MC-data Comparison with Fake-Lepton Estimation in the ttHbb Semi-Leptonic Final State

Shunsuke Honda (U Tsukuba, JAPAN)

Koji Nakamura (KEK), Yuji Enari (U Tokyo), Yasuyuki Horii (Nagoya U)

Report

- fake estimation study is going on in the tthbb analysis.
- For looking into early run2 datasets, we use the simple template estimation of top-fakes(QCD-BG).
- In this talk, we report about
  - fake template estimation using MTW simply
  - with Wjets MC modeling study
  - run2 MC-data comparisons with fakes
- We are going to estimate fakes by the Matrix-Method.

Technical Information

AnalysisTop-2-3-20 + TTHbbLeptonic-02-03-15 package

For looser tth signal-region, our own fake estimation can be needed.
no MET, MTW cut

- GRL / GOODCALO
- LEP_N ==1
- JET CLEAN LOOSEBAD

Jet Object Selection

- pT>25GeV
- |η|<2.5
- JVT>0.64 if [pT<50GeV && |η|<2.4]
- AntiKt4EMTopo
- btagging MV2C20 77%OP.

Lepton Object Selection

- pT>25GeV
- |η|<2.4 && veto LAr-crack region
- ID_{el} = TightLH with isol.
- $ID_{el,loose}$ = LooseLH w/o isol.
- ID_{mu} = Tight with isol.
- $ID_{mu,loose}$ = Loose w/o isol.

- fake-lepton estimation, Wjets modeling study
- data-MC comparison with 78 /pb run2-data.

Dataset

- mc15_13TeV & DAOD_TOPQ1 derivation -

- **ttbar** *410000.PowhegPythiaEvtGen*

  nonallhad

- **sg-top** *410011-16,410025-26.PowhegPythiaEvtGen*

  Wt-ch + t-ch + s-ch

- **Wjets** *361100-05.PowhegPythia8EvtGen*

  enu + munu + taunu

  > also prepare 361300-29,361333-71.Sherpa for MC-modeling

- **Zjets** *361106-08.PowhegPythia8EvtGen*

  ee + mumu + tautau

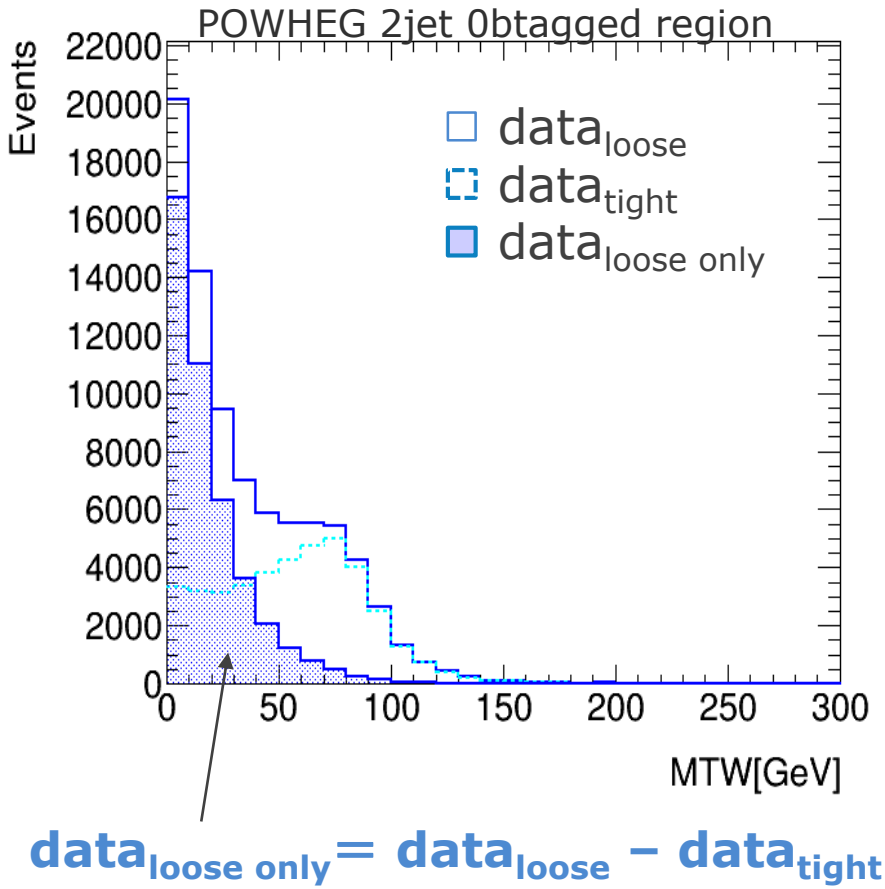
- data15_13TeV & DAOD_TOPQ1 derivation -

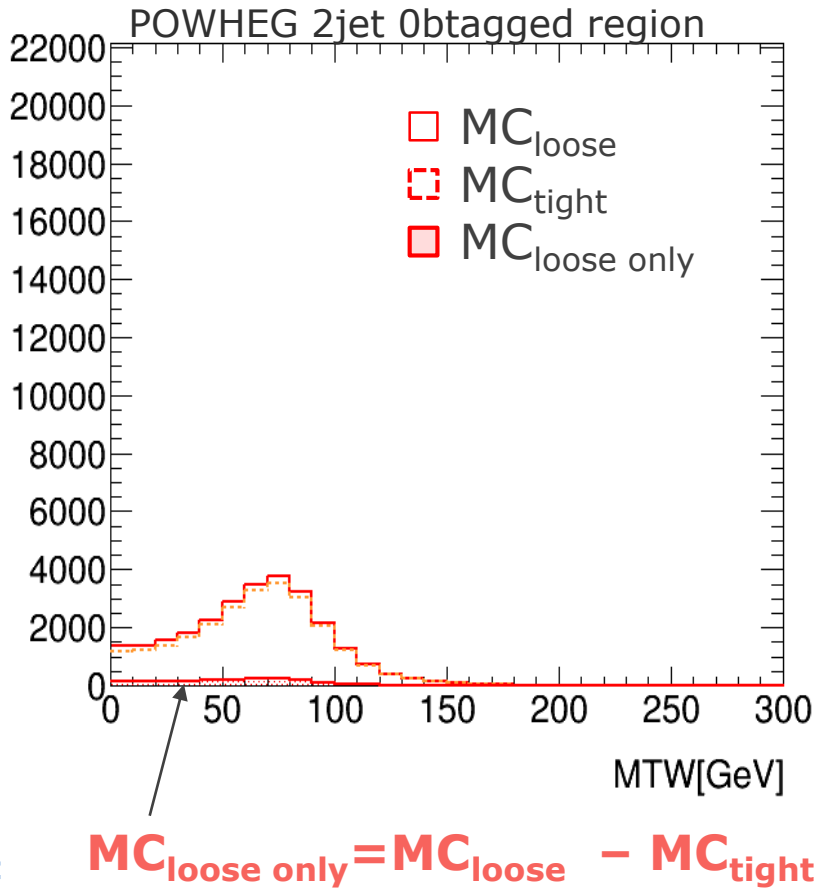
- **periodA-C (~78 pb^{-1})**

Fake-Lepton Template

Using MTW distribution for each jetbin and bjetbin

$$\text{fake-lepton}_{\text{template}} = \text{data}_{\text{loose only}} - \text{MC}_{\text{loose only}}$$





$\text{data}_{\text{loose only}} = \text{data}_{\text{loose}} - \text{data}_{\text{tight}}$

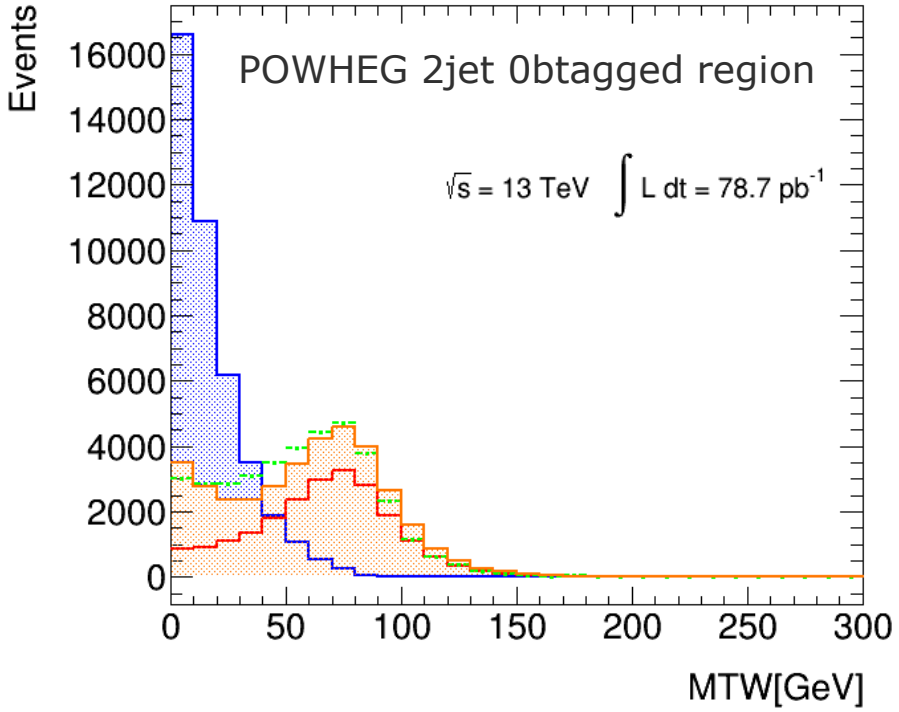
$\text{MC}_{\text{loose only}} = \text{MC}_{\text{loose}} - \text{MC}_{\text{tight}}$

Fake-Lepton and Wjets SF

Chi-squire bin-by-bin fitting

SF_{Wjets} x $Wjets_{tight}$ + SF_{fake} x $fake_{template}$ to

$data\text{-}MC(but\ not\ Wjets)_{tight}$: data(fake+Wjets) using TMinuit

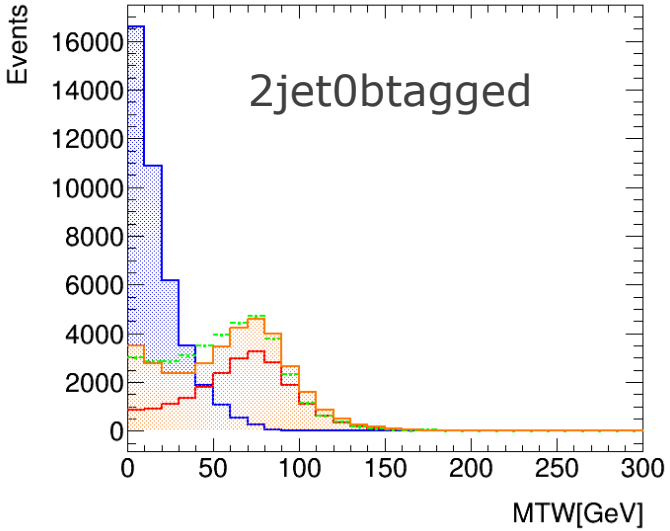


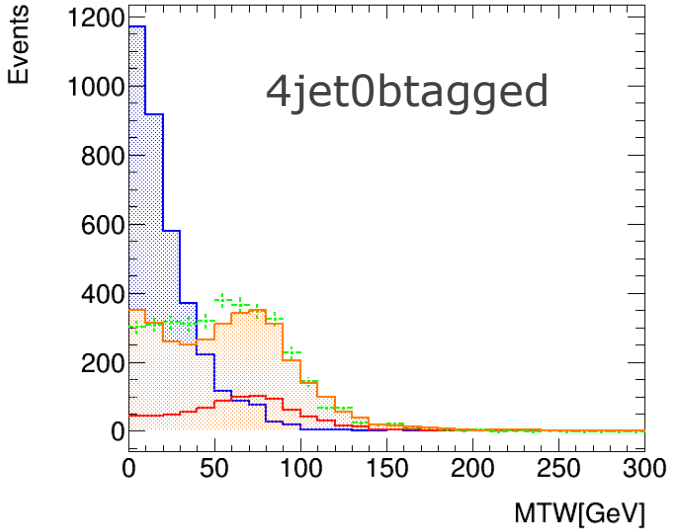
- $fake_{template}$
- $Wjets_{tight}$
- $data\text{-}MC(but\ not\ Wjets)_{tight}$
- $1.404 \times Wjets_{tight} + 0.137 \times fake_{template}$

MTW Distributions for SFs

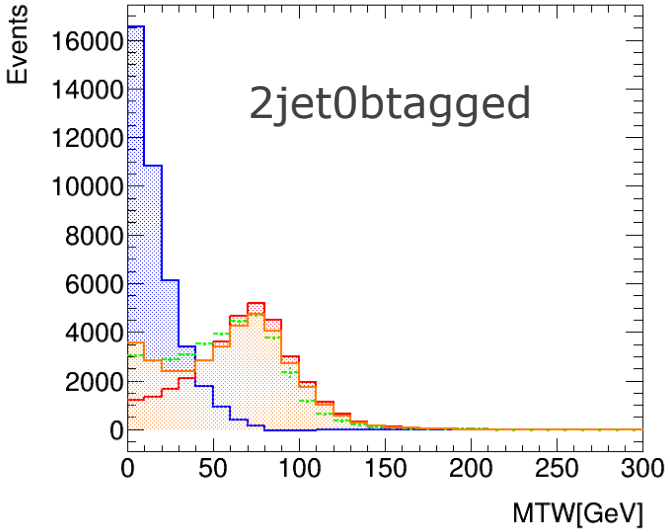
mu-ch

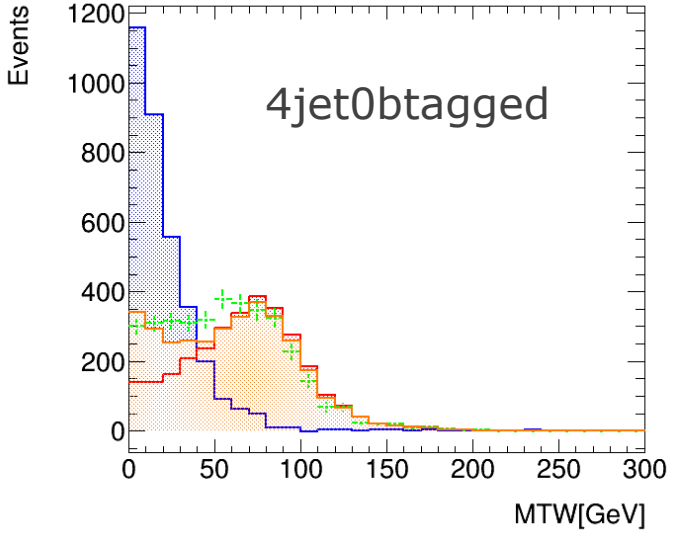
Powheg Pythia





Sherpa





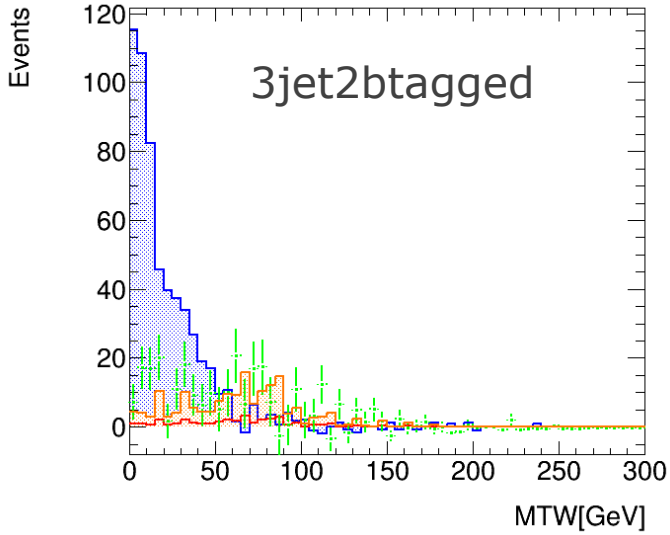
fake$_{template}$ data-MC(but not Wjets)$_{tight}$

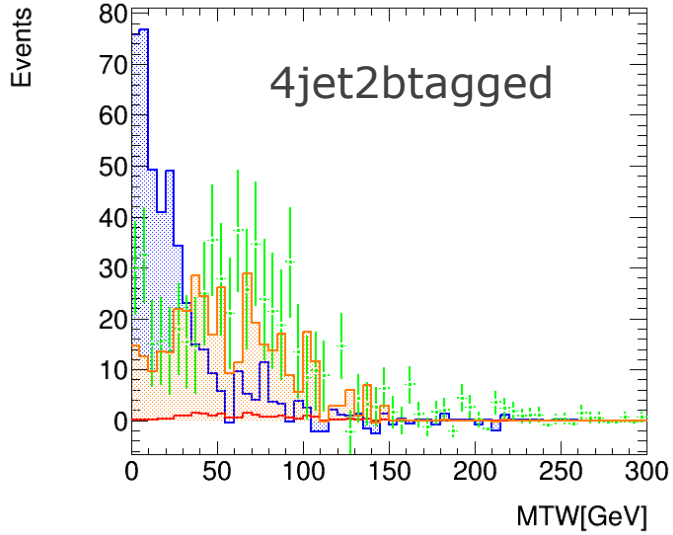
Wjets$_{tight}$ SFwjets x Wjets$_{tight}$ + SFfake x fake$_{template}$

MTW Distributions for SFs

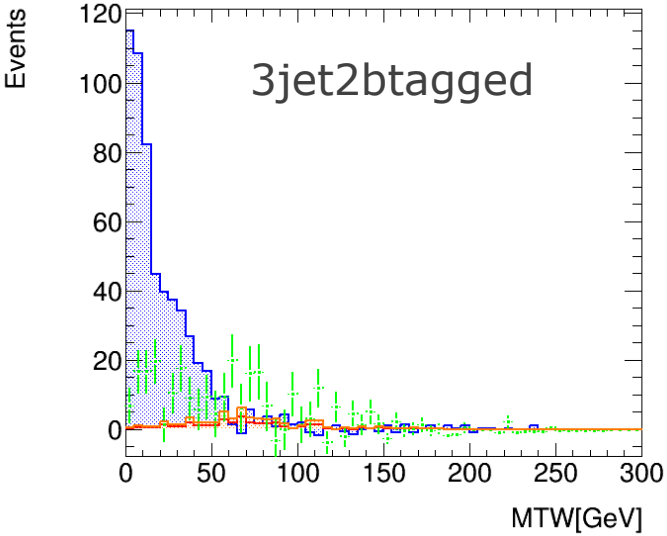
mu-ch

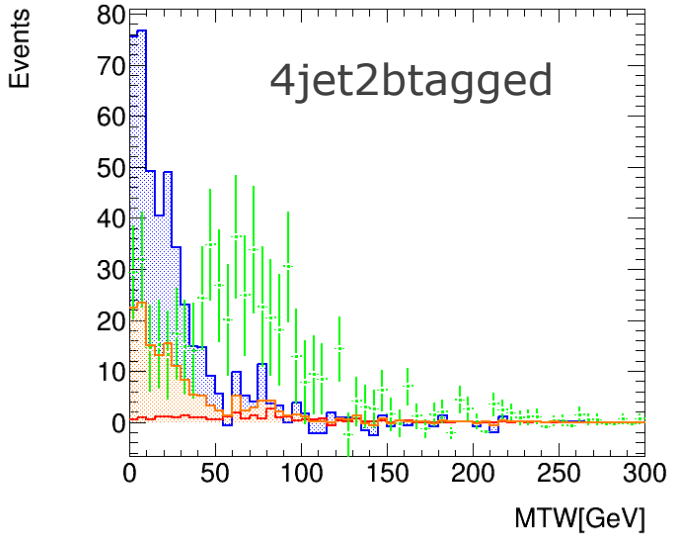
Powheg Pythia





Sherpa





Due to the low statistics with 2btagged region,
W peak are not and SFs are not reliable… :

Determined SFs

muon ch only
(electron ch is now in going)

PowhegPythia

Jetbin	bjetbin	SF in fake	SF in Wj
0	0	0.173	0.919
1	0	0.165	0.950
1	1	0.103	1.159
2	0	0.137	1.523
2	1	0.100	1.968
2	2	0.008	3.331
3	0	0.149	2.331
3	1	0.074	3.200
3	2 incl.	1.23×10^{-9}	4.931
4 incl.	0	0.175	3.631
4 incl.	1	0.119	5.150
4 incl.	2 incl.	0.142	18.380

Sherpa

Jetbin	bjetbin	SF in fake	SF in Wj
0	0	0.152	1.011
1	0	0.159	1.017
1	1	0.100	1.567
2	0	0.145	1.069
2	1	0.100	1.618
2	2	0.058	1,226
3	0	0.165	1.054
3	1	0.073	1.590
3	2 incl.	0.015	2.271
4 incl.	0	0.177	1.103
4 incl.	1	0.139	1.355
4 incl.	2 incl.	0.336	3.735

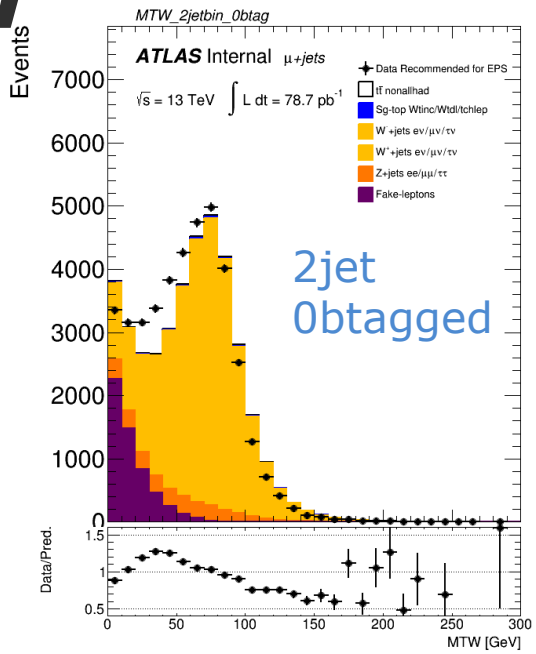
SF of the Wjets increases at higher jet multiplicity events in POWHEG.
This must be because of POWHEG mis-modeling.

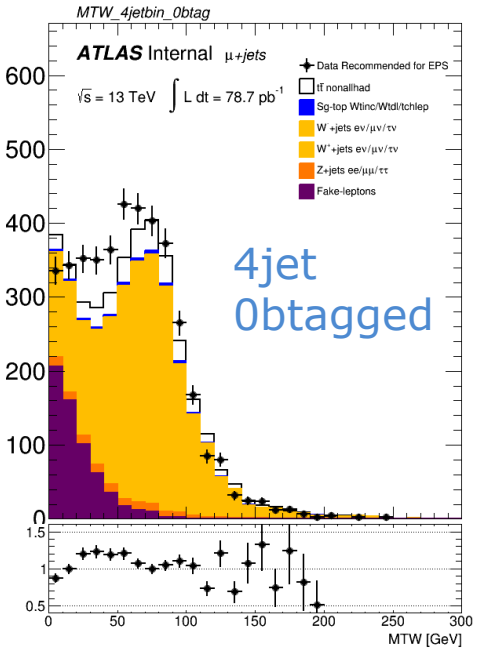
MTW

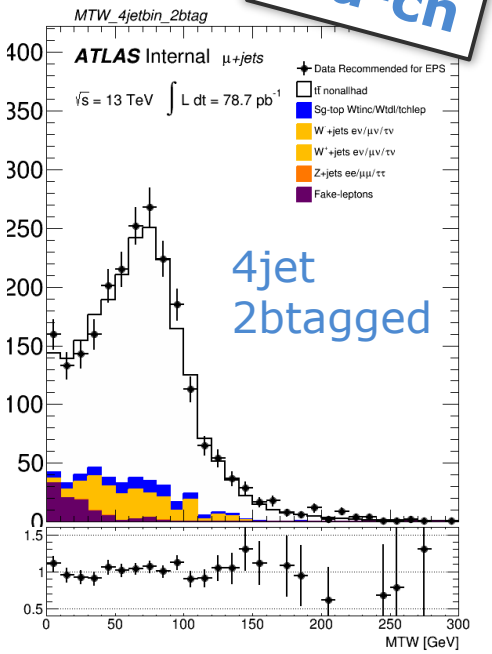
mu-ch

Powheg Pythia

with SF_{fake} and SF_{wjets}

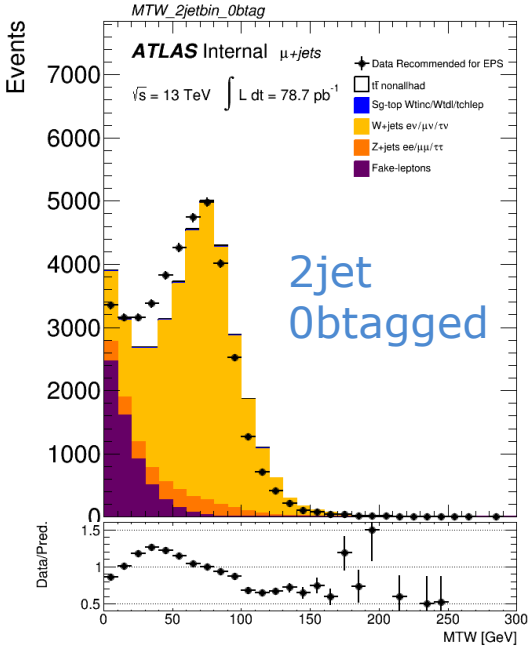


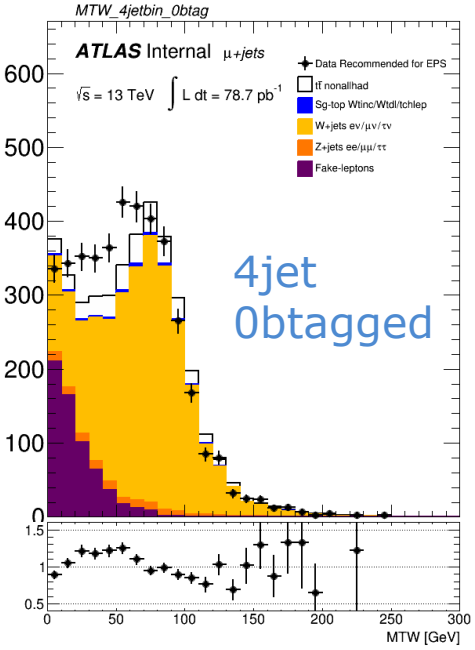


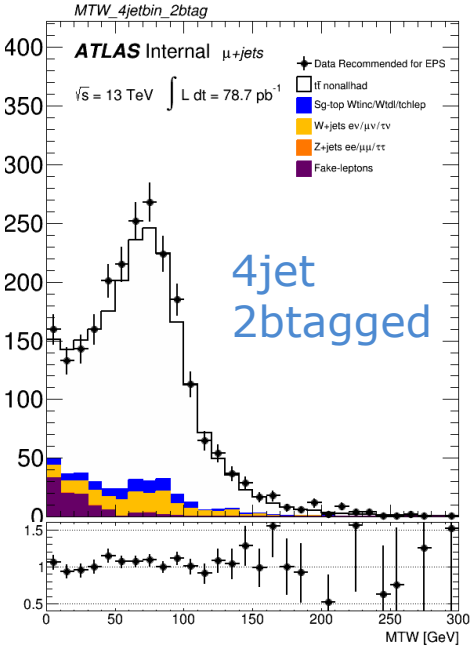


Sherpa

with SF_{fake} and SF_{wjets}



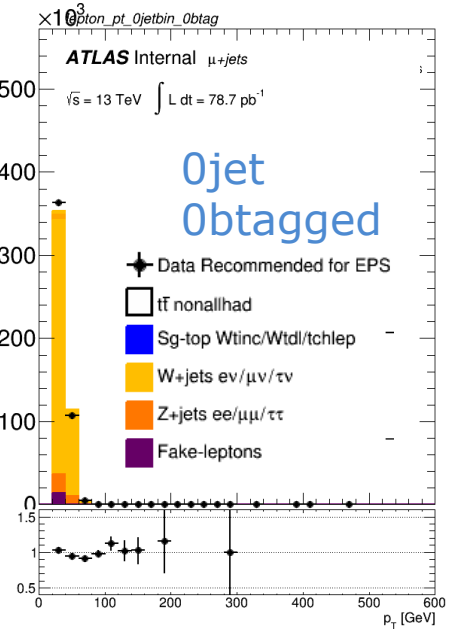


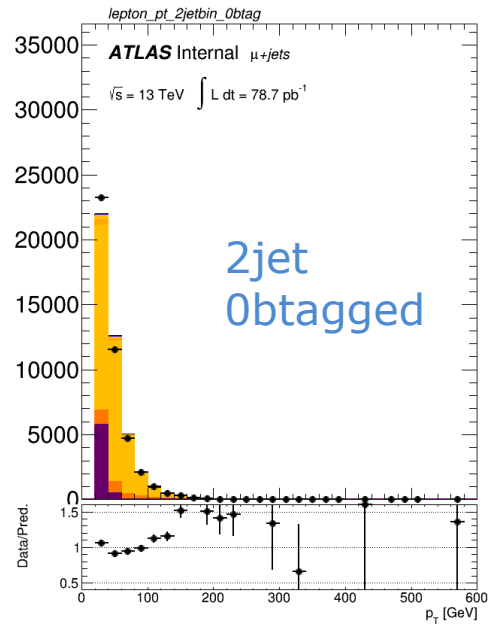


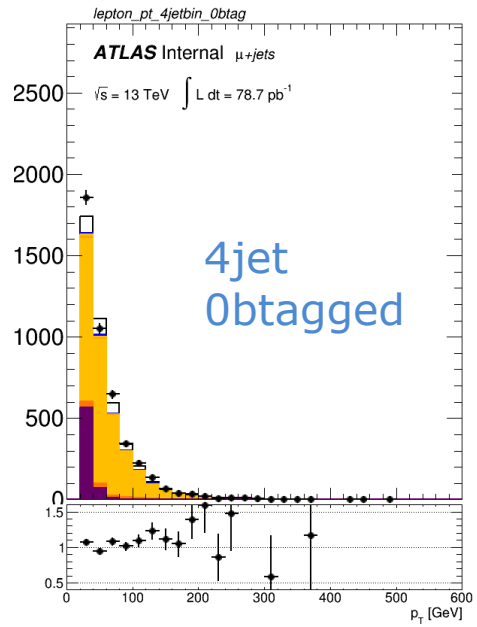
MC-Data Comparison with SF_{fake} only

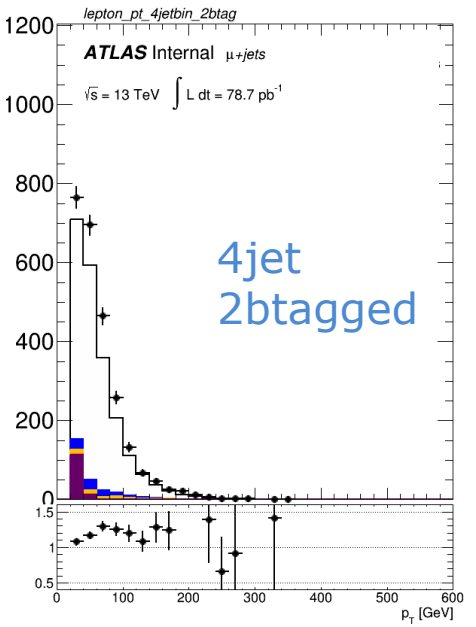
lepton

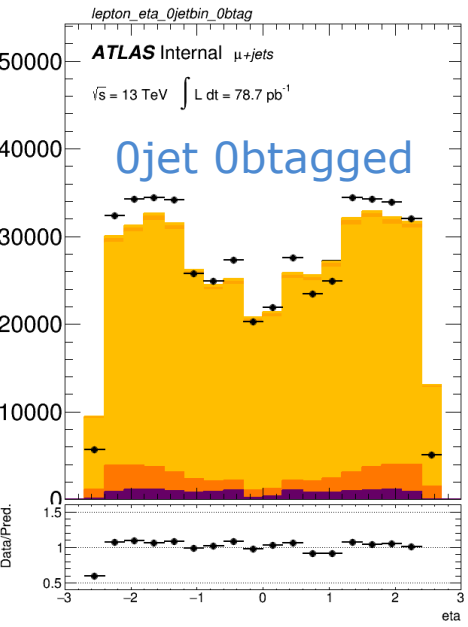
mu-ch, Wmunu:Sherpa
weight = lep*btag*mc

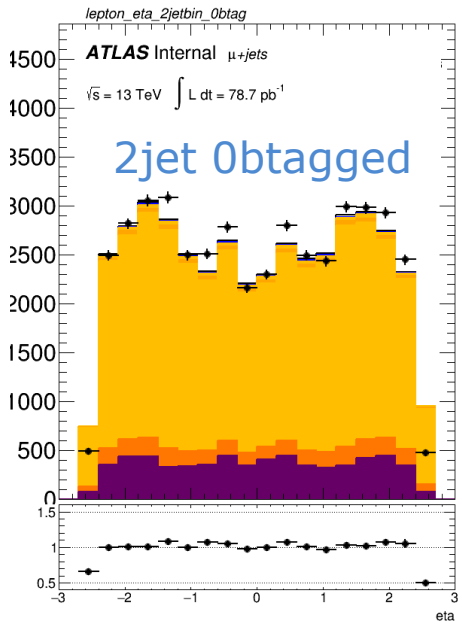


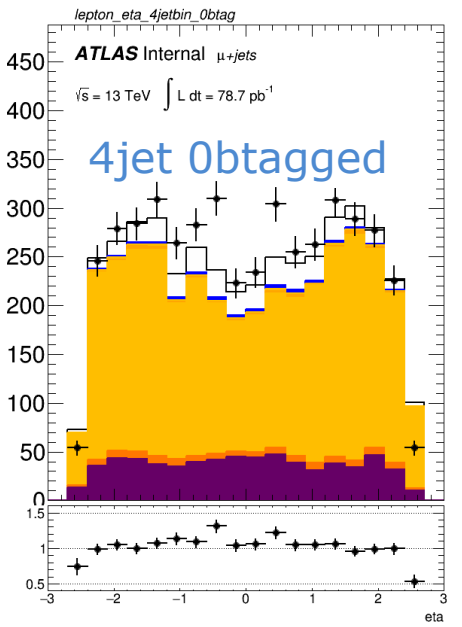


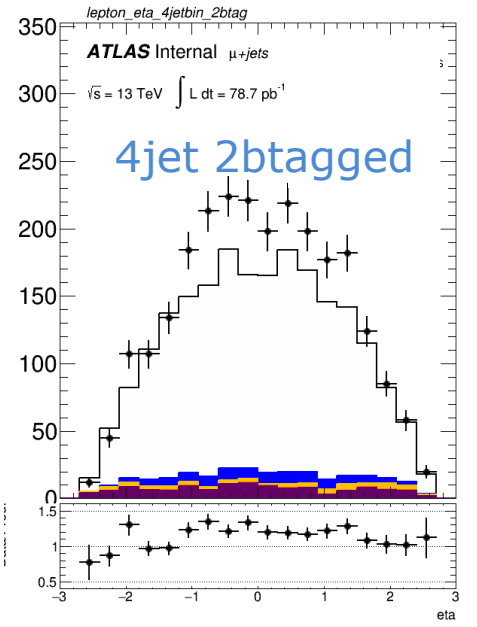








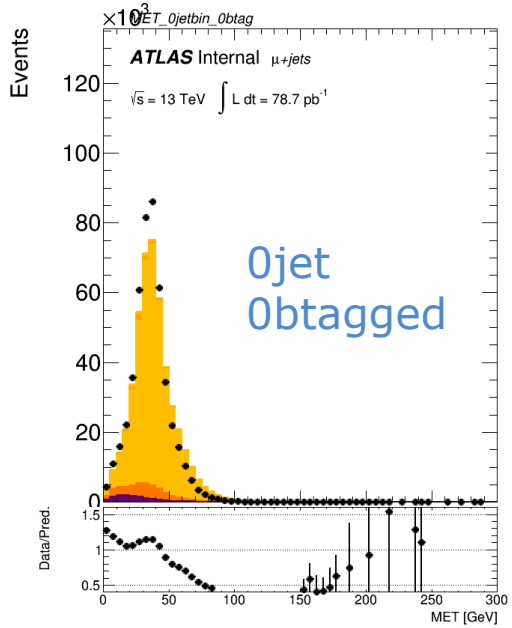


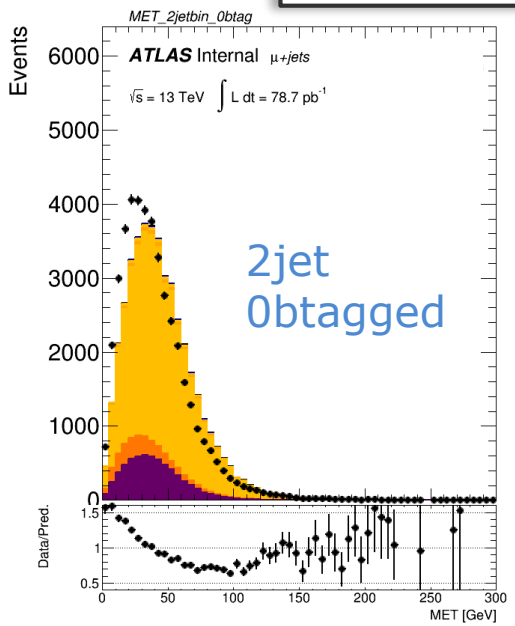


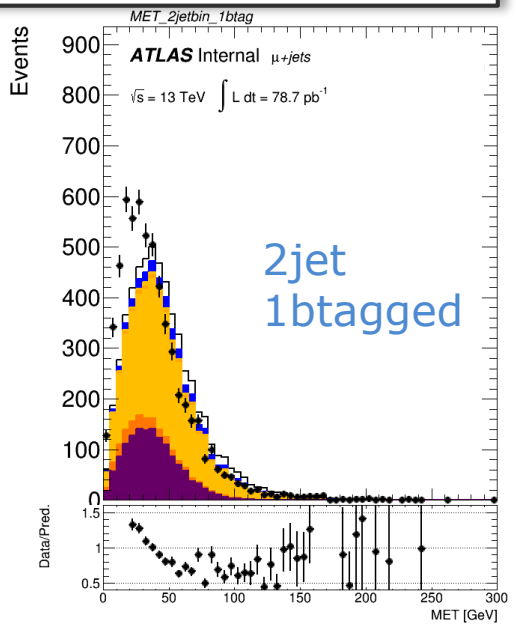
MC-Data Comparison with SF_{fake} only

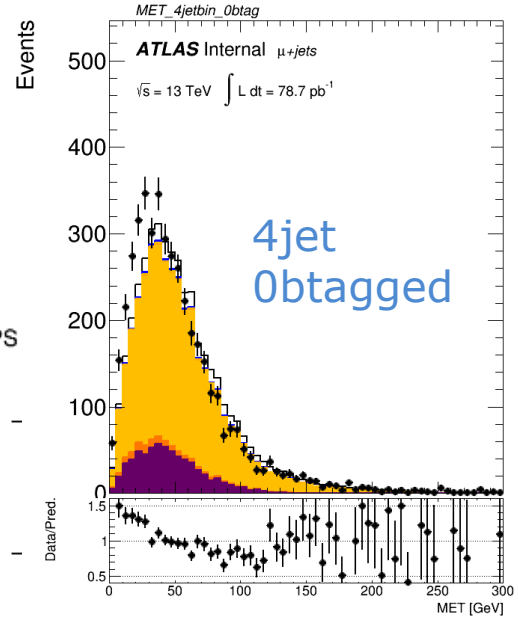
MET

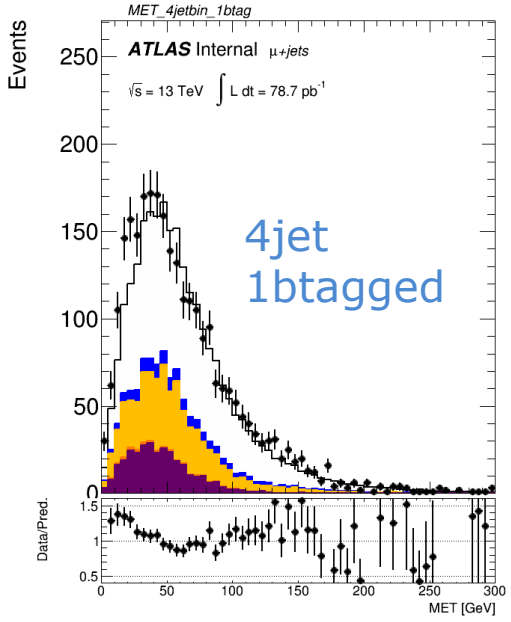
mu-ch, Wmunu:Sherpa
weight = lep*btag*mc

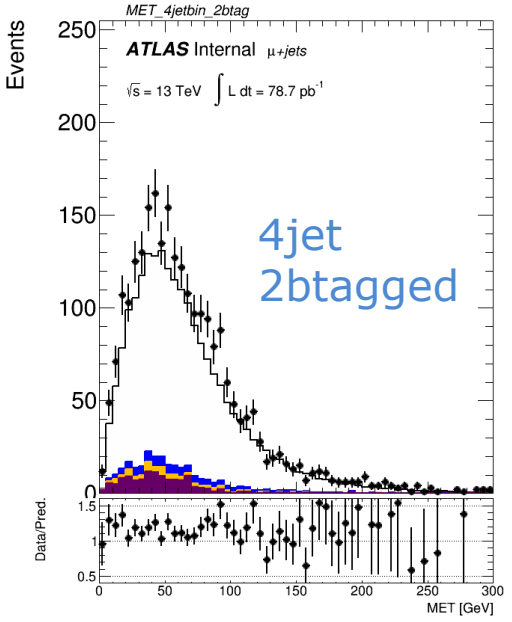










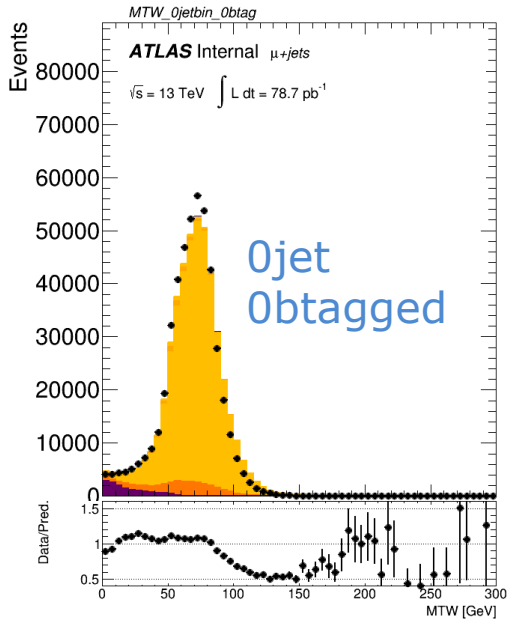


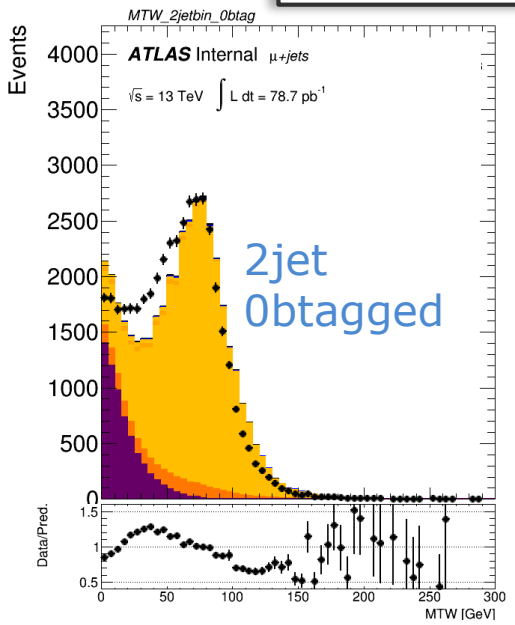
- Data Recommended for EPS
- tt̄ nonallhad
- Sg-top Wtinc/Wtdl/tchlep
- W+jets eν/μν/τν
- Z+jets ee/μμ/ττ
- Fake-leptons

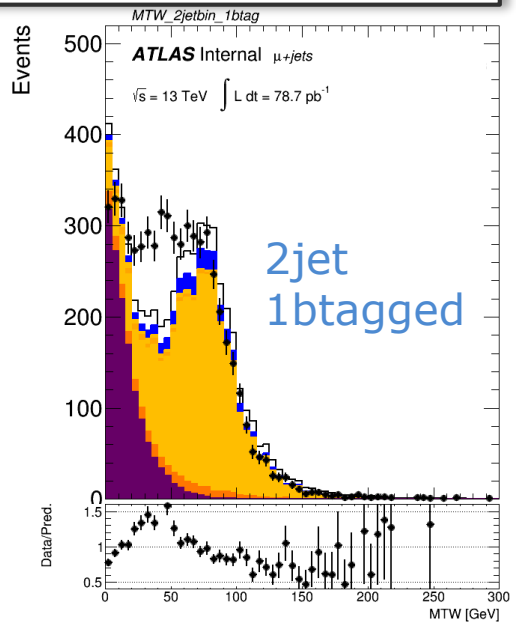
MC-Data Comparison with SF_{fake} only

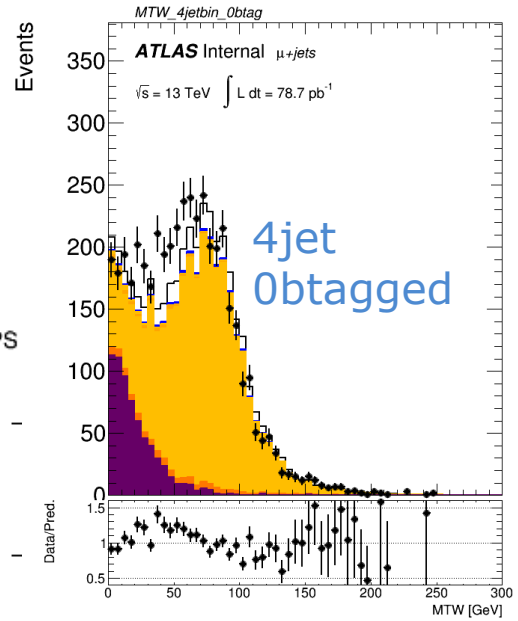
MTW

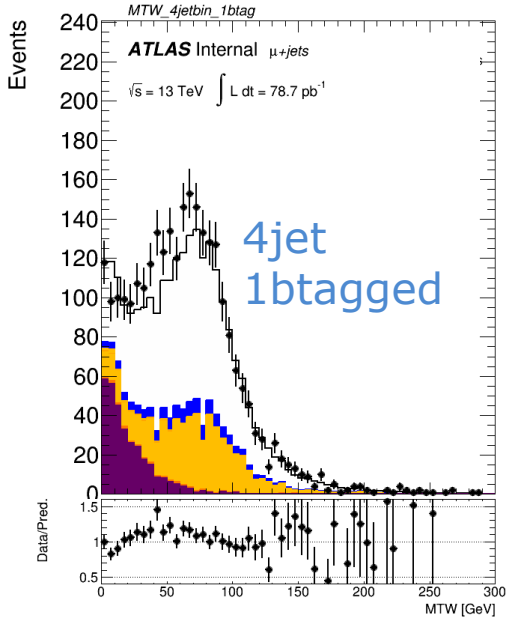
mu-ch, Wmunu:Sherpa
weight = lep*btag*mc

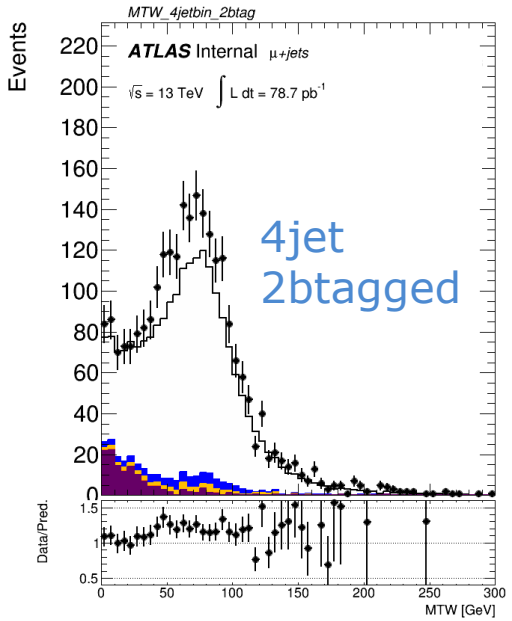










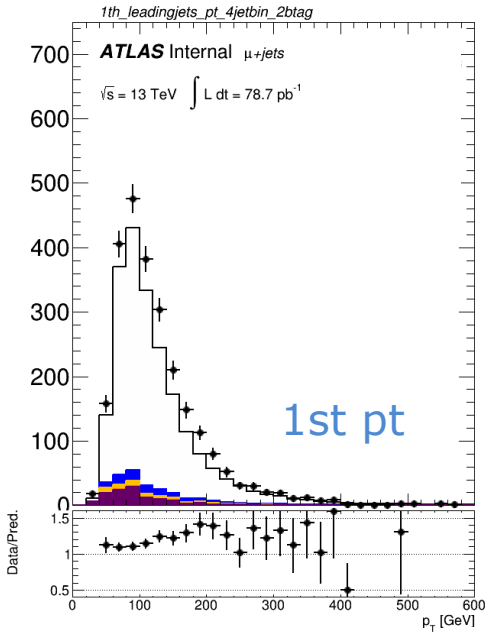


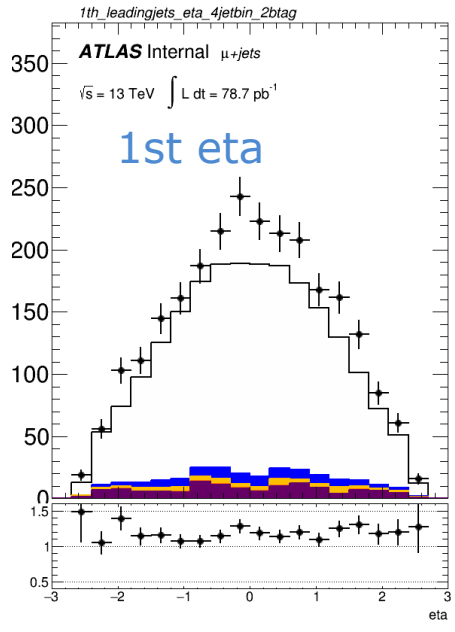
Data Recommended for EPS
tt̄ nonallhad
Sg-top Wtinc/Wtdl/tchlep
W+jets eν/μν/τν
Z+jets ee/μμ/ττ
Fake-leptons

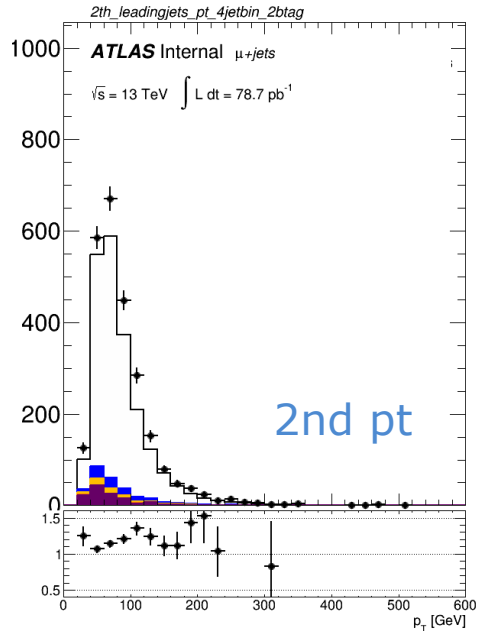
MC-Data Comparison with SF_{fake} only

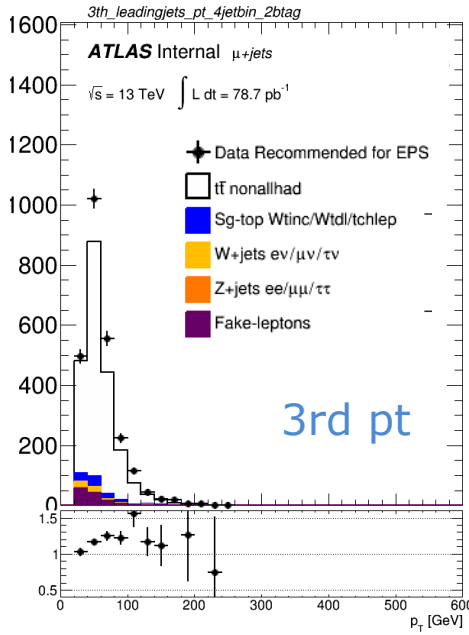
leading jets 4j2b

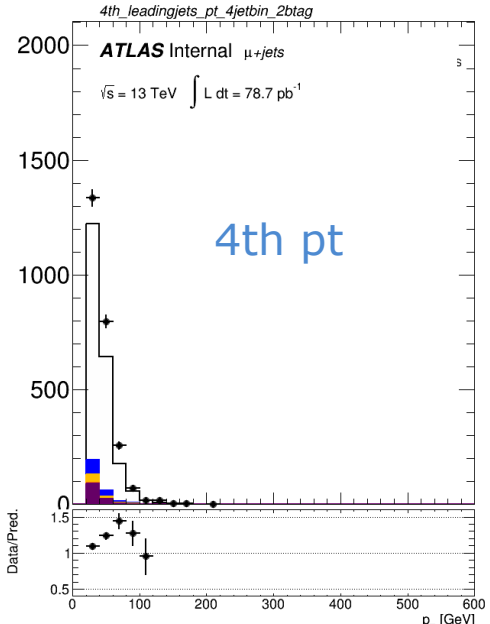
mu-ch, Wmunu:Sherpa
weight = lep*btag*mc

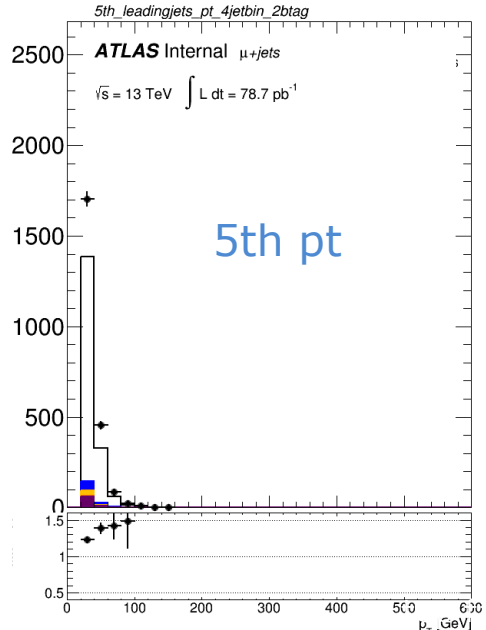


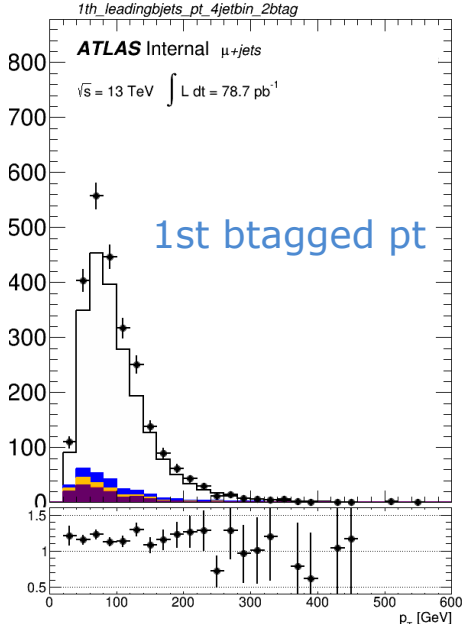


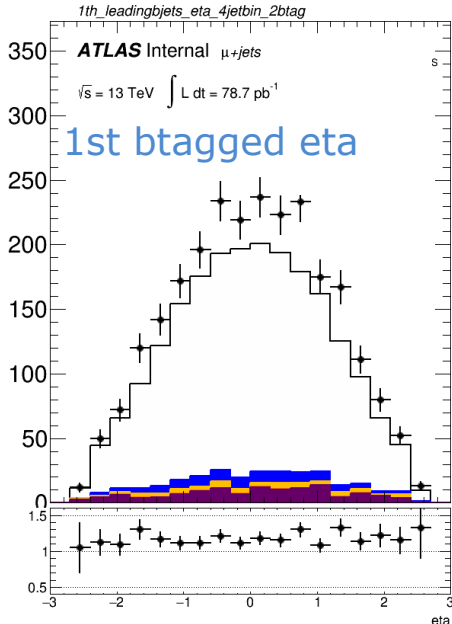










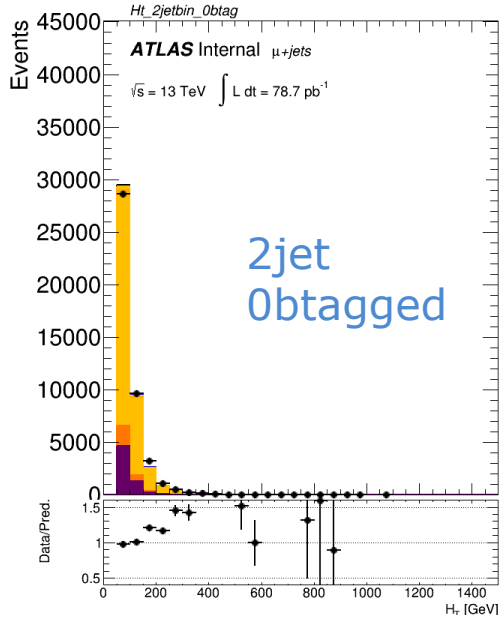


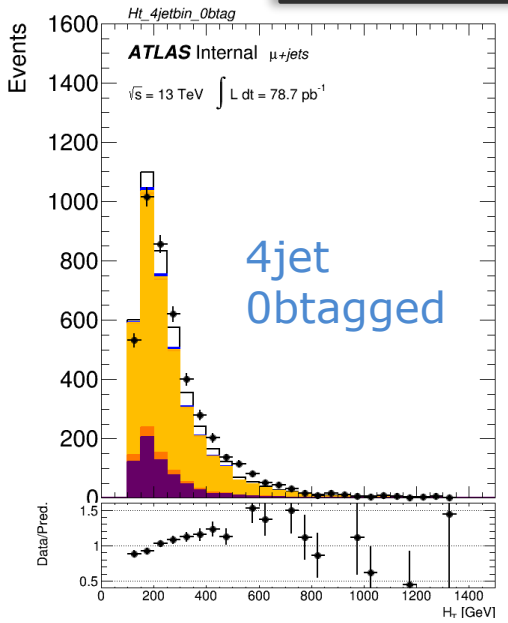
MC-Data Comparison with SF_{fake} only

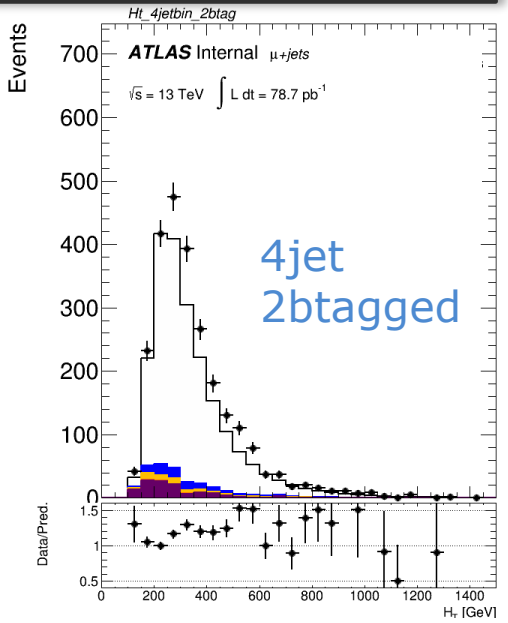
Scalar Sum

mu-ch, Wmunu:Sherpa
weight = lep*btag*mc

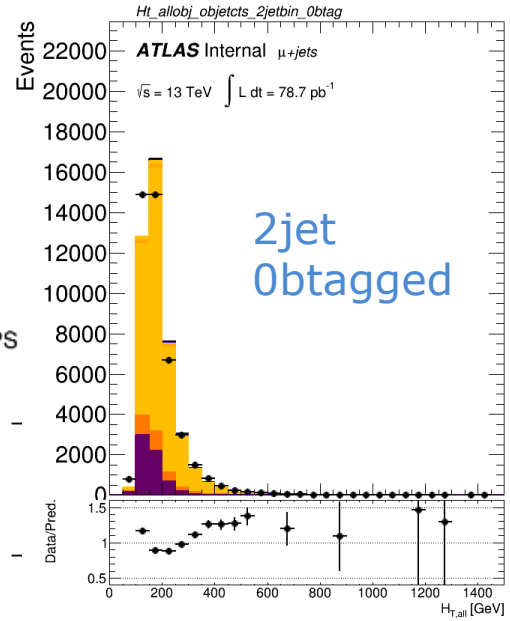
HT
scalar sum of jet-pTs

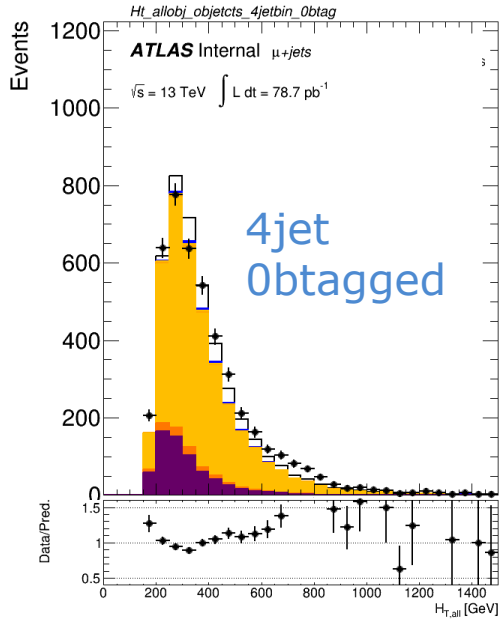


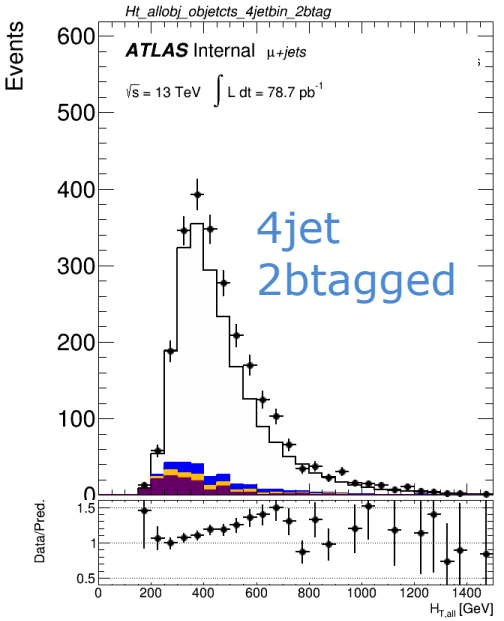




HTall
scalar sum of jet-pTs, lepton-pT, MET



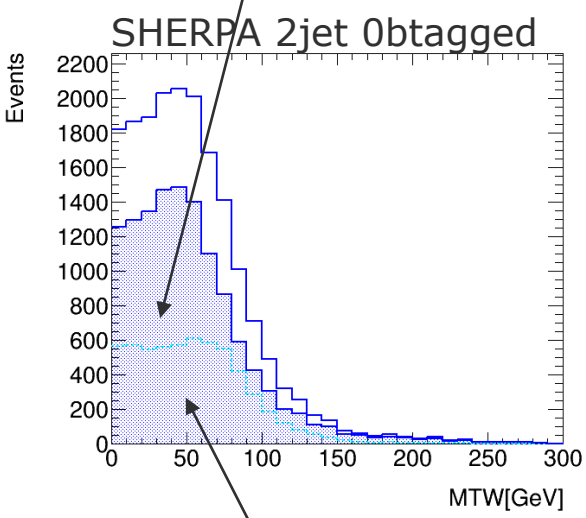


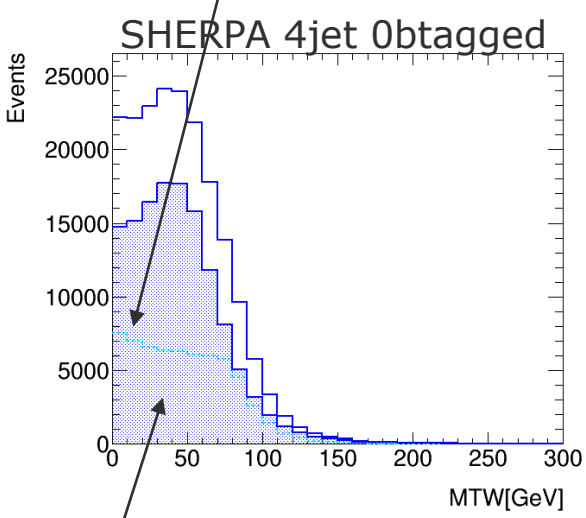


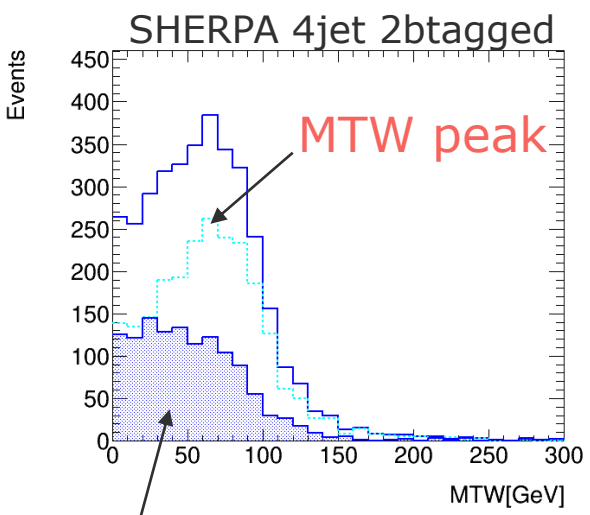
- Data Recommended for EPS
- $t\bar{t}$ nonallhad
- Sg-top Wtinc/Wtdl/tchlep
- W+jets eν/μν/τν
- Z+jets ee/μμ/ττ
- Fake-leptons

Difficulty of Electron Channel

can’t see MTWpeak in the data with tight selection







$data_{loose}$
$data_{tight}$
$data_{loose\ only}$

MWT peak contamination is large in the template

good template

Summary

- **tthbb needs looser signal-region**
  ⇒looser fake estimation are needed.

- **Good agreement of MC/data as first comparison with fake(QCD) estimate.**
  - Fakes were estimated using the MTW simply.
  - Electron channel analysis is now in going.
  - It is better to use Sherpa or multi-leg geenrators for Wjets-MC.

BACK UP >>

Lepton-ID in the derivations

Step 1 – define lepton selections

To estimate fake efficiency, we define two lepton selections (all with p_T > 25 GeV):

Tight (nominal)

- TightLH electron with gradient isolation
- Medium muon with gradient isolation

Loose

- LooseLH electron without isolation requirements
- Loose muon without isolation or ID hits requirements

Unfortunately, it turns out our derivation framework skims out events with loose leptons:

```
EL20 = "(Electrons.pt > 20*GeV && abs(Electrons.eta) < 2.5 && (Electrons.Medium || Electrons.DFCommonElectronsLHMedium))"
MU20 = "(Muons.pt > 20*GeV && abs(Muons.eta) < 2.5 && Muons.muonType == 0 && Muons.DFCommonGoodMuon)"
if TOPQname == 'TOPQ1':
  TOPQ_Selection_lep = "((count("+MU20+") >= 1) || (count("+EL20+") >= 1))"
```

combined muon

from this previous meeting:
https://indico.cern.ch/event/407085/#preview:1613760