tau-ID-efficiency uncertainty at high p_{T}

Dirk Duschinger

September 23, 2013

Intention

- tau-ID-efficiency uncertainty is one of the largest systematic in $Z' \to \tau\tau$-analysis
- uncertainty is measured in $Z \to \tau\tau$ decays only considering taus with $p_{\mathrm{T}} < 100$ GeV
- no applicable Standard Model source of real high p_{T} taus
- need to know the ID-efficiency uncertainty up to p_{T} of 1000 GeV
- question is now, whether the fidelity is any poorer at high p_{T} than at low p_{T}
- assumption: data-mismodeling comes from tau decay modeling and detector response
- recipe:
  - compare ID-efficiency in systematically varied MC-Samples
  - compare di-jet events in data and MC to check that detector response doesn't deteriorate at high p_{T}

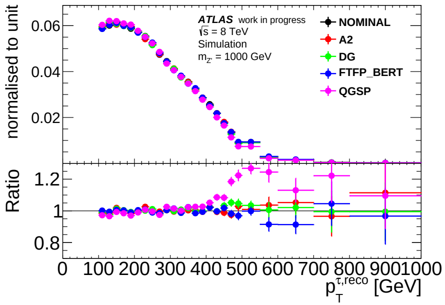
Impact of shower model on ID-efficiency

Definition and samples

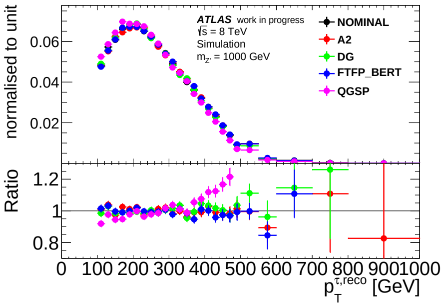
- PYTHIA generated $Z'(1000) \rightarrow \tau^+\tau^-$-samples with different configurations:
  - NOMINAL:
    QGSP_BERT shower-model
  - FTFP_BERT:
    FTFP_BERT shower-model
  - QGSP:
    QGSP shower-model
  - Detector-Geometry:
    QGSP_BERT shower-model, different detector geometry
  - A2:
    QGSP_BERT shower-model, different ATLAS-tune
- $\epsilon_{\tau ID}^{\text{n-prong}} = \frac{\#\tau^{\text{n-prong}}_{\text{reconstructed, truth matched, passed ID}}}{\#\tau^{\text{n-prong}}_{\text{reconstructed, truth matched}}}$

p_{T} distributions before ID

1 prong

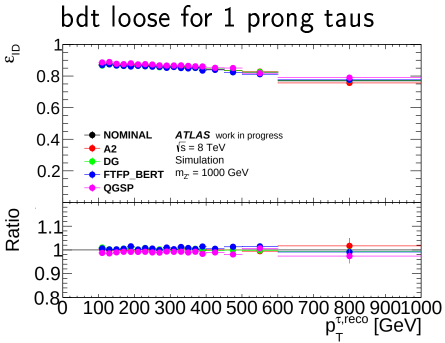


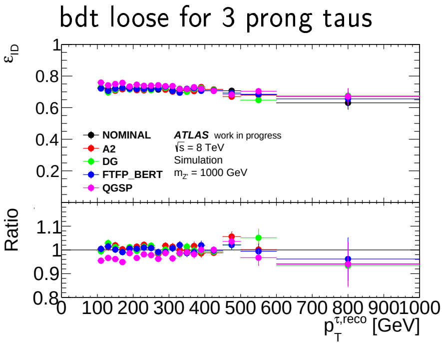
3 prong



- good agreement between samples up to 400 GeV

Comparison of ID-efficiencies





- no large trend in p_T between the different samples within the statistical uncertainty
- simulated shower model has no impact on p_T dependence of the ID-efficiency for signal taus

Detector response comparison in data and MC

Overview

- compare fake-tau-ID-efficiencies in di-jet events in data and MC
- use a tag and a probe style analysis

Tag Selection	Probe Selection
$p_T > 100$ GeV	$p_T > 100$ GeV
$\lvert\eta\rvert < 1.37, 1.52 < \lvert\eta\rvert < 2.5$	$\lvert\eta\rvert < 1.37, 1.52 < \lvert\eta\rvert < 2.5$
$N_{\text{track}} > 1$	$N_{\text{track}} = 1$ or 3
	loose BDT electron veto
	ϕ not in $[-2.5, -2.0]$

- back-to-back: $\cos\Delta\phi(\text{tag,probe}) < -0.95$
- p_{T}-balance: $\left|\frac{p_T^{\text{probe}} - p_T^{\text{tag}}}{p_T^{\text{tag}}}\right| < 0.05$
- jet trigger

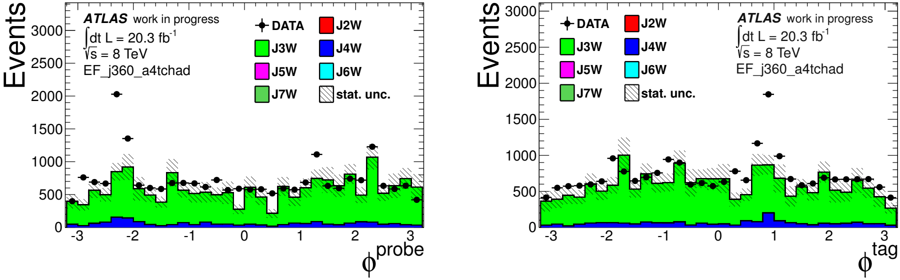
Used trigger and samples

Trigger	int. Luminosity $[pb^{-1}]$
EF_j110_a4tchad	9.84371
EF_j145_a4tchad	36.3830
EF_j180_a4tchad	79.0276
EF_j220_a4tchad	262.236
EF_j360_a4tchad	20340.0
EF_j460_a4tchad	20340.0

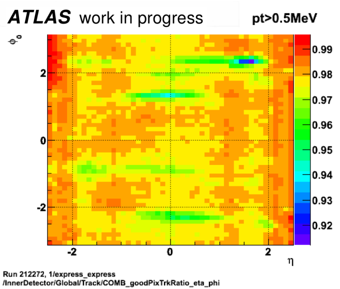
- Samples:
  - data12: periods A-L
  - mc12: J2W-J7W

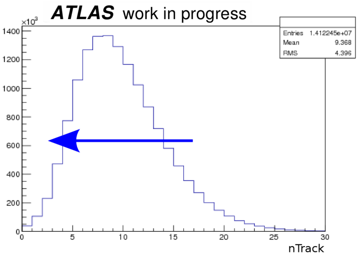
sample	generator p_T range [GeV]	cross section $[pb]$	filter efficiency	generated events
J2W	$80 - 200$	2.6357e07	3.9939E-03	5999034
J3W	$200 - 500$	5.4407e05	1.2187E-03	5922214
J4W	$500 - 1000$	6.4454e03	7.0821E-04	5862258
J5W	$1000 - 1500$	3.9742e01	2.1521E-03	2921084
J6W	$1500 - 2000$	4.1604e-01	4.6843E-03	2968652
J7W	$2000 - \infty$	4.0630e-02	1.4600E-02	2991955

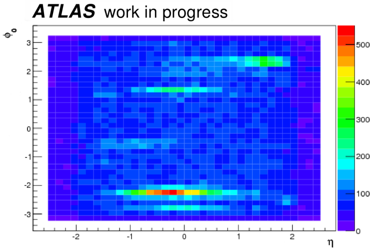
ϕ-distribution



- discrepancy between data and MC was observed in 2011 $Z' \to \tau\tau$ analysis
- also visible in 2012 analysis



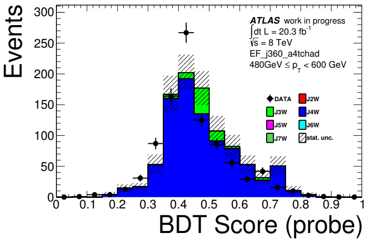




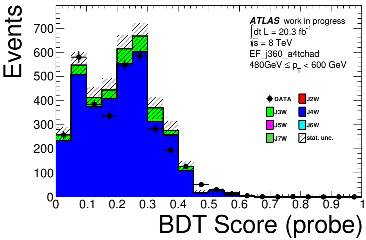
- regions in pixel detector with lower track resulution yields in track migration from jets with more to jets with less tracks
- this feature is not well modeled in MC

BDTJetScore-distribution for events passing the EF_j360_a4tchad trigger

1 track



3 track



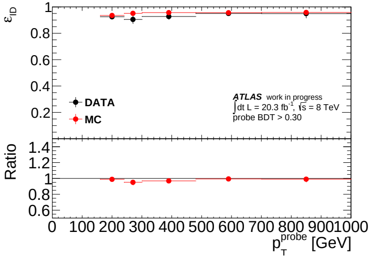
Trigger	p_T-range [GeV]
EF_j110_a4tchad	160-185
EF_j145_a4tchad	185-240
EF_j180_a4tchad	240-300
EF_j220_a4tchad	300-480
EF_j360_a4tchad	480-600
EF_j460_a4tchad	600-∞

- events in a p_{T} region are only considered when they pass the corresponding trigger

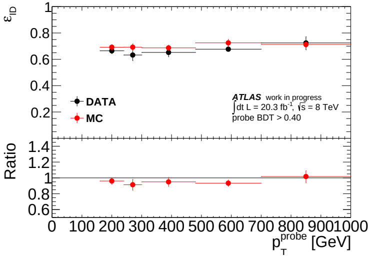
ID-efficiency for 1 track jets

- ID-efficiencies are computed for events in the p_T regions, passing the corresponding jet trigger
- $\epsilon_{\tau_{ID}} = \frac{\#\text{probe jets}_{\text{passed ID}}}{\#\text{probe jets}}$

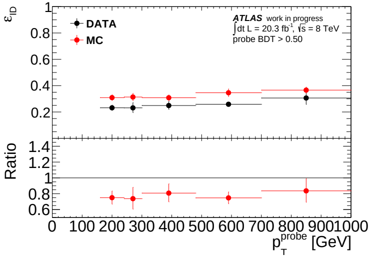
1 track, BDT$>$ 0.30



1 track, BDT$>$ 0.40



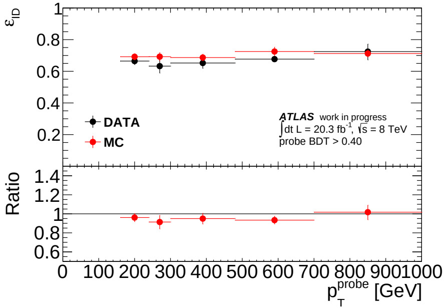
1 track, BDT$>$ 0.50



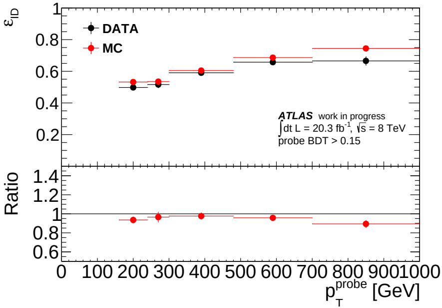
ID-efficiency for 1 and 3 track jets

- for further analysis BDT requirements are chosen, such that the fake ID-efficiency is comparable to ID-efficiency of signal taus in $Z' \rightarrow \tau\tau$
- data/MC ratios of ID-efficiencies are sufficiently flat versus p_{T}

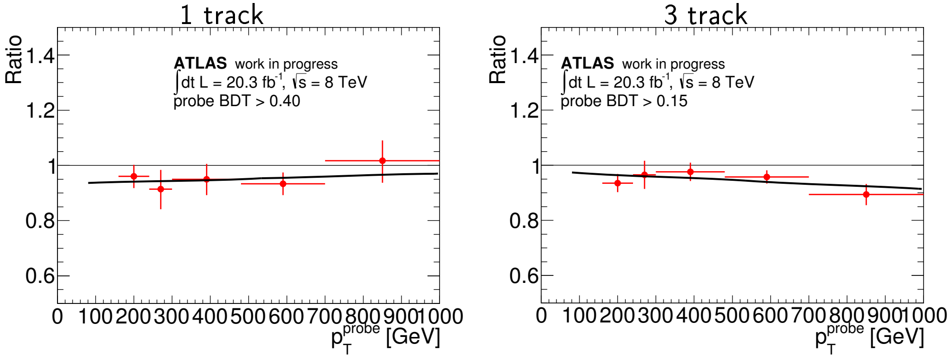
1 track, BDT $>$ 0.40



3 track, BDT $>$ 0.15



Fitting for 1 and 3 track jets



linear fit	$p1 =$	$8.8 \times 10^{-5} \pm 9.9 \times 10^{-5}$	$-1.5 \times 10^{-5} \pm 7.1 \times 10^{-5}$
constant fit	$p0 =$	$0.91 \pm 2.1 \times 10^{-2}$	$0.95 \pm 1.5 \times 10^{-2}$

- uncertainty of linear fit for inflation of ID-efficiency uncertainty
- relative uncertainty of constant fit is used as an upper bound

Conclusion

- shower shape variation has little impact on ID-efficiency
- tau ID-efficiency is well modeled in MC
- prescription for tau ID-efficiency uncertainty:

$$\begin{aligned}\Delta\varepsilon_{1p}(p_\mathrm{T} > 100\ \mathrm{GeV}) =& \Delta\varepsilon_{1p}(p_\mathrm{T} < 100\ \mathrm{GeV}) \\ &+ 0.0099\%/\mathrm{GeV} \cdot (p_\mathrm{T} - 100\ \mathrm{GeV})\end{aligned}$$

$$\begin{aligned}\Delta\varepsilon_{3p}(p_\mathrm{T} > 100\ \mathrm{GeV}) =& \Delta\varepsilon_{3p}(p_\mathrm{T} < 100\ \mathrm{GeV}) \\ &+ 0.0071\%/\mathrm{GeV} \cdot (p_\mathrm{T} - 100\ \mathrm{GeV})\end{aligned}$$

with an upper bound of

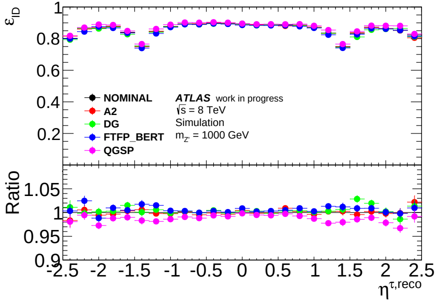
- $\Delta\varepsilon_1^{\mathrm{max}} = 2.3\%$ for 1 prong
- $\Delta\varepsilon_3^{\mathrm{max}} = 1.5\%$ for 3 prong

BACKUP

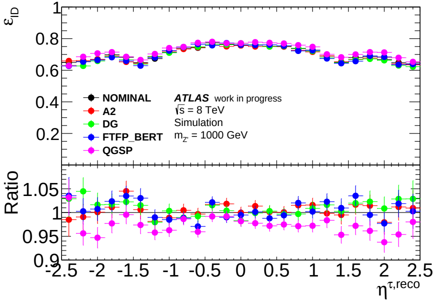
n-track	BDT threshold	linear fit p_1 [GeV^{-1}]	constant fit p_0
1	0.25	$3.8 \times 10^{-5} \pm 2.9 \times 10^{-5}$	0.99 ± 0.006
	0.30	$2.8 \times 10^{-5} \pm 3.9 \times 10^{-5}$	0.98 ± 0.008
	0.35	$4.5 \times 10^{-5} \pm 6.0 \times 10^{-5}$	0.96 ± 0.013
	0.40	$8.8 \times 10^{-5} \pm 9.9 \times 10^{-5}$	0.91 ± 0.021
	0.45	$-3.0 \times 10^{-4} \pm 1.5 \times 10^{-4}$	0.77 ± 0.031
	0.50	$-1.6 \times 10^{-4} \pm 1.9 \times 10^{-4}$	0.72 ± 0.040
	0.55	$-4.1 \times 10^{-4} \pm 2.5 \times 10^{-4}$	0.67 ± 0.051
	0.60	$-2.6 \times 10^{-4} \pm 3.5 \times 10^{-4}$	0.68 ± 0.068
3	0.05	$1.7 \times 10^{-5} \pm 3.1 \times 10^{-5}$	0.99 ± 0.006
	0.10	$7.5 \times 10^{-6} \pm 5.4 \times 10^{-5}$	0.96 ± 0.011
	0.15	$-1.5 \times 10^{-5} \pm 7.1 \times 10^{-5}$	0.95 ± 0.015
	0.20	$4.9 \times 10^{-6} \pm 8.9 \times 10^{-5}$	0.95 ± 0.019
	0.25	$-7.6 \times 10^{-5} \pm 1.2 \times 10^{-4}$	0.92 ± 0.025
	0.30	$-6.4 \times 10^{-5} \pm 1.7 \times 10^{-4}$	0.91 ± 0.037
	0.35	$-2.2 \times 10^{-4} \pm 2.3 \times 10^{-4}$	0.93 ± 0.051
	0.40	$-3.9 \times 10^{-4} \pm 2.9 \times 10^{-4}$	0.93 ± 0.071

Comparison of ID-efficiencies

bdt loose for 1 prong taus



bdt loose for 3 prong taus



- no large trend in η between the different samples within the statistical uncertainty