JetSubStructure

overview on tools, cuts and methods

Denis Rathjens

Institut für Experimentalphysik, Universität Hamburg

13. January 2011

Table of Contents

Table of Contents

Table of Contents

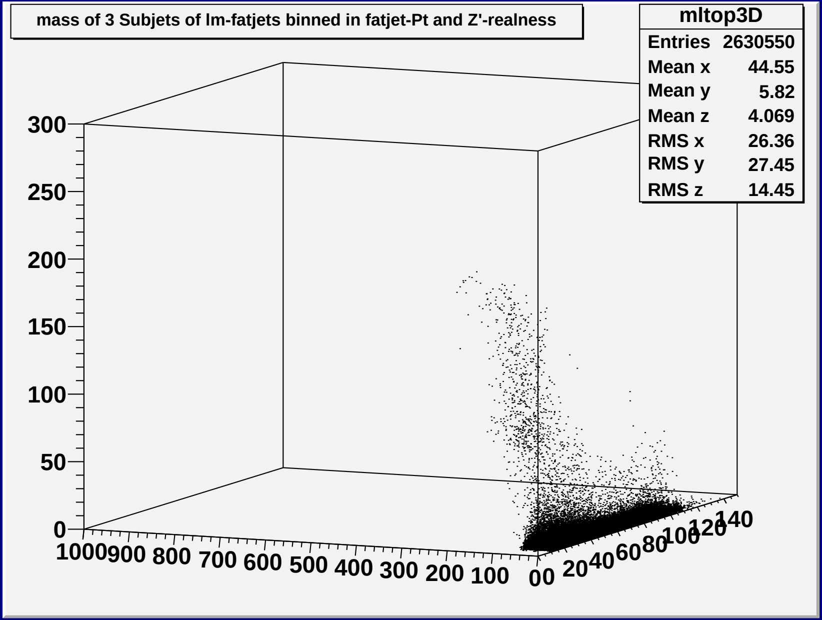
Table of Contents

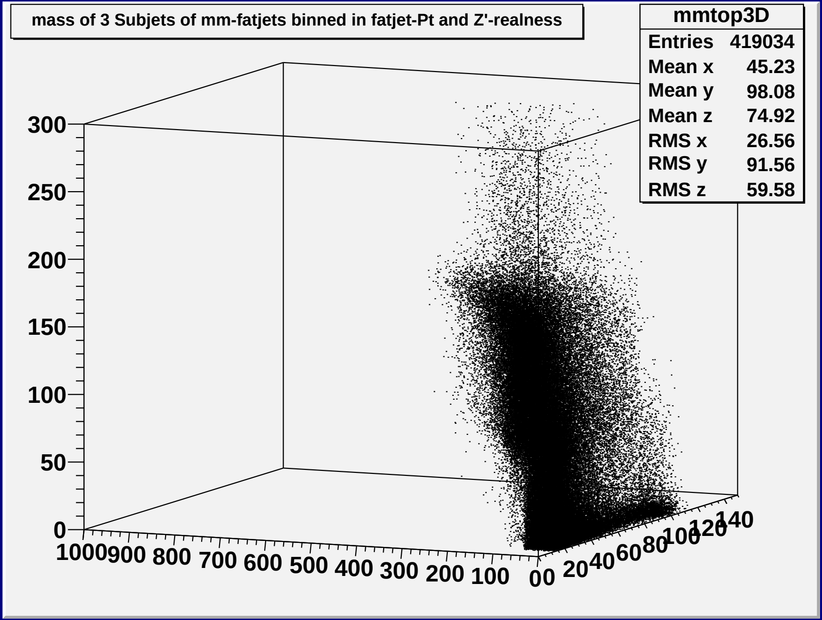
- resolve highly boosted jets
- main problem: jet merging
- solution: decluster fatjets
- additional: groom (top down) or trim (bottom up) fatjets

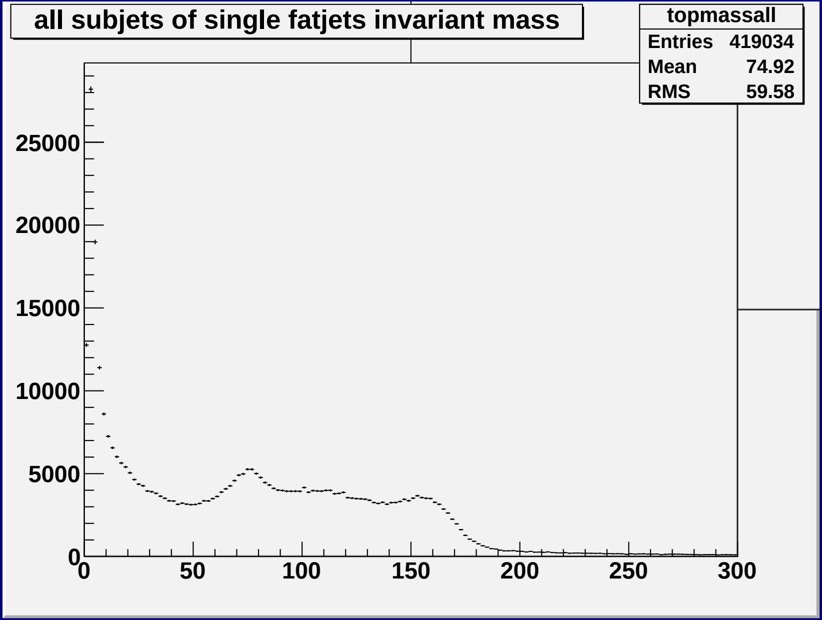
- basically three tools available
- Boosted Higgs (H$\rightarrow b\bar{b}$) - background discrimination efforts
- Boosted top (Z'$\rightarrow t\bar{t}$) - highly collimated topology ($p_T > 500$ GeV)
- Jet Pruning (e.g. boosted W) - general tool with integrated trimming
- but: general principle is identical
- cluster fatjet in CA $\rightarrow$ undo last cluster steps

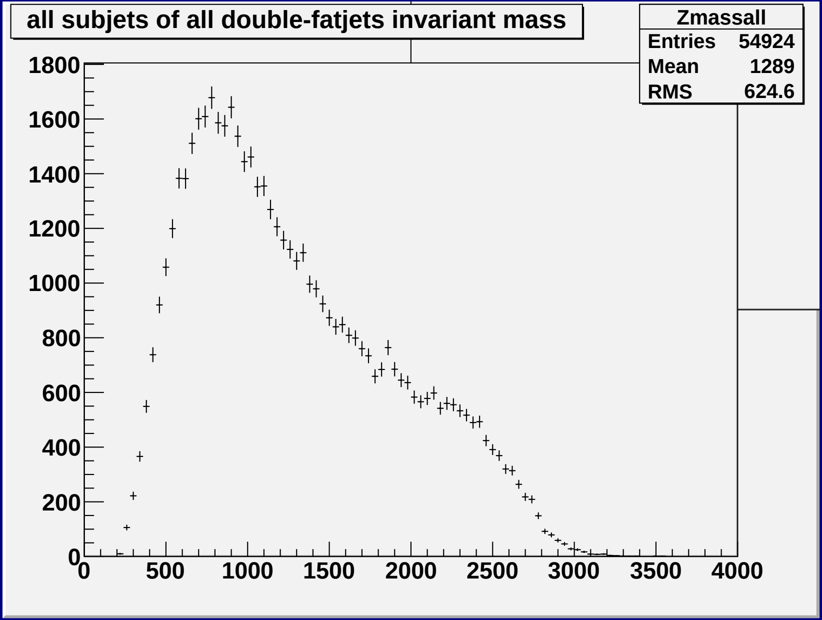
- used: Boosted Higgs-algorithm
- run on /Zprime_M3TeV_W300GeV-madgraph/Spring10-START3X_V26_S09-v1/GEN-SIM-RECO
- no p_T-cut, standard-parameters elsewhise (BoostedHiggsAlgorithm on twiki)

- everything soft in the detector IS clustered into a fatjet, without p_T-cuts
- mean multiplicity with cut $p_T = 100$ GeV: 2
- mean multiplicity without cut $p_T = 100$ GeV: 5.5
- variable opening angle depending on the analysis (dR=0.8-1.2)
- fatjet-masses are a rough estimate for their content
- for a rough but useful selection: use the 2 most massive fatjets in the event


mass of 3 Subjets of lm-fatjets binned in fatjet-Pt and Z'-realness
mltop3D
Entries 2630550
Mean x 44.55
Mean y 5.82
Mean z 4.069
RMS x 26.36
RMS y 27.45
RMS z 14.45
0
50
100
150
200
250
300
1000
900
800
700
600
500
400
300
200
100
0
0
20
40
60
80
100
120
140

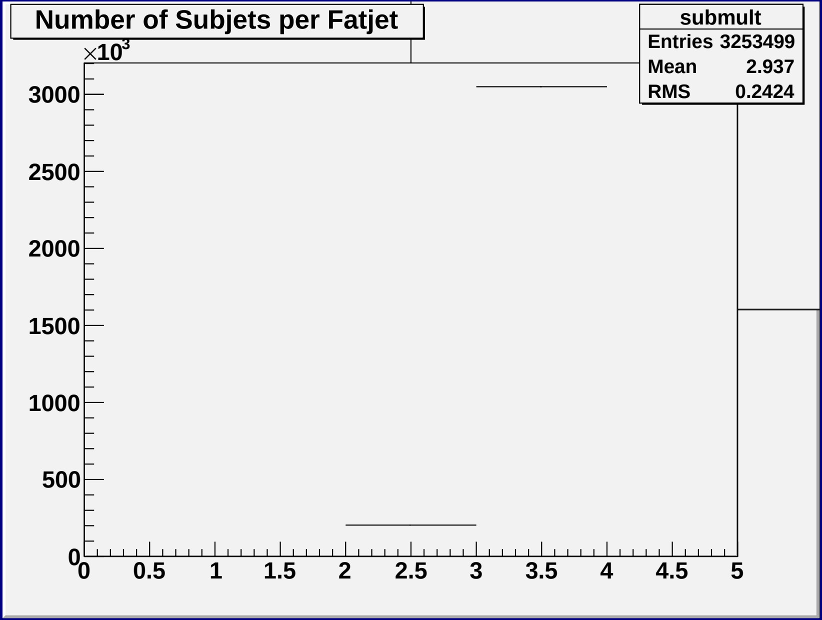

mass of 3 Subjets of mm-fatjets binned in fatjet-Pt and Z'-realness
mmtop3D
Entries 419034
Mean x 45.23
Mean y 98.08
Mean z 74.92
RMS x 26.56
RMS y 91.56
RMS z 59.58
300
250
200
150
100
50
0
1000 900 800 700 600 500 400 300 200 100 0
0 20 40 60 80 100 120 140


all subjets of single fatjets invariant mass
topmassall
Entries 419034
Mean 74.92
RMS 59.58
25000
20000
15000
10000
5000
0
0
50
100
150
200
250
300

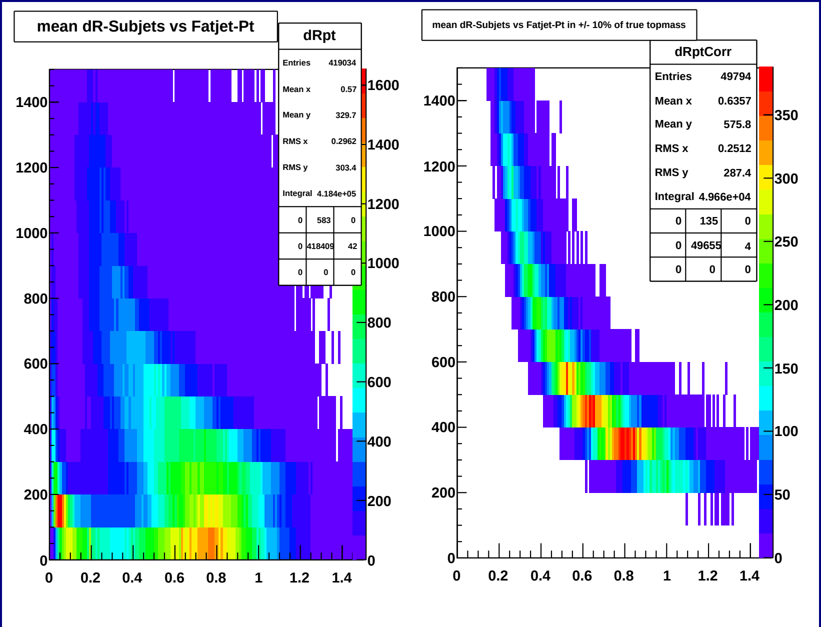


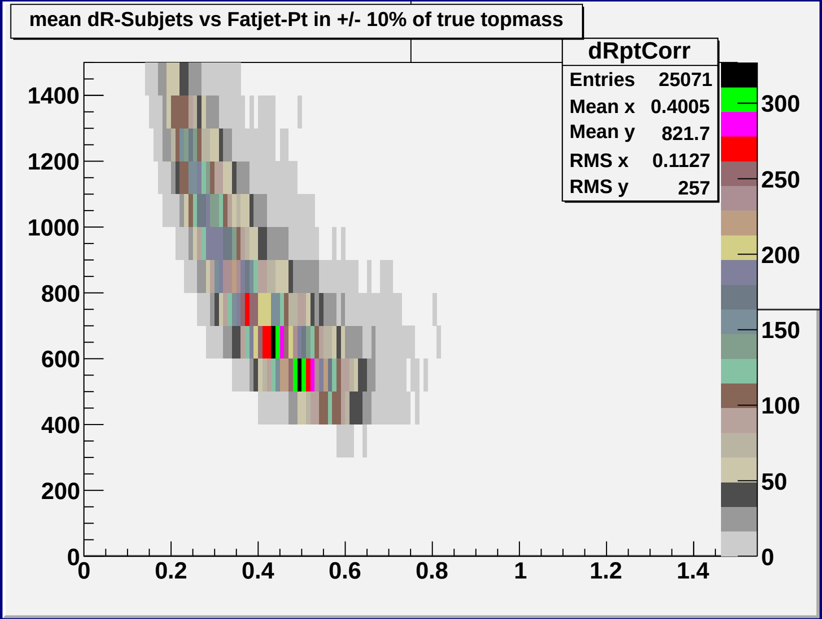
all subjets of all double-fatjets invariant mass
Zmassall
Entries 54924
Mean 1289
RMS 624.6
0
200
400
600
800
1000
1200
1400
1600
1800
0
500
1000
1500
2000
2500
3000
3500
4000

- for boosted Higgs algorithm: $n \leq$ variable
- for boosted top tagging: $n \leq 4$
- for JetPruning algorithm: n fixed

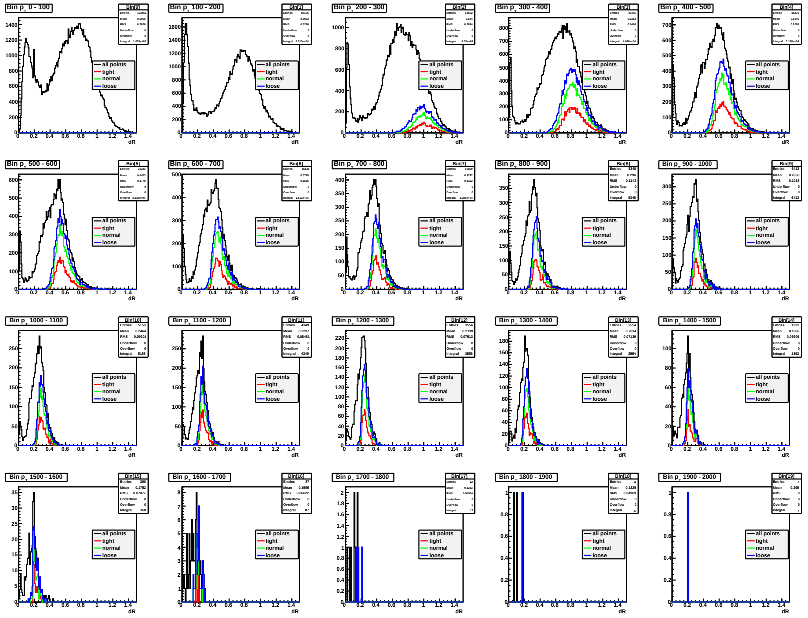

Number of Subjets per Fatjet
submult
Entries 3253499
Mean 2.937
RMS 0.2424
×10^3
3000
2500
2000
1500
1000
500
0
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

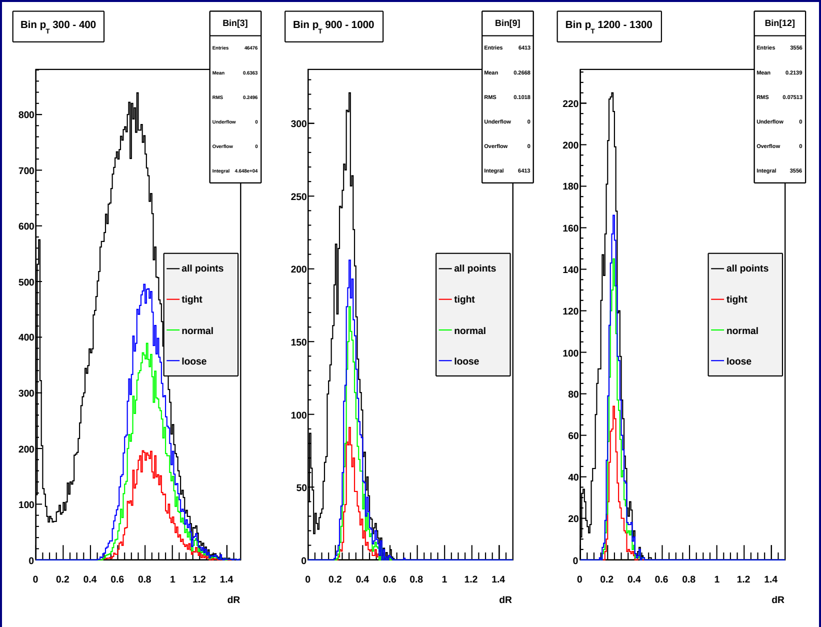
- three general limits for true events
- 1st: subjets must be boosted enough to not overlap
- 2nd: single subjets mustn't be too highly boosted, otherwise calo-granularity and tracker fail
- 3rd: system must be collimated enough to fit into the fatjet


mean dR-Subjets vs Fatjet-Pt
dRpt
Entries 419034
Mean x 0.57
Mean y 329.7
RMS x 0.2962
RMS y 303.4
Integral 4.184e+05
0 583 0
0 418409 42
0 0 0
mean dR-Subjets vs Fatjet-Pt in +/- 10% of true topmass
dRptCorr
Entries 49794
Mean x 0.6357
Mean y 575.8
RMS x 0.2512
RMS y 287.4
Integral 4.966e+04
0 135 0
0 49655 4
0 0 0

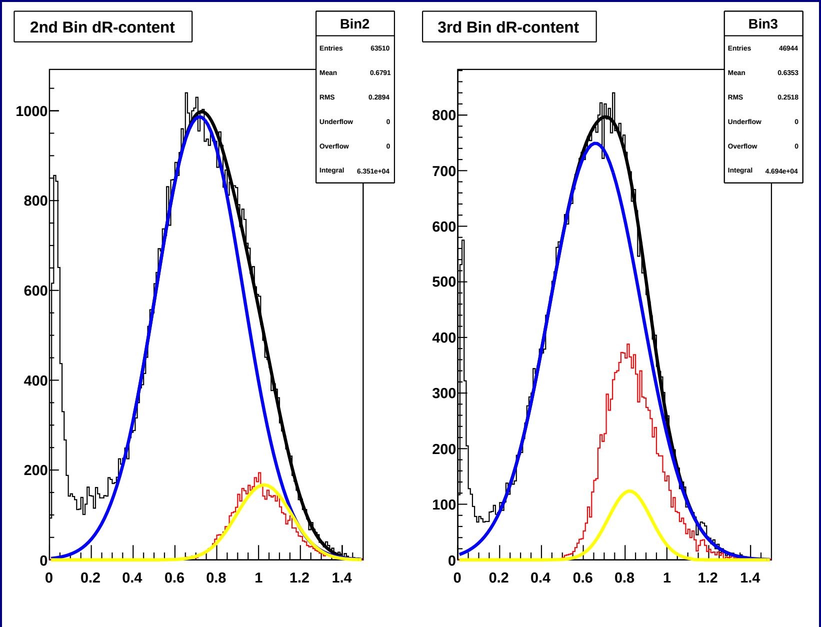


mean dR-Subjets vs Fatjet-Pt in +/- 10% of true topmass
dRptCorr
Entries 25071
Mean x 0.4005
Mean y 821.7
RMS x 0.1127
RMS y 257
0
200
400
600
800
1000
1200
1400
0
0.2
0.4
0.6
0.8
1
1.2
1.4
0
50
100
150
200
250
300

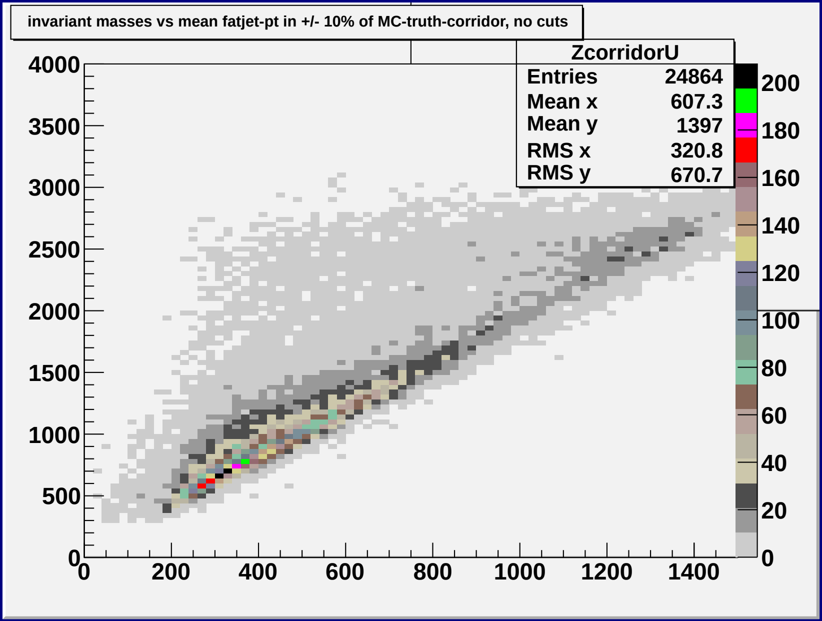

Bin pT 0 - 100
Bin pT 100 - 200
Bin pT 200 - 300
Bin pT 300 - 400
Bin pT 400 - 500
Bin pT 500 - 600
Bin pT 600 - 700
Bin pT 700 - 800
Bin pT 800 - 900
Bin pT 900 - 1000
Bin pT 1000 - 1100
Bin pT 1100 - 1200
Bin pT 1200 - 1300
Bin pT 1300 - 1400
Bin pT 1400 - 1500
Bin pT 1500 - 1600
Bin pT 1600 - 1700
Bin pT 1700 - 1800
Bin pT 1800 - 1900
Bin pT 1900 - 2000
all points
tight
normal
loose
dR



Bin p_T 300 - 400
Bin[3]
Entries 46476
Mean 0.6363
RMS 0.2496
Underflow 0
Overflow 0
Integral 4.648e+04
all points
tight
normal
loose
dR
Bin p_T 900 - 1000
Bin[9]
Entries 6413
Mean 0.2668
RMS 0.1018
Underflow 0
Overflow 0
Integral 6413
all points
tight
normal
loose
dR
Bin p_T 1200 - 1300
Bin[12]
Entries 3556
Mean 0.2139
RMS 0.07513
Underflow 0
Overflow 0
Integral 3556
all points
tight
normal
loose
dR

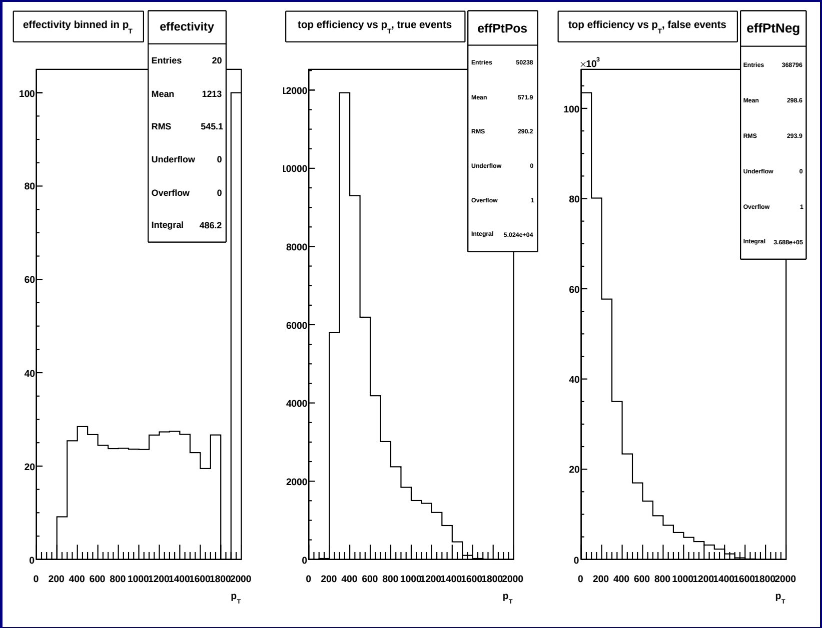

2nd Bin dR-content
Bin2
Entries 63510
Mean 0.6791
RMS 0.2894
Underflow 0
Overflow 0
Integral 6.351e+04
0
200
400
600
800
1000
0 0.2 0.4 0.6 0.8 1 1.2 1.4
3rd Bin dR-content
Bin3
Entries 46944
Mean 0.6353
RMS 0.2518
Underflow 0
Overflow 0
Integral 4.694e+04
0
100
200
300
400
500
600
700
800
0 0.2 0.4 0.6 0.8 1 1.2 1.4

- boosted top tagging uses $dR > 0.4 - 0.0004 * p_T$ for removing overlapping substructure
- efficiency drop according to boosted top team at p_T of 1 GeV
- take optimized fatjet-size for your analysis
- if fatjet is small enough, JEC from ak5 may be used due to small fluctuations at very high p_T
- general principle 'lean jets have a better resolution' should still hold



invariant masses vs mean fatjet-pt in +/- 10% of MC-truth-corridor, no cuts
ZcorridorU
Entries 24864
Mean x 607.3
Mean y 1397
RMS x 320.8
RMS y 670.7
0 500 1000 1500 2000 2500 3000 3500 4000
0 200 400 600 800 1000 1200 1400
0 20 40 60 80 100 120 140 160 180 200

- currently used for searches: boosted Higgs, boosted Top, boosted W
- primary assets: very good background-reduction (cuts on multiplicity, topology, p_T, mass window)
- fake-rates (measured on data by boosted top group) 10 to 15% for W-bosons, 4% for tops
- efficiencys after optimisation typically 30 to 40%
- generally applicable on any boosted n-body-system



effectivity binned in p_T
effectivity
Entries 20
Mean 1213
RMS 545.1
Underflow 0
Overflow 0
Integral 486.2
p_T
top efficiency vs p_T, true events
effPtPos
Entries 50238
Mean 571.9
RMS 290.2
Underflow 0
Overflow 1
Integral 5.024e+04
p_T
top efficiency vs p_T, false events
effPtNeg
×10^3
Entries 368796
Mean 298.6
RMS 293.9
Underflow 0
Overflow 1
Integral 3.688e+05
p_T

- https://twiki.cern.ch/twiki/bin/viewauth/CMS/BoostedHiggsAlgorit
- BOOST2010 workshop http://arxiv.org/pdf/1012.5412
- PAS JME-10-013 currently pre-approved
- https://twiki.cern.ch/twiki/bin/viewauth/CMS/BoostedTop
- https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetPruningAlgorithm
  (already available in CMSSW)