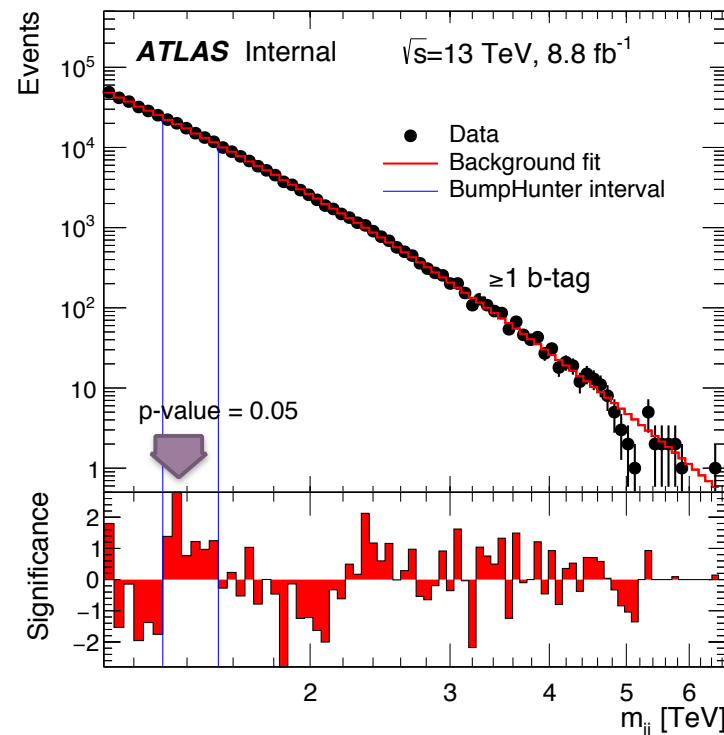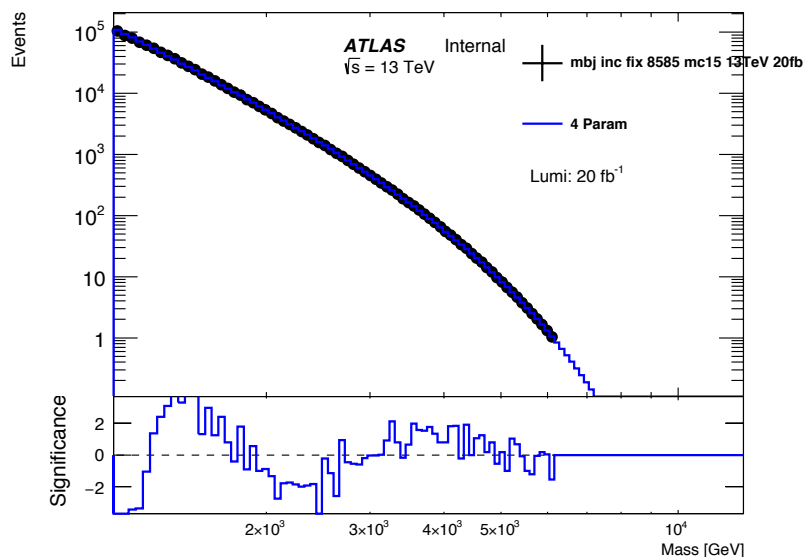# Fit Studies

Laurie McClymont,

Di-b-jet Analysis Team

<u>Ed Board</u>

14 July 2016

- **Evidence that fit is performing badly in mbj case at low masses - Seen before**

- **Change mjj range used**
  - Shown previously that at 1341 GeV fit is improved

- **Where should we put the cut**
  - Study p-values against mjj cut off in data and MC
  - Look for plateau in p-values

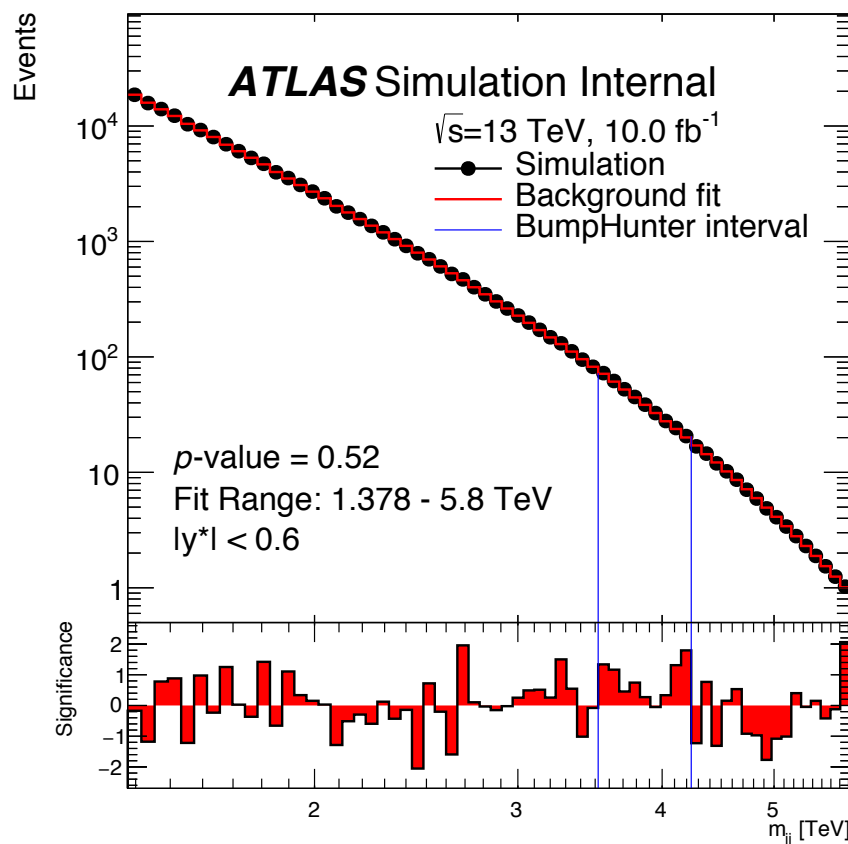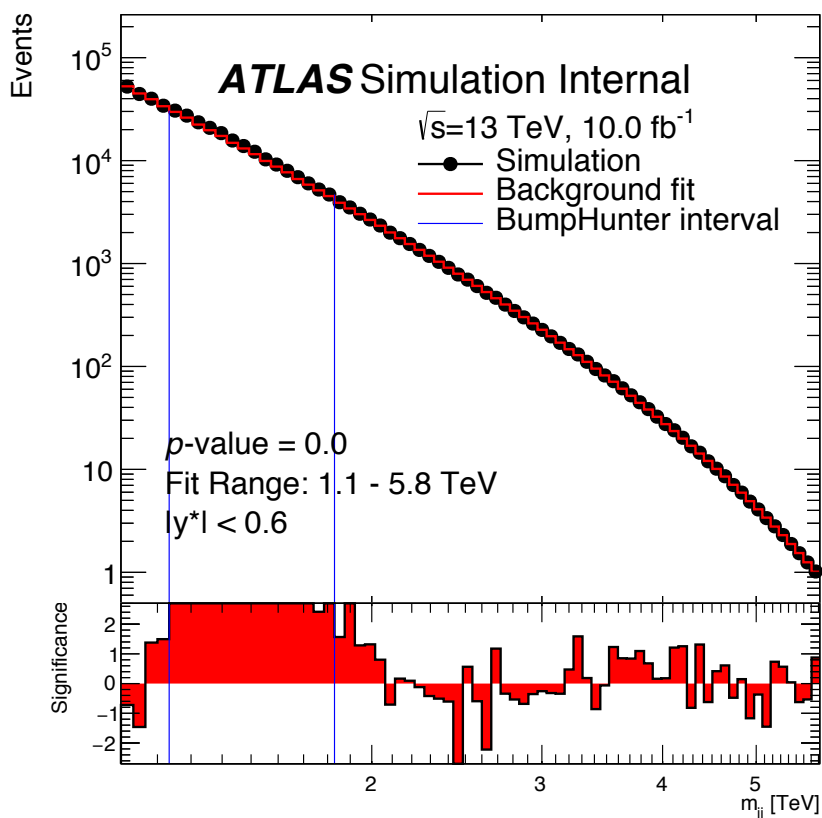- **Currently looking at mbj_inc in MC**

- **Fit to MC**
  - MC errors, number of MC entries rather than poisson errors
  - 'Short' - cut off where we expect one event (limits upper mass range)
  - Fit using search phase.
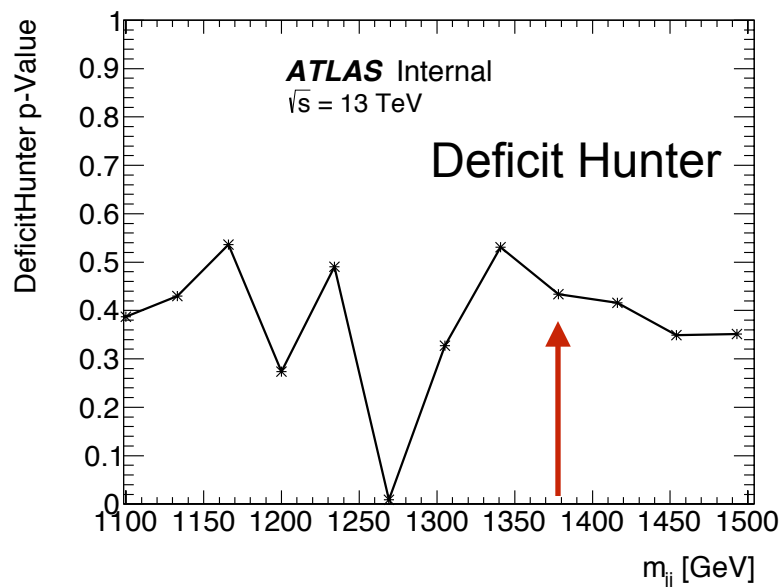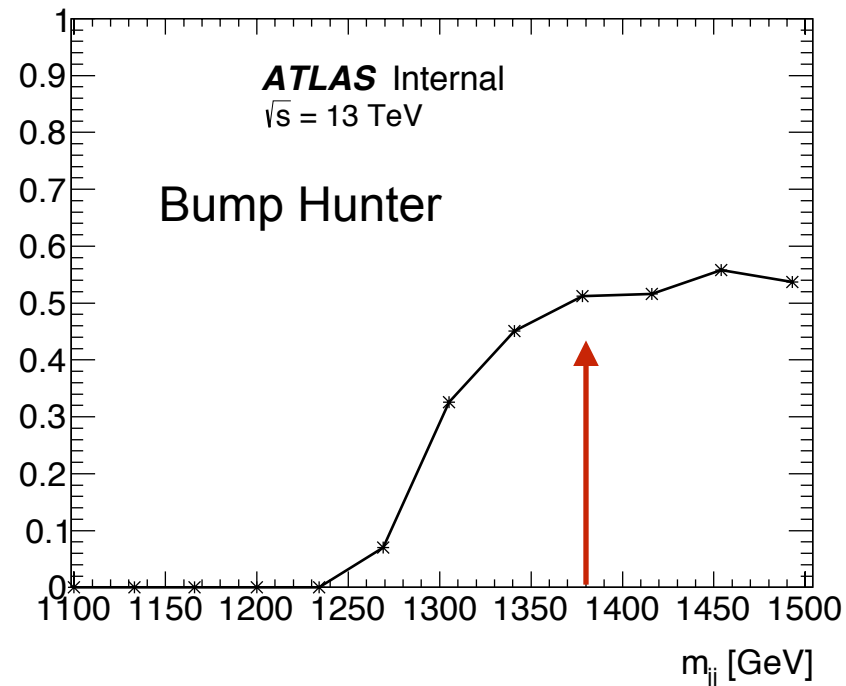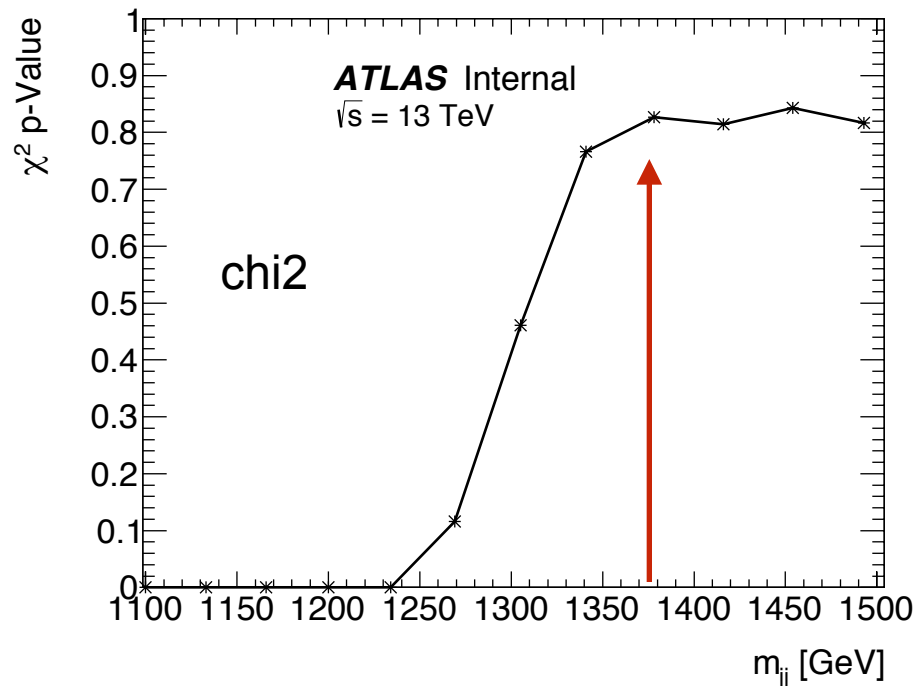 - MC - 20160712
    => Fixed b-tagging bug that was discussed in EB meeting
    => Old scale factors - Updated SFs will be used for ATLAS circulation

UCL

chi2
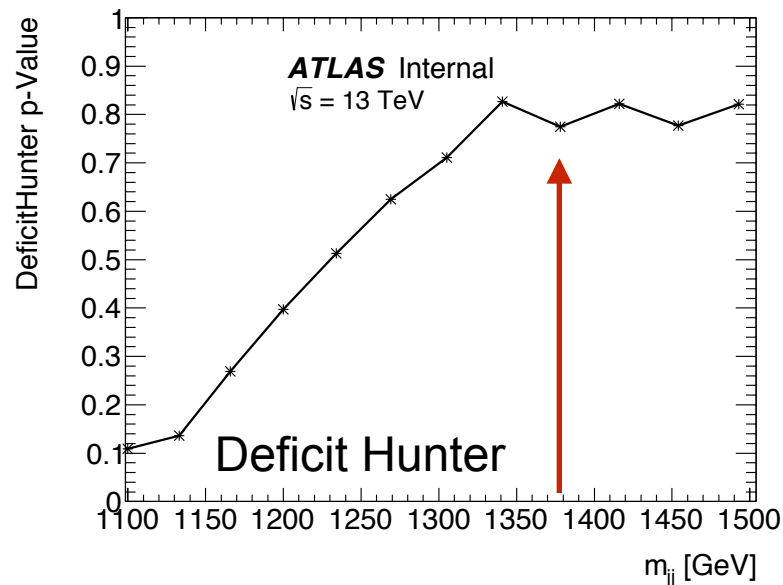
ATLAS Internal
√s = 13 TeV

χ^2 p-Value

m_{jj} [GeV]

Bump Hunter

ATLAS Internal
√s = 13 TeV

m_{jj} [GeV]
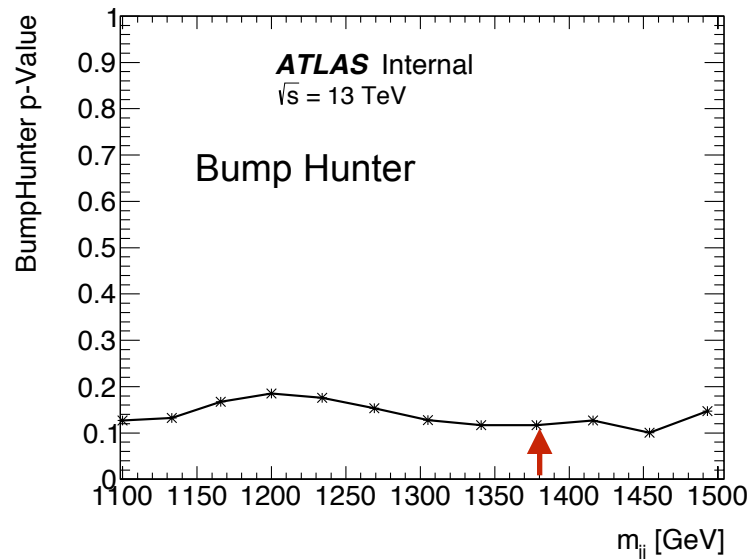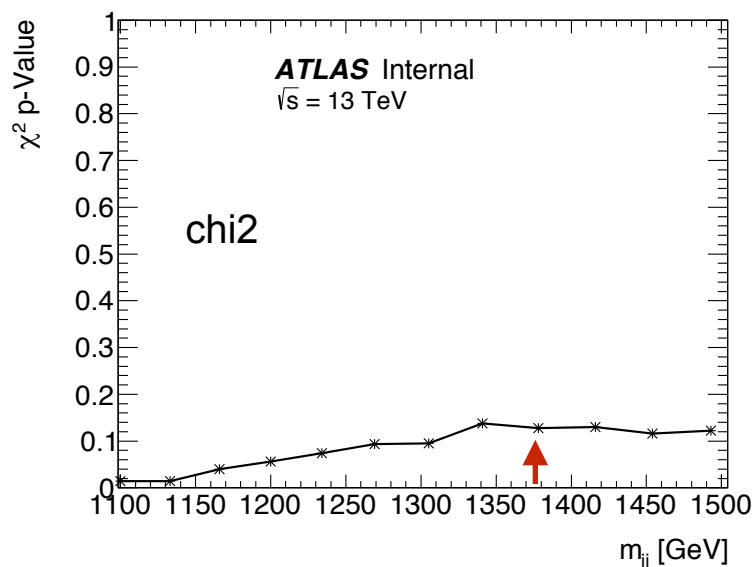

Deficit Hunter

ATLAS Internal
√s = 13 TeV

DeficitHunter p-Value

m_{jj} [GeV]

MC - 20160712
4 parameter fit function
Inclusive 1 b-tag category
Short @ 10 fib (cut off at 1 event)

Plateau at 1378 GeV
- Shows that mjj > 1378 GeV
  is a stable fitting region

χ^2 p-Value — chi2 — ATLAS Internal, $\sqrt{s}$ = 13 TeV — m_{jj} [GeV]

BumpHunter p-Value — Bump Hunter — ATLAS Internal, $\sqrt{s}$ = 13 TeV — m_{jj} [GeV]

DeficitHunter p-Value — Deficit Hunter — ATLAS Internal, $\sqrt{s}$ = 13 TeV — m_{jj} [GeV]

MC - 20160712
4 parameter fit function
2 b-tag category
Short @ 10 fib (cut off at 1 event)

- **Why are p-values lower than 1-tag?**
 - Errors given by MC stats
 - We are sensitive to fluctuations in MC production
 - Seems here there are some fluctuations…

- **We are ok in data-like**
 - These flucts are smaller than Poisson flucts
   we will see in data (especially at high mass)
 - Spurious signal study shows good fit quality

Stable fit region by 1378 GeV

MC - 20160712
4 parameter fit function
2 b-tag category
Short @ 10 fib
- (cut off at 1 event)

Plateau- at 1378 GeV
- Shows that mjj > 1378 GeV
  is a stable fitting region

- **Why are p-values lower than 1-tag?**
- Errors given by MC stats
- We are sensitive to fluctuations
  in MC production
- Seems here there are some
  fluctuations…

- **We are ok in data-like**
- These flucts are smaller than
  poisson flucts we will see in data
  (especially at high mass)
- Spurious signal study shows
  good fit quality

mjj > 1.1 TeV
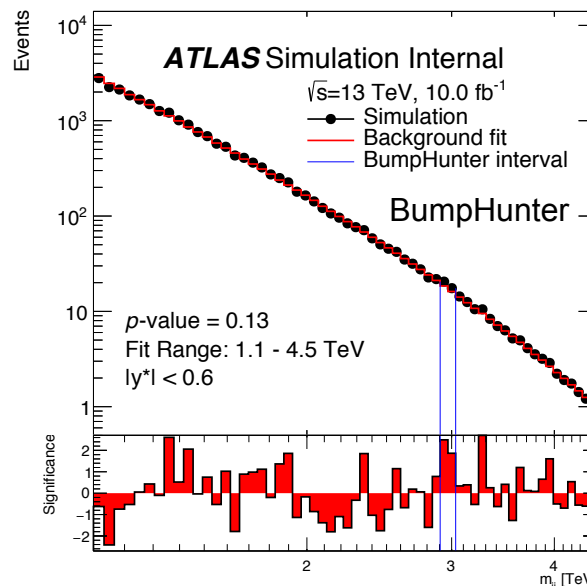
mjj > 1.378 TeV
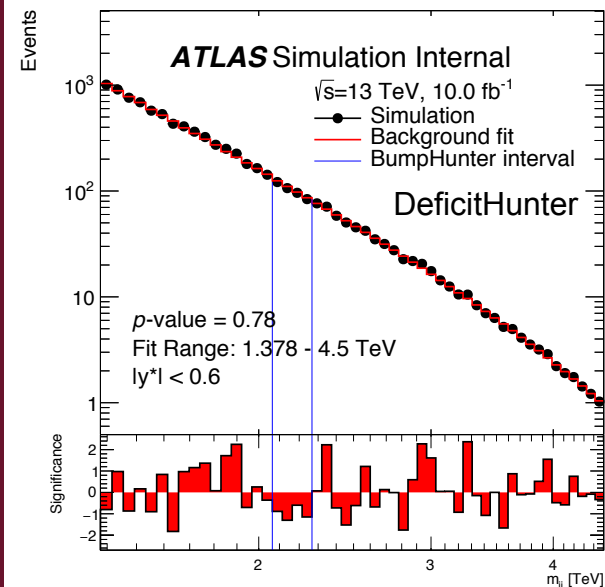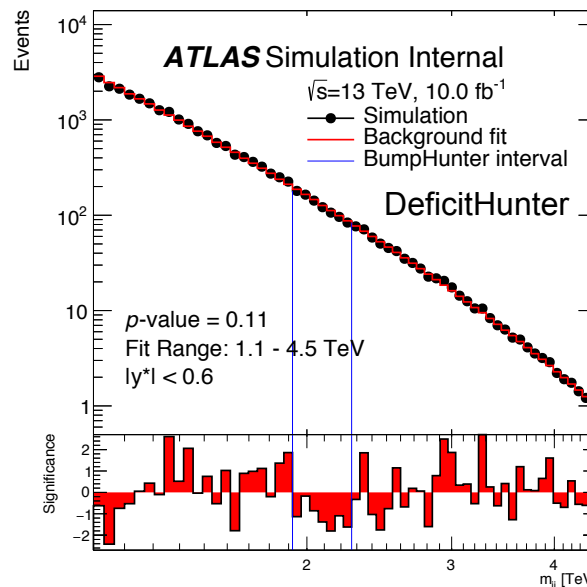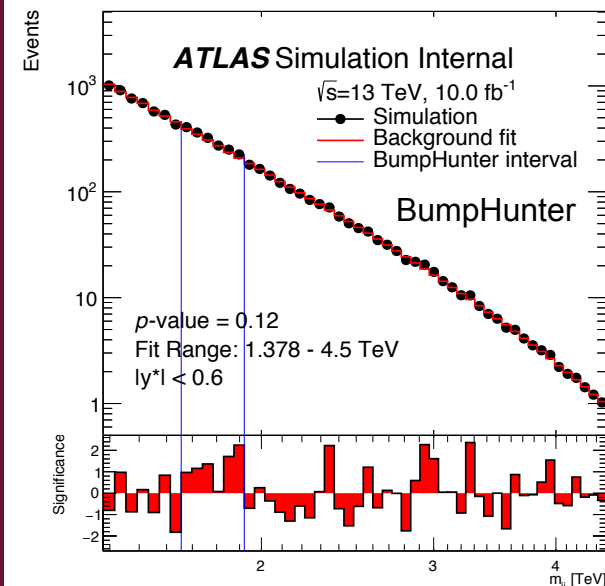
Spurious Signal: >= 1 b-tag

>= 1 b-tag category

Data-like background only distributions
- Taken from MC scaled to 10 ifb
- Apply poison fluctuations
- 200 different fluctuations

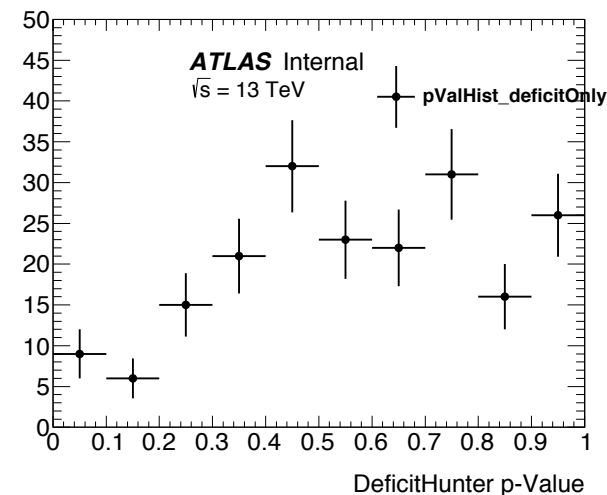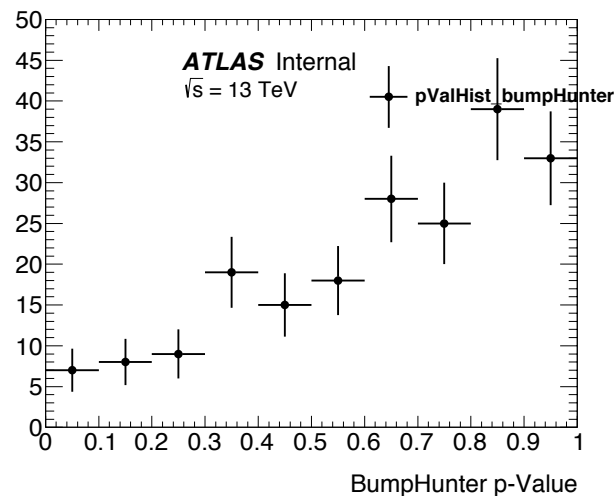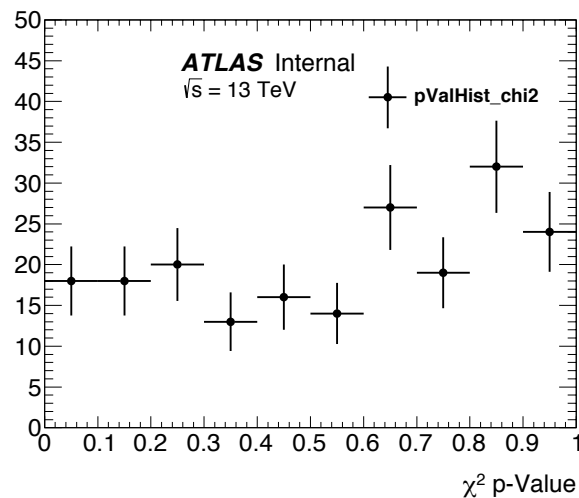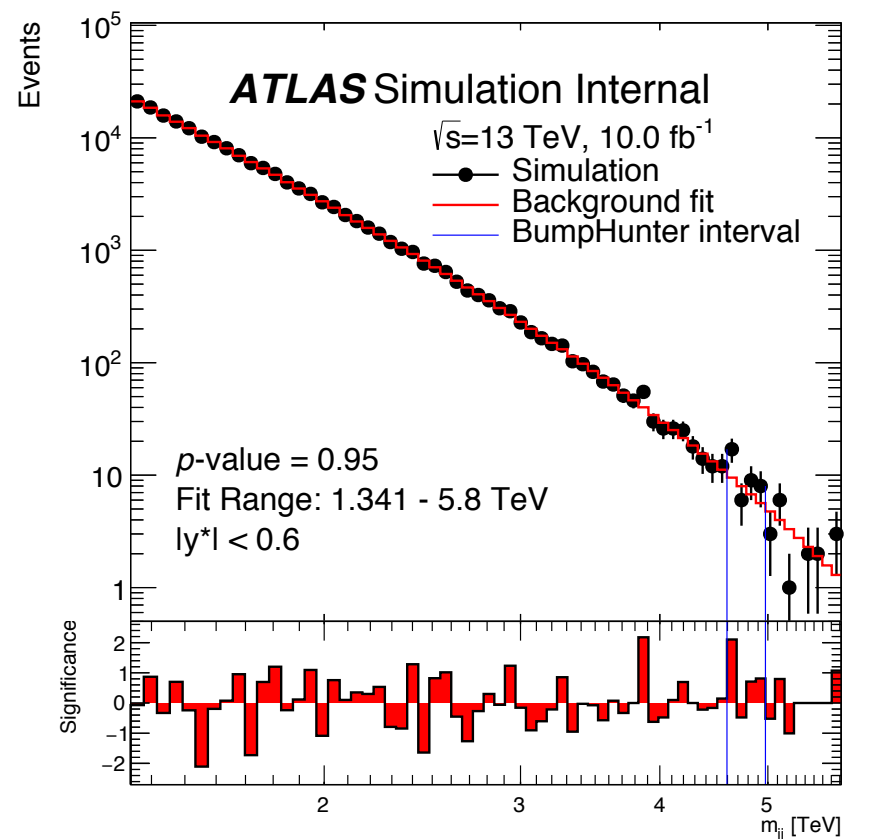Fit Range: 1341 - End of Data
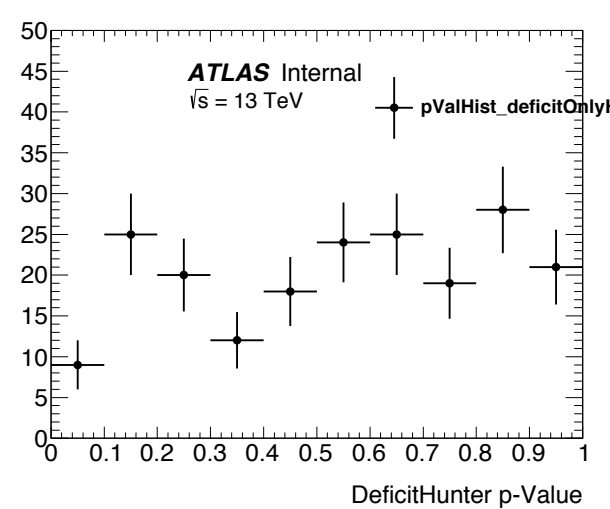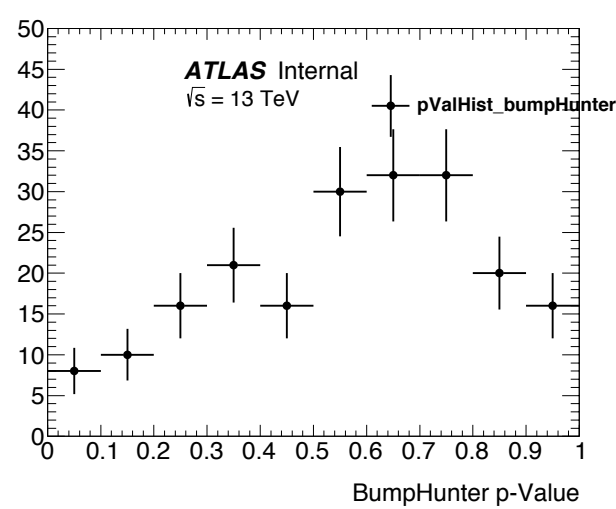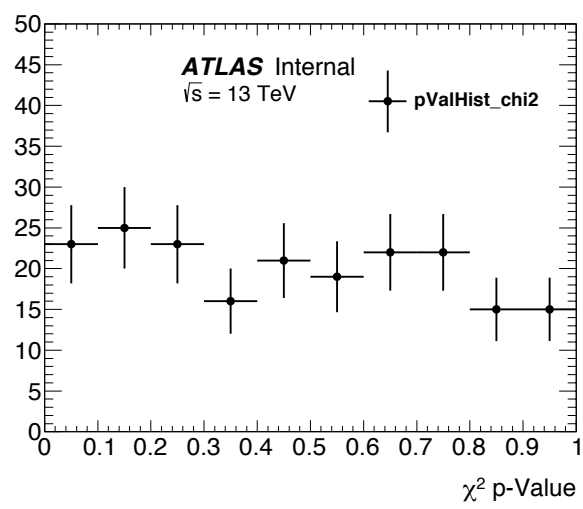=> Region where we have enough precision

2 b-tag category

Data-like background only distributions
- Taken from MC scaled to 10 ifb
- Apply poison fluctuations
- 200 different fluctuations

Fit Range: 1100 - End of Data
=> Original Full Range
=> Use patch to fix bins where we don't have enough precision

ATLAS Simulation Internal
$\sqrt{s}$=13 TeV, 10.0 fb^{-1}
● Simulation
— Background fit
— BumpHunter interval

p-value = 0.54
Fit Range: 1.1 - 4.9 TeV
ly*l < 0.6

m_{jj} [TeV]

ATLAS Internal
$\sqrt{s}$ = 13 TeV
pValHist_chi2

χ^2 p-Value

ATLAS Internal
$\sqrt{s}$ = 13 TeV
pValHist_bumpHunter

BumpHunter p-Value

ATLAS Internal
$\sqrt{s}$ = 13 TeV
pValHist_deficitOnly

DeficitHunter p-Value
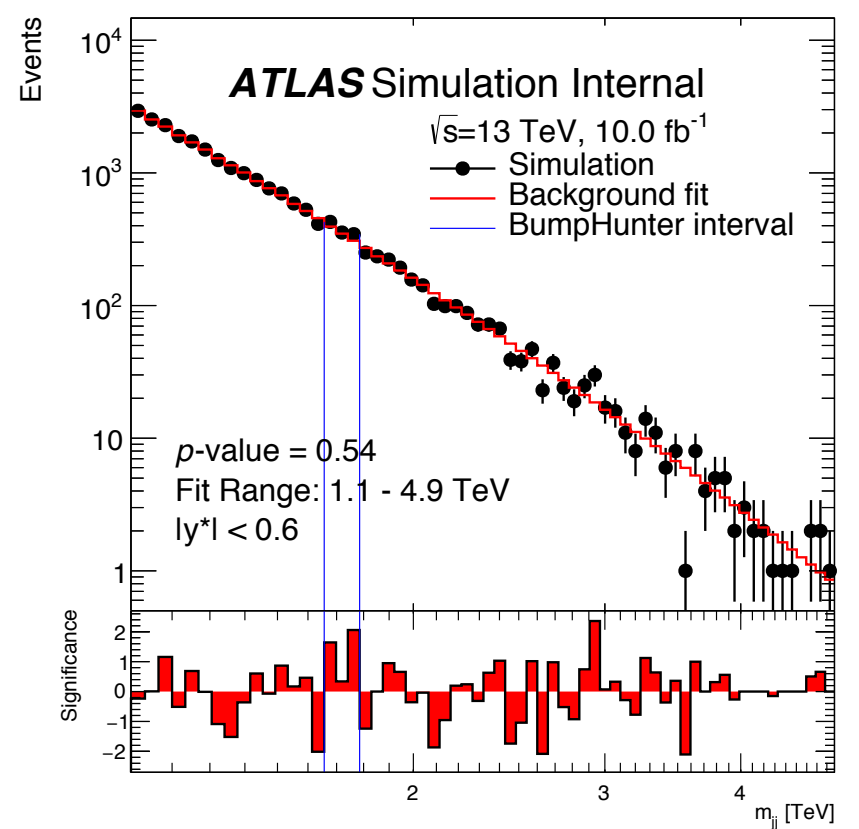
- **MC - 20160712**
=> Fixed b-tagging bug
=> Old scale factors

- **>= 1 b-tag category**
 - **chi2 and bH p-value plateau @ mjj > 1378 GeV**
 - No spurious signal, mjj > 1341 GeV
 - Running 3-parameter now

- **2 b-tag category**
 - chi2 and bH p-value plateau @ mjj > 1378 GeV
 - No spurious signal, mjj > 1.1 TeV

- **MC - 20160713 - Coming Soon**
=> Fixed b-tagging
=> New scale factors
 - Will run now/over weekend
 - Differences should be small

Backup

Event Selection

Pythia8EvtGen MC Di-Jet Sample
- **HLT_j380**
- 2016 MC
- di-b-jet Ntuple production

Scale to 10ifb
- Will update for final lumi
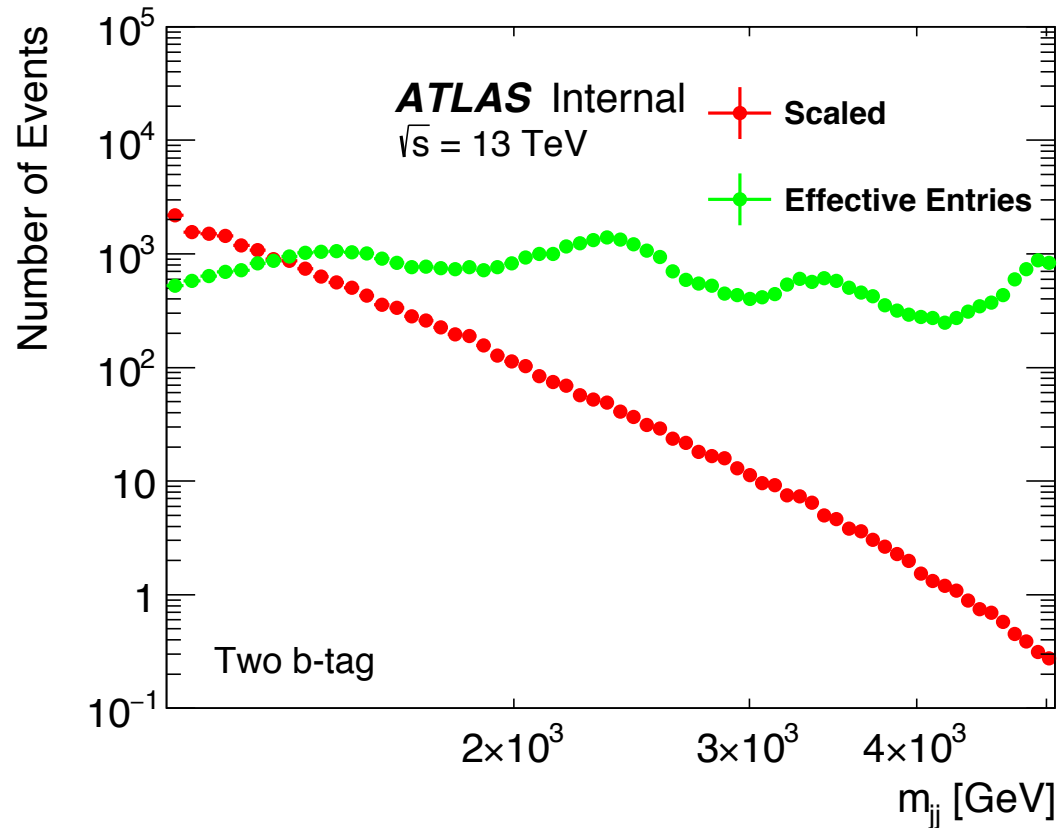
- **Standard Dijet Resonance Cuts**
  - Leading Jet pT > 430 GeV
  - Sublead Jet pT > 60 GeV
  - $|y^*| < 0.6$
  - mjj > 1100 GeV

- **MV2c10**
  - Using fixed cut 85% for both jets
  - mbb_fix_8585
  - mbj_inc_fix_8585

We want data-like distributions, poisson fluctuations applied to precise background estimate

Where Effective Entries > Scaled: We have enough stats for fit tests

Scaled > Eff Ent. : Errors driven by MC, can't make data-like distributions
- We do not have a high precision prediction for background here
- In this case, there are 7 bins where this is true
- Mass range: 1100 - 1341 GeV

Patch 7 bins using fit to full spectrum

Procedure:
- Scale to 20ifb => Large range
- **Cut off => Consider up to 1**
           => True MC errors
- Fit using **4 para fit function**
- Use this fit for 7 discrepant bins
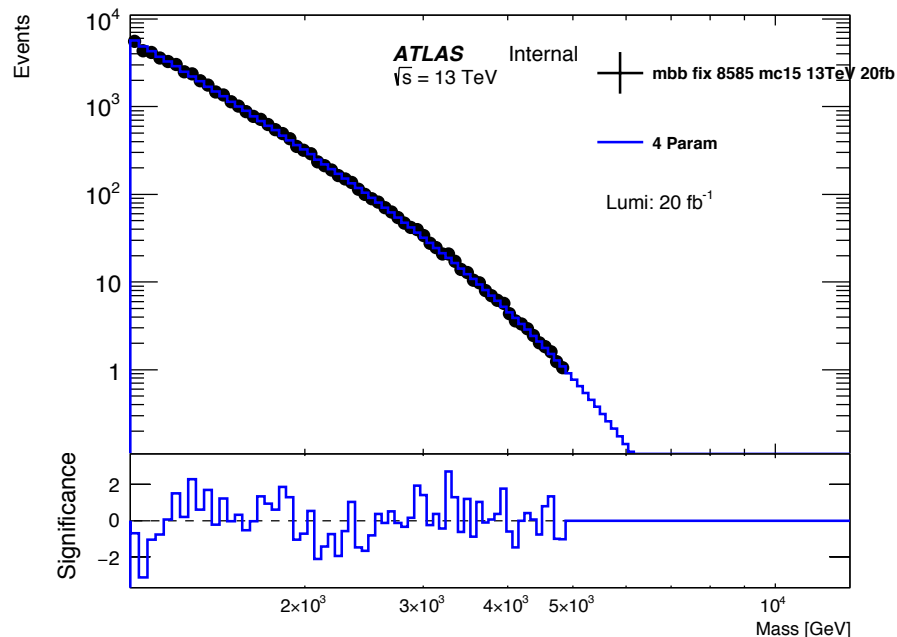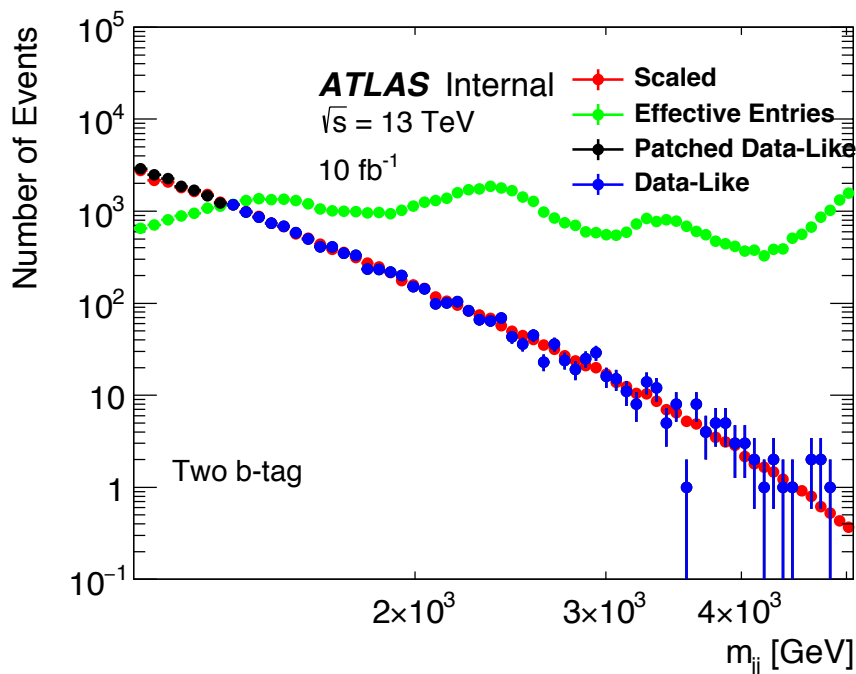
Chi2: 93.585   NdF: 115.000

Make Data Like Distribution

Procedure:
- Use Scaled distribution from MC
- (Use patched fit for those 7 bins)
- Apply poisson fluctuations
- This can be done for many seeds