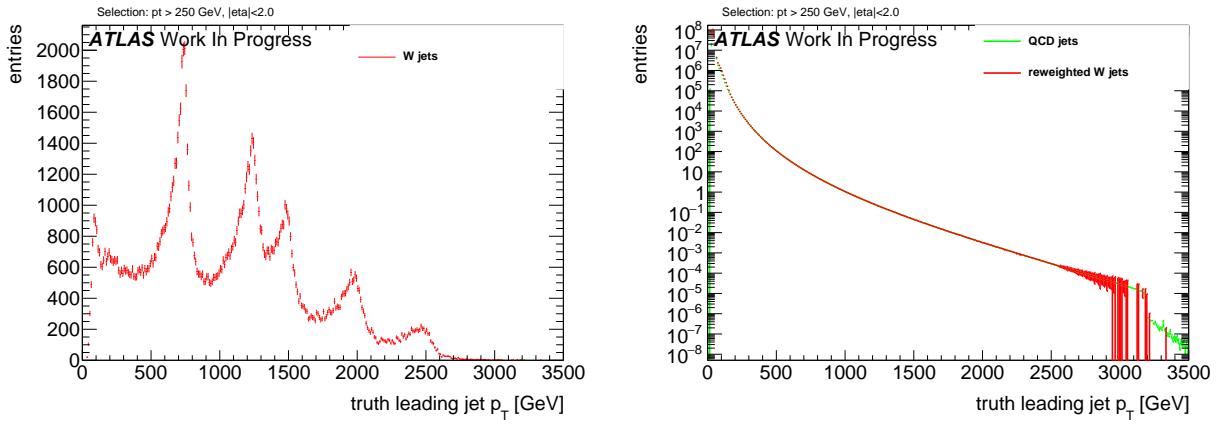


Figure 268: Exemplary  $p_T$  distributions of  $W$  jets (left) and QCD jets from multi-jet events with reweighted  $W$  boson events (right).

## 877 K $p_T$ reweighting

## 878 L $p_T$ Dependence of Substructure Observables

879 Due to the low weights for high  $p_T$ , the correlation plots are divided into the six  $p_T$  regions. For C2, see  
 880 Figure 269, observed is a strong trend to lower values (signal and background) for clusters and TAS. The  
 881 TAS distributions concentrate at lower values compared to calorimeter counterparts.

882 For D2, Figure 270, and  $\tau_{21}$ , Figure 271, there is a slight upward trend of the calorimeter variables in the  
 883 lower  $p_T$  regions. With rising boost this slows down and ends in a broader distribution for  $\tau_{21}$ . This verifies  
 884 the higher  $p_T$  dependence of the C2 variable in comparison to D2 and  $\tau_{21}$ . The TAS counterparts feature  
 885 an even more robust signal with the background moving to higher values, hence improving separation.  
 886 The  $p_T$  dependence of variables calculated with tracks is very similar to the ones with TAS, therefore they  
 are omitted.

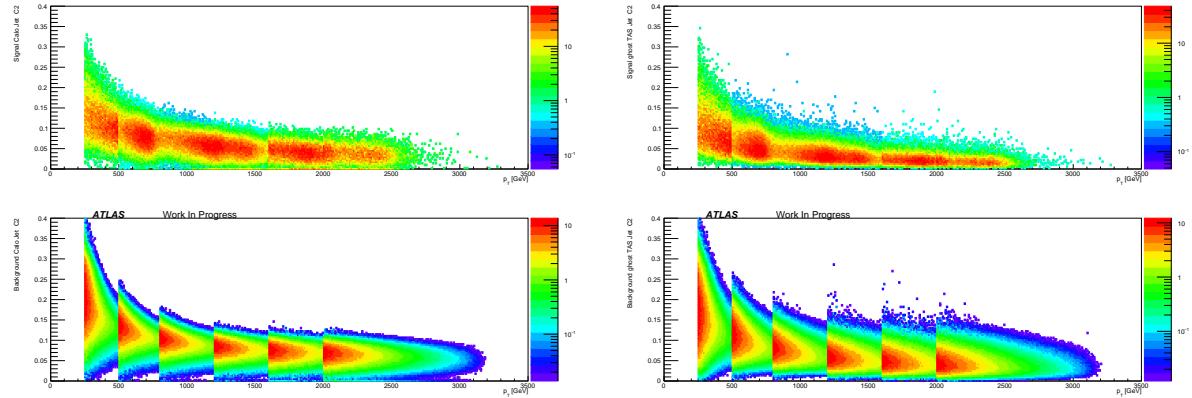


Figure 269: Correlation between C2 at  $\beta = 1$  and  $p_T$  applied on  $W$  boson signal (above) and QCD background (below) for calorimeter (left) and TAS (right).

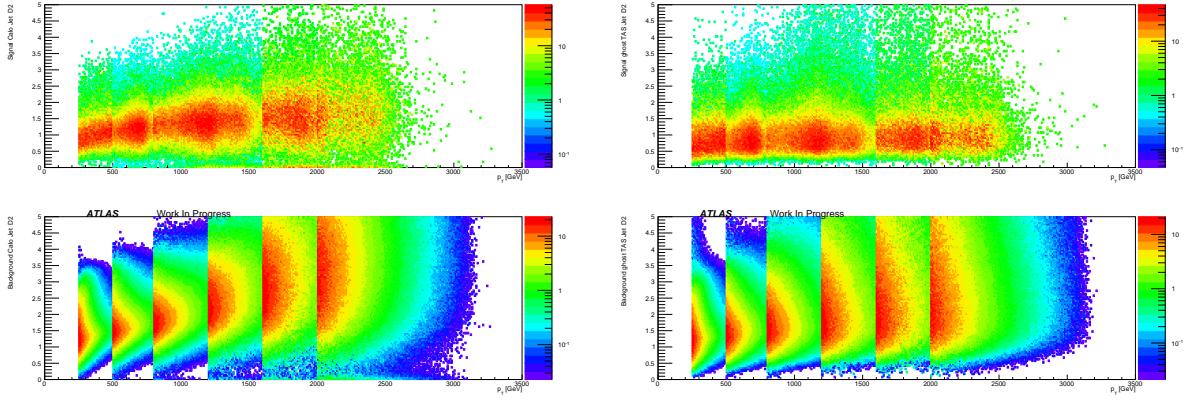


Figure 270: Correlation between D2 at  $\beta = 1$  and  $p_T$  applied on  $W$  boson signal (above) and QCD background (below) for calorimeter (left) and TAS (right).

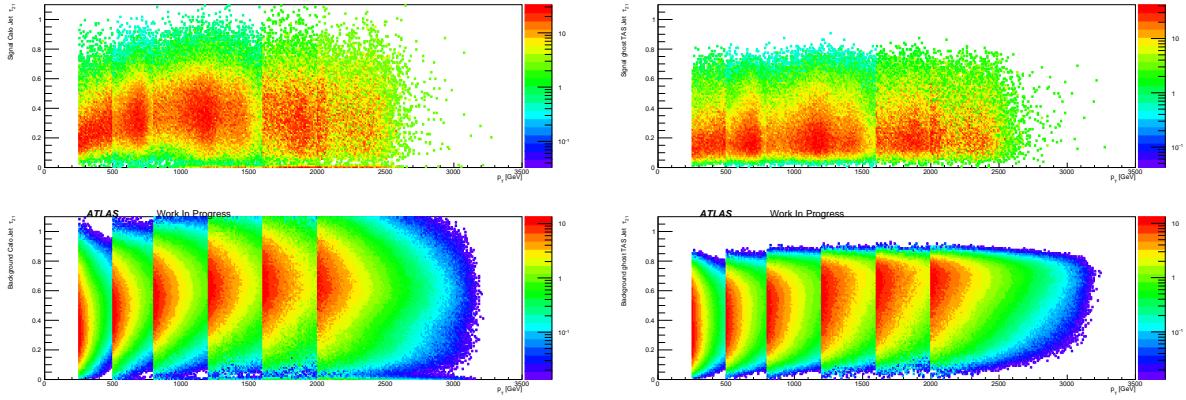


Figure 271: Correlation between  $\tau_{21}$  at  $\beta = 1$  and  $p_T$  applied on  $W$  boson signal (above) and QCD background (below) for calorimeter (left) and TAS (right).

## 888 M Results of $\beta$ Optimisation

### 889 M.1 Performance for $W$ tagging

Calorimeter	$p_T$ [GeV]	C2			D2			$\tau_{21}$		
		$\beta = 0.5$	1	2	$\beta = 0.5$	1	2	$\beta = 0.5$	1	2
250 - 500	29.7(1.5)	31.7(1.9)	31.4(1.6)	30.7(1.9)	28.5(1.4)	27.2(2.0)	35.0(2.0)	33.0(1.8)	25.7(1.2)	27.6(1.3)
500 - 800	44.2(1.8)	50.1(2.0)	49.6(1.9)	48.6(1.8)	42.6(1.9)	40.3(2.2)	55.3(2.6)	56.3(2.4)	39.3(1.3)	49.4(2.0)
800 - 1200	32.0(1.5)	37.5(1.7)	35.4(1.5)	33.4(1.5)	26.8(0.9)	34.0(2.1)	41.1(2.0)	38.5(1.6)	34.9(1.3)	41.1(1.4)
1200 - 1600	30.1(1.3)	34.4(1.8)	29.4(1.3)	26.8(1.0)	20.7(0.8)	34.1(1.8)	38.1(1.9)	31.4(1.4)	27.6(1.2)	19.3(0.5)
1600 - 2000	20.9(1.3)	22.4(1.5)	18.2(1.2)	16.5(0.9)	12.9(0.6)	26.4(1.7)	25.4(1.3)	19.3(1.1)	16.9(0.9)	11.9(0.5)
> 2000	16.9(1.4)	18.7(1.4)	14.1(0.9)	12.6(0.8)	9.9(0.7)	23.3(1.9)	21.9(1.7)	15.7(1.1)	13.5(0.9)	9.2(0.4)

TAS	$p_T$ [GeV]	C2			D2			$\tau_{21}$		
		$\beta = 0.5$	1	1.7	$\beta = 0.5$	1	1.7	$\beta = 0.5$	1	1.7
250 - 500	29.4(1.9)	30.1(1.9)	28.9(1.5)	28.5(1.3)	27.7(1.3)	28.6(2.0)	37.7(2.1)	35.4(2.3)	29.4(1.2)	31.5(1.6)
500 - 800	48.2(2.0)	55.5(2.7)	58.6(2.6)	59.1(2.7)	56.8(2.0)	42.8(2.3)	67.2(3.1)	67.6(3.2)	63.7(3.0)	52.6(2.3)
800 - 1200	31.0(1.2)	44.6(1.9)	54.6(2.8)	55.2(2.8)	53.0(3.2)	26.1(1.3)	47.6(2.3)	54.9(2.4)	52.6(2.8)	43.1(1.5)
1200 - 1600	20.9(0.7)	39.1(1.9)	53.8(2.6)	55.1(3.0)	50.1(1.6)	22.7(1.4)	42.1(2.4)	50.8(1.8)	49.6(2.3)	41.1(1.2)
1600 - 2000	16.7(0.7)	36.9(2.9)	50.9(4.3)	50.3(4.4)	42.2(2.4)	18.7(1.7)	32.7(3.3)	37.8(2.0)	36.1(2.4)	28.7(1.2)
> 2000	11.6(0.6)	31.2(3.2)	46.1(4.7)	45.5(5.2)	35.5(3.8)	17.8(2.0)	33.0(4.0)	36.3(2.0)	34.0(2.5)	27.4(1.3)

Tracks	$p_T$ [GeV]	C2			D2			$\tau_{21}$		
		$\beta = 0.5$	1	1.7	$\beta = 0.5$	1	1.7	$\beta = 0.5$	1	1.7
250 - 500	27.1(1.2)	28.1(1.5)	28.7(1.9)	28.2(1.7)	21.6(1.2)	28.9(2.0)	29.5(1.8)	29.1(1.6)	28.1(1.3)	28.7(1.8)
500 - 800	46.5(1.9)	52.9(2.4)	57.7(2.6)	58.1(2.7)	55.8(2.5)	30.1(1.8)	46.8(2.4)	53.4(2.2)	52.1(2.3)	46.6(1.7)
800 - 1200	30.3(1.1)	44.5(2.2)	54.8(2.8)	56.4(3.0)	53.7(3.6)	24.5(1.5)	42.3(2.3)	48.6(2.5)	47.5(1.2)	34.5(1.6)
1200 - 1600	20.7(0.6)	39.0(1.9)	54.2(2.7)	55.5(3.3)	50.9(1.7)	22.7(1.3)	41.0(2.2)	50.0(1.6)	47.6(2.2)	41.4(1.2)
1600 - 2000	16.6(0.7)	36.7(2.3)	51.7(5.2)	51.6(4.0)	43.1(2.3)	18.5(1.7)	32.1(3.0)	37.0(1.9)	35.9(2.3)	29.3(1.2)
> 2000	11.6(0.5)	31.5(3.0)	46.8(5.7)	46.0(4.2)	36.1(4.3)	17.8(2.2)	33.0(3.3)	35.9(2.1)	34.2(2.6)	28.1(1.0)

Table 7: Listing of the QCD rejection for  $W$  jets achieved with C2, D2 and  $\tau_{21}$  calculated with varying angular weightings  $\beta$  and constituents. The highest achieved background rejection per energy range is highlighted in red.

<sup>890</sup> **M.2 Performance for Higgs tagging**

Calorimeter	$p_T [GeV]$	C2			D2			$\tau_{21}$		
		$\beta = 0.5$	1	2	$\beta = 0.5$	1	2	$\beta = 0.5$	1	2
250 - 500	4.6(0.1)	5.0(0.1)	5.2(0.1)	5.3(0.1)	5.5(0.1)	5.7(0.1)	7.3(0.2)	8.4(0.2)	8.4(0.2)	7.6(0.2)
500 - 800	15.7(0.3)	16.7(0.4)	17.0(0.4)	16.9(0.4)	16.2(0.4)	13.6(0.3)	16.9(0.4)	17.7(0.4)	17.2(0.4)	15.2(0.3)
800 - 1200	22.1(0.5)	23.8(0.5)	25.0(0.6)	25.0(0.6)	23.4(0.5)	18.4(0.4)	23.7(0.6)	26.3(0.6)	22.3(0.5)	22.8(0.5)
1200 - 1600	24.0(0.6)	26.0(0.8)	26.4(0.8)	25.9(0.7)	23.0(0.6)	19.3(0.6)	24.9(0.7)	27.0(0.8)	26.1(0.7)	21.9(0.5)
1600 - 2000	12.1(0.7)	13.9(0.8)	14.3(0.7)	14.0(0.7)	12.3(0.6)	11.1(0.7)	14.1(0.9)	14.9(0.8)	14.2(0.6)	11.8(0.5)

TAS	$p_T [GeV]$	C2			D2			$\tau_{21}$		
		$\beta = 0.5$	1	2	$\beta = 0.5$	1	2	$\beta = 0.5$	1	2
250 - 500	4.8(0.1)	5.2(0.1)	5.5(0.1)	5.6(0.1)	5.8(0.1)	5.9(0.1)	7.6(0.2)	8.5(0.2)	8.6(0.2)	8.5(0.2)
500 - 800	16.1(0.4)	17.3(0.4)	17.7(0.4)	17.7(0.4)	17.6(0.4)	14.0(0.3)	18.2(0.4)	18.7(0.4)	18.3(0.4)	16.9(0.4)
800 - 1200	20.6(0.5)	23.5(0.5)	26.2(0.6)	26.9(0.7)	27.7(0.6)	18.8(0.4)	25.6(0.6)	28.5(0.7)	28.4(0.7)	26.8(0.6)
1200 - 1600	18.6(0.4)	22.6(0.6)	27.4(0.7)	28.7(0.8)	30.0(0.7)	17.9(0.4)	24.3(0.7)	28.9(0.7)	29.3(0.6)	28.1(0.7)
1600 - 2000	8.0(0.3)	11.3(0.5)	15.4(0.9)	16.5(1.0)	17.8(0.7)	10.0(0.5)	14.0(0.8)	17.7(0.8)	18.1(0.9)	17.9(0.6)

Tracks	$p_T [GeV]$	C2			D2			$\tau_{21}$		
		$\beta = 0.5$	1	2	$\beta = 0.5$	1	2	$\beta = 0.5$	1	2
250 - 500	4.9(0.1)	5.2(0.1)	5.5(0.1)	5.6(0.1)	5.9(0.1)	5.8(0.1)	7.4(0.2)	8.3(0.2)	8.3(0.2)	7.4(0.2)
500 - 800	15.6(0.3)	17.2(0.4)	17.8(0.4)	17.9(0.4)	17.7(0.4)	13.5(0.3)	17.1(0.4)	17.9(0.4)	17.7(0.4)	16.8(0.4)
800 - 1200	20.1(0.5)	24.0(0.5)	26.9(0.6)	27.7(0.7)	28.4(0.6)	18.8(0.4)	25.3(0.6)	28.0(0.7)	28.0(0.7)	26.9(0.6)
1200 - 1600	18.5(0.5)	23.8(0.6)	28.8(0.8)	30.0(0.8)	31.1(0.7)	19.4(0.5)	26.3(0.7)	30.3(0.8)	29.2(0.7)	28.0(0.5)
1600 - 2000	8.0(0.3)	11.7(0.5)	16.1(0.9)	17.1(0.9)	18.3(0.9)	11.0(0.7)	15.5(0.7)	18.5(0.8)	18.7(0.8)	18.4(0.6)

Table 8: Listing of the QCD rejection for Higgs jets achieved with C2, D2 and  $\tau_{21}$  calculated with varying angular weightings  $\beta$  and constituents. The highest achieved background rejection per energy range is highlighted in red.

---

**891 M.3 Performance for Top tagging**

<b>Calorimeter</b>	$\tau_{32}$			
$p_T [GeV]$	$\beta = 1$	1.7	2	3
250 - 500	<b><math>9.7 \pm 0.2</math></b>	$9.5 \pm 0.2$	$9.5 \pm 0.4$	$9.4 \pm 0.2$
500 - 800	$20.1 \pm 0.5$	$22.2 \pm 0.6$	<b><math>22.4 \pm 0.6</math></b>	$22.0 \pm 0.6$
800 - 1200	$17.3 \pm 0.4$	$20.3 \pm 0.5$	<b><math>20.6 \pm 0.5</math></b>	$20.3 \pm 0.5$
1200 - 1600	$14.3 \pm 0.3$	$16.4 \pm 0.4$	<b><math>16.6 \pm 0.5</math></b>	$16.1 \pm 0.5$
1600 - 2000	$11.7 \pm 0.3$	<b><math>13.3 \pm 0.4</math></b>	<b><math>13.3 \pm 0.4</math></b>	$12.6 \pm 0.3$
> 2000	$9.6 \pm 0.3$	$11.0 \pm 0.4$	$10.9 \pm 0.4$	$10.1 \pm 0.3$

<b>TAS</b>	$\tau_{32}$			
$p_T [GeV]$	$\beta = 1$	1.7	2	3
250 - 500	<b><math>10.7 \pm 0.2</math></b>	$10.1 \pm 0.2$	$9.9 \pm 0.2$	$9.6 \pm 0.2$
500 - 800	<b><math>22.8 \pm 0.6</math></b>	<b><math>22.8 \pm 0.6</math></b>	$22.5 \pm 0.6$	$21.6 \pm 0.6$
800 - 1200	<b><math>23.6 \pm 0.6</math></b>	$24.1 \pm 0.6$	$23.6 \pm 0.6$	$22.2 \pm 0.5$
1200 - 1600	$22.0 \pm 0.6$	<b><math>22.3 \pm 0.6</math></b>	$21.7 \pm 0.6$	$19.8 \pm 0.6$
1600 - 2000	$18.9 \pm 0.6$	$18.8 \pm 0.6$	$17.9 \pm 0.5$	$16.0 \pm 0.5$
> 2000	$16.5 \pm 0.7$	$15.7 \pm 0.7$	$15.2 \pm 0.7$	$13.1 \pm 0.6$

<b>Tracks</b>	$\tau_{32}$			
$p_T [GeV]$	$\beta = 1$	1.7	2	3
250 - 500	<b><math>10.5 \pm 0.2</math></b>	$9.8 \pm 0.2$	$9.6 \pm 0.2$	$9.4 \pm 0.2$
500 - 800	$20.6 \pm 0.5$	<b><math>21.3 \pm 0.6</math></b>	$21.1 \pm 0.5$	$20.3 \pm 0.5$
800 - 1200	$21.8 \pm 0.6$	$22.9 \pm 0.6$	$22.6 \pm 0.6$	$21.4 \pm 0.6$
1200 - 1600	$21.7 \pm 0.6$	<b><math>22.1 \pm 0.6</math></b>	$21.6 \pm 0.6$	$19.5 \pm 0.6$
1600 - 2000	$19.3 \pm 0.6$	$19.0 \pm 0.6$	$18.2 \pm 0.6$	$16.0 \pm 0.5$
> 2000	$16.8 \pm 0.7$	$15.8 \pm 0.7$	$15.1 \pm 0.7$	$13.0 \pm 0.5$

Table 9: Listing of the QCD rejection for Top jets achieved with  $\tau_{32}$  calculated with varying angular weightings  $\beta$  and constituents. The highest achieved background rejection per energy range is highlighted in red.



## 892 N Signal and Background Distributions

### 893 N.1 W Distributions

894  $\beta = 0.5$

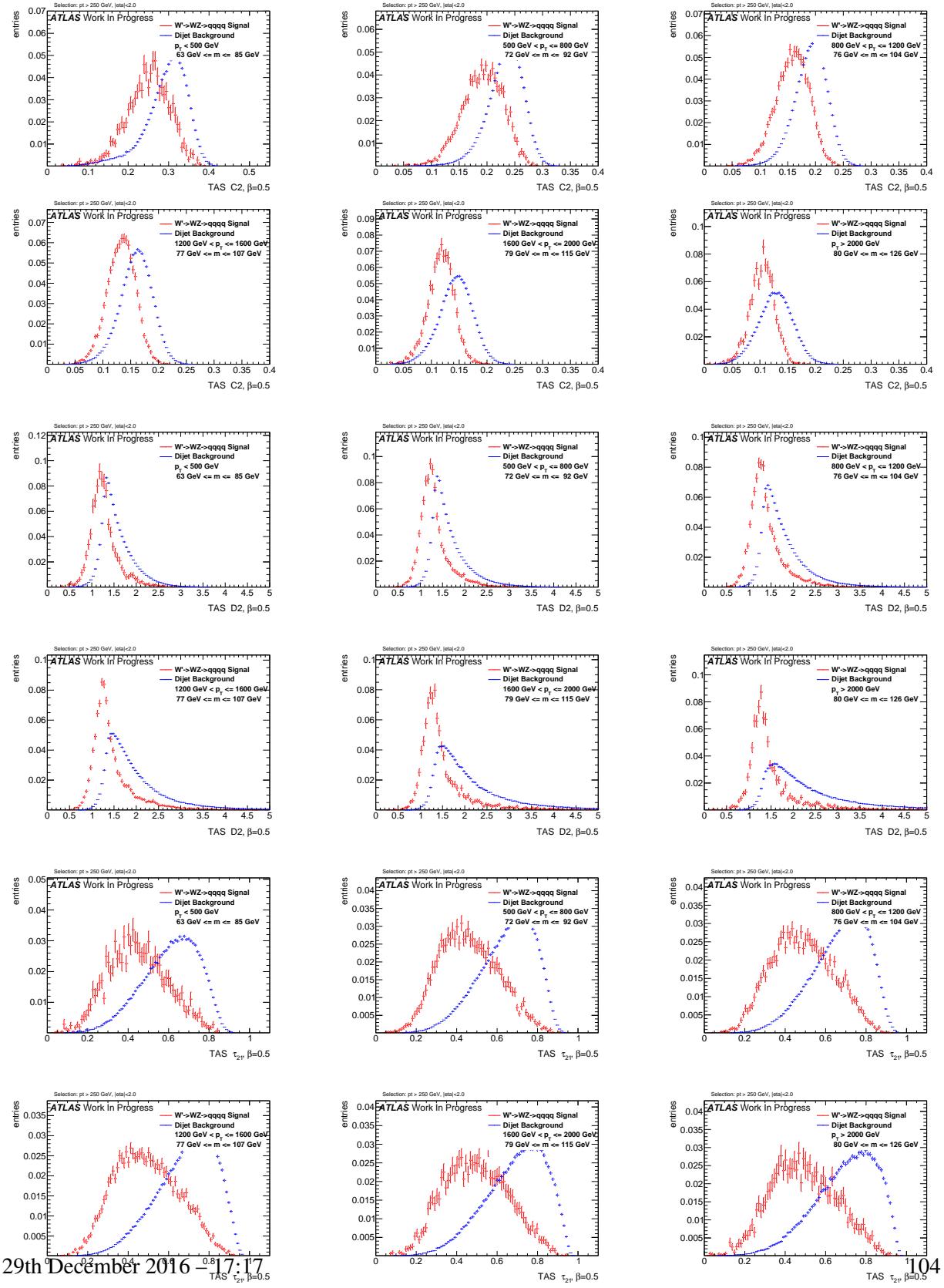
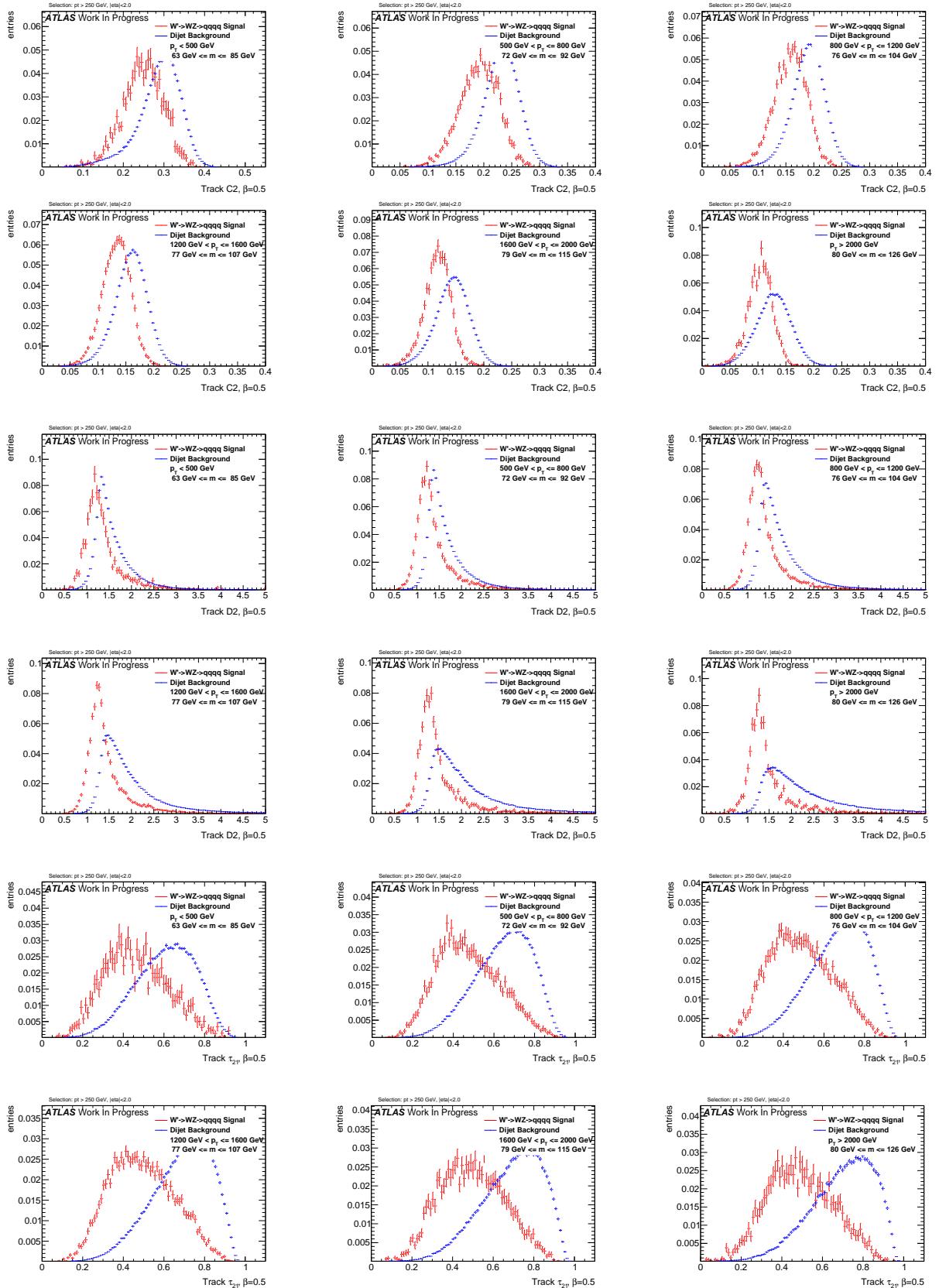
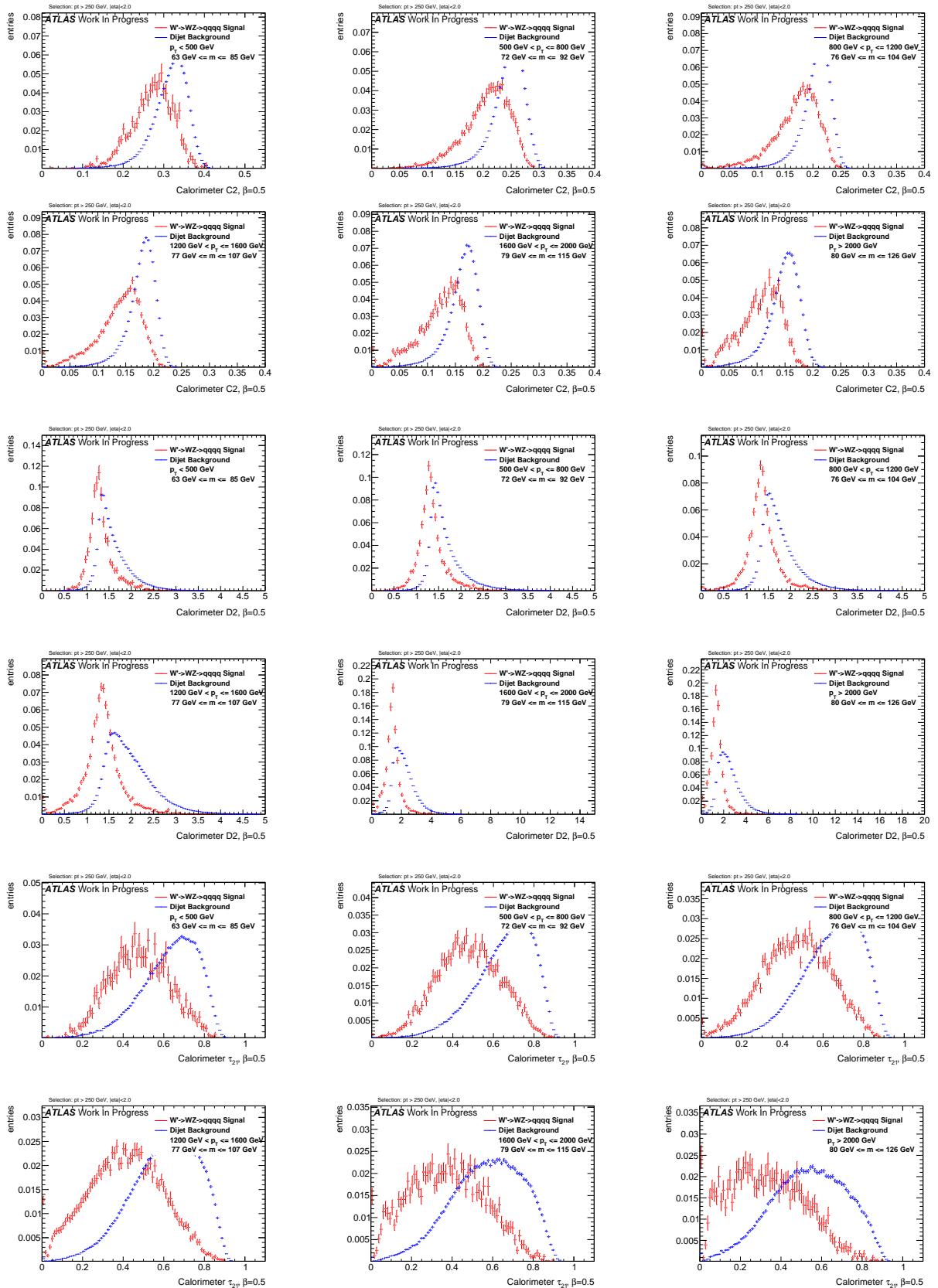


Figure 272: Distributions for W boson tagging using TAS  $\beta = 0.5$ . C2, D2,  $\tau_{21}$  top down.

Figure 273: Distributions for  $W$  boson tagging using tracks  $\beta = 0.5$ . C2, D2,  $\tau_{21}$  top down.

Figure 274: Distributions for  $W$  boson tagging using calorimeter clusters  $\beta = 0.5$ . C2, D2,  $\tau_{21}$  top down.



895  $\beta = 1$

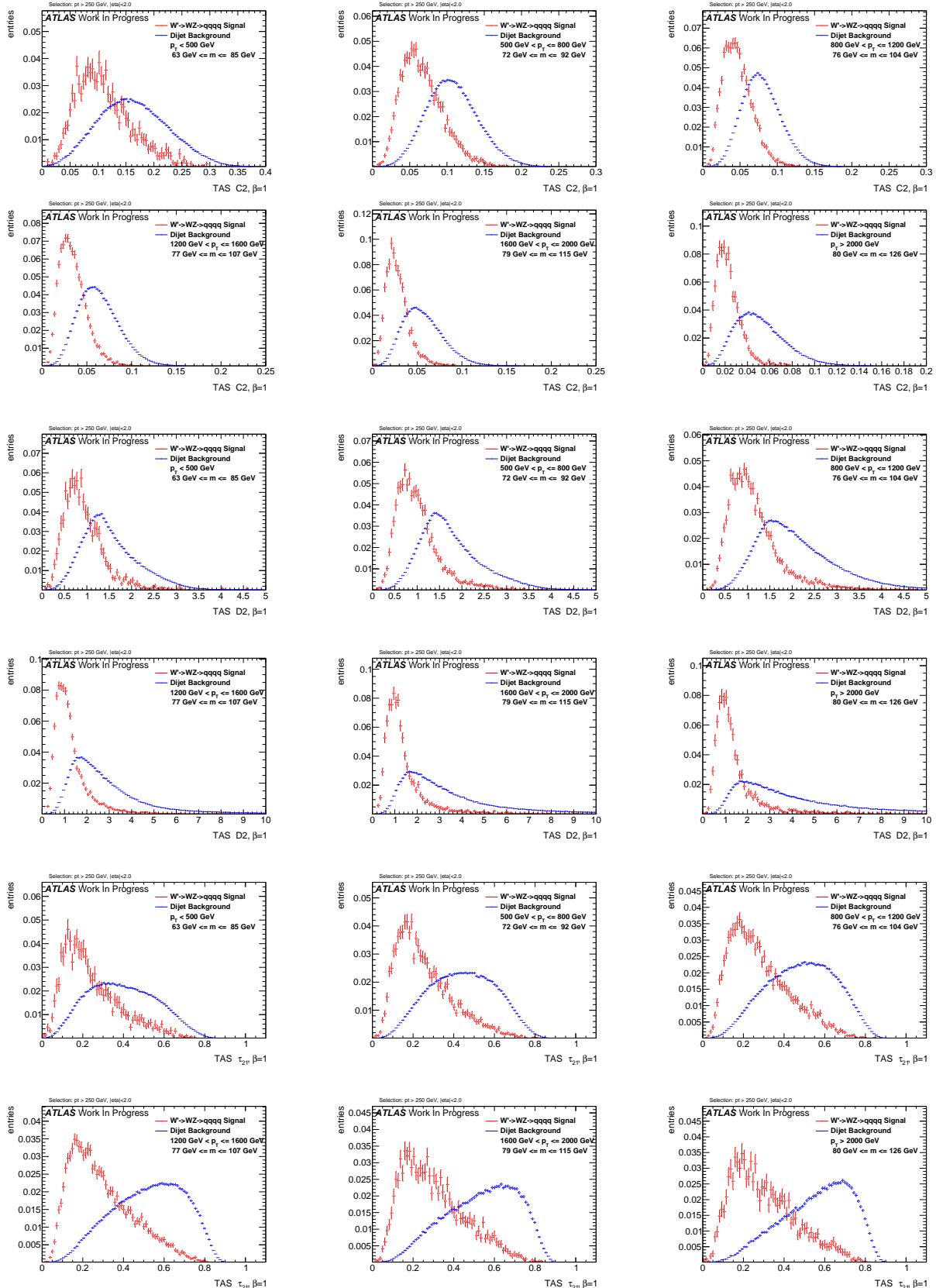
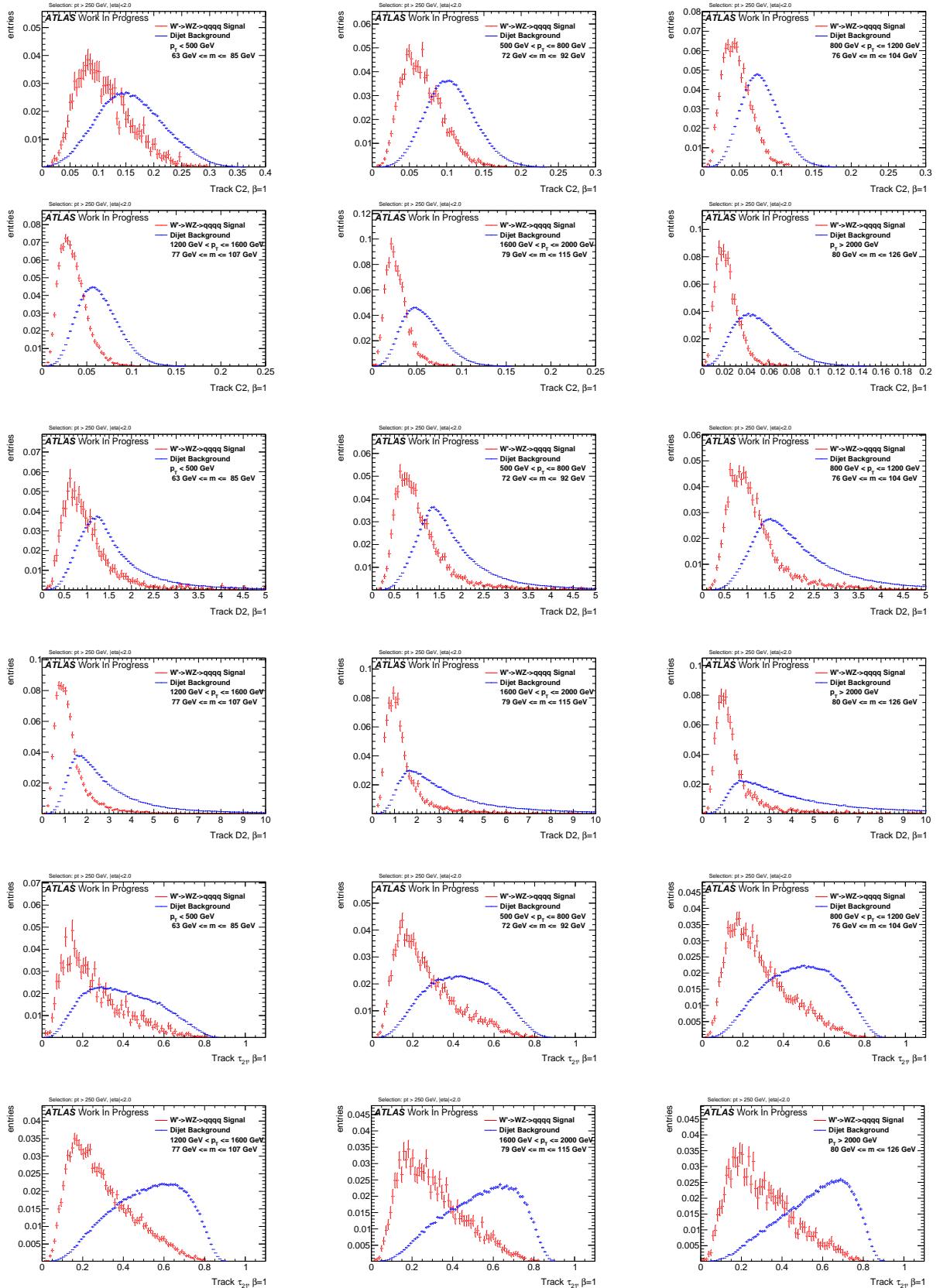
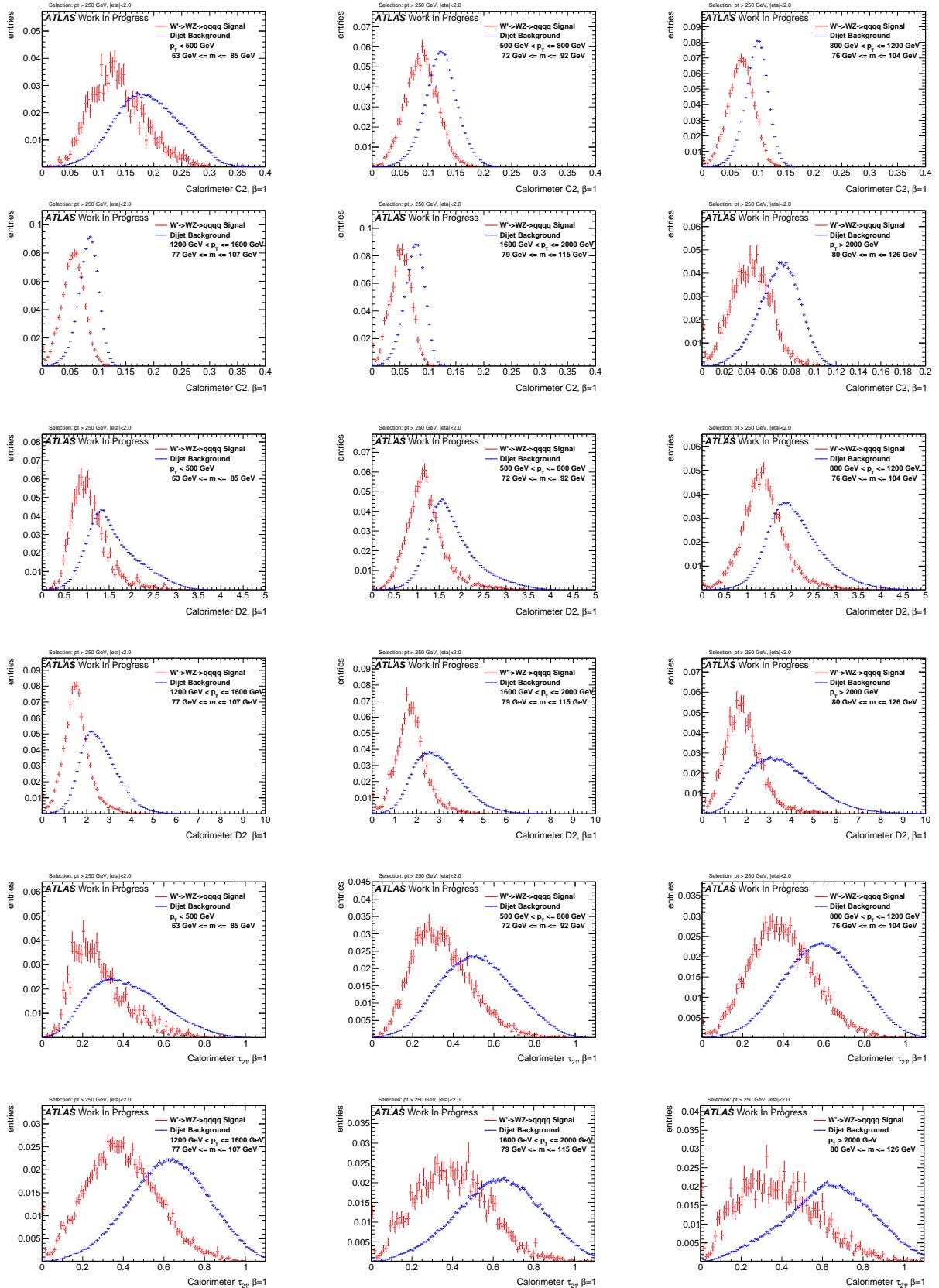
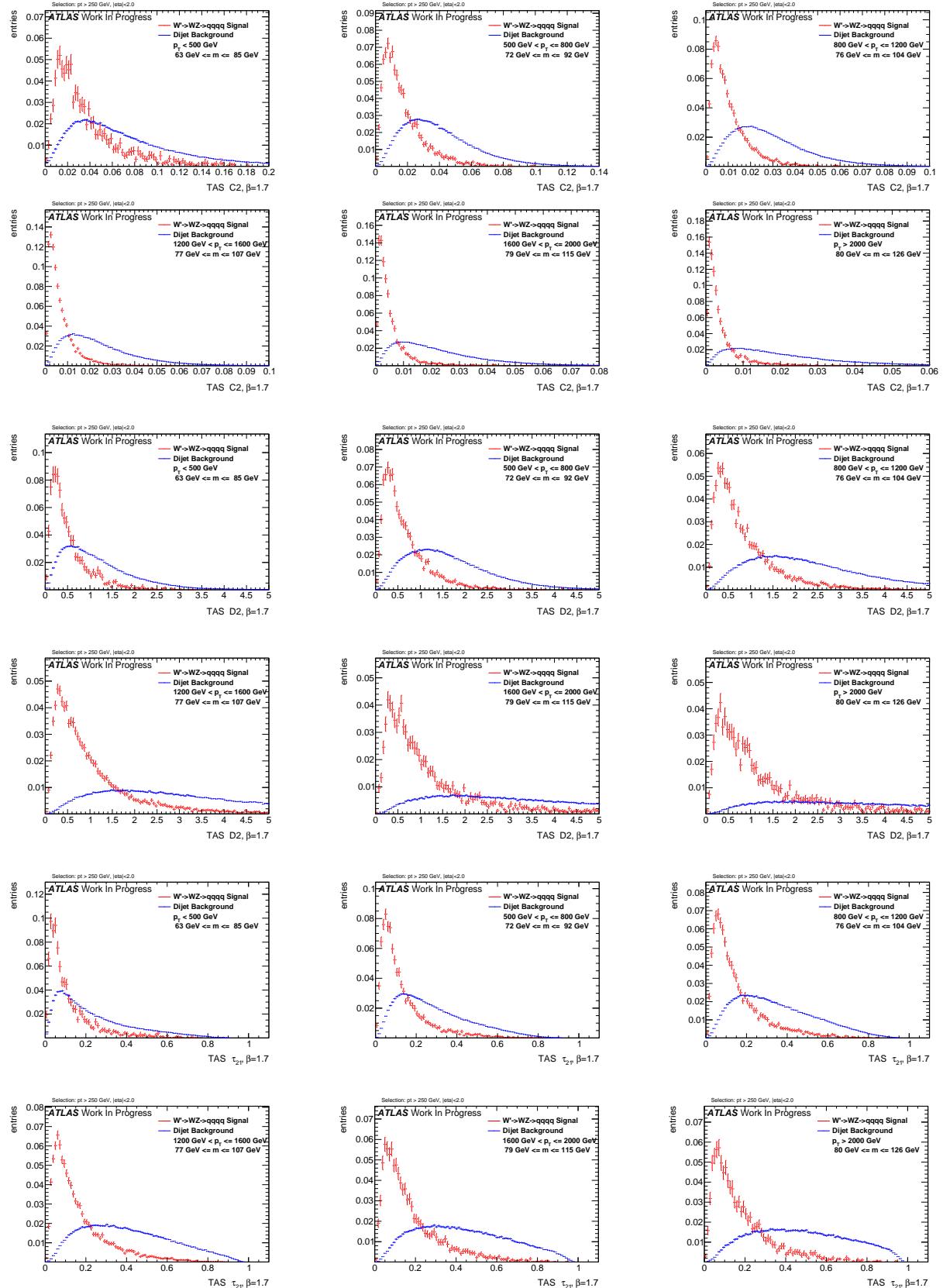


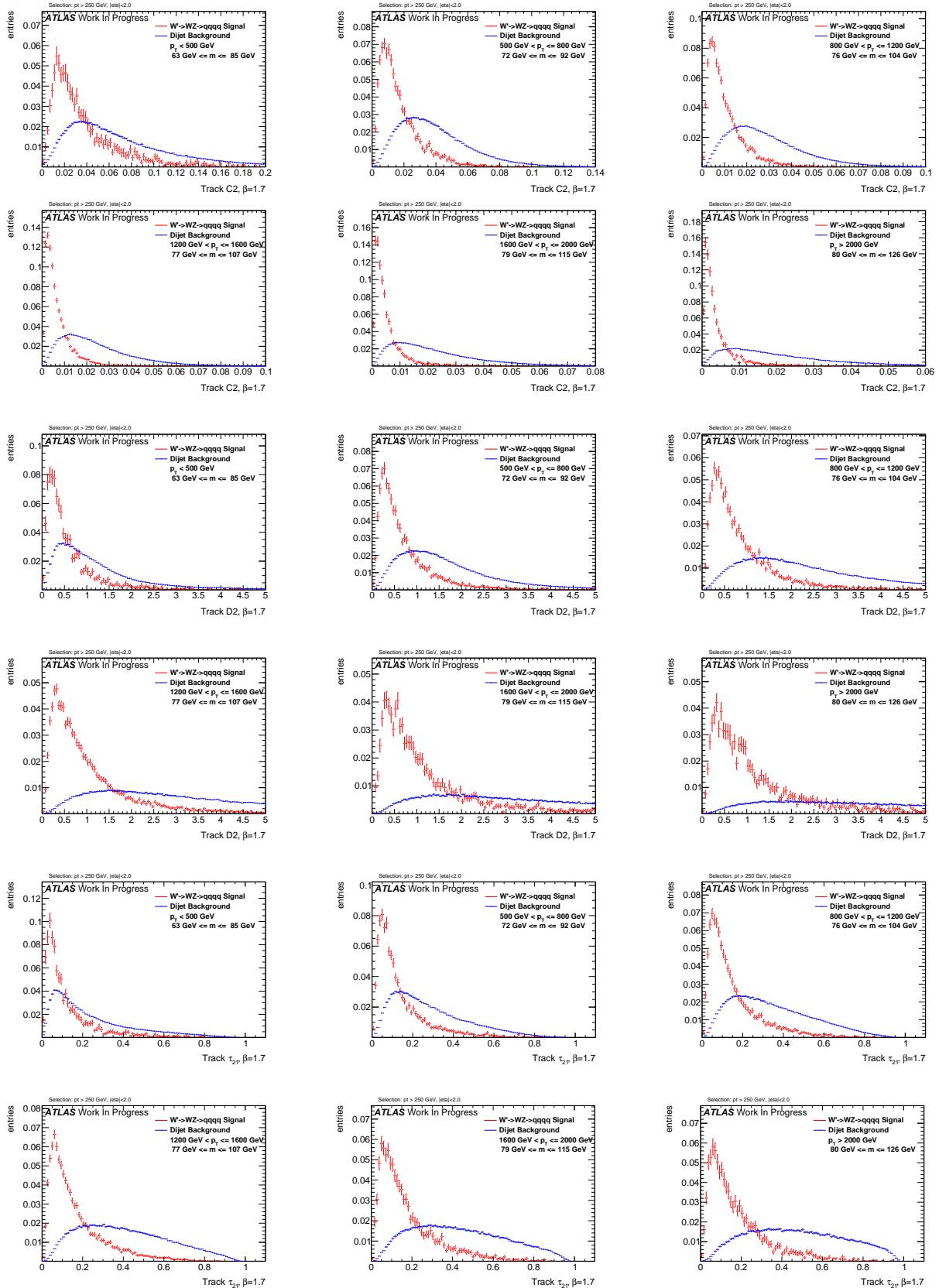
Figure 275: Distributions for  $W$  boson tagging using TAS  $\beta = 1$ . C2, D2,  $\tau_{21}$  top down.  
29th December 2016 – 17:17

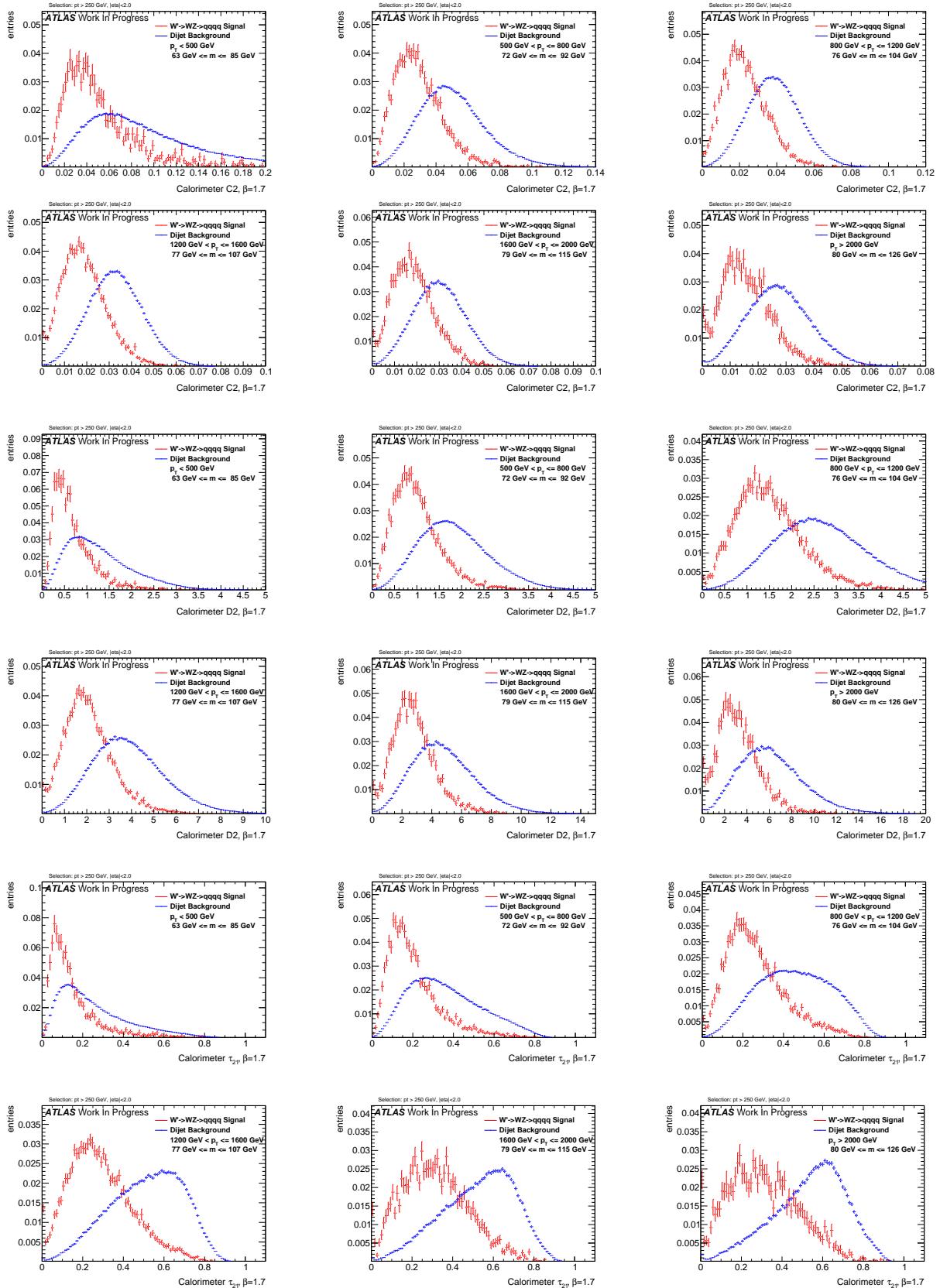
Figure 276: Distributions for  $W$  boson tagging using tracks  $\beta = 1$ . C2, D2,  $\tau_{21}$  top down.

Figure 277: Distributions for  $W$  boson tagging using calorimeter clusters  $\beta = 1$ . C2, D2,  $\tau_{21}$  top down.

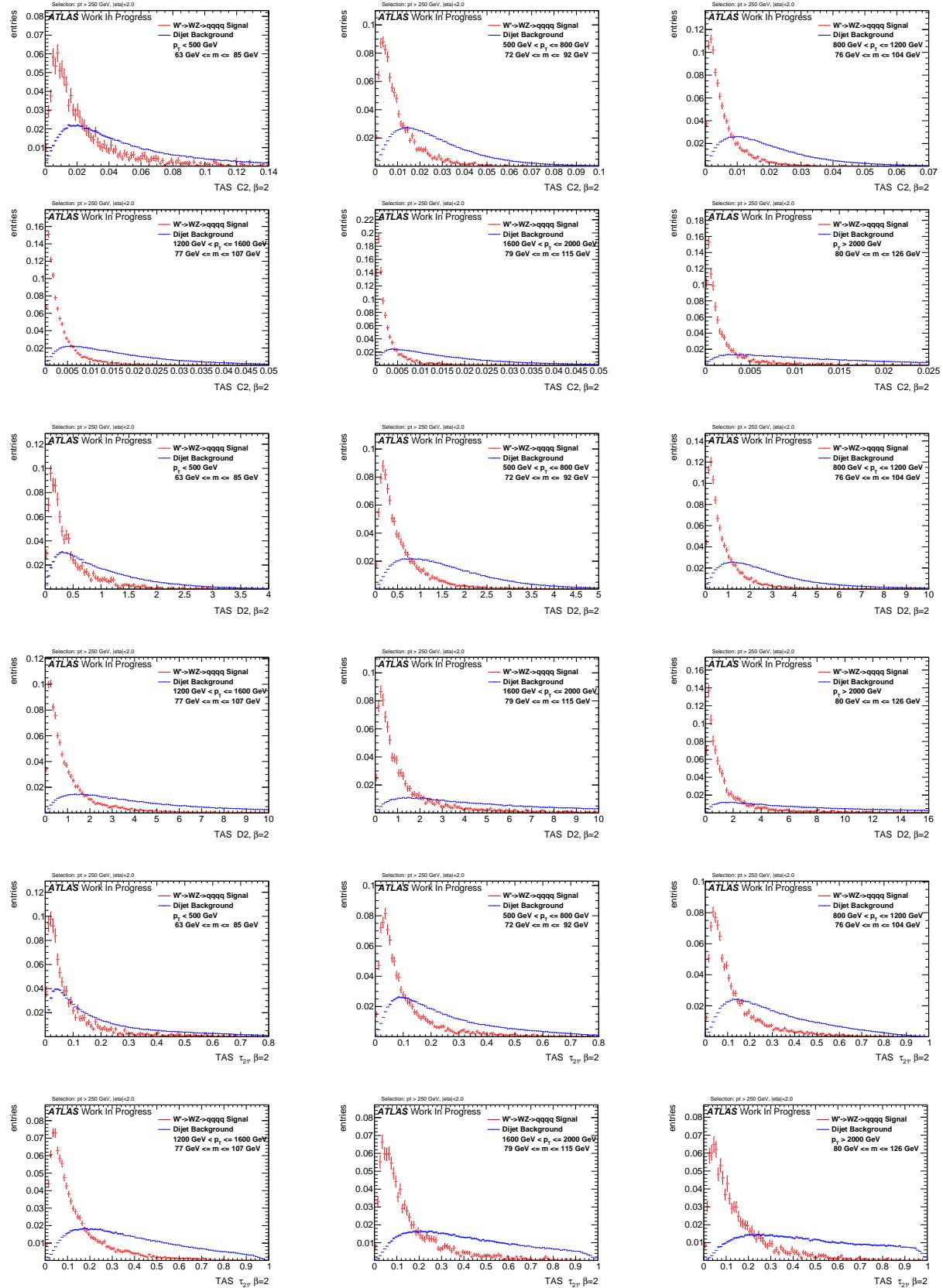


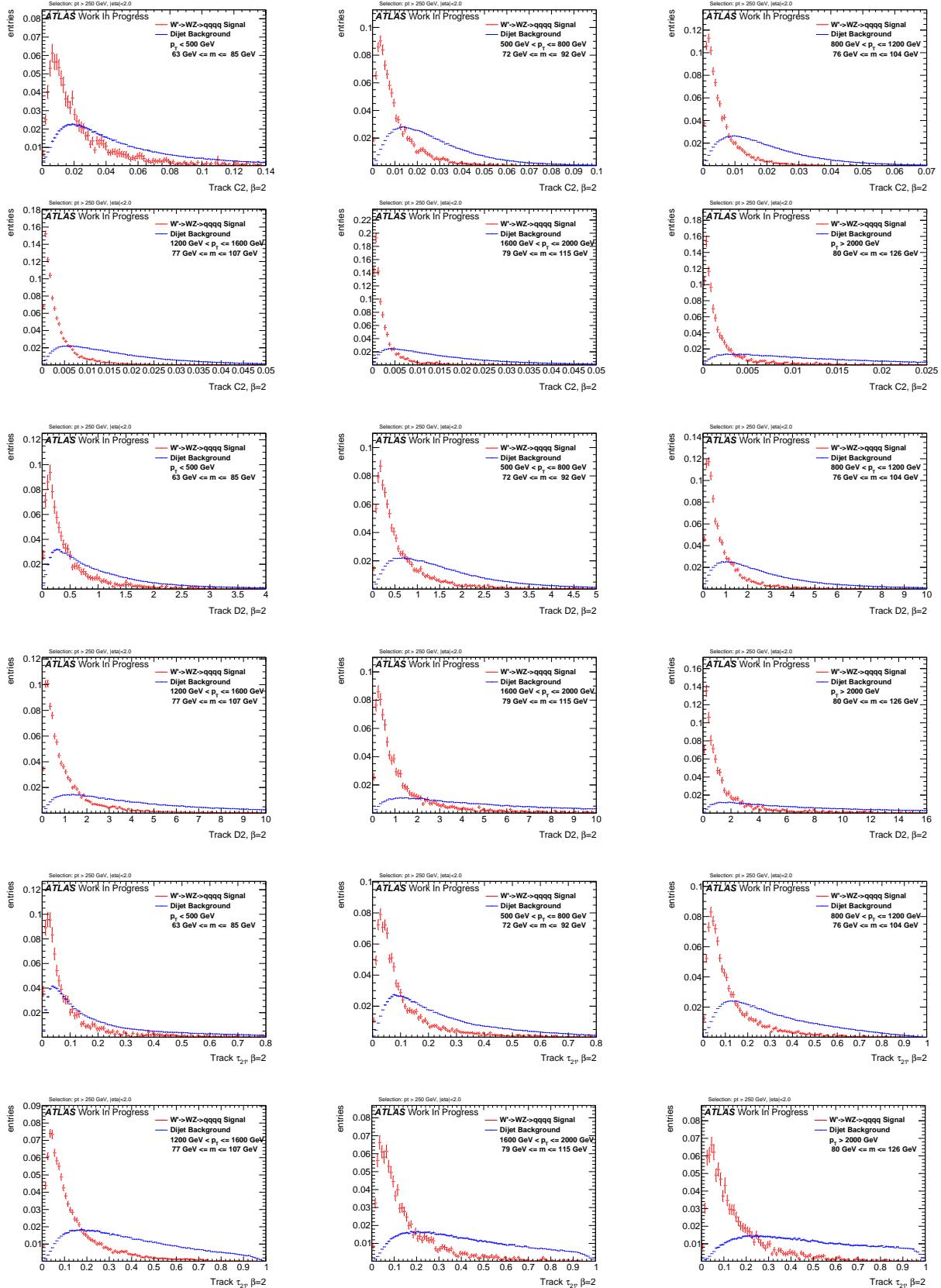
896  $\beta = 1.7$ 

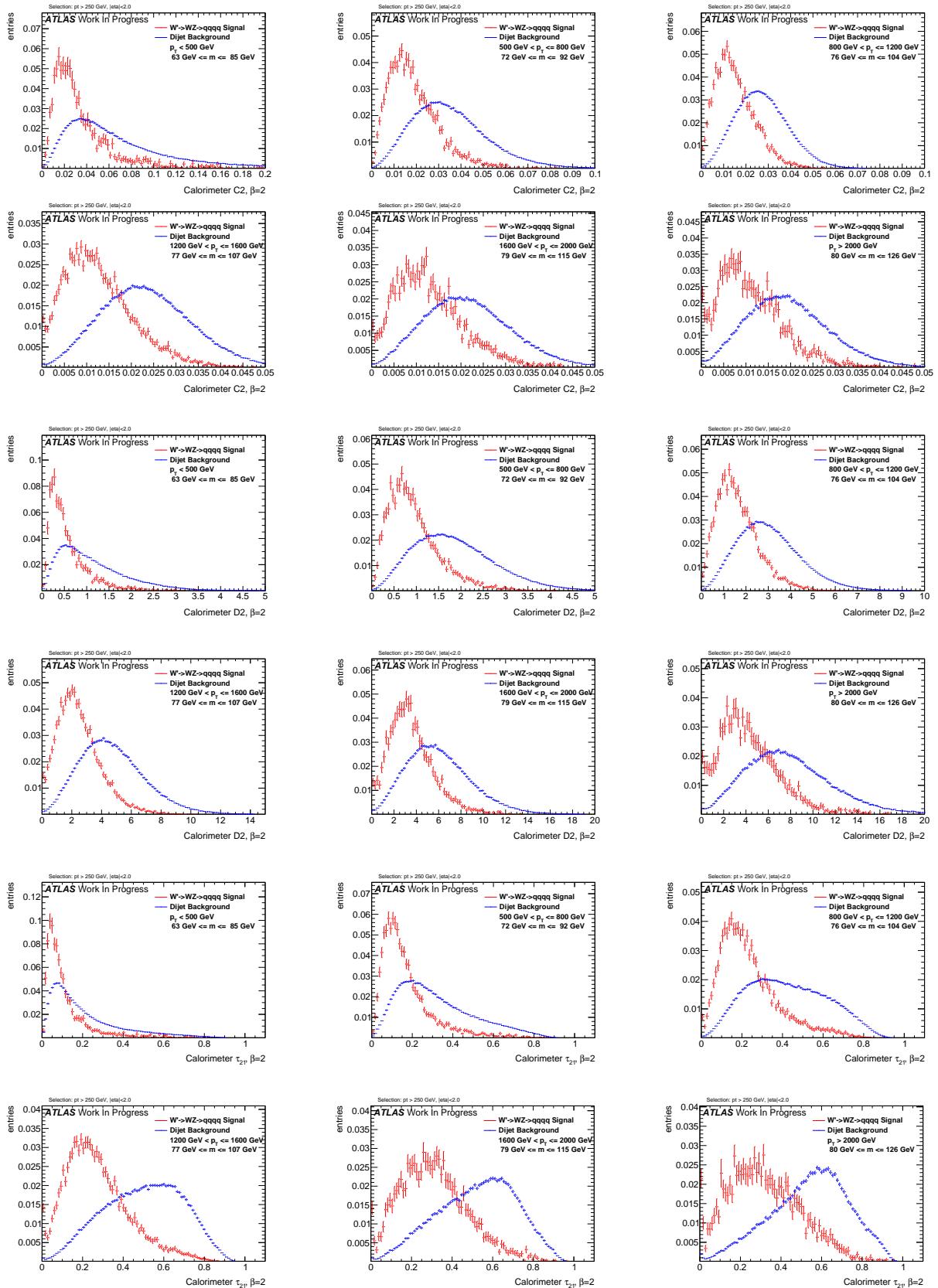
Figure 279: Distributions for W boson tagging using tracks  $\beta = 1.7$ . C2, D2,  $\tau_{21}$  top down.

Figure 280: Distributions for  $W$  boson tagging using calorimeter clusters  $\beta = 1.7$ . C2, D2,  $\tau_{21}$  top down.

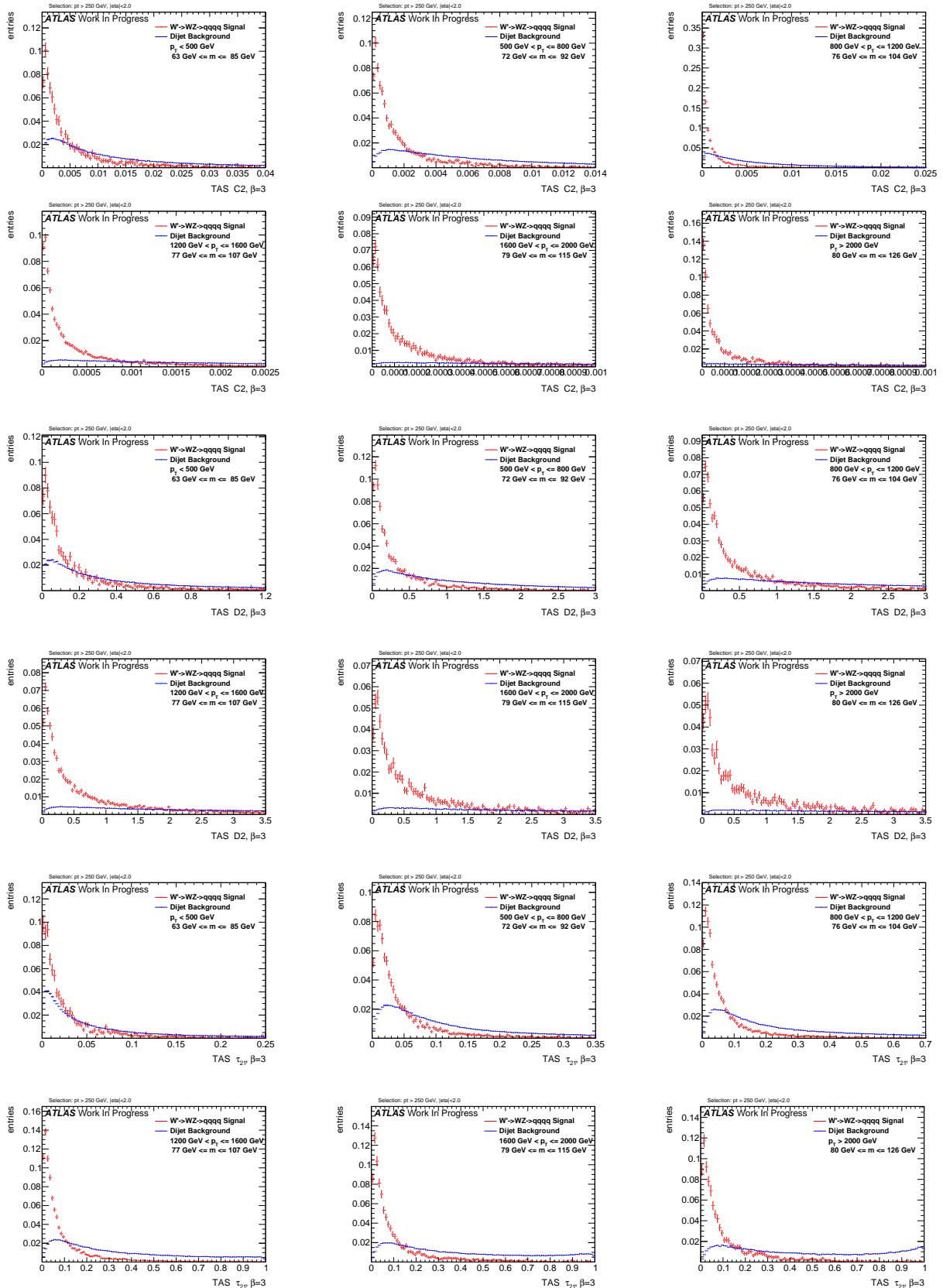


897  $\beta = 2$ 

Figure 282: Distributions for  $W$  boson tagging using tracks  $\beta = 2$ . C2, D2,  $\tau_{21}$  top down.

Figure 283: Distributions for  $W$  boson tagging using calorimeter clusters  $\beta = 2$ . C2, D2,  $\tau_{21}$  top down.

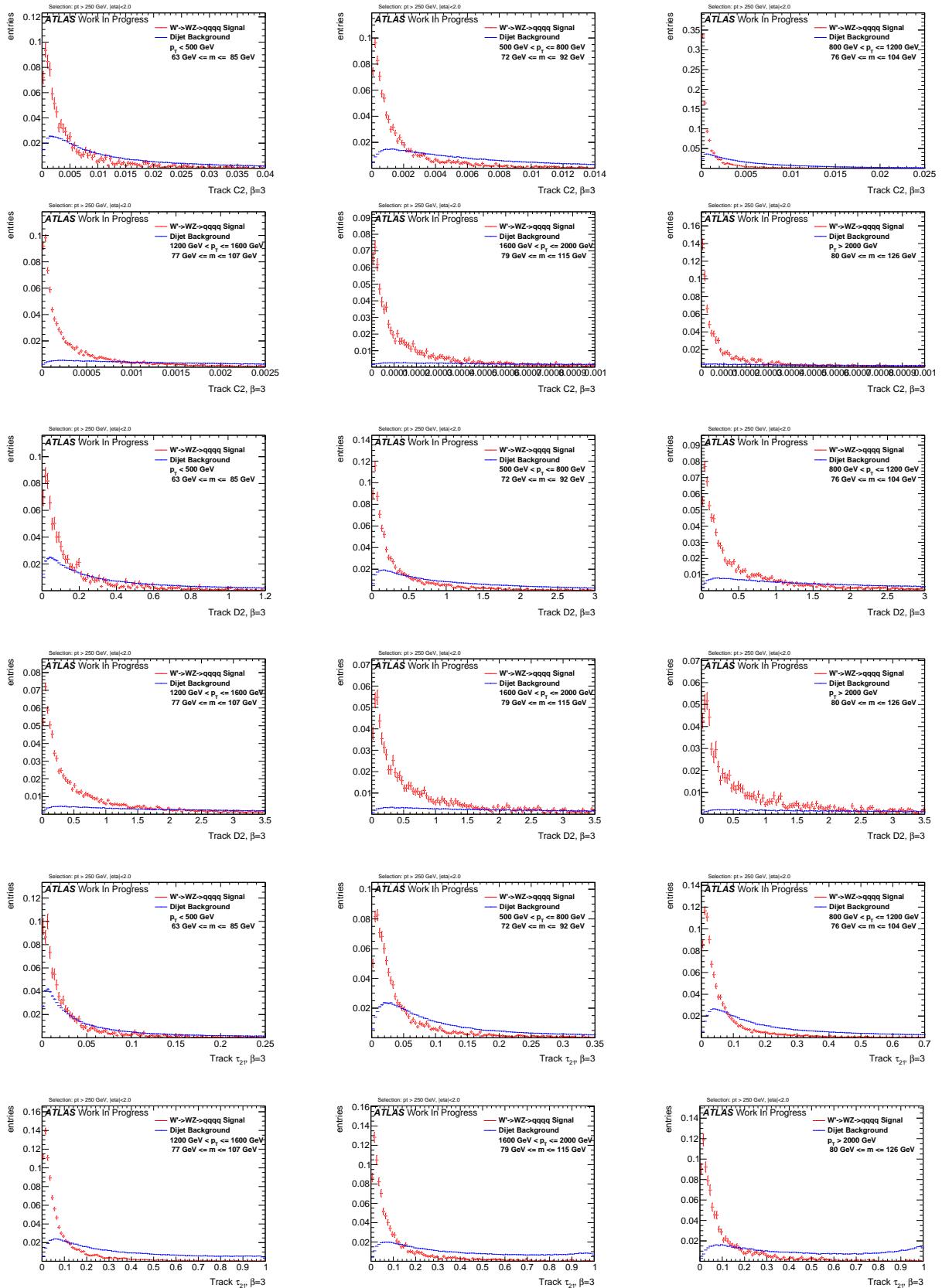


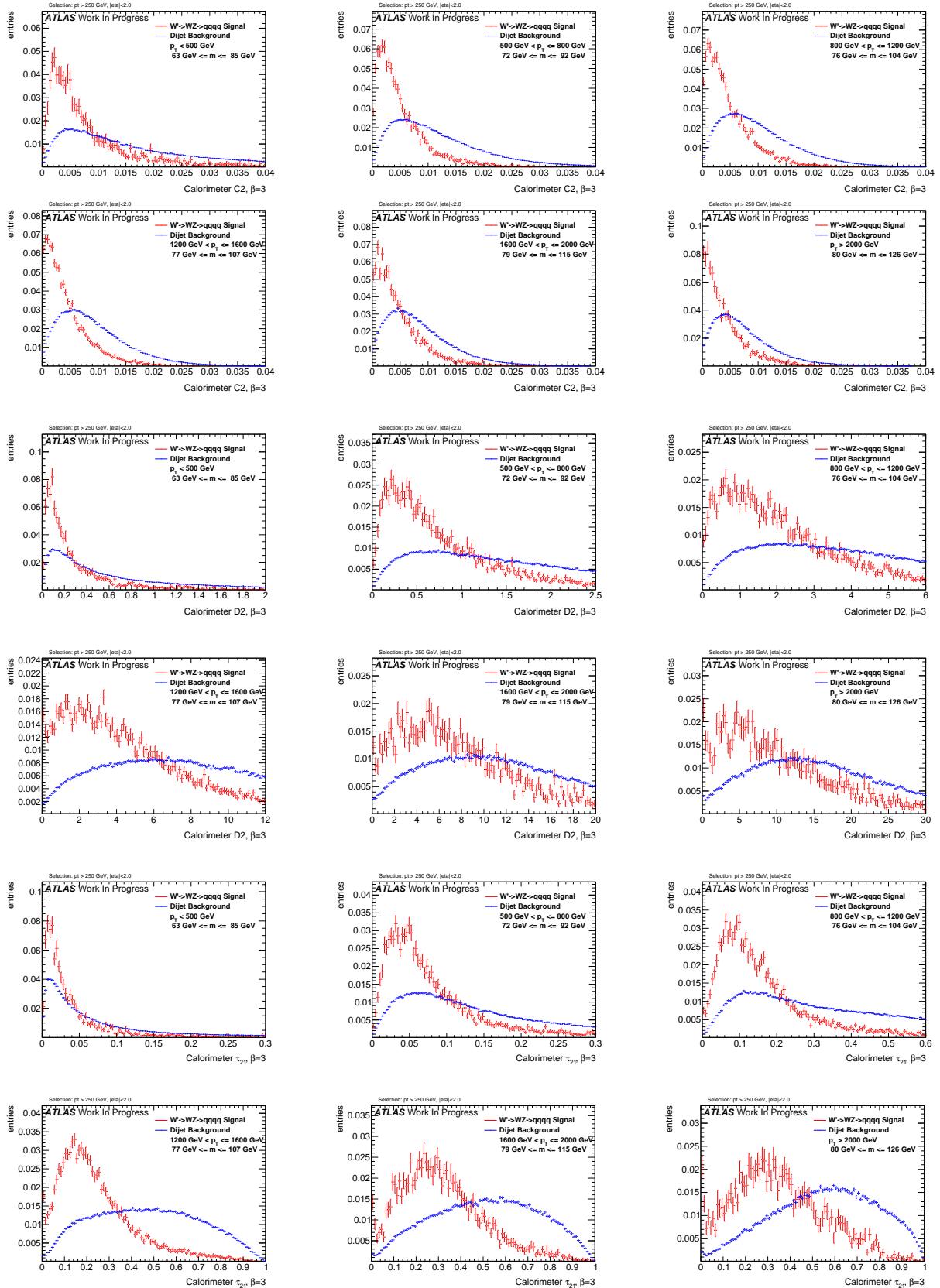
898  $\beta = 3$ 

29th December 2016 - 17:17

Figure 284: Distributions for W boson tagging using TAS  $\beta = 3$ . C2, D2,  $\tau_{21}$  top down.

120

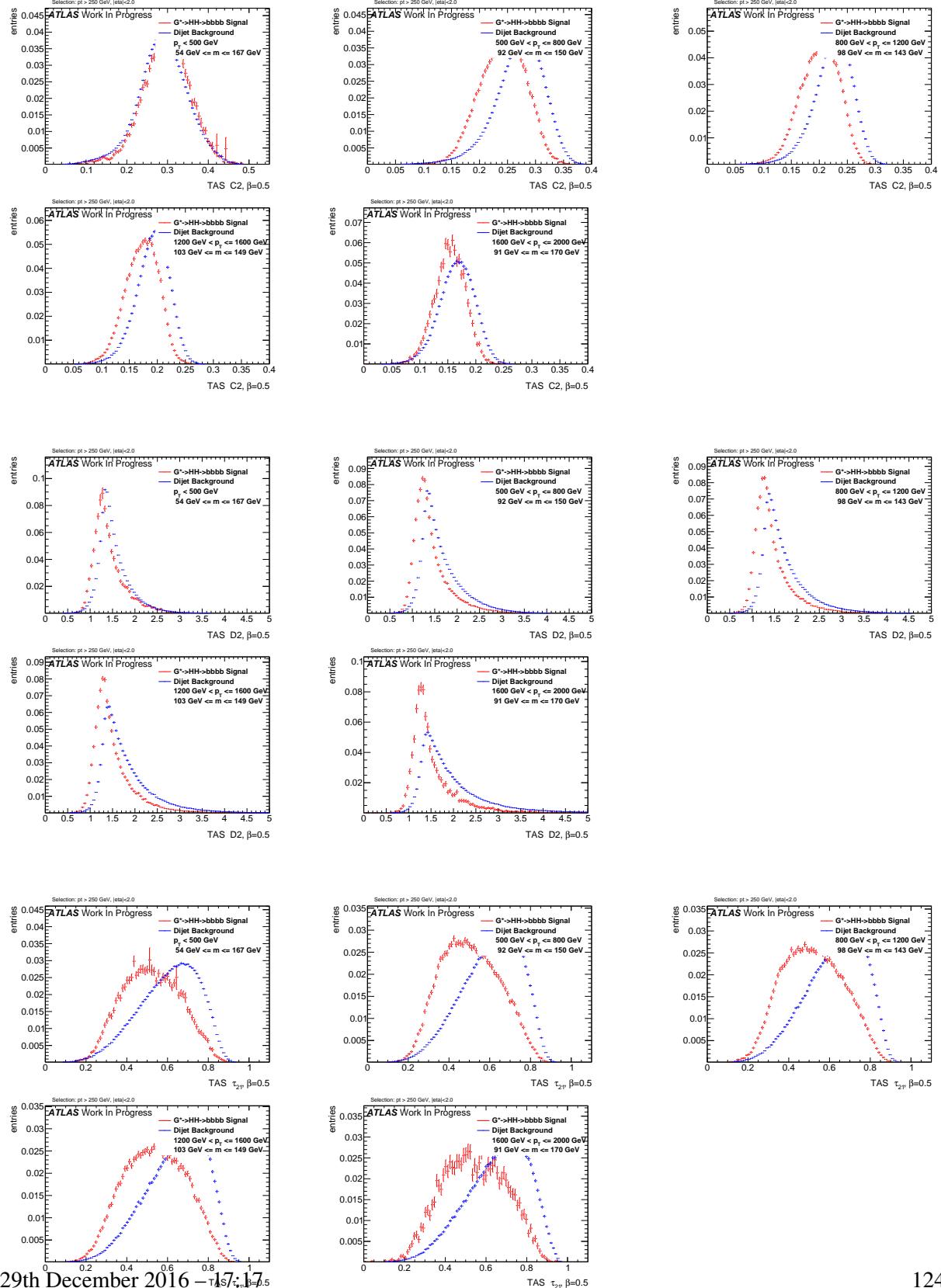
Figure 285: Distributions for  $W$  boson tagging using tracks  $\beta = 3$ . C2, D2,  $\tau_{21}$  top down.

Figure 286: Distributions for  $W$  boson tagging using calorimeter clusters  $\beta = 3$ . C2, D2,  $\tau_{21}$  top down.



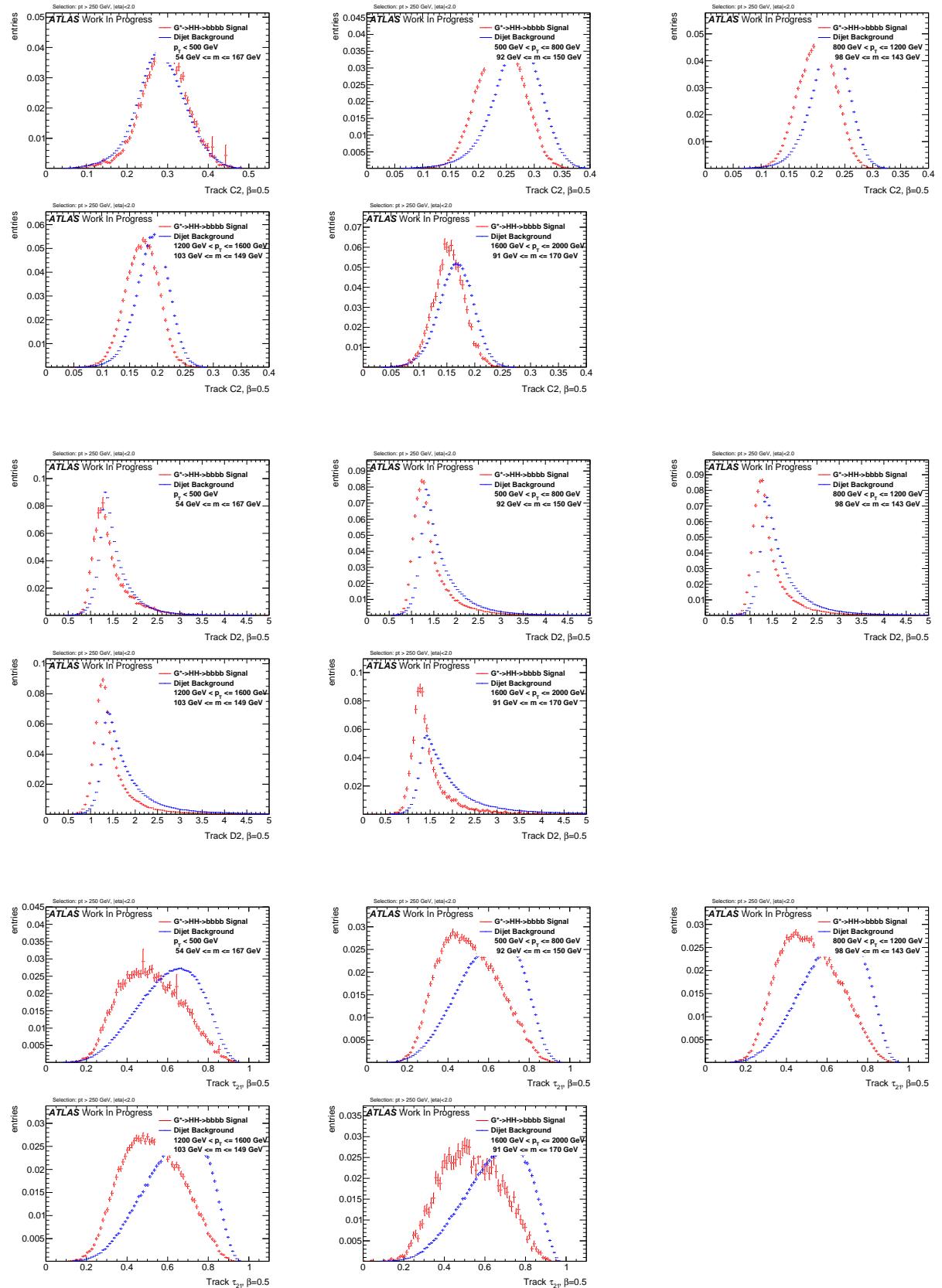
## 899 N.2 Higgs Distributions

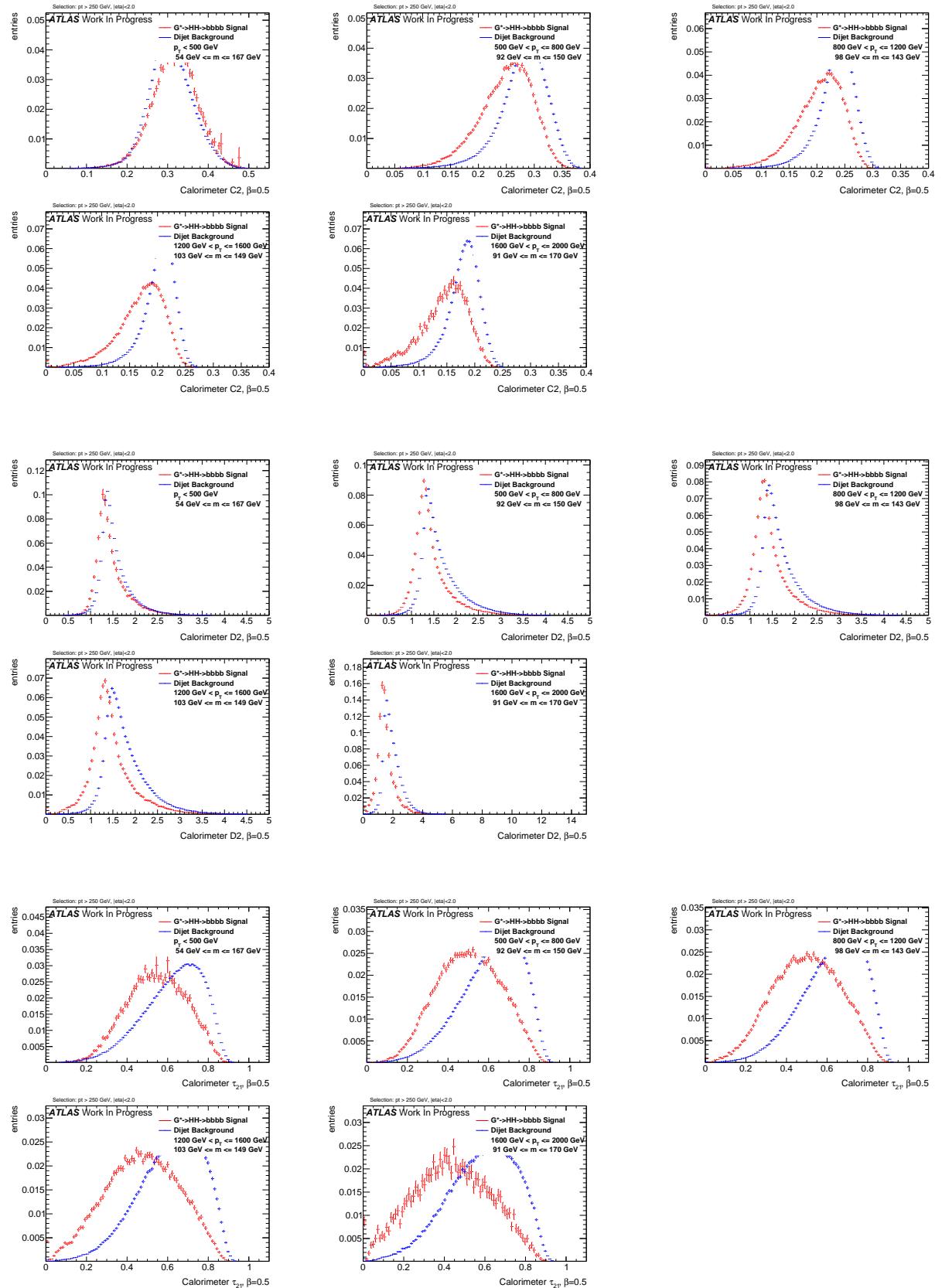
900  $\beta = 0.5$



29th December 2016 - 17:17:15

Figure 287: Distributions for Higgs boson tagging using TAS  $\beta = 0.5$ . C2, D2,  $\tau_{2\gamma}$  top down.

Figure 288: Distributions for Higgs boson tagging using tracks  $\beta = 0.5$ . C2, D2,  $\tau_{21}$  top down.

Figure 289: Distributions for Higgs boson tagging using calorimeter clusters  $\beta = 0.5$ . C2, D2,  $\tau_{21}$  top down.



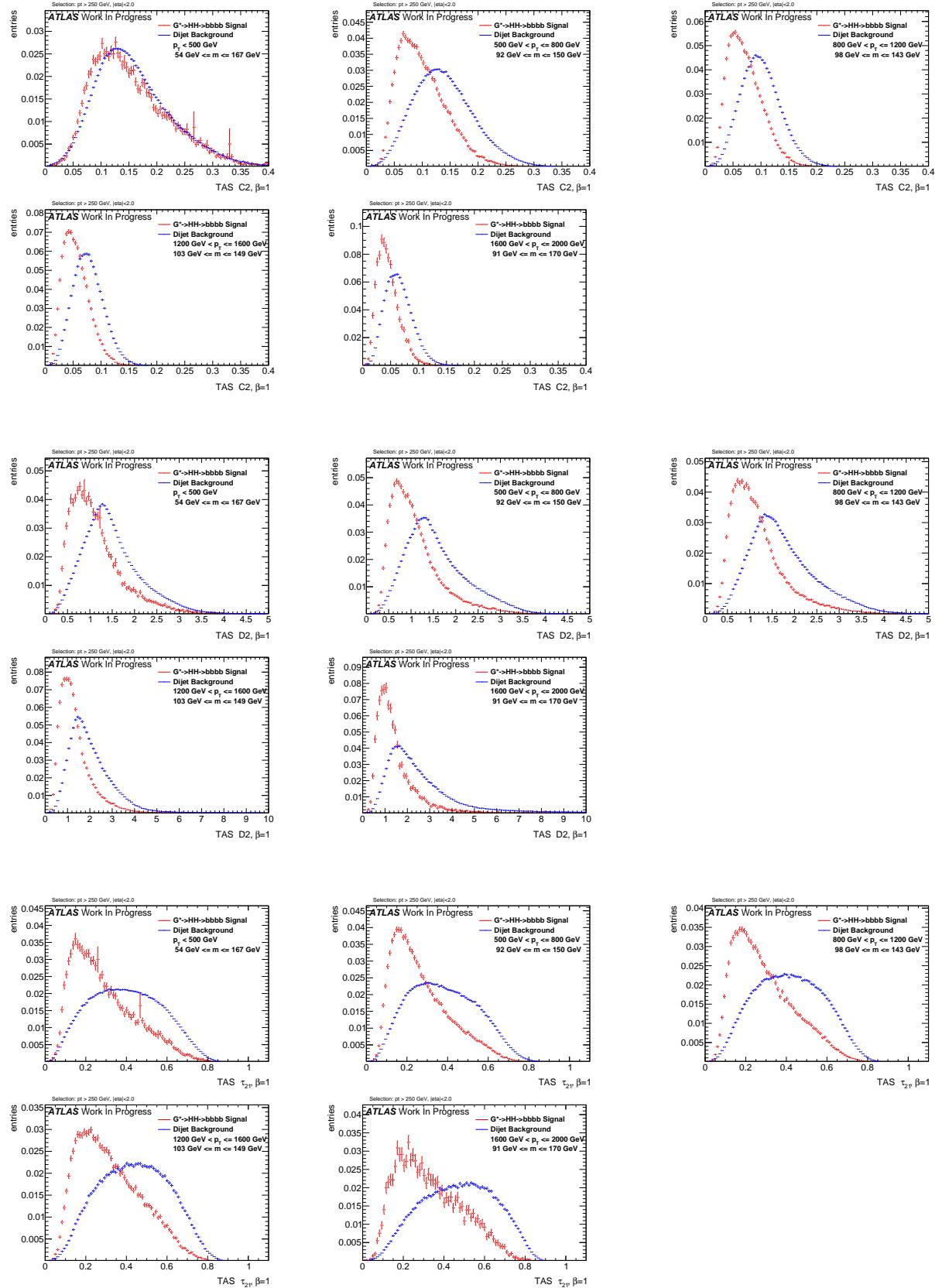
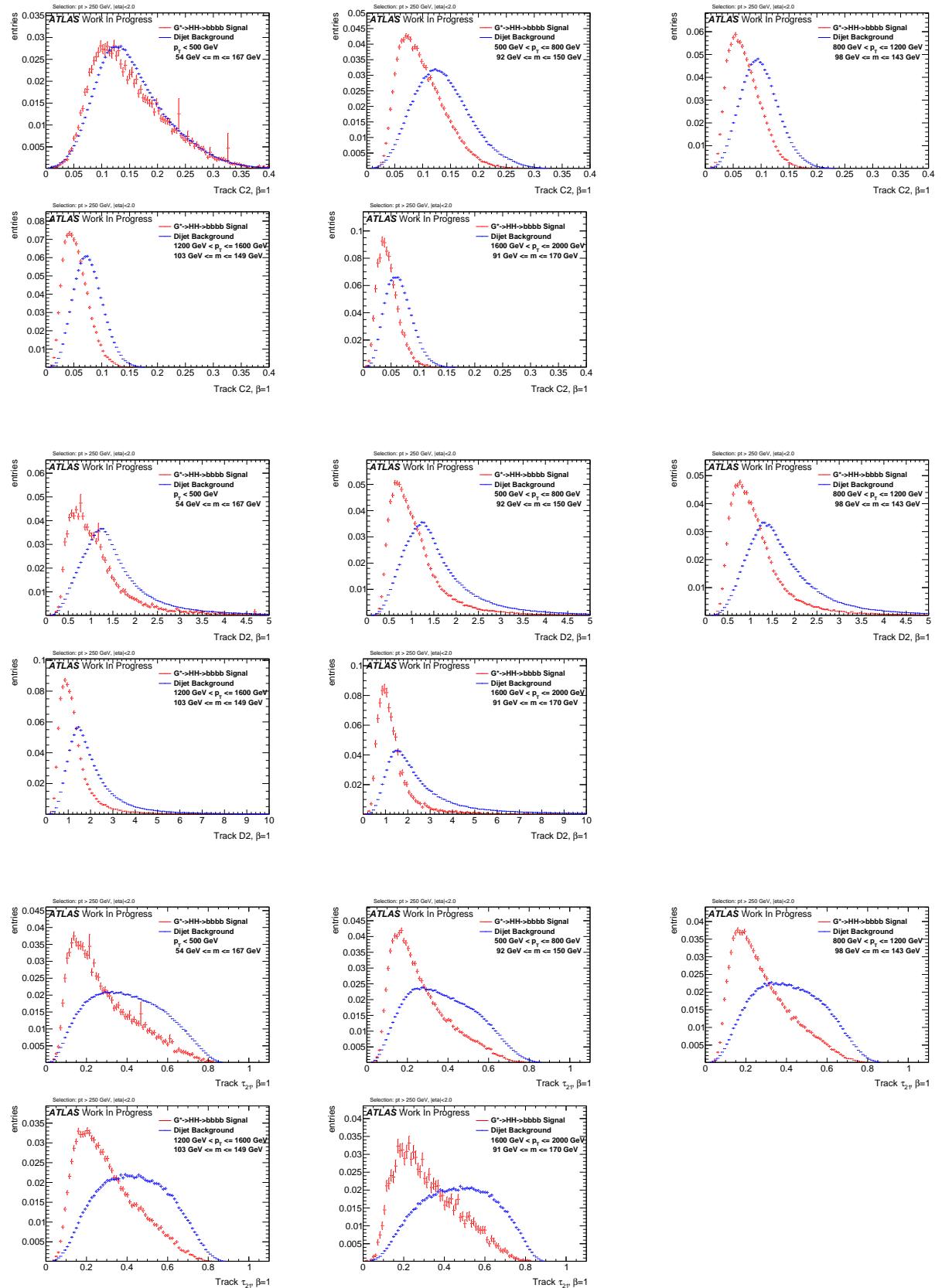
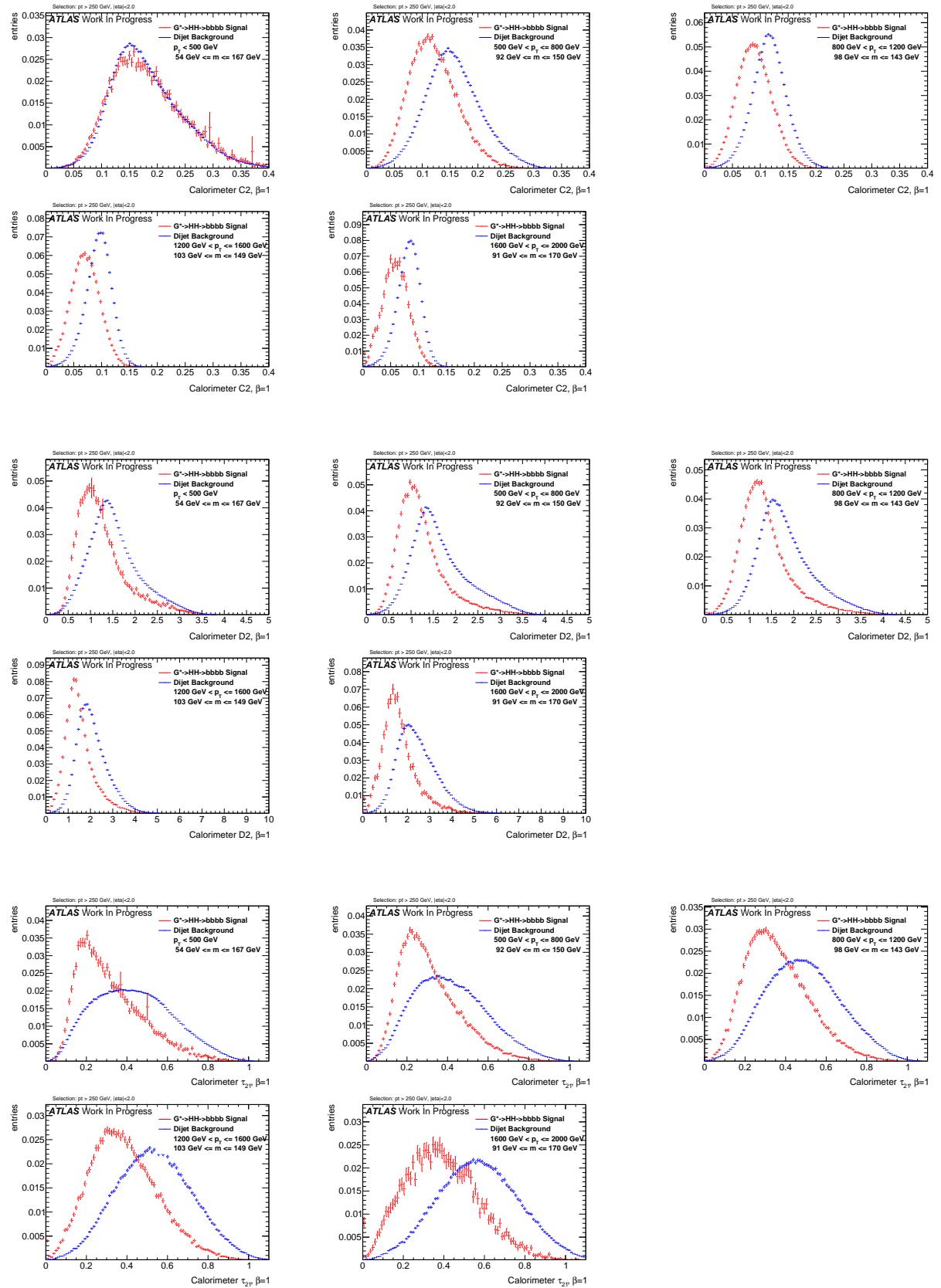
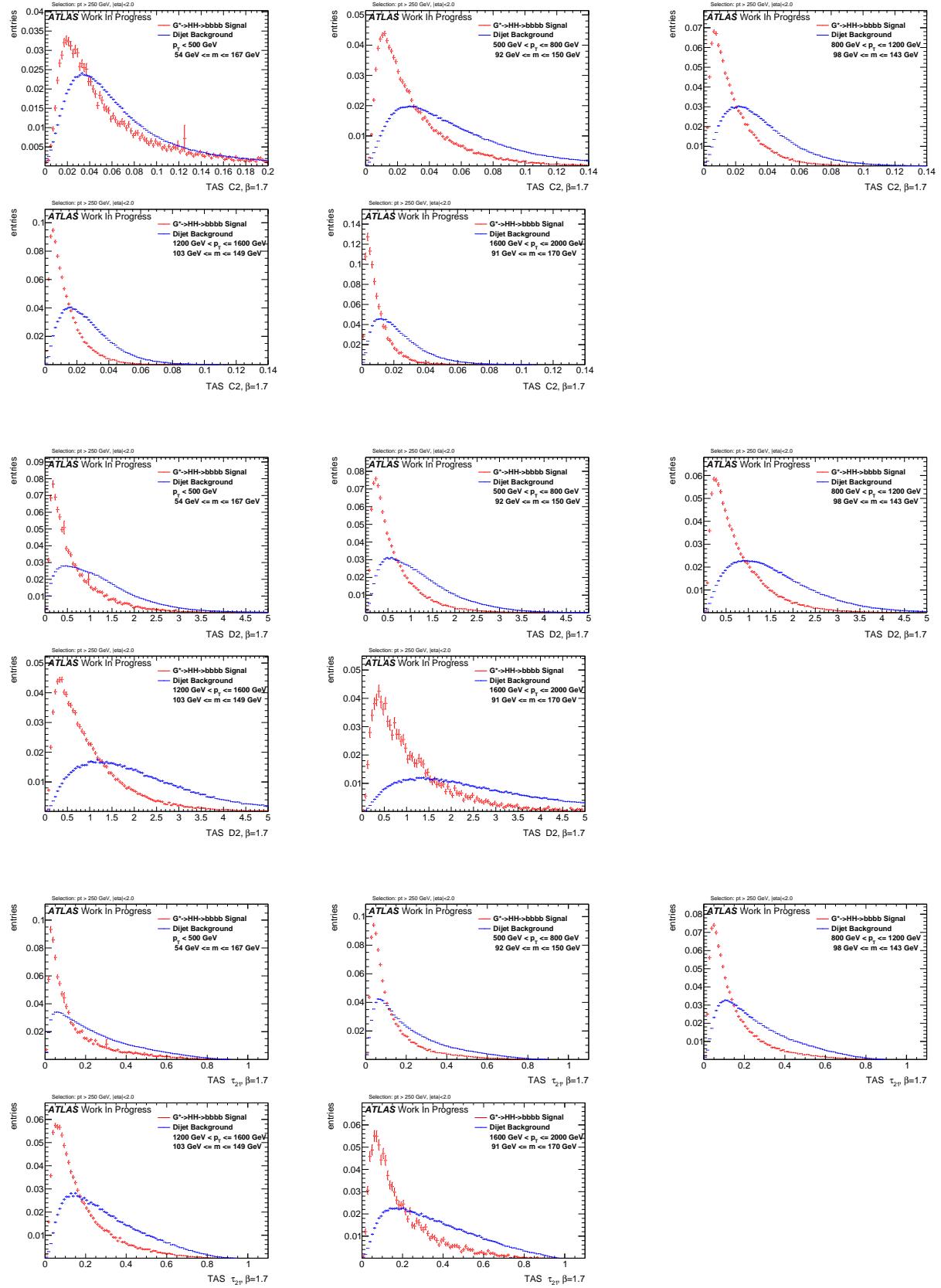
901  $\beta = 1$ 

Figure 290: Distributions for Higgs boson tagging using TAS  $\beta = 1$ . C2, D2,  $\tau_{21}$  top down.  
29th December 2016 – 17:17

Figure 291: Distributions for Higgs boson tagging using tracks  $\beta = 1$ . C2, D2,  $\tau_{21}$  top down.

Figure 292: Distributions for Higgs boson tagging using calorimeter clusters  $\beta = 1$ . C2, D2,  $\tau_{21}$  top down.

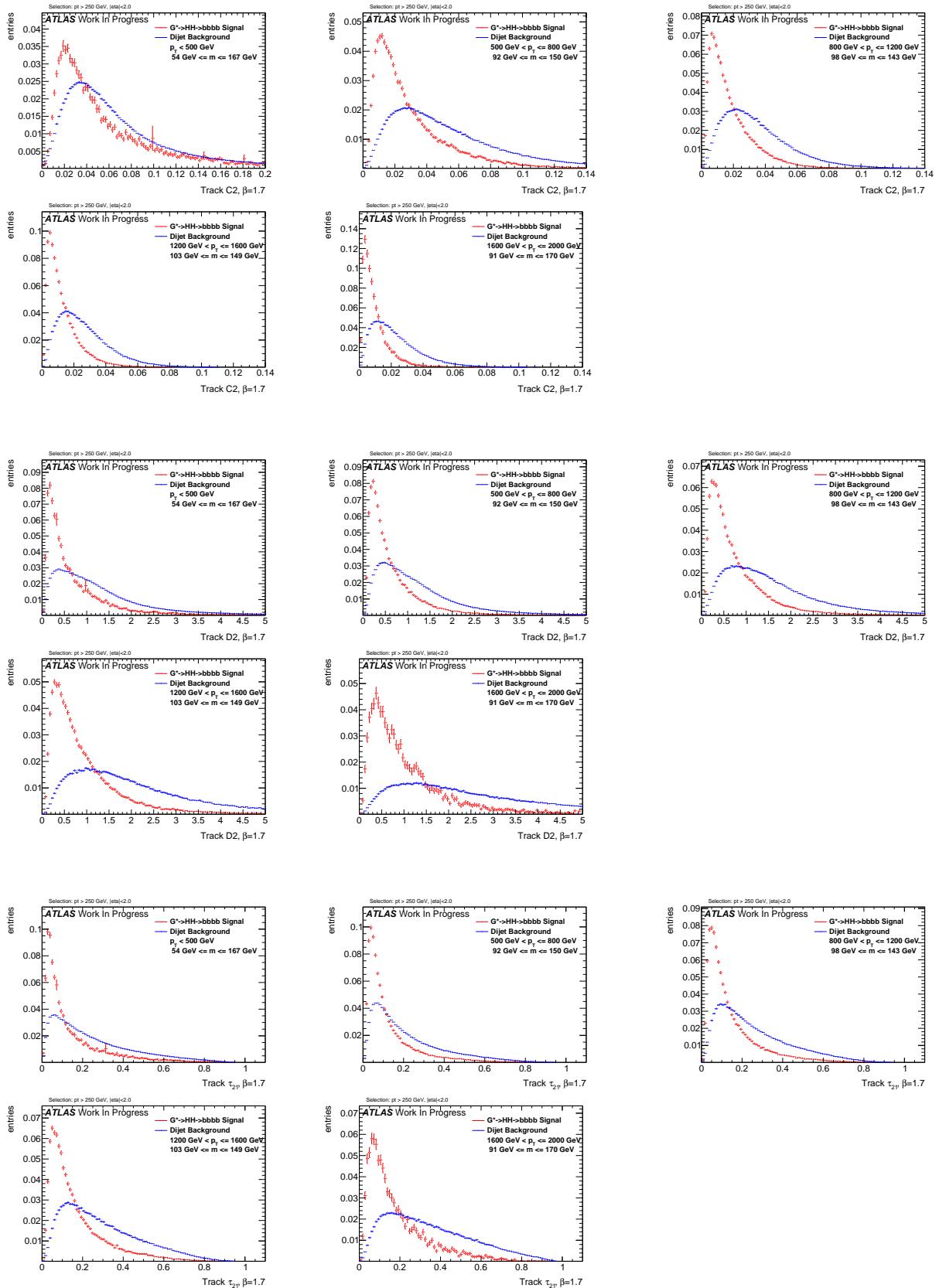


902  $\beta = 1.7$ 

29th December 2016 – 17:17

Figure 293: Distributions for Higgs boson tagging using TAS  $\beta = 1.7$ . C2, D2,  $\tau_{21}$  top down.

132

Figure 294: Distributions for Higgs boson tagging using tracks  $\beta = 1.7$ . C2, D2,  $\tau_{21}$  top down.

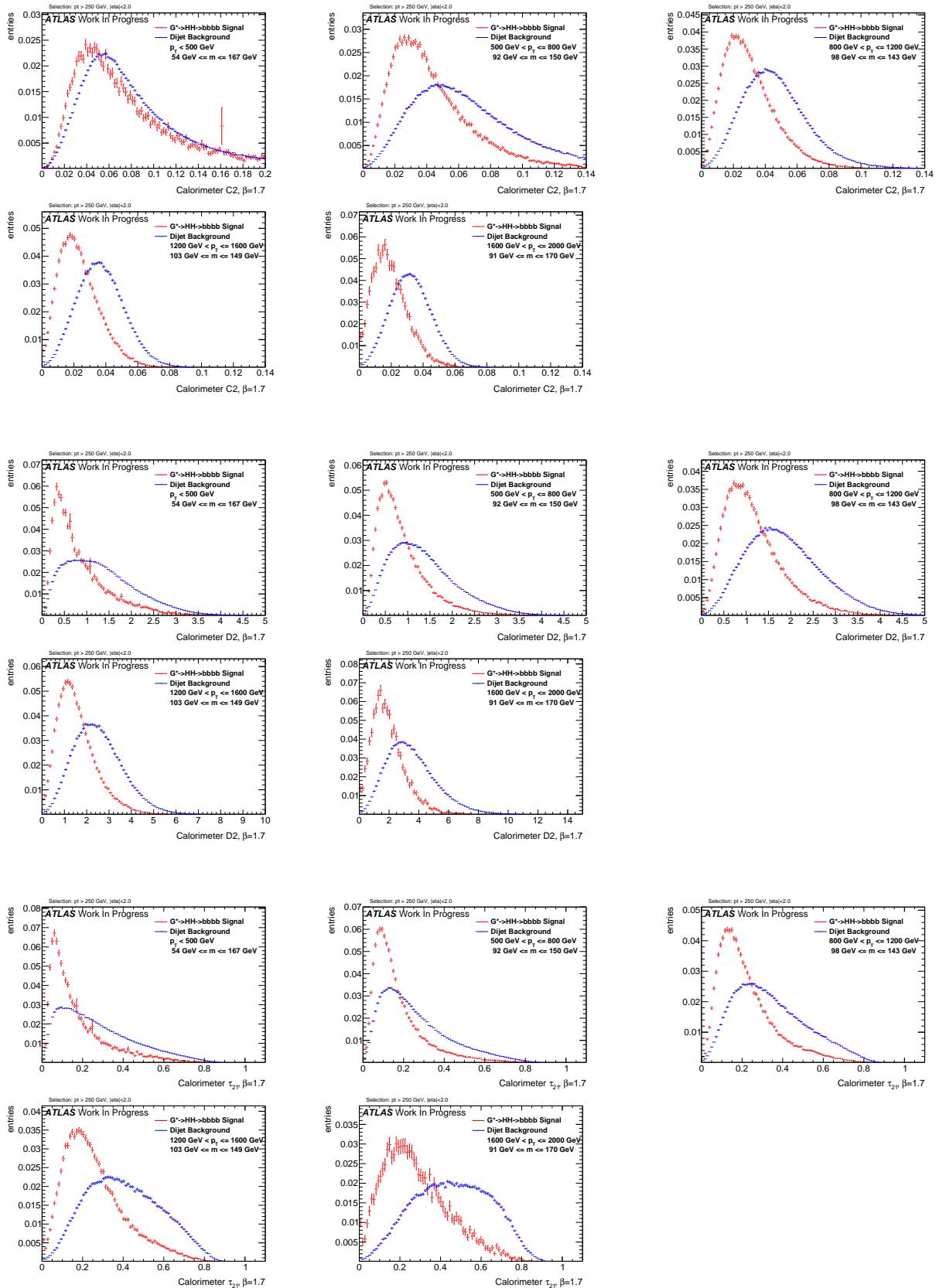
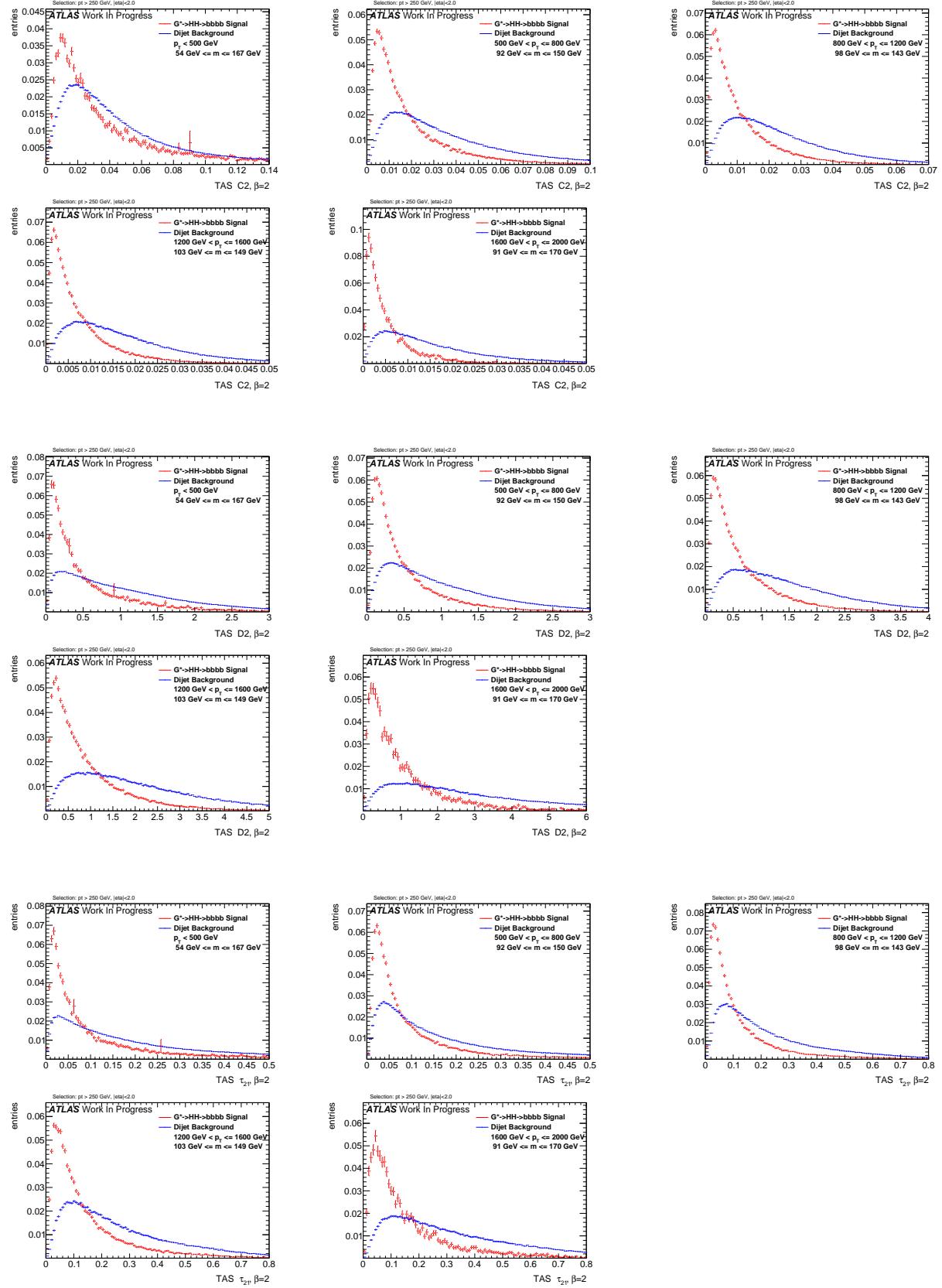


Figure 295: Distributions for Higgs boson tagging using calorimeter clusters  $\beta = 1.7$ . C2, D2,  $\tau_{21}$ , top down.

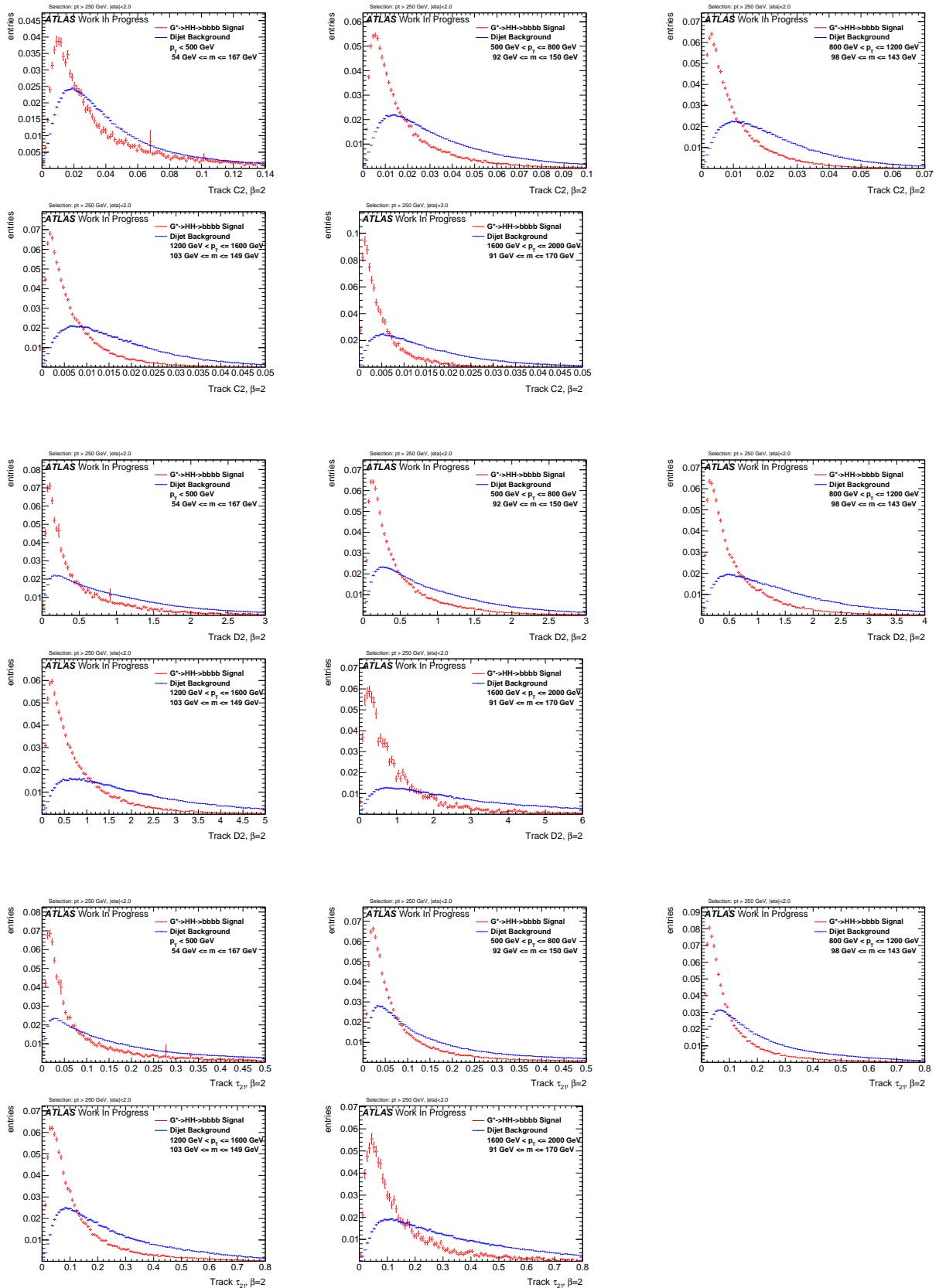


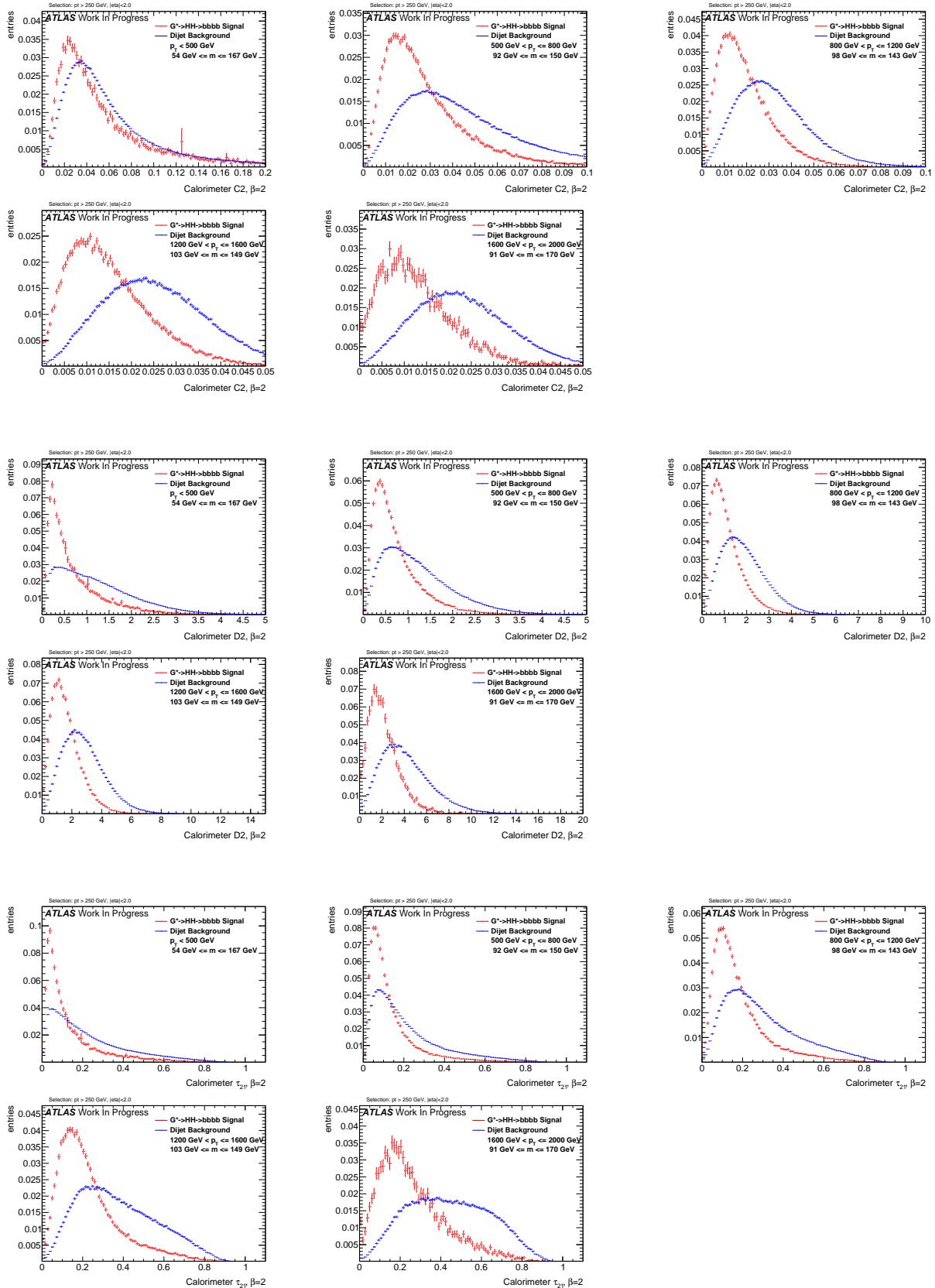
903  $\beta = 2$ 

29th December 2016 17:17

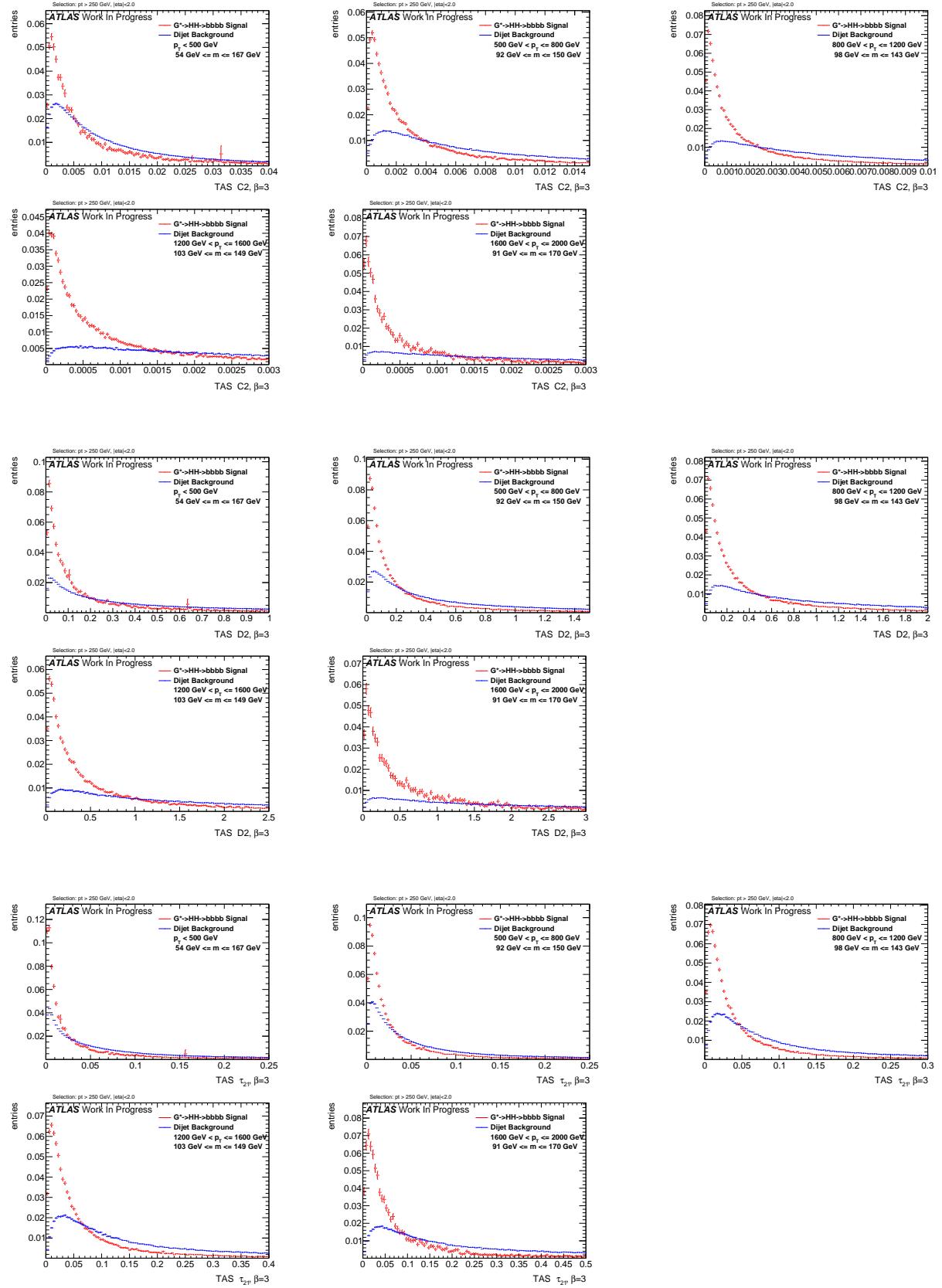
Figure 296: Distributions for Higgs boson tagging using TAS  $\beta = 2$ . C2, D2,  $\tau_{21}$  top down.

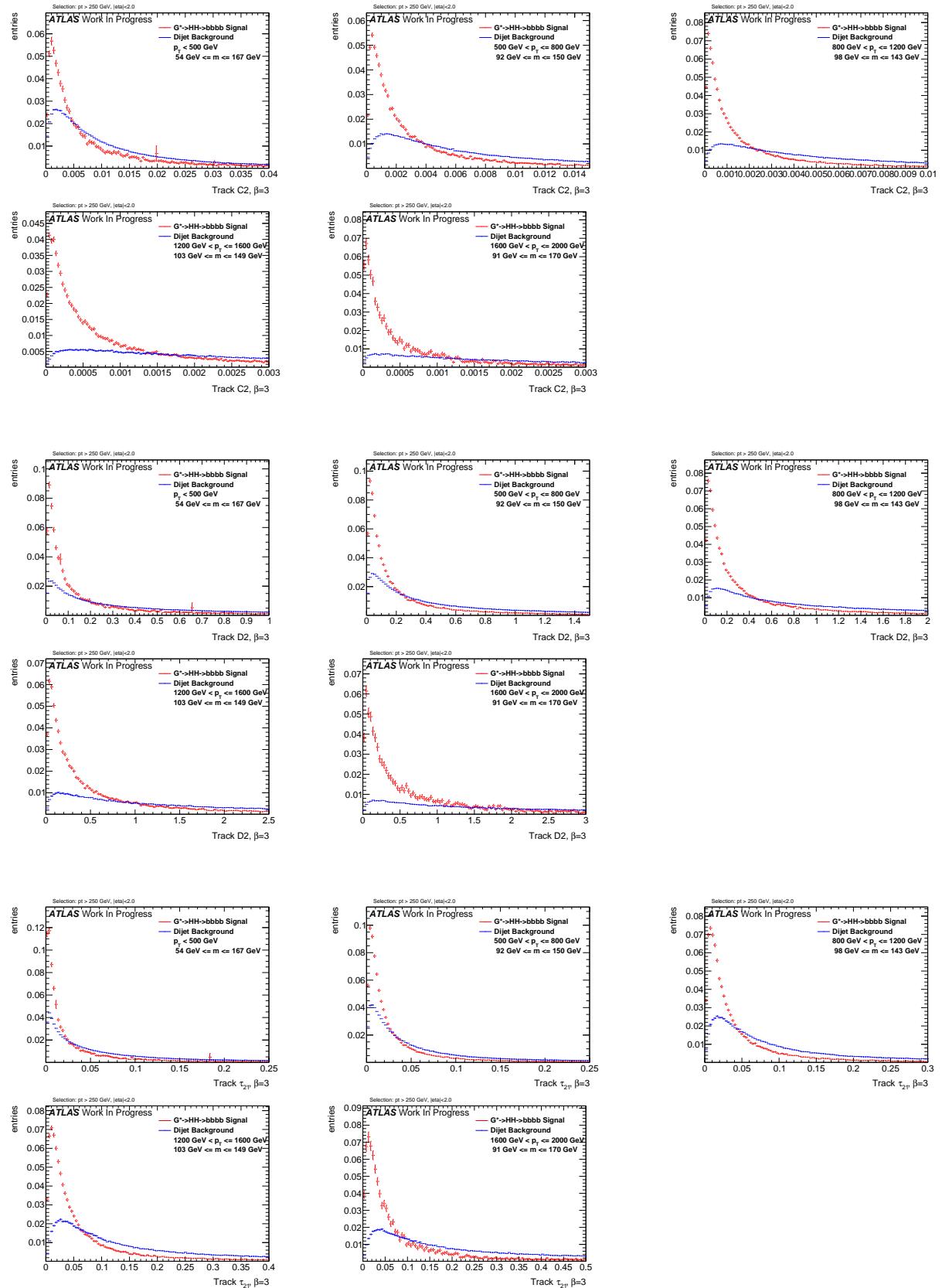
136

Figure 297: Distributions for Higgs boson tagging using tracks  $\beta = 2$ . C2, D2,  $\tau_{21}$  top down.

Figure 298: Distributions for Higgs boson tagging using calorimeter clusters  $\beta = 2$ . C2, D2,  $\tau_{21}$  top down.



904  $\beta = 3$ 

Figure 300: Distributions for Higgs boson tagging using tracks  $\beta = 3$ . C2, D2,  $\tau_{21}$  top down.

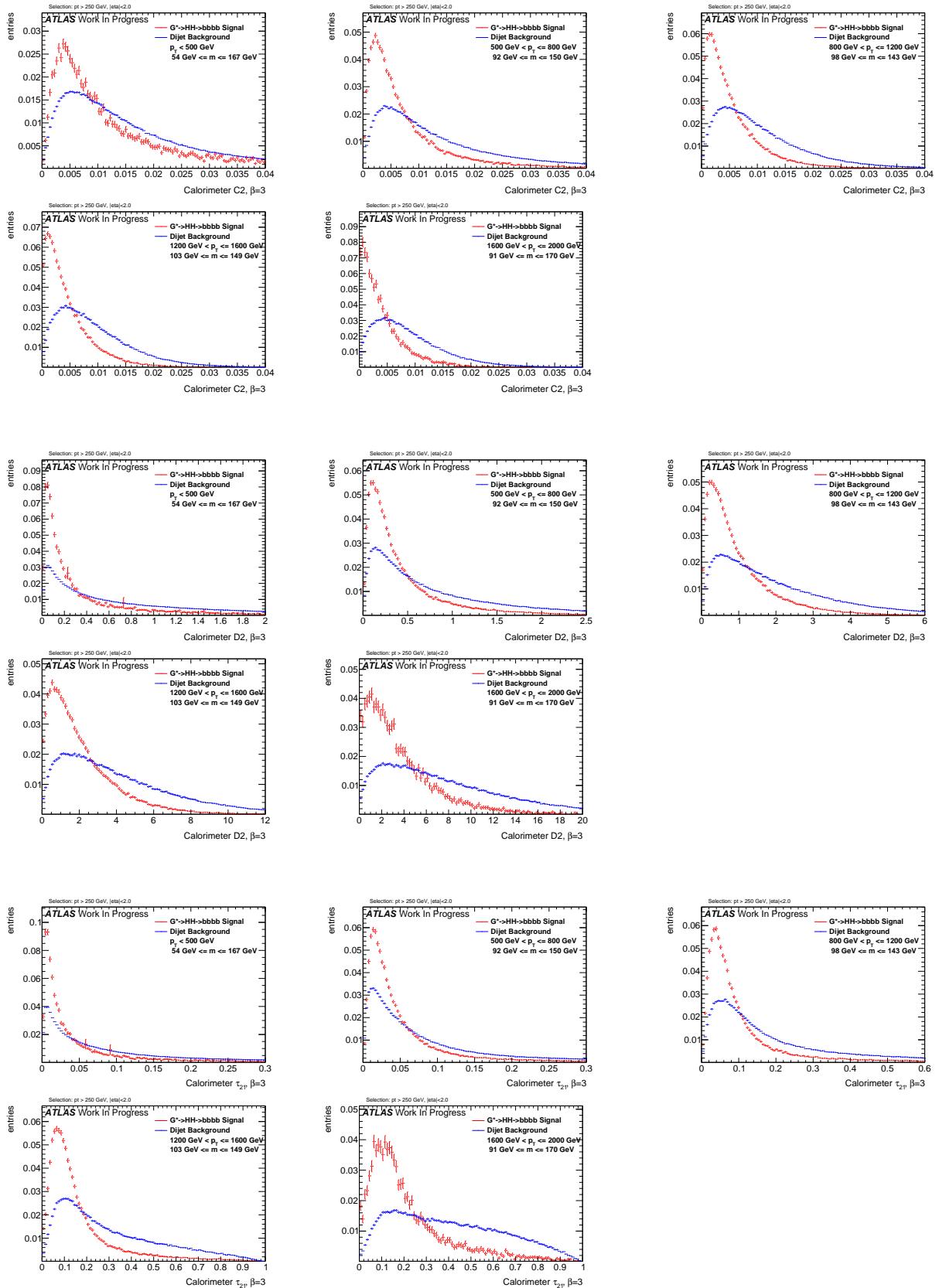
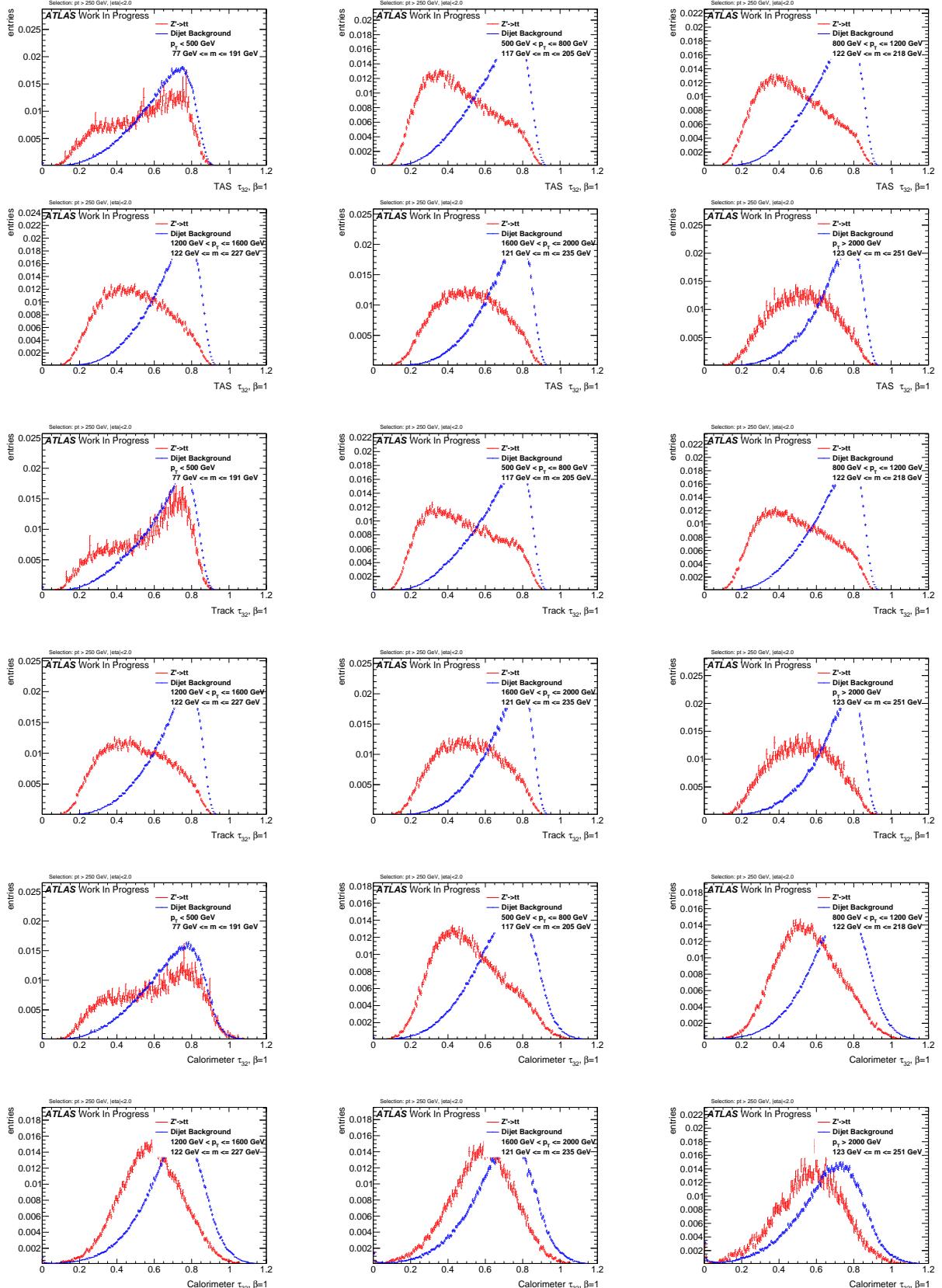


Figure 301: Distributions for Higgs boson tagging using calorimeter clusters  $\beta = 3$ . C2, D2,  $\tau_{21}$  top down.



### 905 N.3 Top Distributions

906  $\beta = 1$

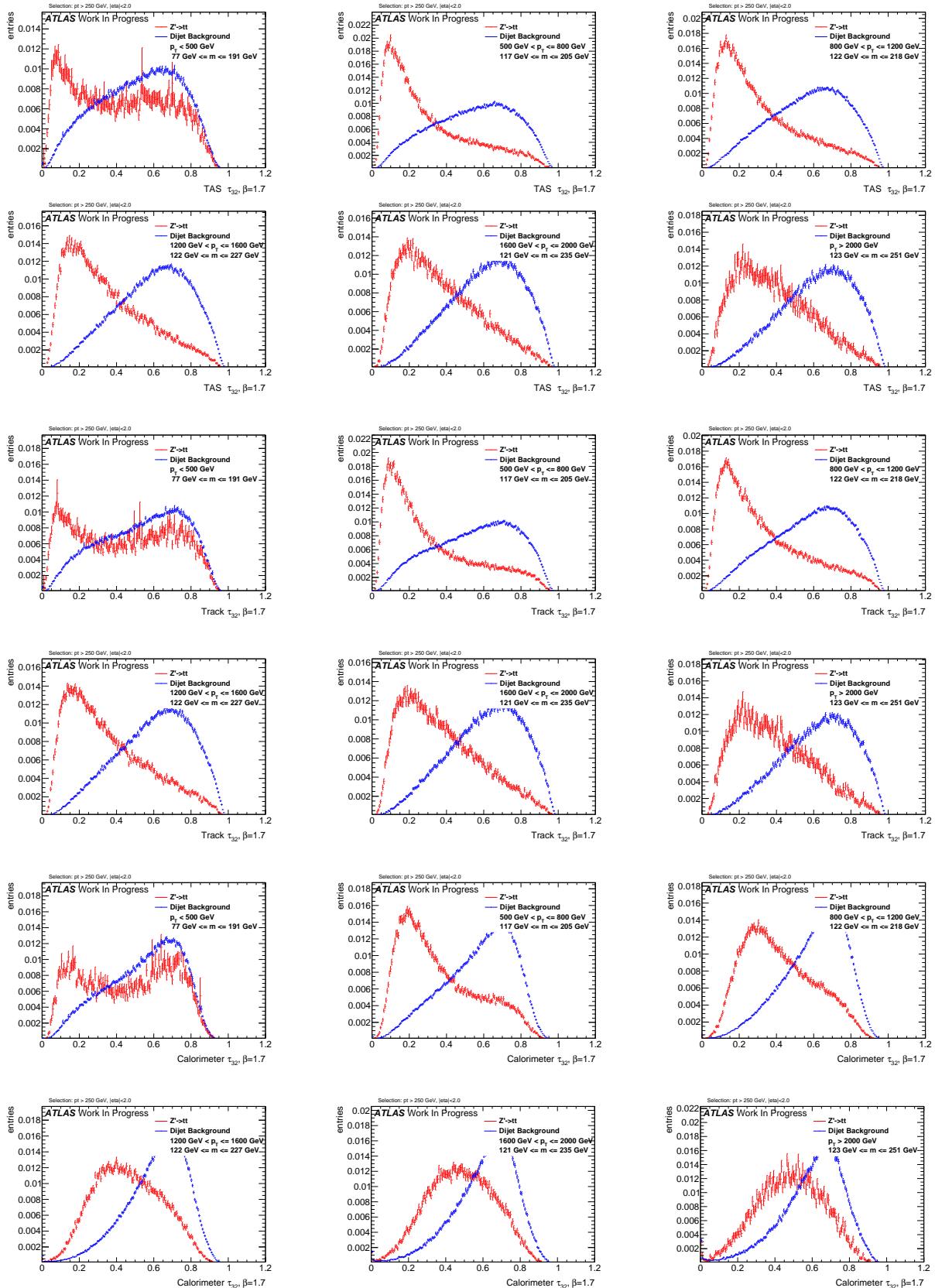


29th December 2016 – 17:17

Figure 302: Distributions for Top tagging using  $\tau_{32}$  ( $\beta = 1$ ) using TAS, tracks and calorimeter clusters top down.

144



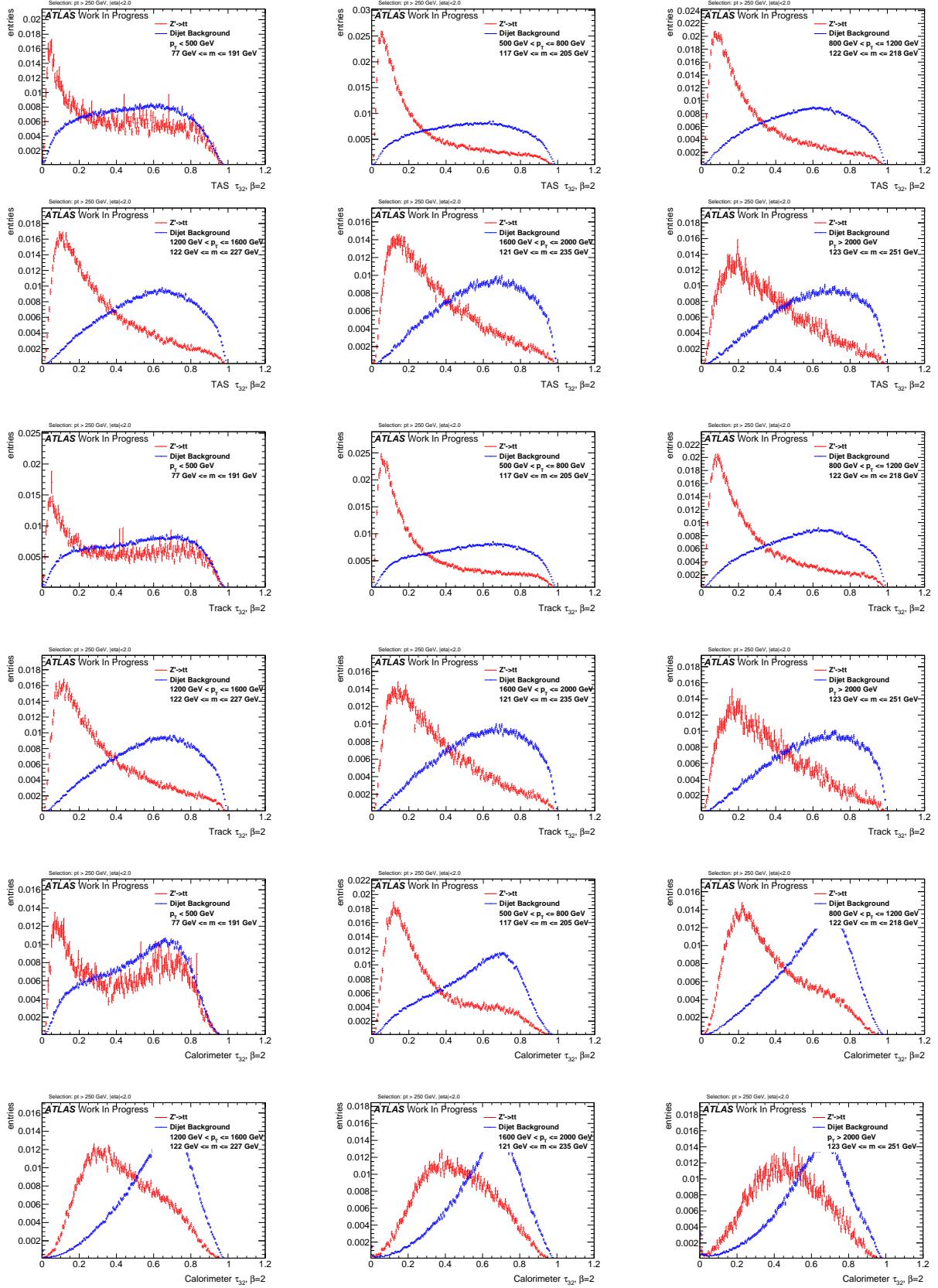
907  $\beta = 1.7$ 

29th December 2016 – 17:17

Figure 303: Distributions for Top tagging using  $\tau_{32}$  ( $\beta = 1.7$ ) using TAS, tracks and calorimeter clusters top down.

146



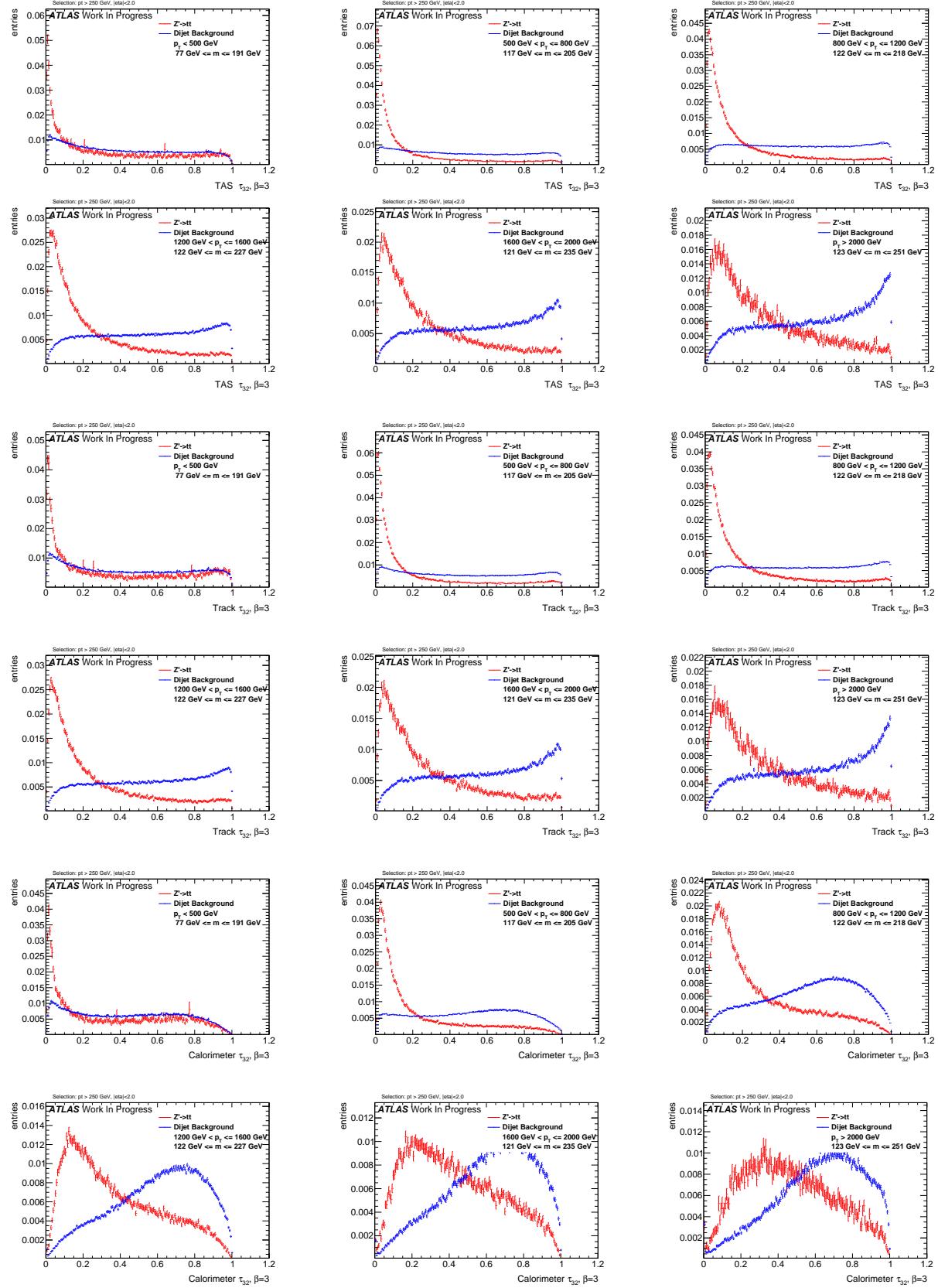
908  $\beta = 2$ 

29th December 2016 – 17:17

Figure 304: Distributions for Top tagging using  $\tau_{32}$  ( $\beta = 2$ ) using TAS, tracks and calorimeter clusters top down.

148



909  $\beta = 3$ 

29th December 2016 – 17:17

Figure 305: Distributions for Top tagging using  $\tau_{32}$  ( $\beta = 3$ ) using TAS, tracks and calorimeter clusters top down.

150

## 910 References

- 911 [1] Particle Data Group Collaboration, K. Olive *et al.*, *Review of Particle Physics*, Chinese Physics C 38  
 912 (2014) 090001.
- 913 [2] Paul Langacker, *The Physics of Heavy Z Gauge Bosons*. In: Rev. Mod. Phys. 81 (2009), pp. 1199–1228.  
 914 doi: 10.1103/RevModPhys.81.1199. arXiv: 0801.1345 [hep-ph].
- 915 [3] Martin Schmaltz and Christian Spethmann, *Two Simple W' Models for the Early LHC*. In: JHEP 07  
 916 (2011), p. 046. doi: 10.1007/JHEP07(2011)046. arXiv: 1011. 5918 [hep-ph].
- 917 [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:  
 918 10.1016/j.physletb.2016.02.015. arXiv: 1512.05099 [hep-ex].
- 920 [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  
 921 [hep-ex].
- 923 [6] C. Patrignani *et al.*, Particle Data Group, Chin. Phys. C, 40, 100001 (2016).  
 924 <http://pdg.lbl.gov/2015/reviews/rpp2015-rev-wprime-searches.pdf>
- 925 [7] C. Patrignani *et al.*, Particle Data Group, Chin. Phys. C, 40, 100001 (2016).  
 926 <http://pdg.lbl.gov/2015/reviews/rpp2015-rev-zprime-searches.pdf>
- 927 [8] L. Randall and R. Sundrum, *A Large Mass Hierarchy from a Small Extra Dimension*, Phys. Rev. Lett.  
 928 83 (1999) 3370, arXiv: hep-ph/9905221.
- 929 [9] H. Davoudiasl, J.L. Hewett and T.G. Rizzo, *Warped Phenomenology*, Stanford Linear Accelerator  
 930 Center Stanford CA 94309, USA, arXiv:hep-ph/9909255v1 6 Sep 1999.
- 931 [10] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1 & 2*  
 932 , Tech. Rep. ATLAS TDR 14, CERN/LHCC 99-14, Geneva, 1999.
- 933 [11] L. Evans and P. Bryant, *LHC Machine*, JINST 3 (2008) S08001.
- 934 [12] C. Patrignani *et al.*, Particle Data Group, Chin. Phys. C, 40, 100001 (2016).  
 935 <http://pdg.lbl.gov/2010/reviews/rpp2010-rev-qcd.pdf>
- 936 [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*, CERN-PH-EP/2013-037 2015/05/04 arXiv:hep-ph/14121633.
- 939 [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.
- 941 [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).
- 943 [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].
- 945 [17] Matteo Cacciari, Gavin P. Salam, Gregory Soyez, *The anti- $k_t$  jet clustering algorithm*, arXiv:0802.1189v2 [hep-ph].

- [18] F. Halzen, A.D. Martin and D.M. Scott, Phys. Rev. D25 (1982) 754, Phys. Lett. B112 (1982) 160.
- [19] Torbjorn Sjostrand, *Monte Carlo Generators*, hep-ph/0611247, CERN-LCGAPP-2006-06.
- [20] W. Lampl *et al.*, *Calorimeter Clustering Algorithms: Description and Performance*, ATL-LARG-PUB-2008-002.
- [21] The ATLAS Collaboration, *Jet energy resolution in proton-proton collisions at  $\sqrt{s} = 7$  TeV recorded in 2010 with the ATLAS detector*, CERN-PH-EP-2012-191 arXiv:1210.6210v1 [hep-ex].
- [22] GEANT4 Collaboration (S. Agostinelli *et al.*), *GEANT4: A Simulation toolkit*, Nucl.Instrum.Meth. A506 (2003) 250-303 DOI: 10.1016/S0168-9002(03)01368-8 SLAC-PUB-9350, FERMILAB-PUB-03-339.
- [23] Ariel Schwartzman, *Jet energy calibration at the LHC*, SLAC National Accelerator Laboratory, 2575 Sand Hill Road Menlo Park, California 94025, USA, arXiv:1509.05459v1 [hep-ex] 17 Sep 2015.
- [24] The ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035 19 July 2016.
- [25] The ATLAS Collaboration, *Boosted Higgs ( $\rightarrow b\bar{b}$ ) Boson Identification with the ATLAS Detector at  $\sqrt{s} = 13$  TeV*, ATLAS-CONF-2016-039 30th July 2016.
- [26] The ATLAS Collaboration, *Search for high-mass diboson resonances with boson-tagged jets in proton-proton collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector*, arXiv:1506.00962v3 [hep-ex].
- [27] The CMS Collaboration, *Combination of searches for WW, WZ, ZZ, WH, and ZH resonances at  $\sqrt{s} = 8$  and 13 TeV*, CMS PAS B2G-16-007.
- [28] The ATLAS Collaboration, *Search for heavy particles decaying to pairs of highly-boosted top quarks using lepton-plus-jets events in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS-CONF-2016-014.
- [29] The ATLAS Collaboration, *Search for resonances with boson-tagged jets in  $15.5 \text{ fb}^{-1}$  of p-p collisions at  $\sqrt{s} = 13$  TeV collected with the ATLAS detector*, ATLAS-CONF-2016-055.
- [30] The ATLAS Collaboration, *Search for pair production of Higgs bosons in the  $b\bar{b}\bar{b}\bar{b}$  final state using proton proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, ATLAS-CONF-2016-049.
- [31] The ATLAS Collaboration, *Performance of jet substructure techniques for large-R jets in proton-proton collisions at  $\sqrt{s} = 7$  TeV using the ATLAS detector*, CERN-PH-EP-2013-069 arXiv:1306.4945 [hep-ex].
- [32] Oleg Zenin, on behalf of the ATLAS Collaboration, *Soft QCD and underlying event measurements at ATLAS*, Nuclear and Particle Physics Proceedings Volumes 273275, April-June 2016, Pages 2053-2058 37th International Conference on High Energy Physics (ICHEP).
- [33] Page title: JetSubstructureECFA2014 AtlasPublic TWiki Website name: Twiki.cern.ch URL:<https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetSubstructureECFA2014> Access date: 27th October 2016.
- [34] E. Lehmann and G. Casella, *Theory of Point Estimation*, Springer Verlag, 1998, isbn: 0387985026.
- [35] Andrew J. Larkoski, Ian Moult, and Duff Neill. *Power Counting to Better Jet Observables*. *JHEP*, 12:009, 2014.

- 985 [36] Andrew J. Larkoski, Ian Moult, and Duff Neill. *Analytic Boosted Boson Discrimination*. *JHEP*,  
986 05:117, 2016.
- 987 [37] Andrew J. Larkoski, Gavin P. Salam, and Jesse Thaler. *Energy Correlation Functions for Jet Sub-*  
988 *structure*. *JHEP*, 06:108, 2013.
- 989 [38] Jesse Thaler and Ken Van Tilburg. *Identifying Boosted Objects with N-subjettiness*. *JHEP*, 03:015,  
990 2011.
- 991 [39] Georges Aad et al. *Identification of boosted, hadronically decaying W bosons and comparisons with*  
992 *ATLAS data taken at  $\sqrt{s} = 8$  TeV*. *Eur. Phys. J.*, C76(3):154, 2016.
- 993 [40] Jet Substructure and jet-by-jet tagging meeting: URL: ht-  
994 [tps://indico.cern.ch/event/495766/contributions/1174422/attachments/1231007/1804520/JSS\\_18Feb.pdf](https://indico.cern.ch/event/495766/contributions/1174422/attachments/1231007/1804520/JSS_18Feb.pdf)

995 **List of contributions**

996

## **Auxiliary material**

997 In an ATLAS paper, auxiliary plots and tables that are supposed to be made public should be collected in  
999 an appendix that has the title ‘Auxiliary material’. This appendix should be printed after the Bibliography.  
1000 At the end of the paper approval procedure, this information can be split into a separate document – see  
1001 `atlas-auxmat.tex`.

1002 In an ATLAS note, use the appendices to include all the technical details of your work that are relevant  
1003 for the ATLAS Collaboration only (e.g. dataset details, software release used). This information should  
1004 be printed after the Bibliography.