Flavour Composition Studies: *Spurious Signal Test, Using 3 Parameter Initialisation*

Laurie McClymont,
di-b-jet Analysis Team

Di-b-jet Ed Board Meet

10 February 2016

Getting the Flavour Fractions

- We want to understand how varying the flavour composition will affect the fitting function.
  => Are the fitting functions robust to changes flavour composition?
  => Vary the amount that different flavour combinations contribute and fit.

- Changes for v1.5 in response to comments @ JDM and Exotic approval meetings
  => Spurious signal checks included.
  => Change to modelling flavour fraction components using 3-parameter fit function.

Details

Pythia8EvtGen MC Di-Jet Sample
- di-b-jet Ntuple production

Standard Dijet Resonance Cuts
- Leading Jet pT > 410 GeV
- Sublead Jet pT > 50 GeV
- $|y^*| < 0.6$
- mjj > 1100 GeV
- $|\eta| < 2.4$

Using fixed cut 85% for both jets.
- mbb_fix_8585

Cone matching truth flavour
- jetHadronConeExclTruthLabelID

Work Flow

phys-exotics/jdm/dijet/inputs/Btag/MC15_DiJet_20151104

Packages used

- DijetHelpersPackage:
=> Create scaled distributions.
=> Vary flavour fractions.
=> Create p-values of fit.

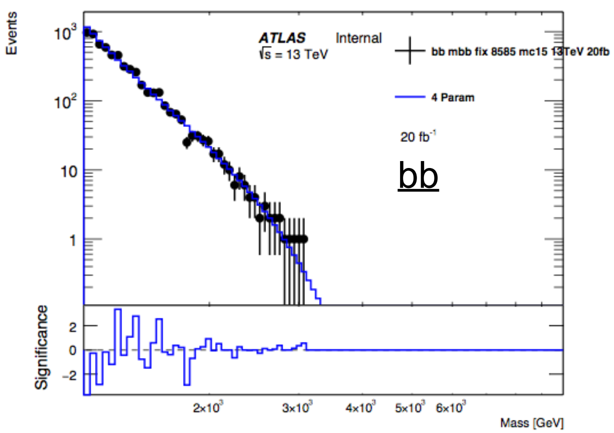
- Dijet Statistical Packages:
=> Using search phase from this package
=> Bumphunter to search for discrepant regions
=> Spurious signal check

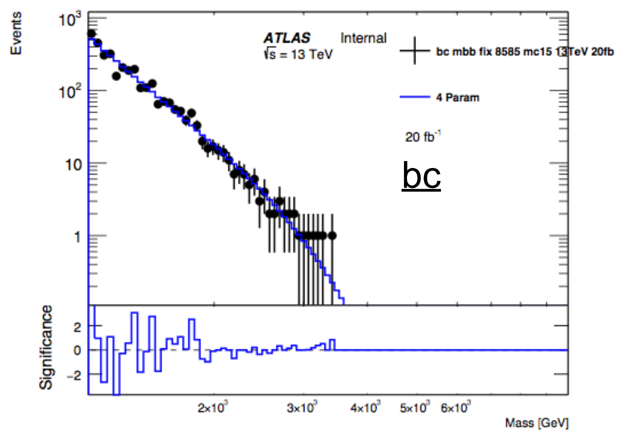
Getting the Flavour Fractions

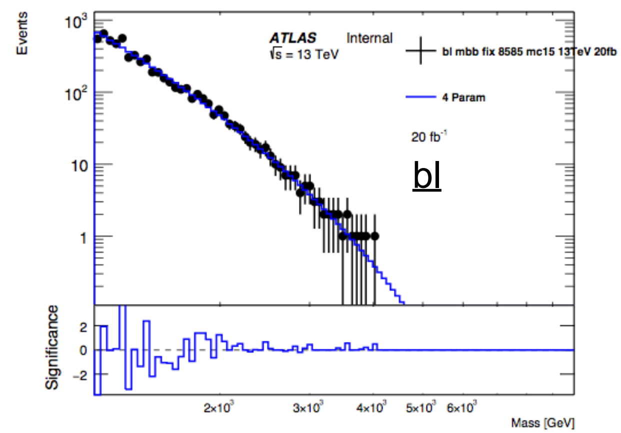
- Flavour fractions are extracted from MC using truth information
- The dijet mass spectrums for these flavour fractions are then scaled to 20fb^{-1}

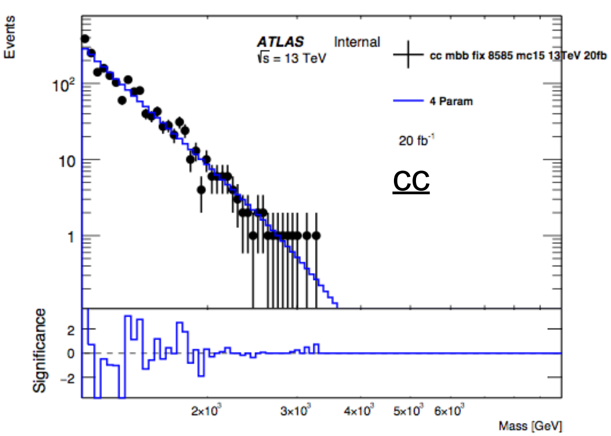
- The dijet mass spectrums are fitted to using the 4-parameter fit function.

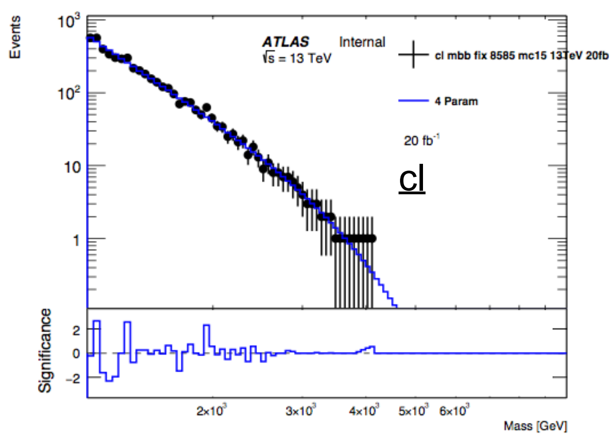
  - *Comment: Why not use 3-parameter fit function*

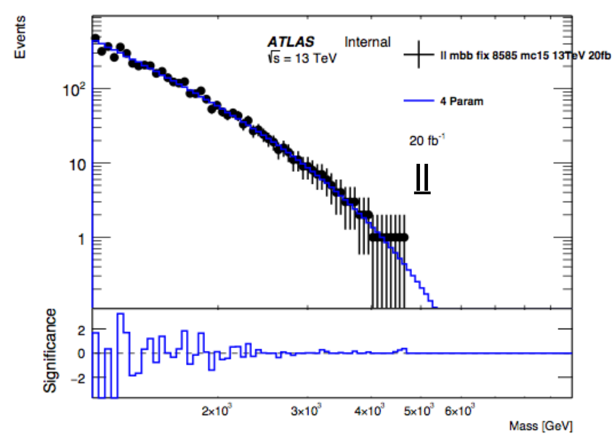








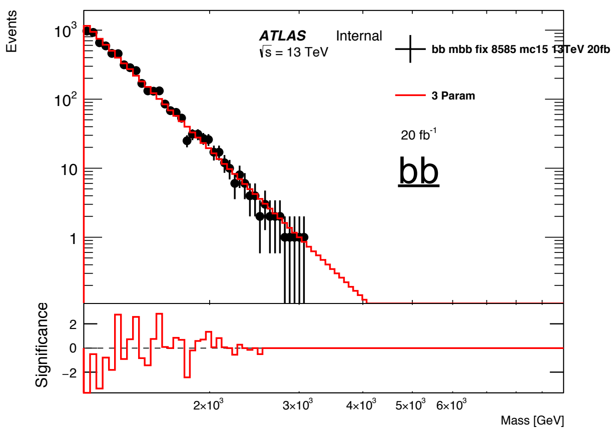


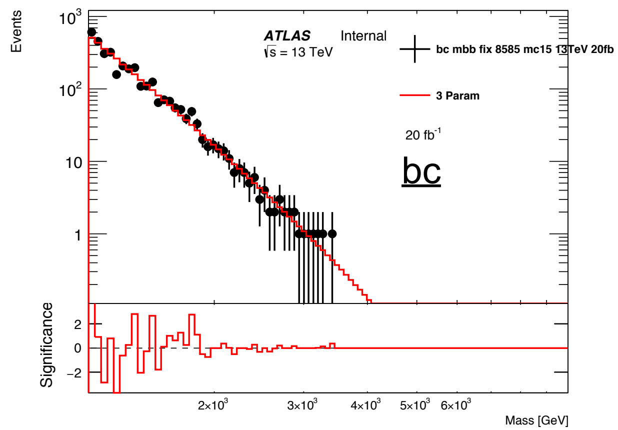


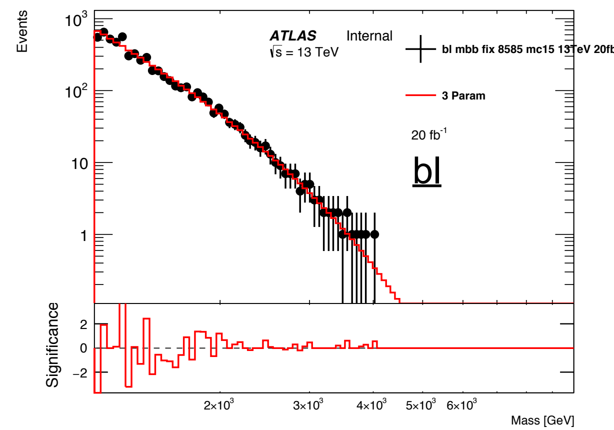
Getting the Flavour Fractions

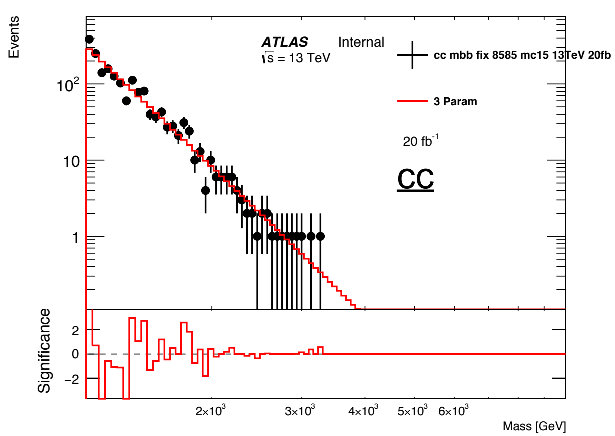
- Flavour fractions are extracted from MC using truth information
- The dijet mass spectrums for these flavour fractions are then scaled to 20fb^{-1}
- The dijet mass spectrums are fitted to using the **3-parameter** fit function.

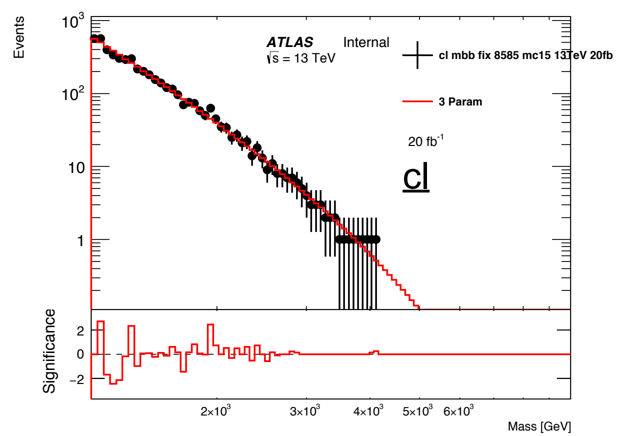
  - *Note: these will be called* *3 Para Init* *and* *4 Para Init* *in this talk* *(init = initialisation)*

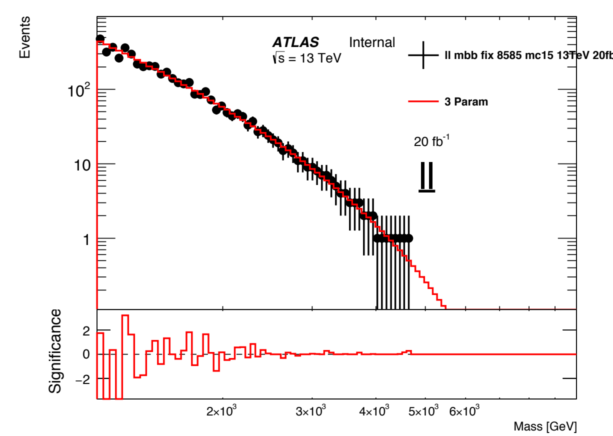






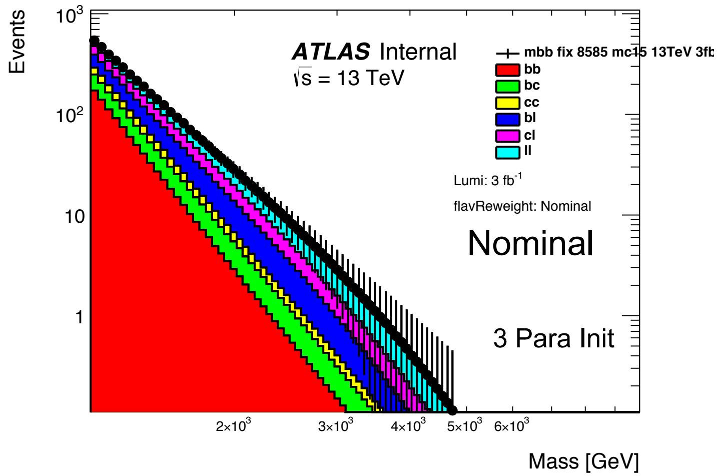


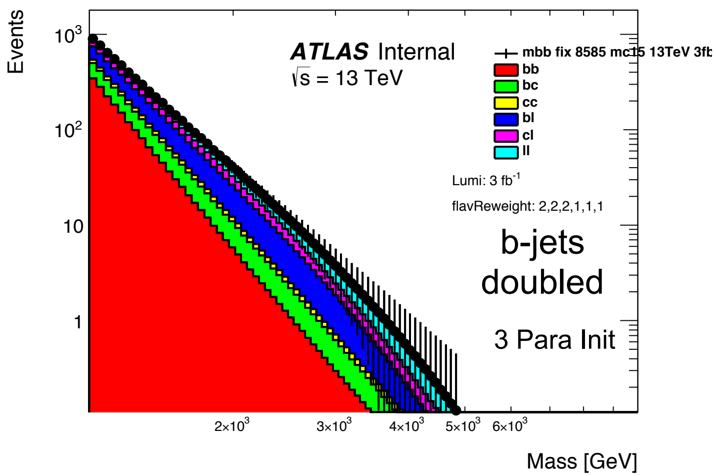


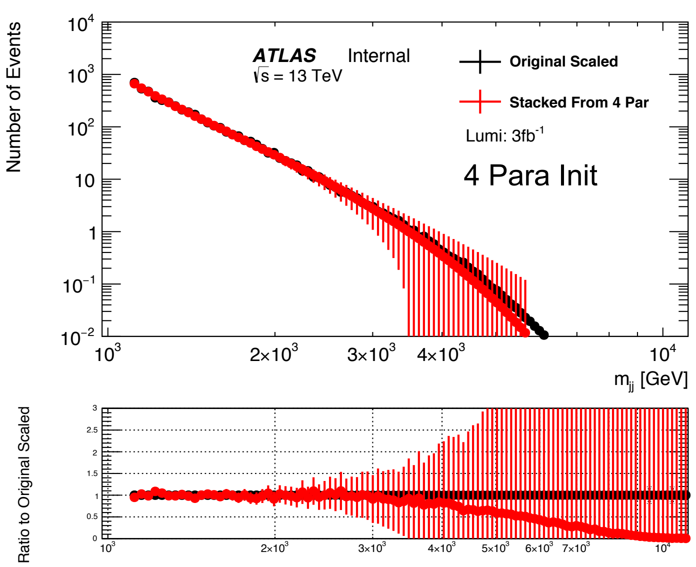


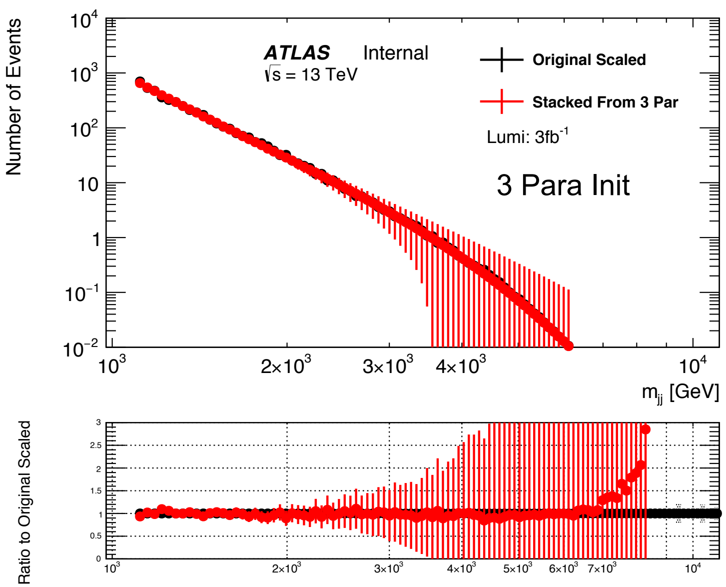
Stacking the Flavour Fractions

- Creates new scaled like distributions.
  => Adding templates from fits to 20 fb^{-1} scaled to 3 fb^{-1}
  => Adding the fractions in different ways to produce various spectra



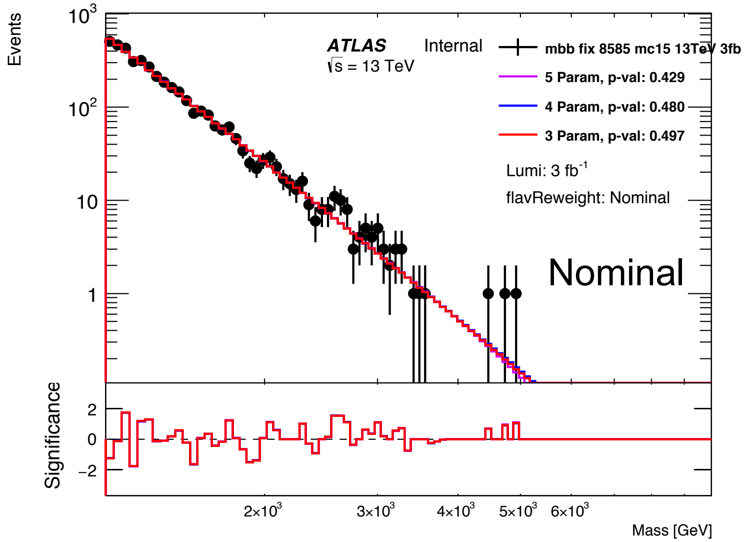


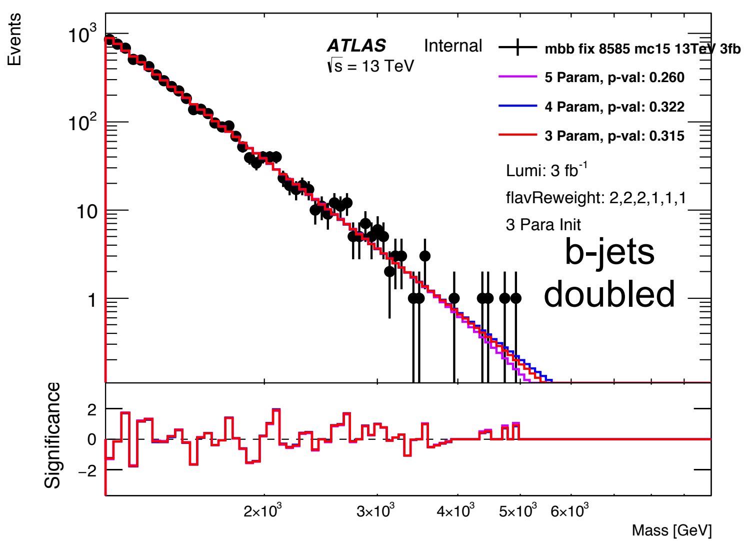




Making Data-Like

- By applying poisson fluctuations we can create 'data-like' distribution
- These are fitted using the 3, 4 and 5 parameter fit function



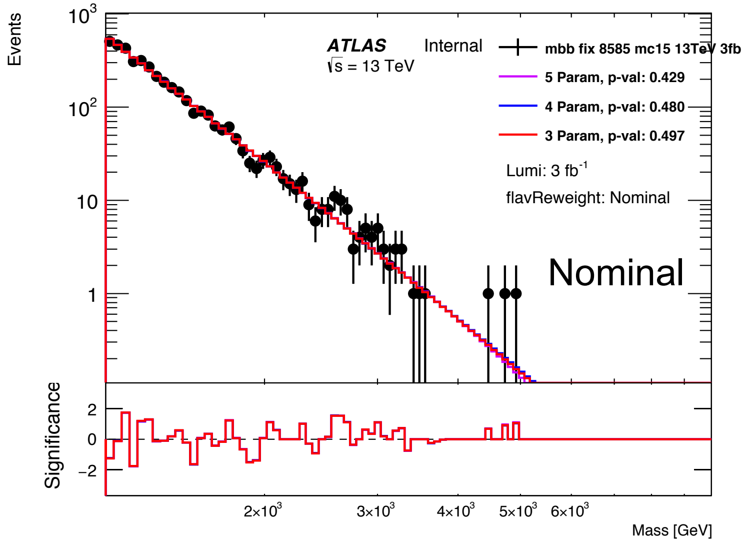


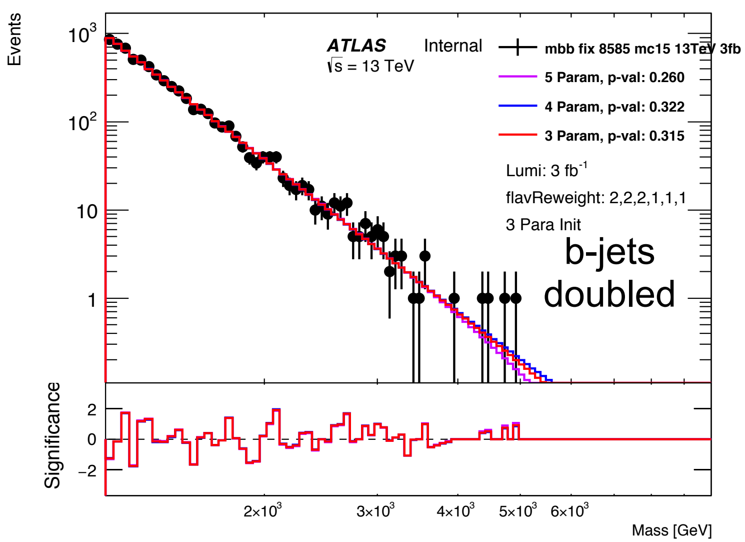
To calculate p-value of a fit:

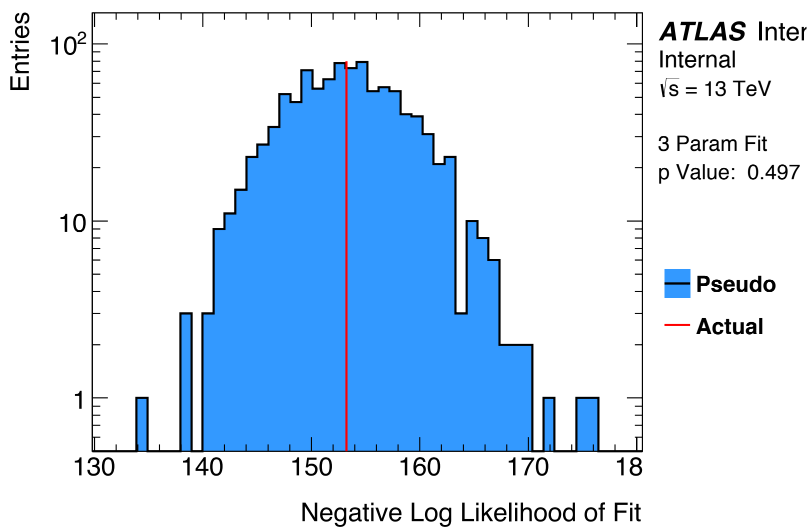
1. Take the fit function and apply poisson fluctuations. (Pseudo-experiment)
2. Re-fit to the pseudo-data using the same fit function.
3. Compare quality of fit to pseudo-experiment to that of the original fit.
   - For a measure quality of fit I use negative log likelihood
4. Repeat 1000 times and count fraction of pseudo-experiments that have a worse quality of fit than the original fit.

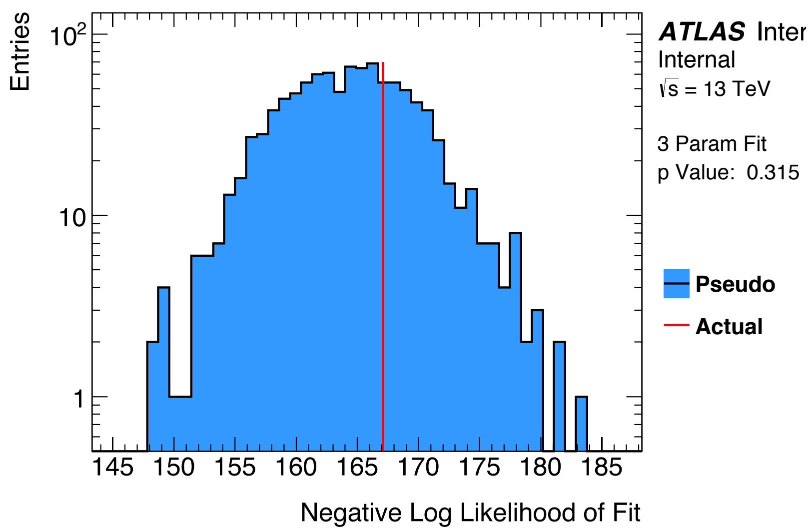
Data-Like p-Values

- By applying poisson fluctuations we can create ‘data-like’ distribution
- These are fitted using the 3, 4 and 5 parameter fit function



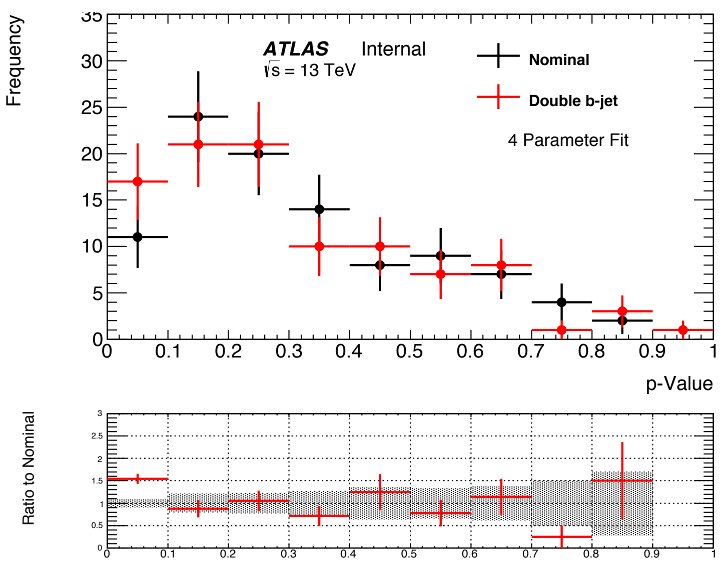






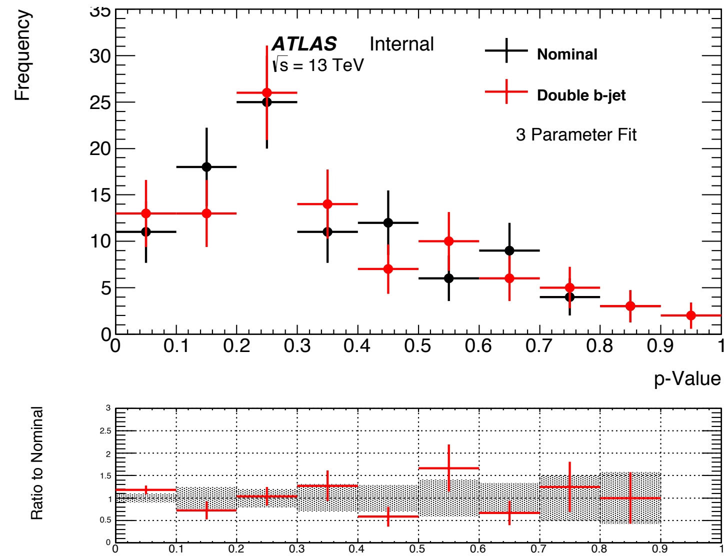
p-Value studies

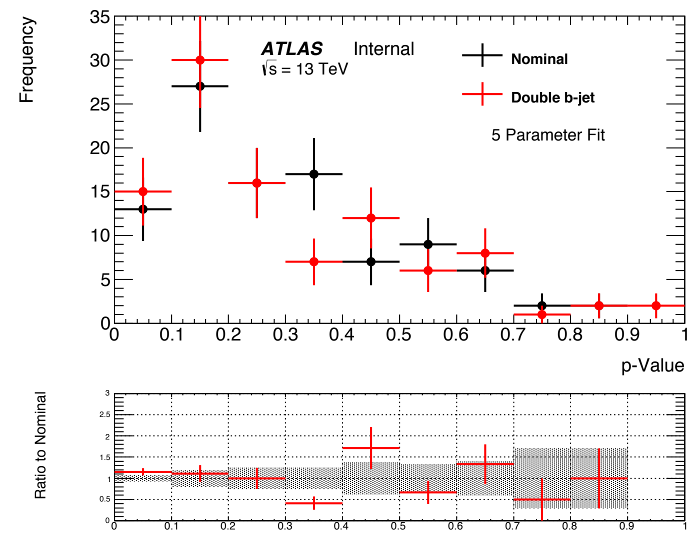
- Different sets of poisson fluctuations means a different ‘data-like’ spectrum
- Each ‘data-like’ dist. can be fitted to, giving a different p-value for each fit variation.
- 100 different data-like distributions have been studied



Mean p-values

	3-Para. Fit	4-Para. Fit	5-Para. Fit
Nominal	0.336 +/- 0.021	0.311 +/- 0.022	0.296 +/- 0.021
b-jet Doubled	0.347 +/- 0.023	0.307 +/- 0.022	0.297 +/- 0.022





Comparison with 4 Para Init.

New: 3 Para used for fractions

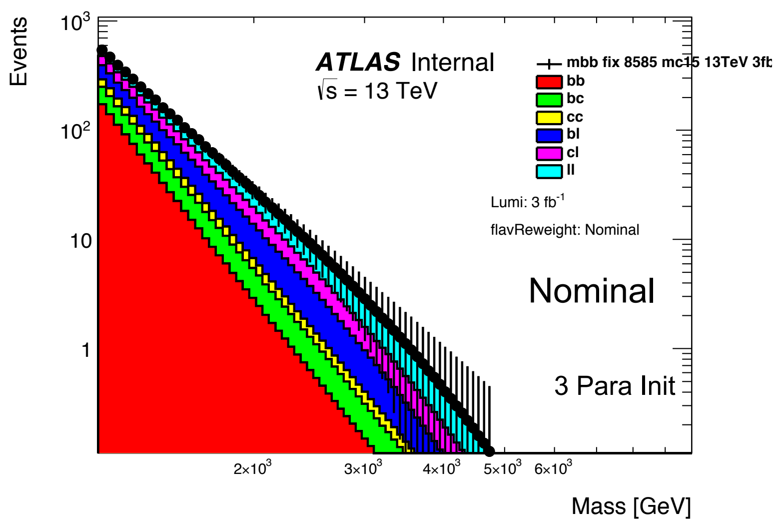
	3-Para. Fit	4-Para. Fit	5-Para. Fit
Nominal	0.336 +/- 0.021	0.311 +/- 0.022	0.296 +/- 0.021
b-jet Doubled	0.347 +/- 0.023	0.307 +/- 0.022	0.297 +/- 0.022

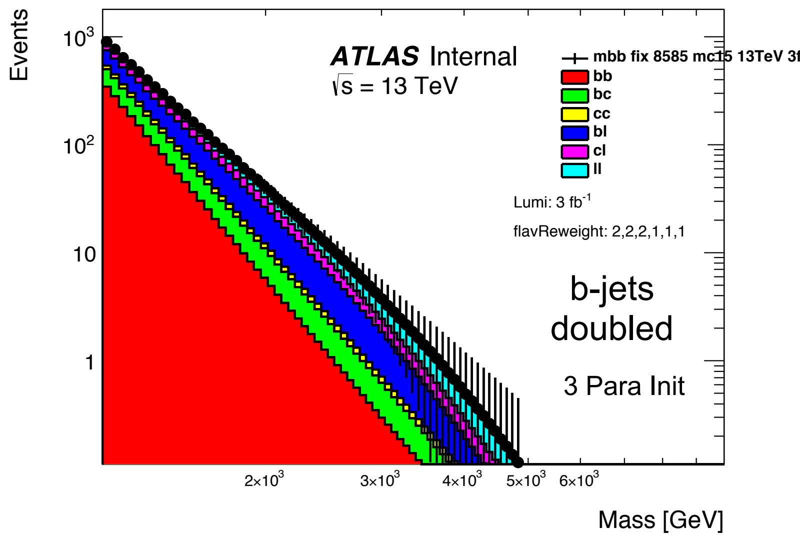
Old: 3 Para used for fractions

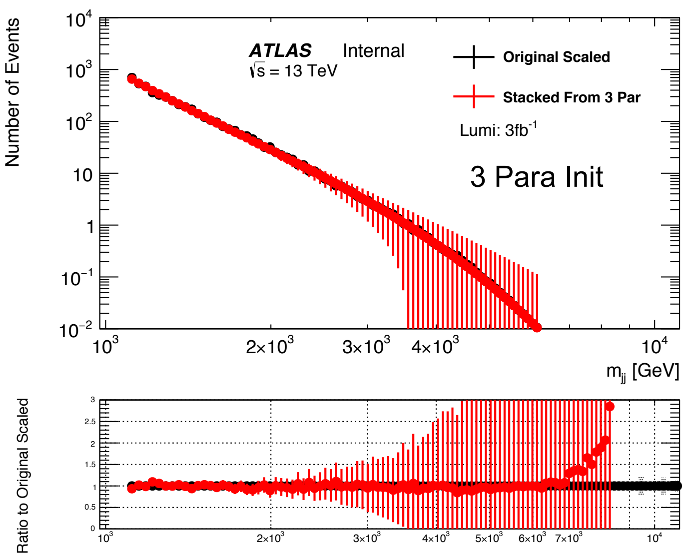
	3-Para. Fit	4-Para. Fit	5-Para. Fit
Nominal	0.325 +/- 0.024	0.280 +/- 0.023	0.283 +/- 0.022
b-jet Doubled	0.308 +/- 0.024	0.267 +/- 0.022	0.276 +/- 0.022

Test for Spurious Signal - Intro

- New for v1.5



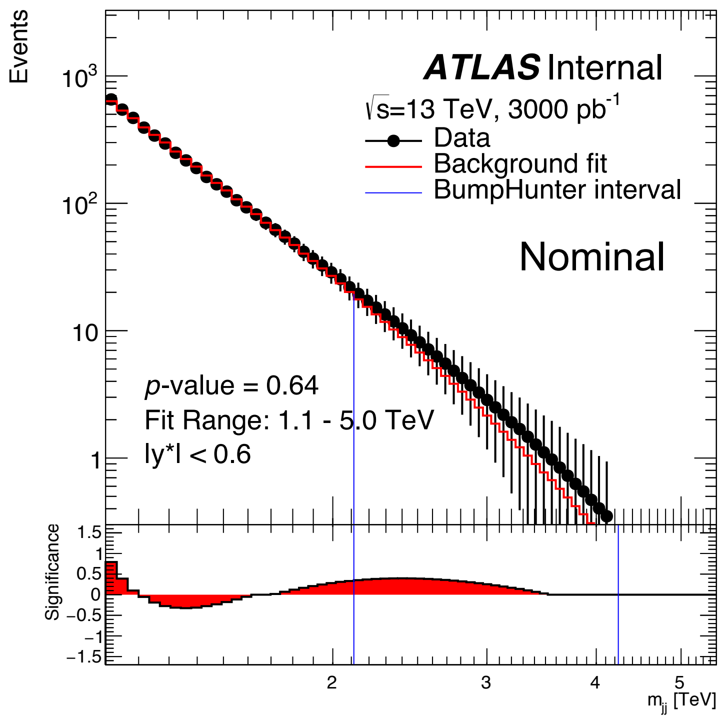


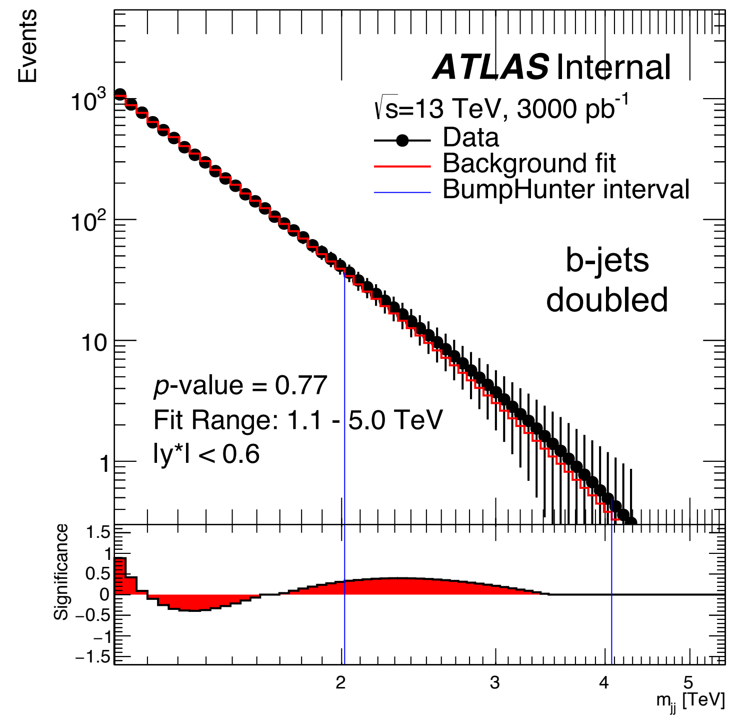


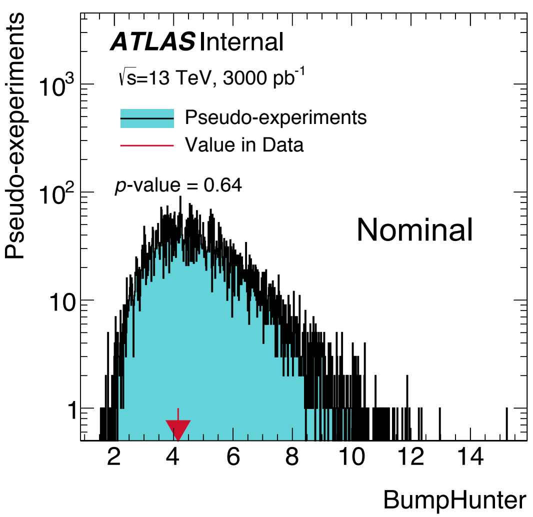
- Test for spurious signal
  => Use scaled spectra before Poisson noise
  => Fit to this spectra using 3 par. function
  => BumpHunter will identify discrepant region
  => BH then can calculate p-value

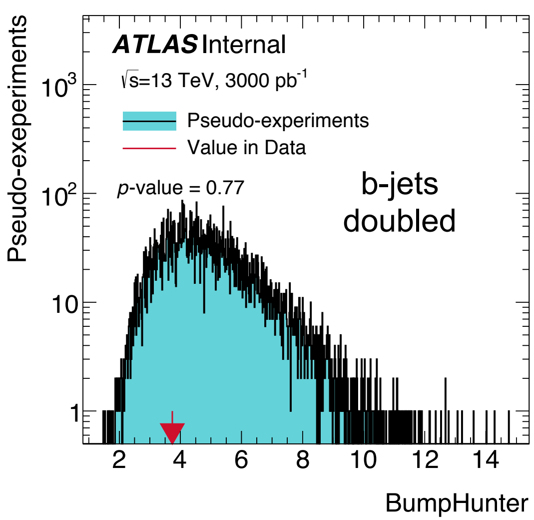
- Mass Range of Fit
  => 1.1 - 5 TeV
  => Larger than mass range in data.

Test for Spurious Signal - Results









- No significant spurious signal found.
- Consistent p-Value in both flav. composition cases
- Wide discrepant region => Unlike benchmark models

Conclusions

- **3 Parameter now used to fit to flavour fractions**
  - Better matches original MC to high masses (well above 5 GeV)
  - No substantial change to p-Values, same conclusions.

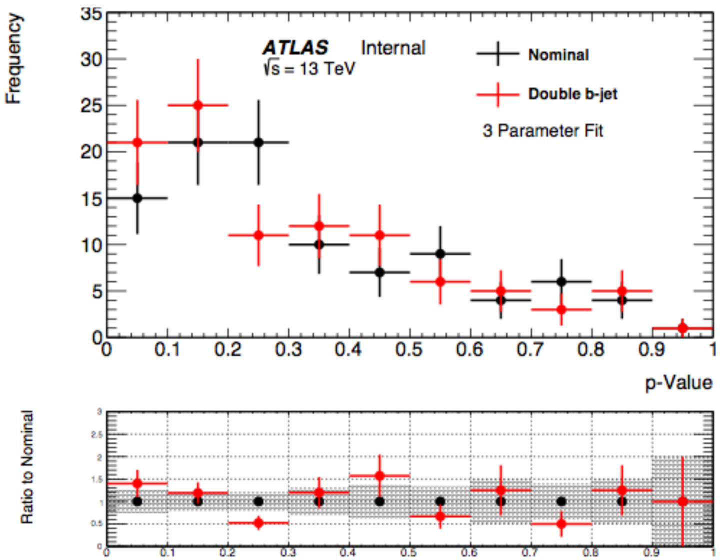
- **p-Value of fitting function**
  - Fitted to 'data-like' distributions.
  - We see no drop in performance (p-value) in the case where b-jet content is doubled.
  - Evidence that fit is robust to flavour fraction.
  - Systematic from fit parameters and fit function choice are enough.

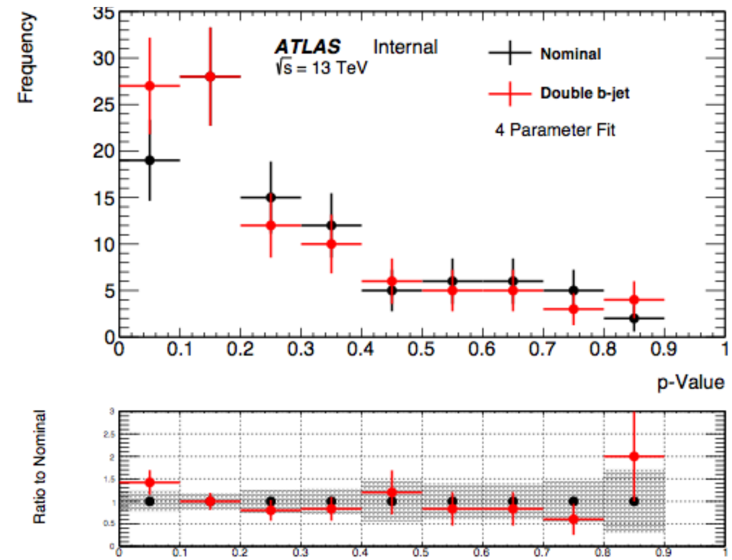
- **Added Spurious Signal Tests**
  - Check scaled dijet mass spectrum for any large deviations.
  - Cutting scaled spectra off at 5 TeV
  - BumpHunter values: Nominal = 0.64, Double b-Jet Content = 0.77
  - No significant spurious signal found.
  - Consistent p-values in both flavour composition cases.
  - Wide signal unlike benchmark models.

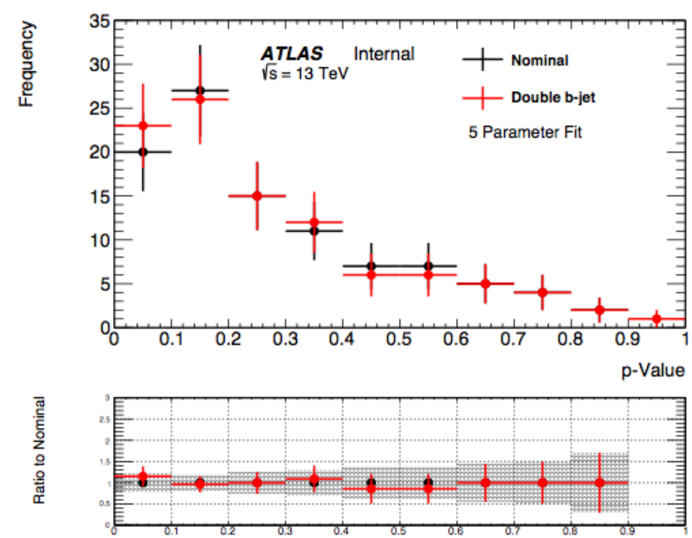
Backup:

p-Values in the Old Case: 4 Para Init.

- Different sets of poisson fluctuations means a different 'data-like' spectrum
- Each 'data-like' dist. can be fitted to, giving a different p-value for each fit variation.
- 100 different data-like distributions have been studied







Mean p-values

	3-Para. Fit	4-Para. Fit	5-Para. Fit
Nominal	0.325 +/- 0.024	0.280 +/- 0.023	0.283 +/- 0.022
b-jet Doubled	0.308 +/- 0.024	0.267 +/- 0.022	0.276 +/- 0.022