Tau fake rate studies

Sylvie Brunet

Working on that:

David, Philip, Sebastian & myself

DESY-Atlas weekly meeting, 23/11/2007

Motivations

- Technical motivation: do a *quick* study (missing so far) to be added in the Tau CSC note

- Physics motivation: present a method to evaluate the **tau fake rate** (by jets) **in data**.

Goal:

We want to evaluate
the fraction of jets
selected by the 2 existing
tau algorithms

Method

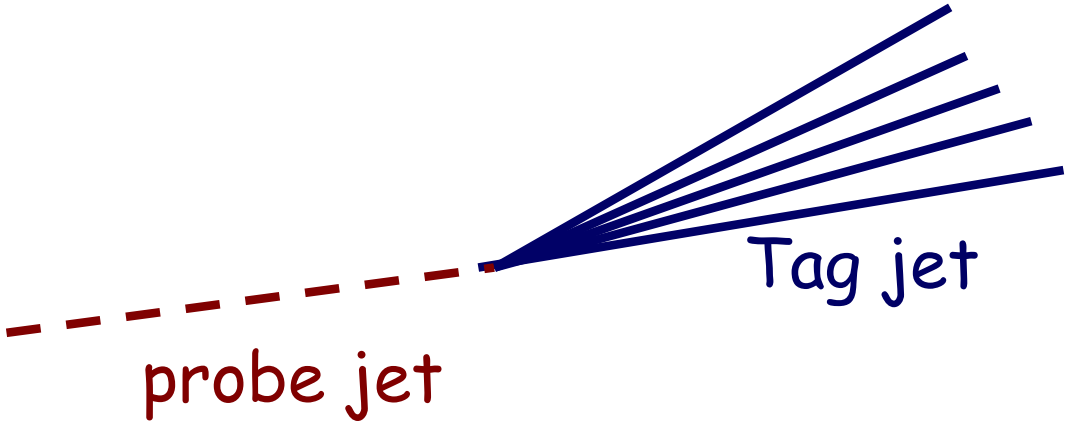
Tag and Probe style

Dijet QCD events will be dominant in data…

The idea is to select a sample of very likely QCD jets (eventually in data) and see how many are (wrongly) identified as taus.

We make predictions using MC:

- 2 back to back objects
- One "nice" jet for the **tag side** (then we are confident that we also have a jet on the other side)
- We use the other side (**probe jet**) to compute the fake rate



$$\text{Fake Rate} = \frac{\text{Nb Probe jets identified as } \tau}{\text{Nb Probe jets}}$$

MC Data Used

- QCD dijets samples (J0-J5) (perform our studies)
- Z to tau tau and W to Tau nu (crosscheck that we will get rid of most of the "real" taus)

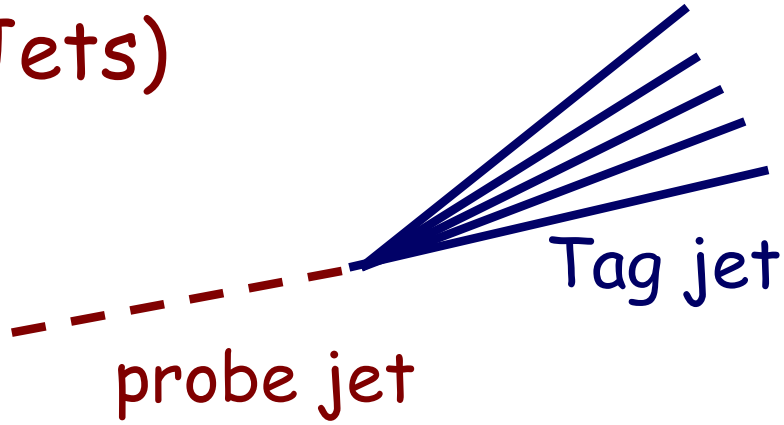
- 2 independent setups:
  - **Sebastian:**
    - 12-series ntuples coming from the SUSYView production
  - **Sylvie:**
    - 13-series ntuples (CBNT) coming from private production (Freiburg)

- Main differences (found so far):
  - Overlap removal (e/gamma) for jets in SUSYView prod.
  - TopoJets vs TowerJets
  - Missing tau1p3p variables in SUSYView prod.
  - Some other minor differences

  - **Normal that both setups give different results**
  - **Very nice to cross-check each other**
  - **In CSC note: 13-series ntuples**

Selections

- 2 back to back jets (Cone4Jets)
  - abs(eta) for each jet <= 2.5
  - abs(DeltaPhi) = pi+/- 0.30
  - min pt (?)



- 1 "nice" tag jet:
  - nTrk in Jet >=4
  - isolated (?), overlap removal (?)

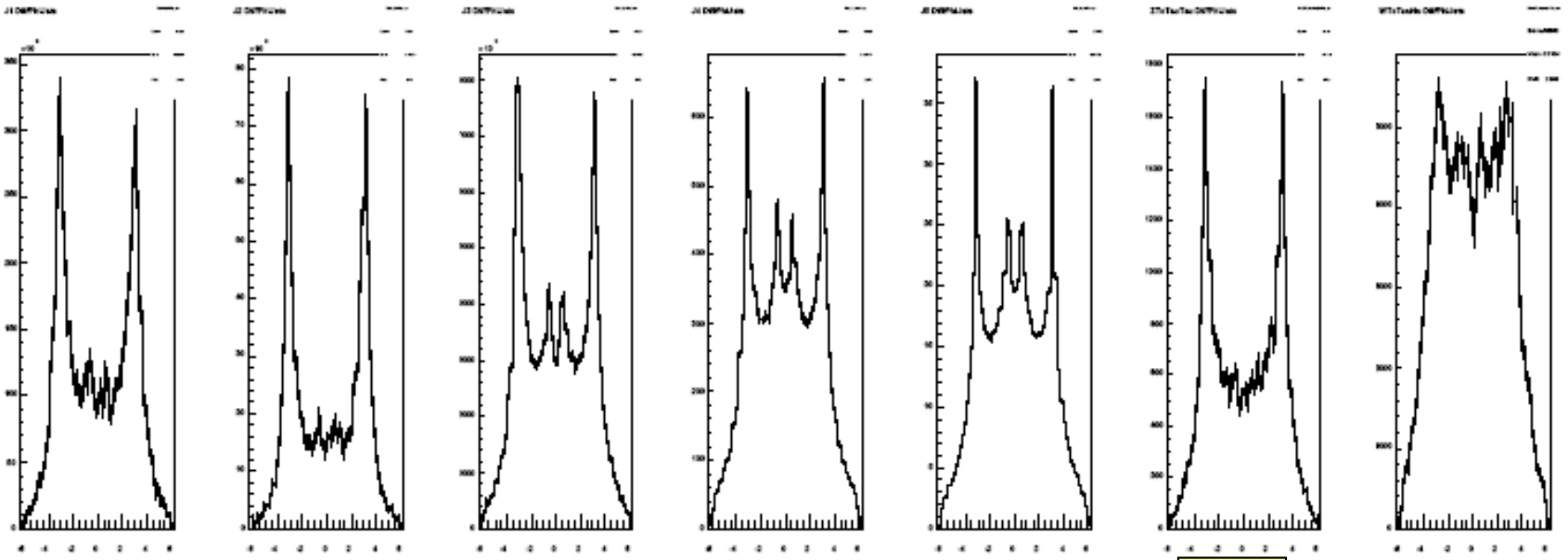
- Probe jet:
  - no further selections (to keep whole spectrum)
  - check if identified as a tau

What we have learned so far

Disclaimer:
WORK IN PROGRESS!!!

Some interesting plots…

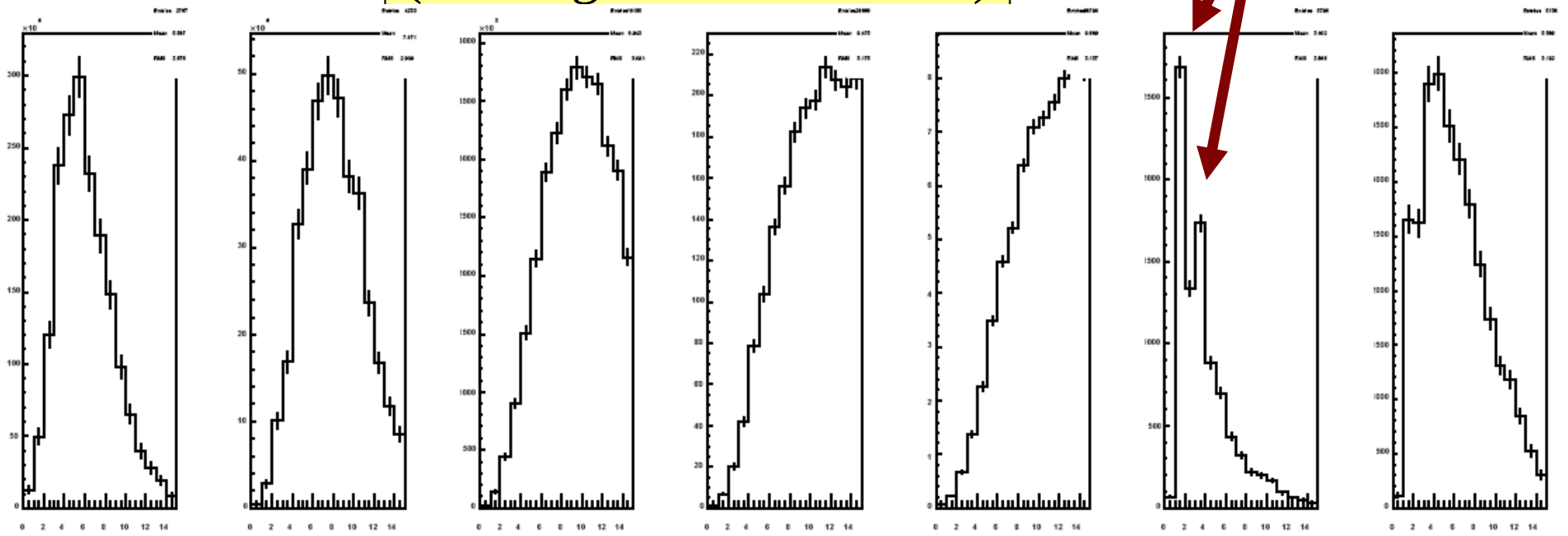
deltaPhi between jets in the event
(xaxis goes from -2pi to +2pi)



J1 J2 J3 J4 J5 ZtoTautau WtoTauNu

Some interesting plots…

tracks per jets
(x axis goes from 0 to 14)



J1 J2 J3 J4 J5 ZtoTautau WtoTauNu

What we have learned so far

- Bad news: We select some real taus in the Z to tau tau and W to tau nu samples *(can increase the nTrk requirement on the tag side if we want to further kill these events).*

- Good News: However, the cross-section is so small compared to the dijets that it will be negligible

(Nevertheless, we want to put a number on this "negligible")

Some expected cross-sections:

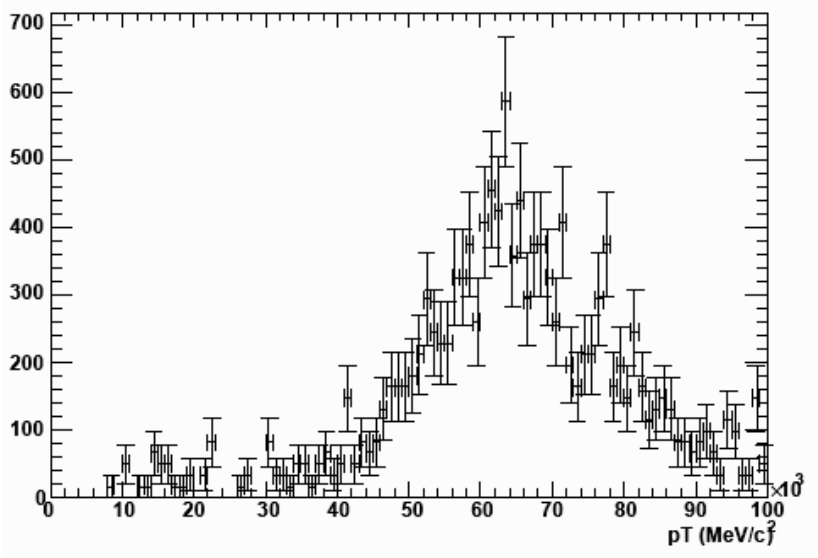
Dijet(J1, 17-35GeV)	1.4 mb
Dijet(J2, 35-70GeV)	93.3 um
Dijet(J3, 70-140GeV)	5.9 um
Z to tau tau	1.6 nb
W to tau nu	17.3 nb

One of the remaining puzzles: pt distribution

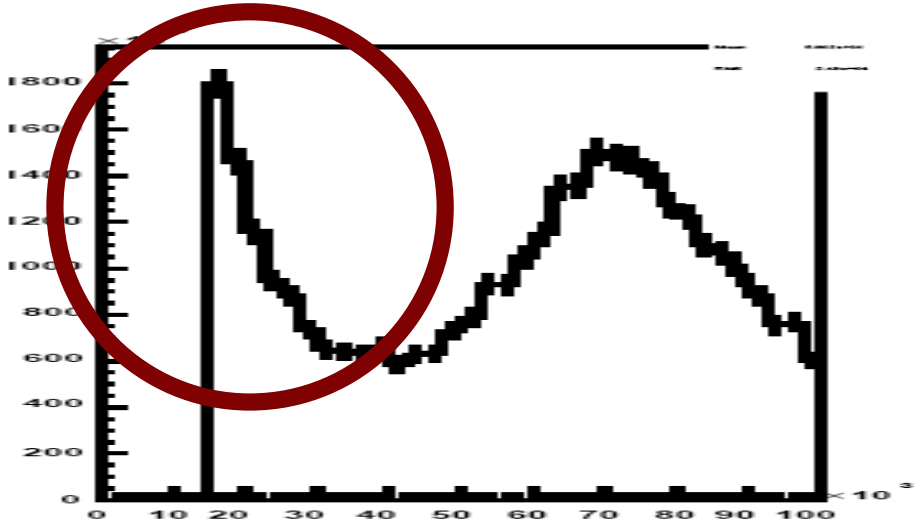
- In Sylvie's ntuples, seems to have some garbage jets remaining… Need more clean-up. Not seen in Sebastian's ntuples…
  - Related to overlap removal/isolation?
  - track quality cuts for the tag side?
  - other cuts made in SUSYView?

Currently Investigating that…

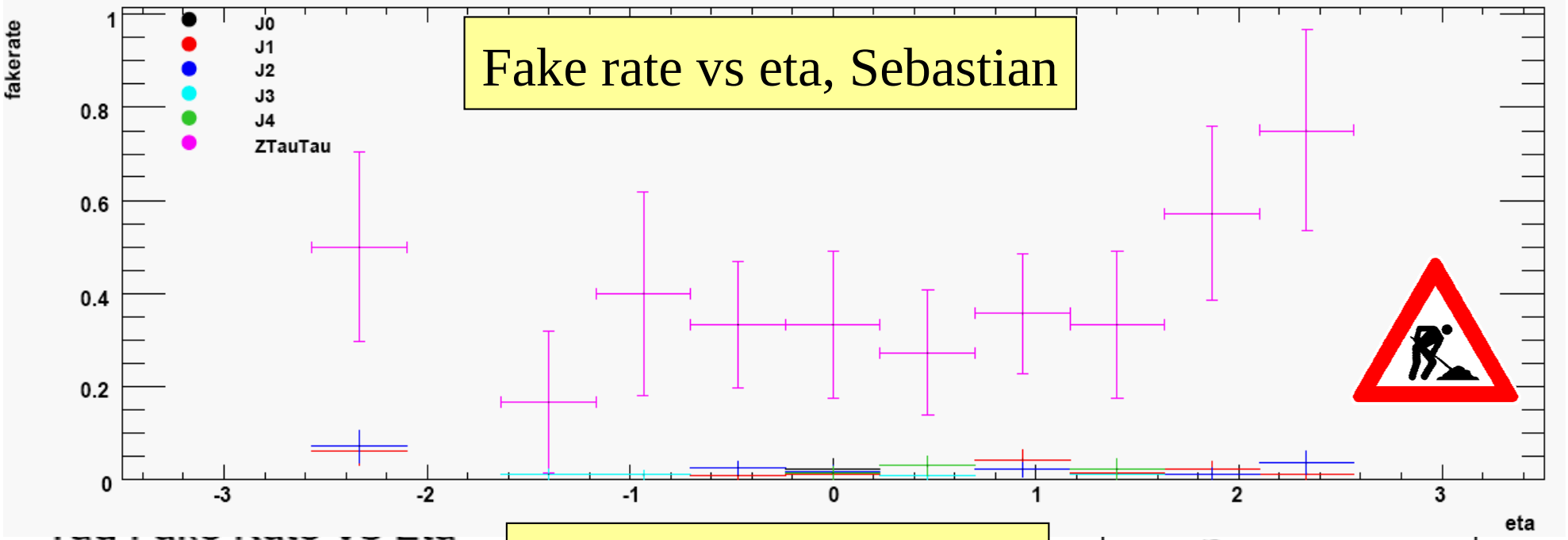
J3, pt, den., Sebastian



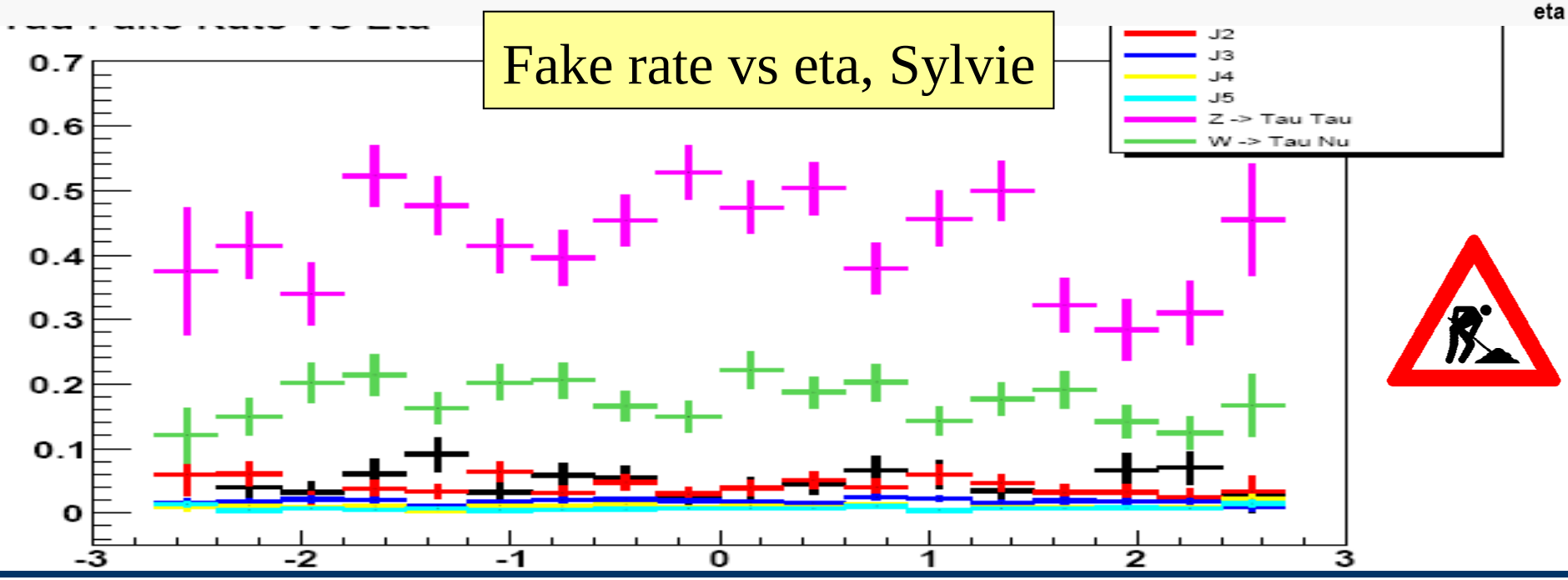
J3, pt, den., Sylvie



What we have learned so far



Fake rate vs eta, Sebastian



Fake rate vs eta, Sylvie

To-do list

- Do it for Tau1p3p as well
- Do it for different integrated luminosities
- Continue crosschecks (mainly using MCTruth)
- Compare our results with the TauWG rejection factors on MC

Summary/Conclusion

- We have a method to evaluate the tau fake rate (from jets) in data. Targeted for the tau CSC note.
- We use a "tag and probe" approach
- We have 2 independent setups, very useful to investigate problems and crosscheck results
- Still things to sort out (main thing is the clean-up of the jets), but progressing rapidly

Backup

Data Type = J1

ScalingFactor for J1 = 663462
Wanted Lumi = 10 pb^-1
We will process 20800 events.

Evts
Before: 1.38e+10
Has Jet Pair: 1.27882e+10 (eff = 92.6683 %, cut flow= 92.6683 %)
Pair in Eta Range: 9.36741e+09 (eff = 67.8798 %, cut flow= 73.2503 %)
Pt for both jets: 3.3538e+09 (eff = 24.3029 %, cut flow= 35.8028 %)
Pair b to b: 1.59032e+09 (eff = 11.524 %, cut flow= 47.4184 %)
Tag nTrk OK: 1.17565e+09 (eff = 8.51923 %, cut flow= 73.9257 %)
Has Identified as tau: 6.17019e+07 (eff = 0.447115 %, cut flow= 5.24831 %)

Data Type = J2

ScalingFactor for J2 = 93768.8
Wanted Lumi = 10 pb^-1
We will process 9950 events.

Evts
Before: 9.33e+08
Has Jet Pair: 9.31969e+08 (eff = 99.8894 %, cut flow= 99.8894 %)
Pair in Eta Range: 7.89252e+08 (eff = 84.593 %, cut flow= 84.6866 %)
Pt for both jets: 6.22625e+08 (eff = 66.7337 %, cut flow= 78.888 %)
Pair b to b: 3.56884e+08 (eff = 38.2513 %, cut flow= 57.3193 %)
Tag nTrk OK: 3.24815e+08 (eff = 34.8141 %, cut flow= 91.0142 %)
Has Identified as tau: 1.45342e+07 (eff = 1.55779 %, cut flow= 4.4746 %)

Data Type = J3

ScalingFactor for J3 = 2325.67
Wanted Lumi = 10 pb^-1
We will process 25283 events.

Evts
Before: 5.88e+07
Has Jet Pair: 5.88e+07 (eff = 100 %, cut flow= 100 %)
Pair in Eta Range: 5.46022e+07 (eff = 92.8608 %, cut flow= 92.8608 %)
Pt for both jets: 4.96996e+07 (eff = 84.5232 %, cut flow= 91.0214 %)
Pair b to b: 3.27873e+07 (eff = 55.7608 %, cut flow= 65.971 %)
Tag nTrk OK: 3.15222e+07 (eff = 53.6091 %, cut flow= 96.1413 %)
Has Identified as tau: 700028 (eff = 1.19052 %, cut flow= 2.22075 %)

Data Type = J4

ScalingFactor for J4 = 0.118307
Wanted Lumi = 10 pb^-1
We will process 26034 events.

Evts
Before: 3080
Has Jet Pair: 3080 (eff = 100 %, cut flow= 100 %)
Pair in Eta Range: 3005.47 (eff = 97.5801 %, cut flow= 97.5801 %)
Pt for both jets: 2904.31 (eff = 94.2959 %, cut flow= 96.6344 %)
Pair b to b: 2224.88 (eff = 72.2363 %, cut flow= 76.606 %)
Tag nTrk OK: 2176.85 (eff = 70.6768 %, cut flow= 97.8411 %)
Has Identified as tau: 27.5655 (eff = 0.894983 %, cut flow= 1.2663 %)

Data Type = J5

ScalingFactor for J5 = 0.00255003
Wanted Lumi = 10 pb^-1
We will process 49019 events.

Evts
Before: 125
Has Jet Pair: 125 (eff = 100 %, cut flow= 100 %)
Pair in Eta Range: 124.125 (eff = 99.3003 %, cut flow= 99.3003 %)
Pt for both jets: 123.016 (eff = 98.4129 %, cut flow= 99.1063 %)
Pair b to b: 106.734 (eff = 85.3873 %, cut flow= 86.7644 %)
Tag nTrk OK: 105.694 (eff = 84.555 %, cut flow= 99.0252 %)
Has Identified as tau: 1.01746 (eff = 0.81397 %, cut flow= 0.962652 %)

Data Type = ZToTauTau

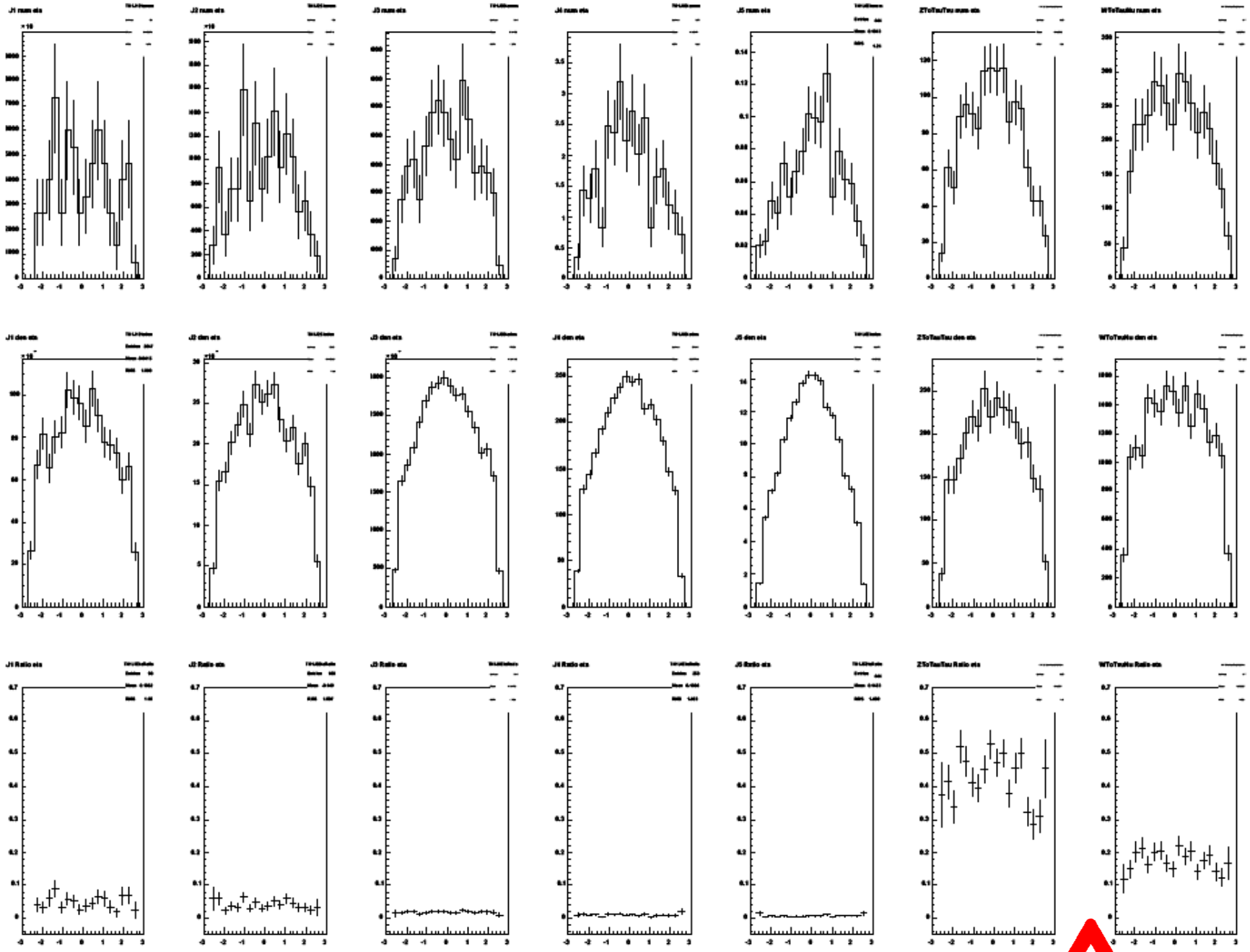
ScalingFactor for ZToTauTau = 1.56938
Wanted Lumi = 10 pb^-1
We will process 10450 events.

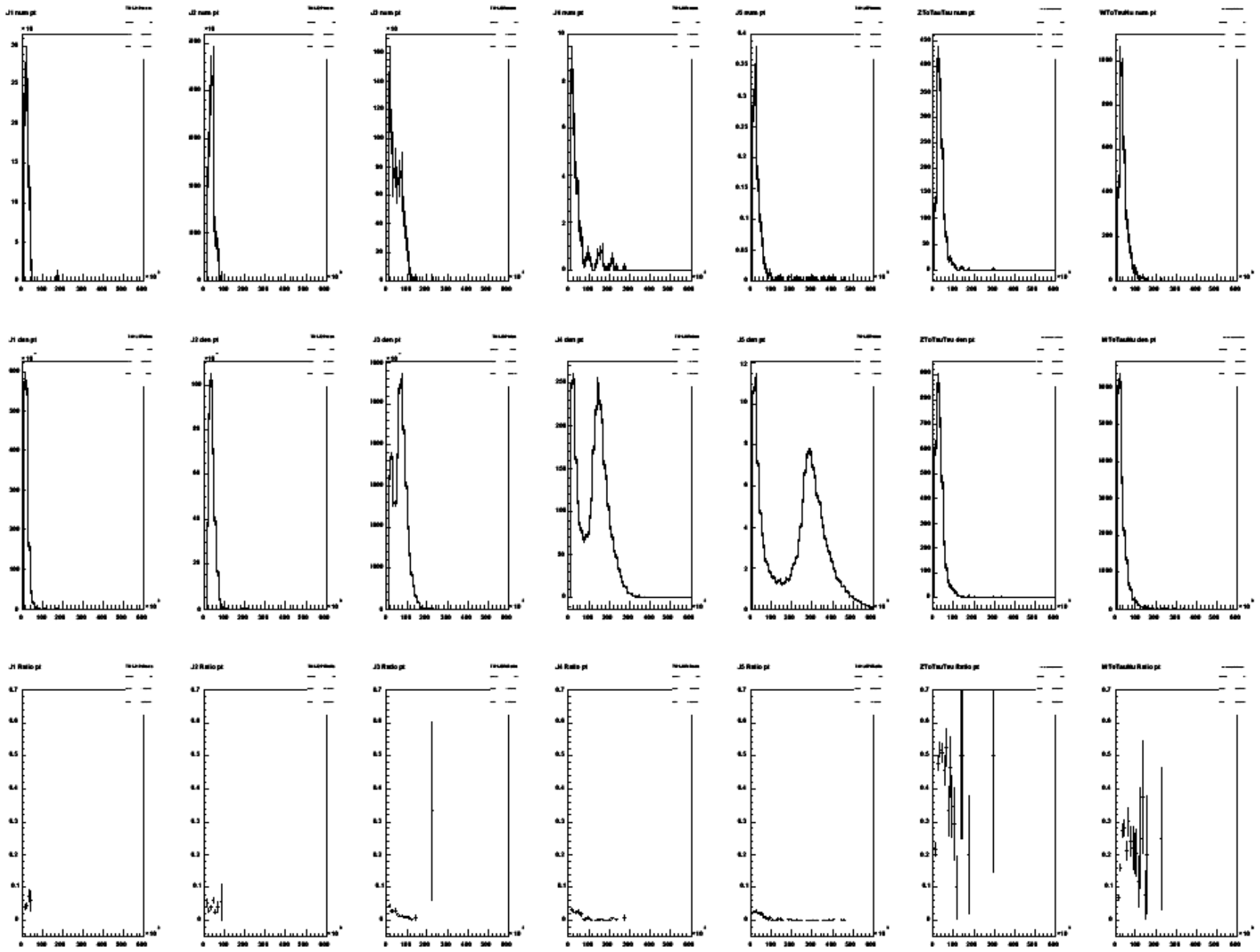
Evts
Before: 16400
Has Jet Pair: 15574.5 (eff = 94.9665 %, cut flow= 94.9665 %)
Pair in Eta Range: 14802.4 (eff = 90.2584 %, cut flow= 95.0423 %)
Pt for both jets: 12828.1 (eff = 78.2201 %, cut flow= 86.6624 %)
Pair b to b: 7611.48 (eff = 46.4115 %, cut flow= 59.3345 %)
Tag nTrk OK: 2636.56 (eff = 16.0766 %, cut flow= 34.6392 %)
Has Identified as tau: 1316.71 (eff = 8.02871 %, cut flow= 49.9405 %)

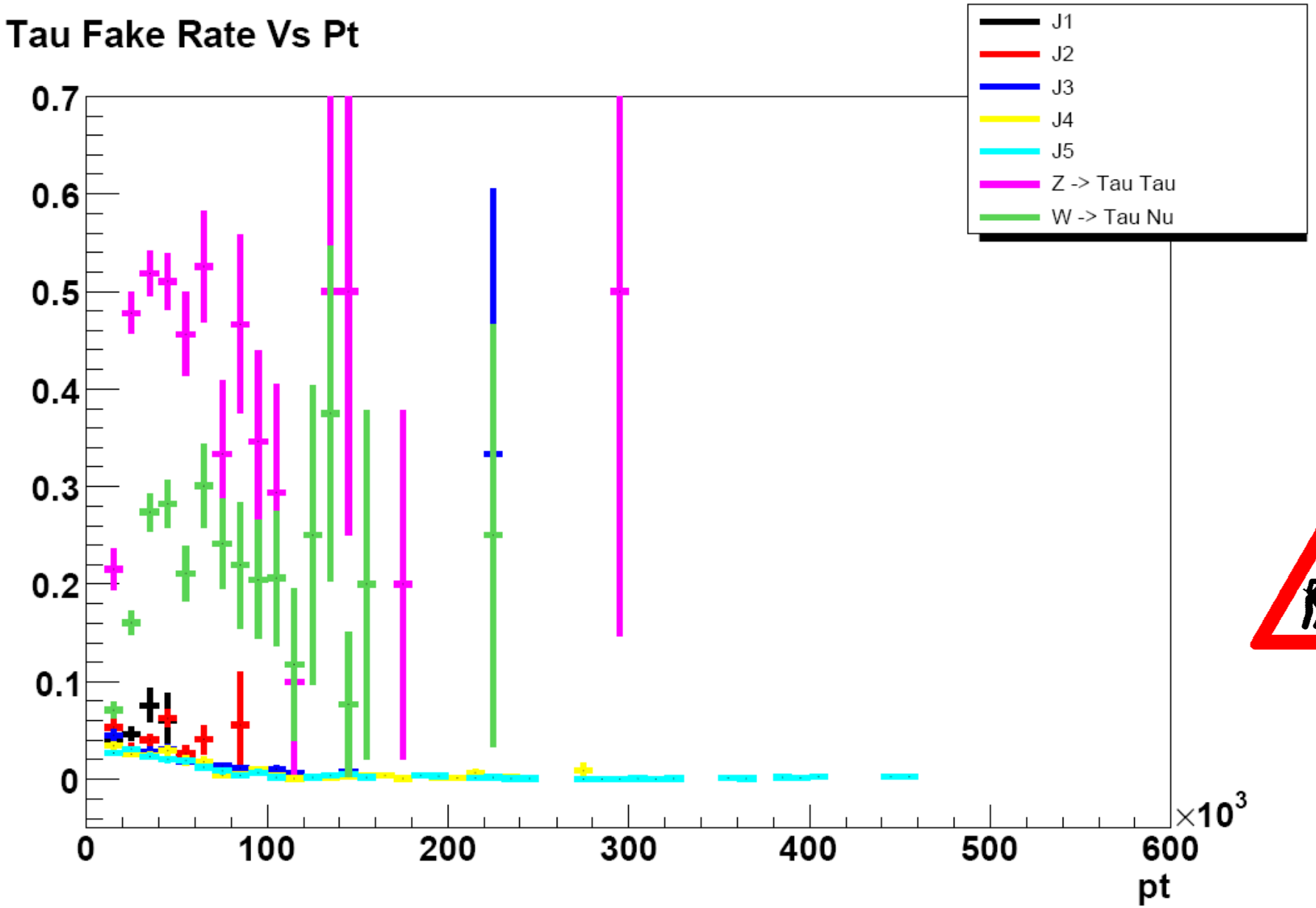
Data Type = WToTauNu

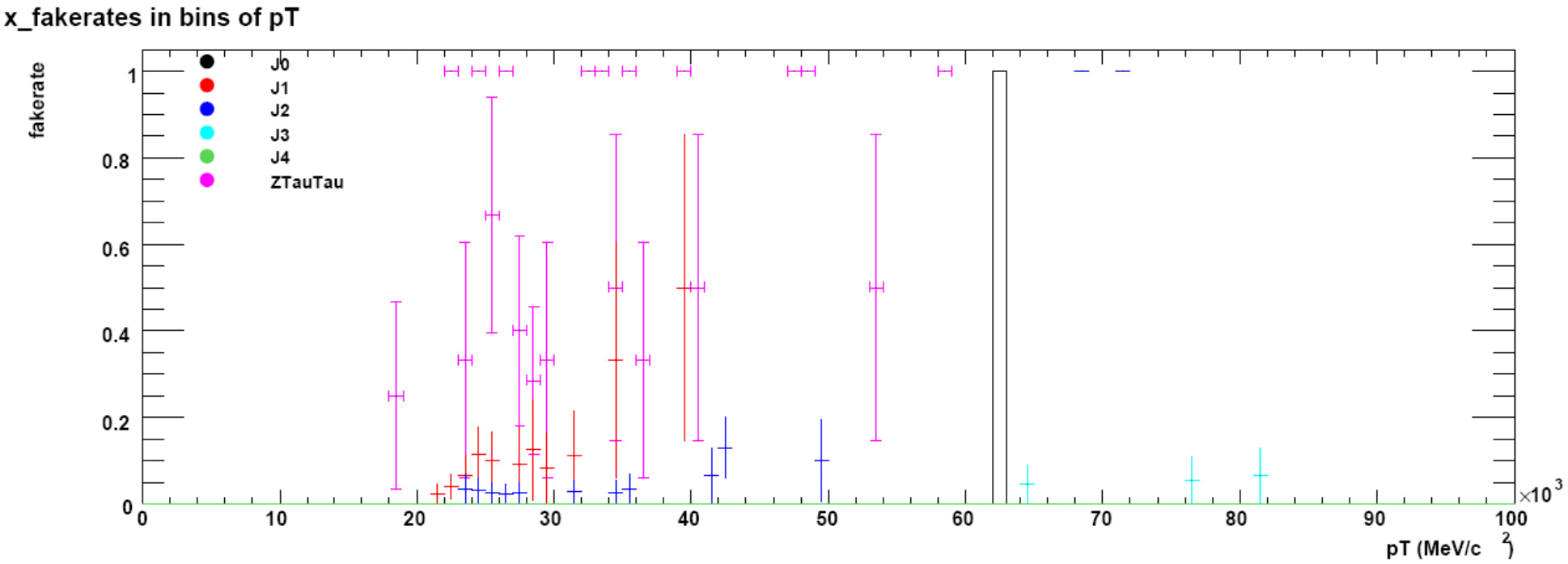
ScalingFactor for WToTauNu = 6.21185
Wanted Lumi = 10 pb^-1
We will process 27850 events.

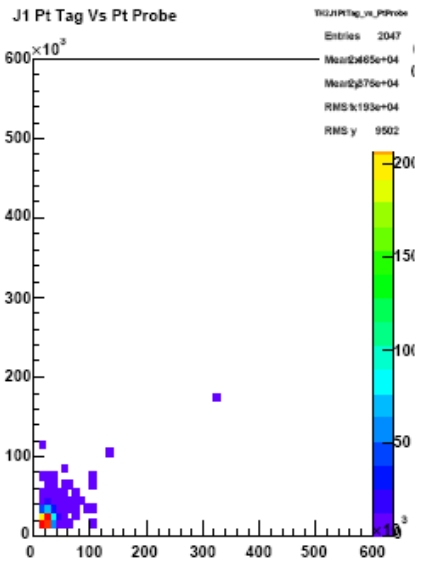
Evts
Before: 173000
Has Jet Pair: 151911 (eff = 87.8097 %, cut flow= 87.8097 %)
Pair in Eta Range: 135294 (eff = 78.2047 %, cut flow= 89.0615 %)
Pt for both jets: 79474.4 (eff = 45.939 %, cut flow= 58.742 %)
Pair b to b: 25002.7 (eff = 14.4524 %, cut flow= 31.4601 %)
Tag nTrk OK: 16890 (eff = 9.76302 %, cut flow= 67.5528 %)
Has Identified as tau: 3702.26 (eff = 2.14004 %, cut flow= 21.9198 %)

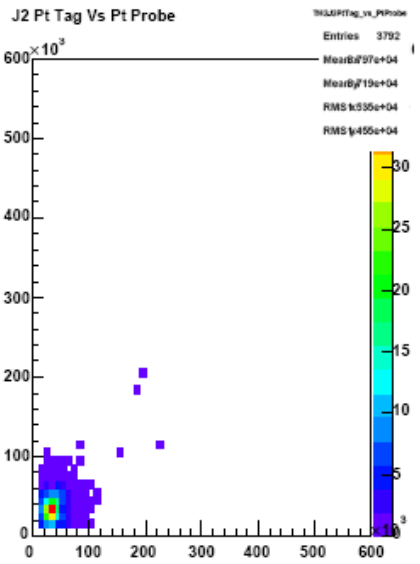


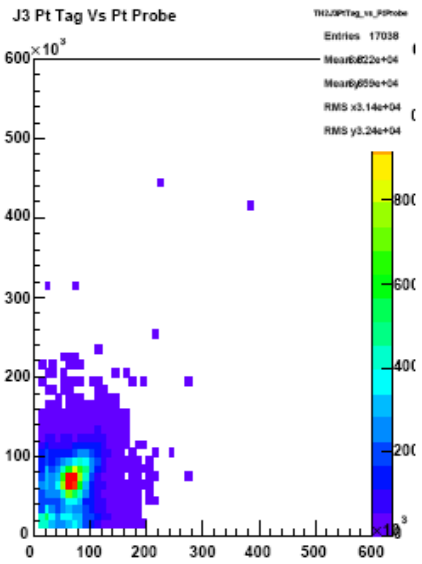


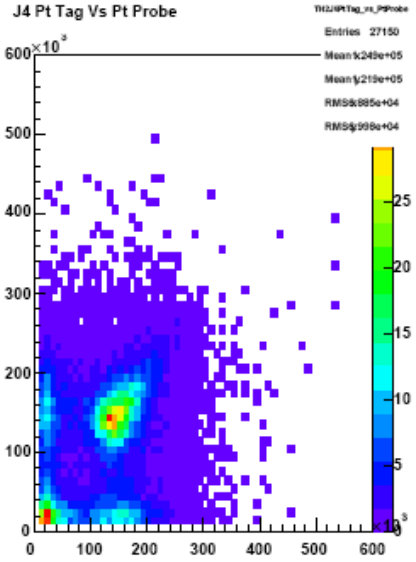

Tau Fake Rate Vs Pt
J1
J2
J3
J4
J5
Z -> Tau Tau
W -> Tau Nu
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
100
200
300
400
500
600
×10^3
pt

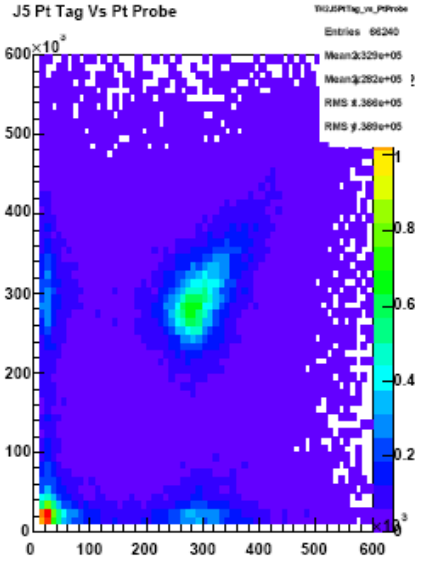

x_fakerates in bins of pT
fakerate
J0
J1
J2
J3
J4
ZTauTau
0
0.2
0.4
0.6
0.8
1
0
10
20
30
40
50
60
70
80
90
100
×10 3
pT (MeV/c 2)

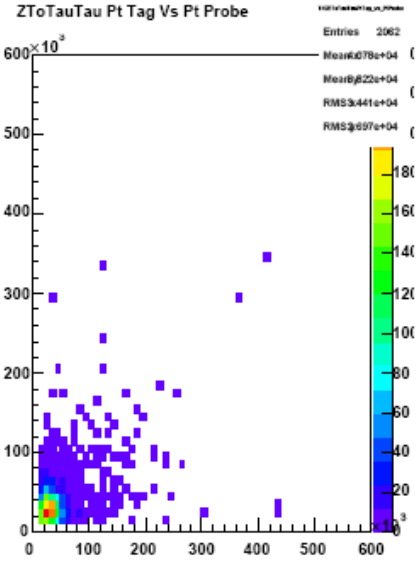

J1 Pt Tag Vs Pt Probe

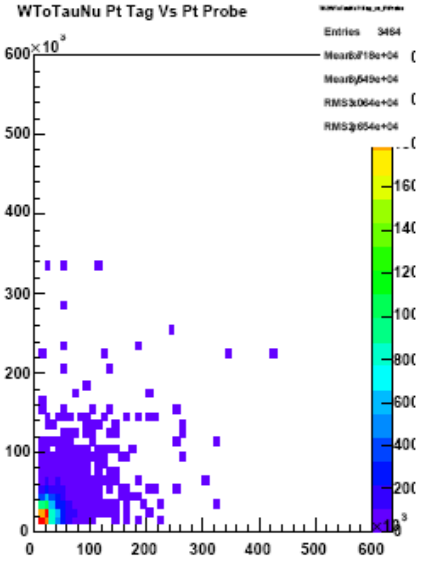

J2 Pt Tag Vs Pt Probe


J3 Pt Tag Vs Pt Probe


J4 Pt Tag Vs Pt Probe


J5 Pt Tag Vs Pt Probe


ZToTauTau Pt Tag Vs Pt Probe


WToTauNu Pt Tag Vs Pt Probe