Flavour Composition Studies:
Quality of Fit Discussion

Laurie McClymont,
di-b-jet Analysis Team

Di-b-jet Ed Board Discussion

12 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.

- **Response to comments @ Ed Board Meeting 10/02/2016**
  => Currently NLL is used as a measure of QoF in calculations of p-value = Not Valid
  => Our proposal, still fit by minimising NLL (Our fit procedure is the same)
  => Use Chi2 as quality of fit, still calculate p-values using toys.

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

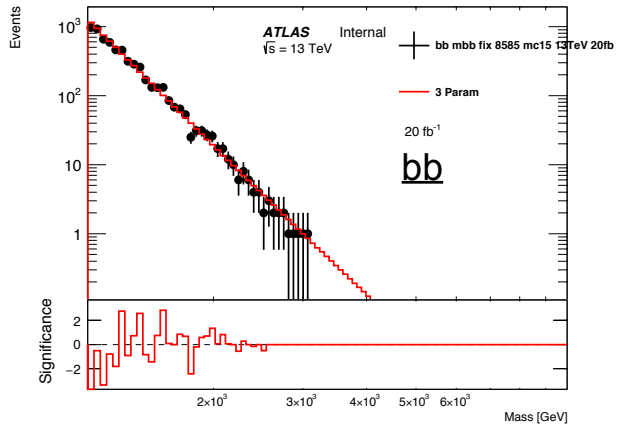
Packages used

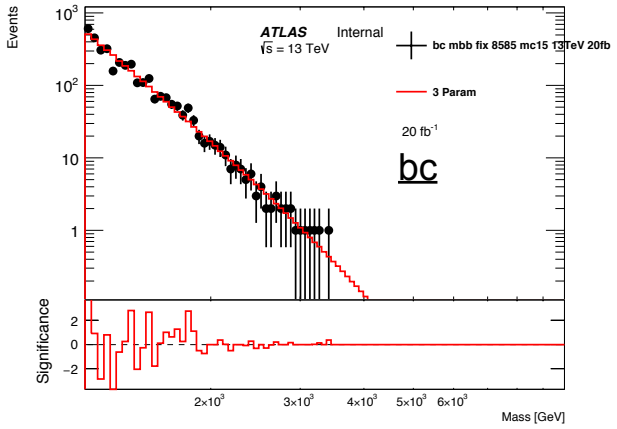
- DijetHelpersPackage:
=> Create scaled distributions.
=> Vary flavour fractions.
=> Create p-values of fit.

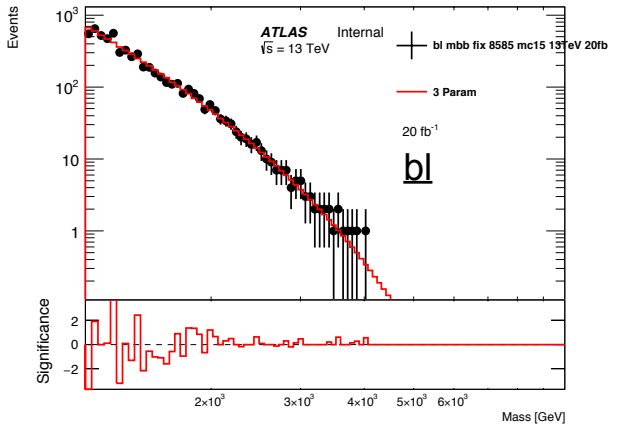
- Dijet Statistical Packages:
=> Using search phase from this package
=> Bumphunter to search for discrepant regions
=> Spurious signal check

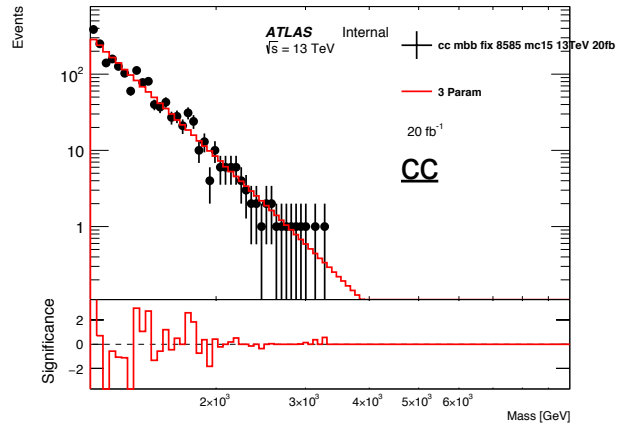
Getting the Flavour Fractions

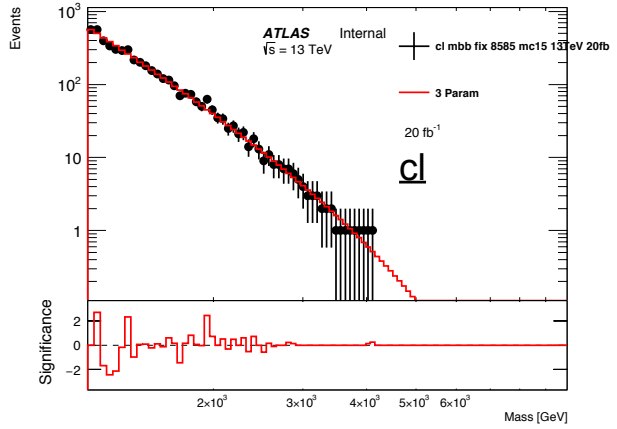
- Flavour fractions are extracted from MC using truth information
- The dijet mass spectrums for these flavour fractions are then scaled to 20fb^{-1}
- The dijet mass spectrums are fitted to using the **3-parameter** fit function.

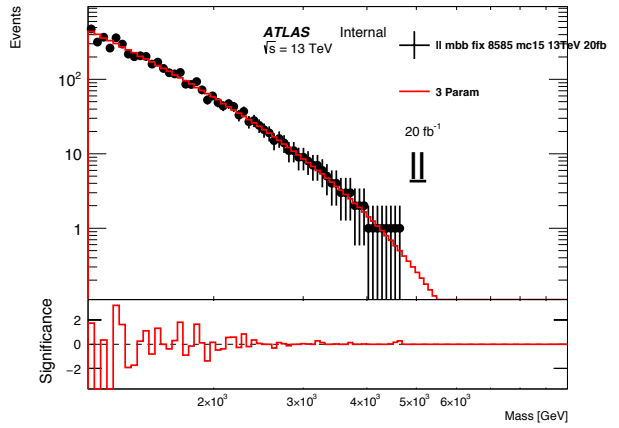






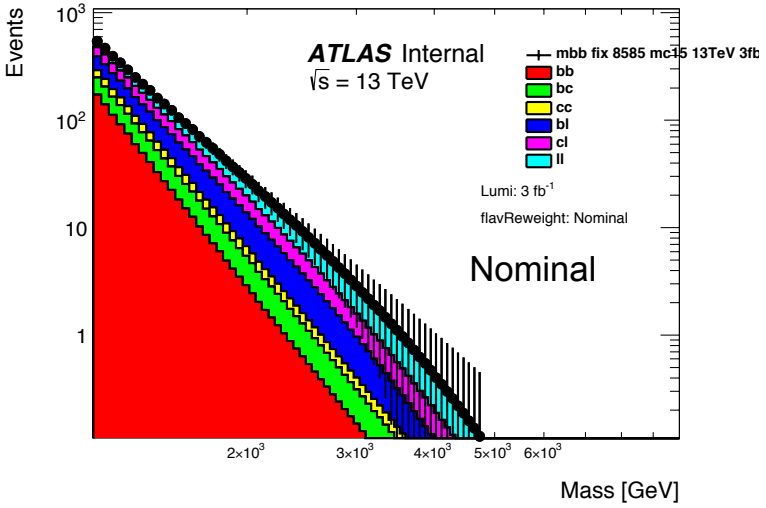


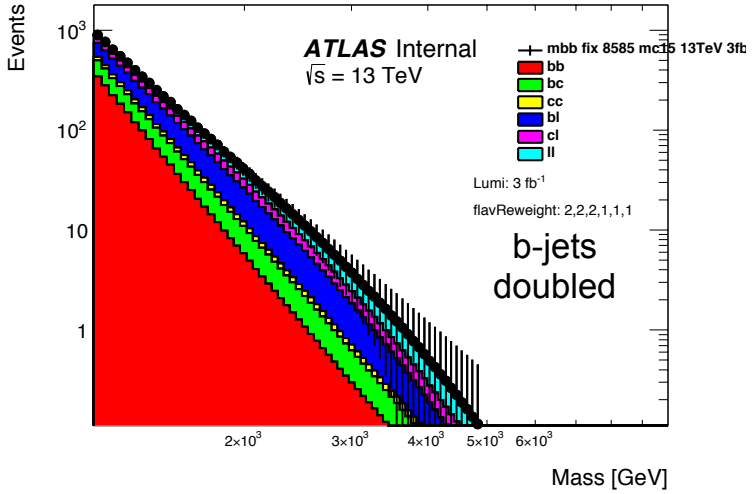


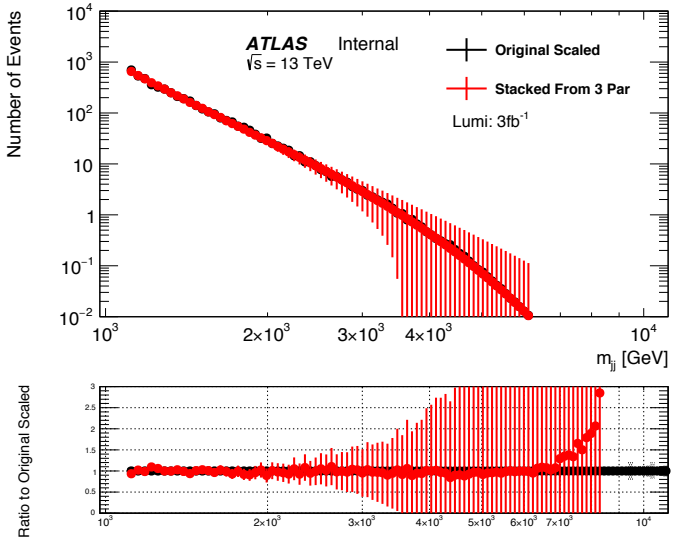


Stacking the Flavour Fractions

- 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



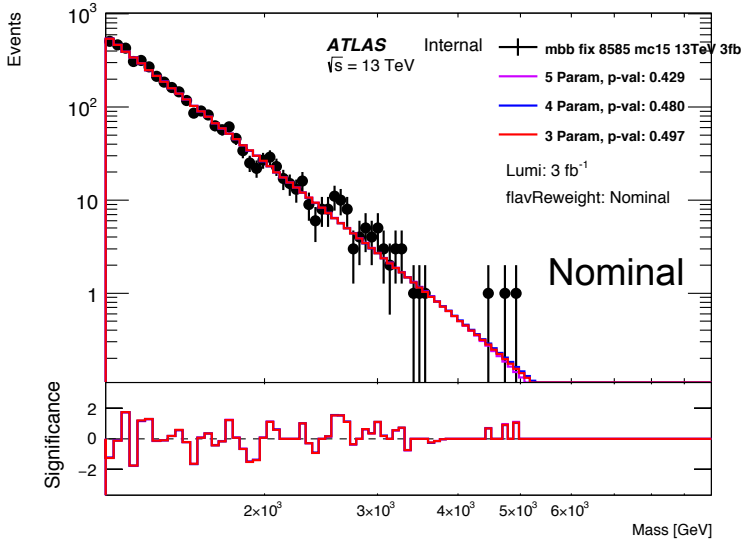


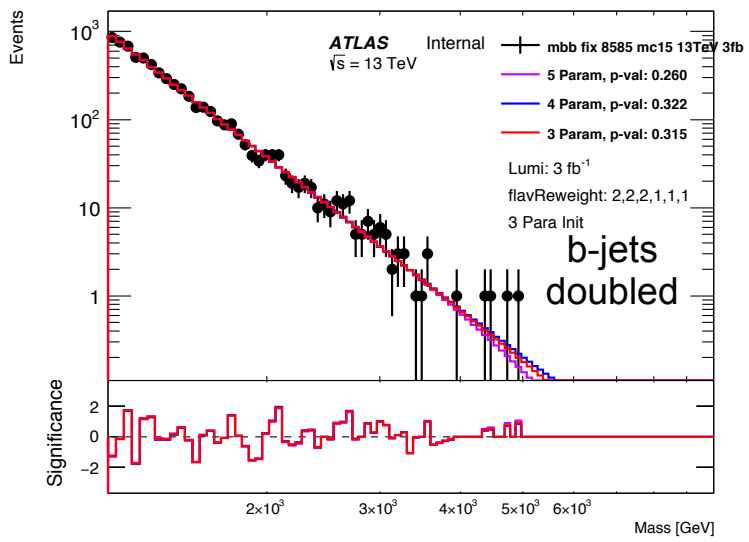


- Able to reproduce MC up to large masses.
- We conclude that the procedure of fitting to flavour fractions then stacking is appropriate.

Making Data-Like - Negative Log Likelihood

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



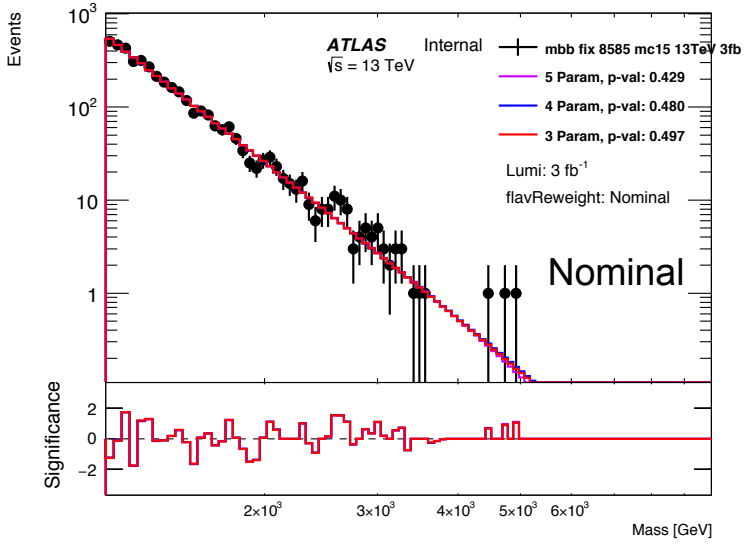


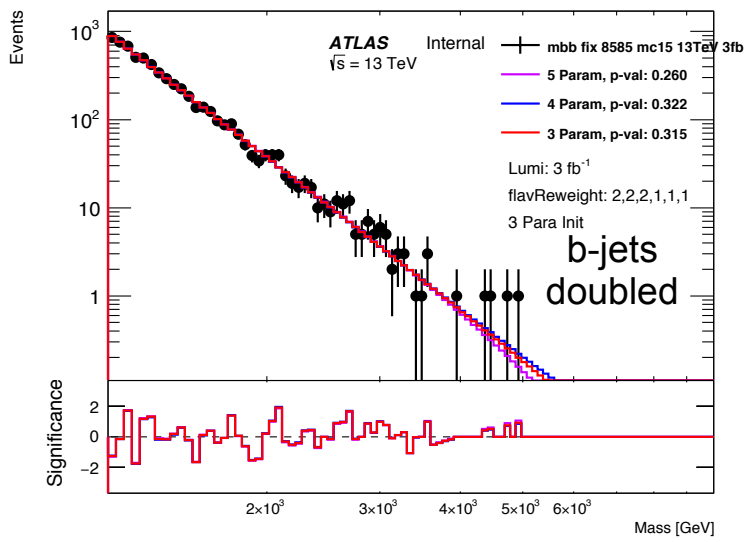
To calculate p-value of a fit:

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 we use negative log likelihood** **(Not Valid)**
4. Repeat 1000 times and count fraction of pseudo-experiments that have a worse quality of fit than the original fit.

Making Data-Like - Now using Chi²

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



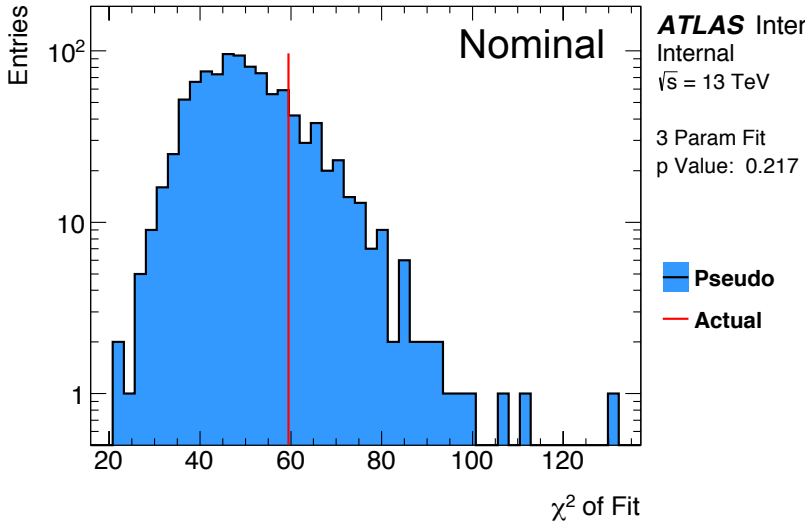


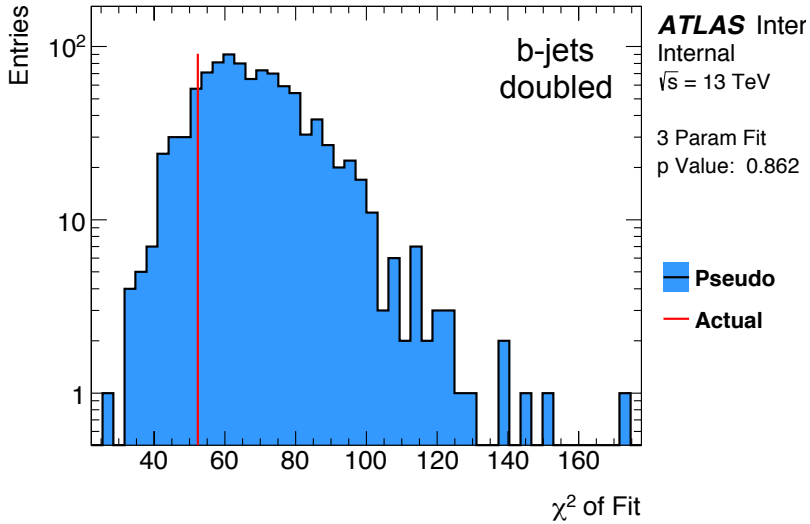
To calculate p-value of a fit:

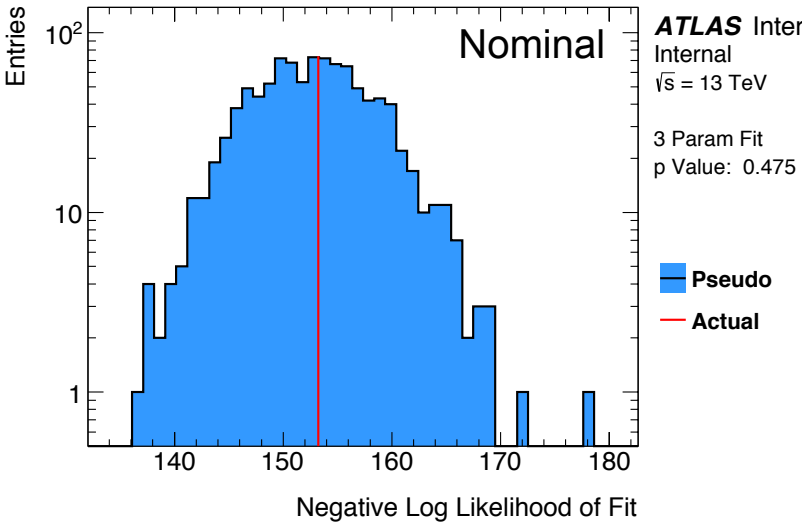
1. Take the fit function and apply poisson fluctuations. (Pseudo-experiment)
2. Re-fit to the pseudo-data using the same fit function.
3. Compare quality of fit to pseudo-experiment to that of the original fit.
   - **For a measure quality of fit we now use Chi²**
4. Repeat 1000 times and count fraction of pseudo-experiments that have a worse quality of fit than the original fit.

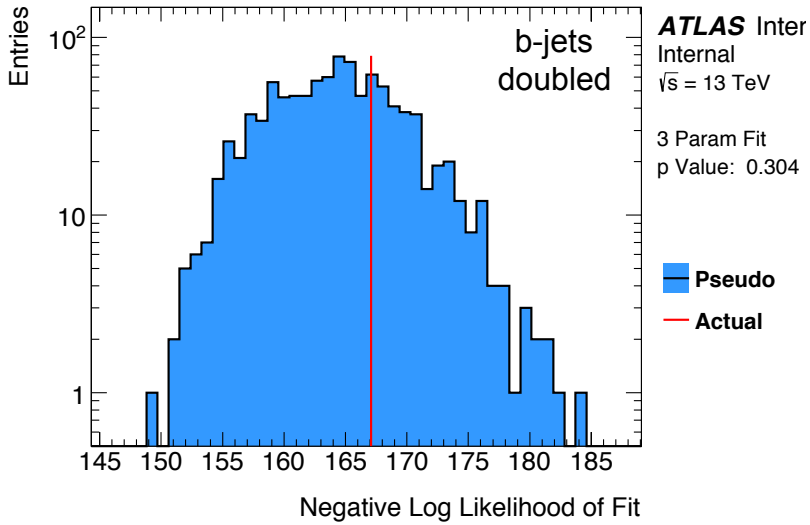
Data-Like p-Values

- By applying poisson fluctuations we can create ‘data-like’ distribution
- Fitted to using the 3 parameter by minimising NLL, this fixes parameters.
- We then use 1000 toys and Chi^2 as QoF test.



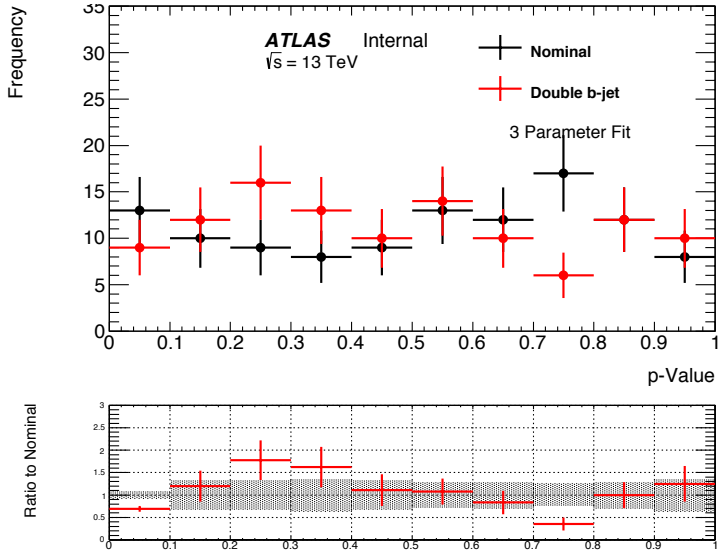


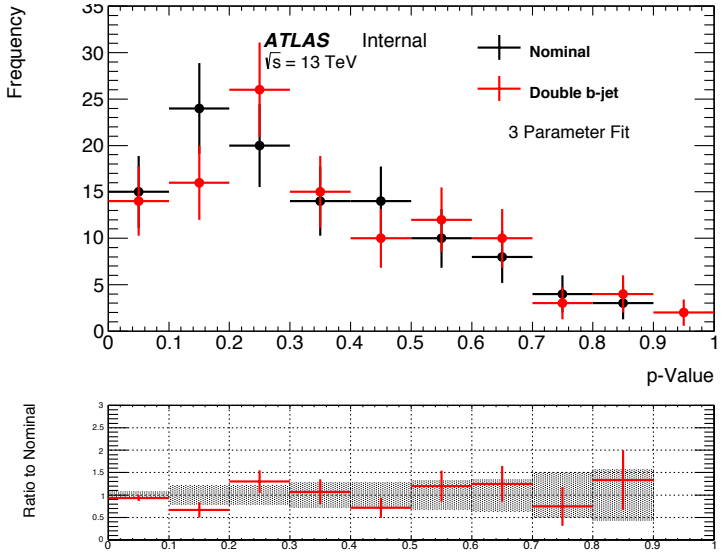




p-Value studies

- 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





Chi2

	3-Para. Fit
Nominal	0.512 +/- 0.027
b-jet Doubled	0.471 +/- 0.026

NLL

	3-Para. Fit
Nominal	0.332 +/- 0.021
b-jet Doubled	0.350 +/- 0.023

Conclusions

- **Changes made**
  - NLL is not a good measure of QoF, so no longer used to calculate p-values.
  - Minimising NLL can be used to choose and fix fitting parameters.
  - We now use Chi^2 as a measure of QoF.
  - p-Values can be calculated by throwing toys and comparing Chi^2 values.

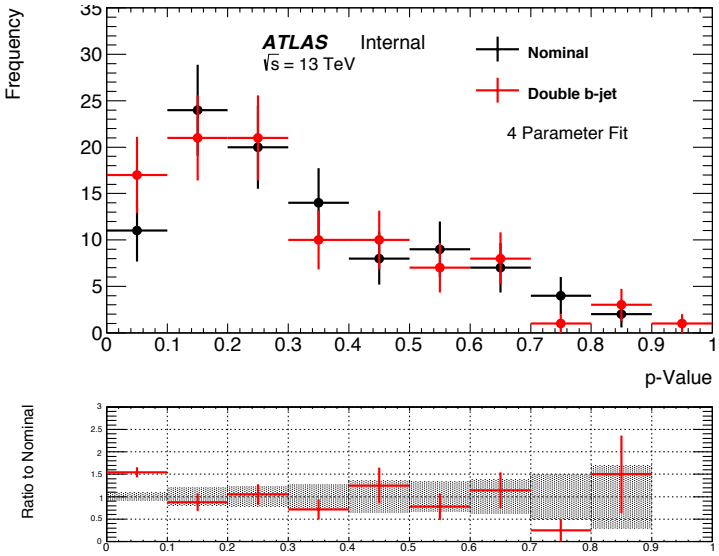
- **Results of preliminary studies**
  - We see that the Chi^2 produces flat distributions of p-values (slide 8)
  - There is no discrepant mean p-values between Nominal and b-Jet Content doubled
  - Conclusions of study should not be changed

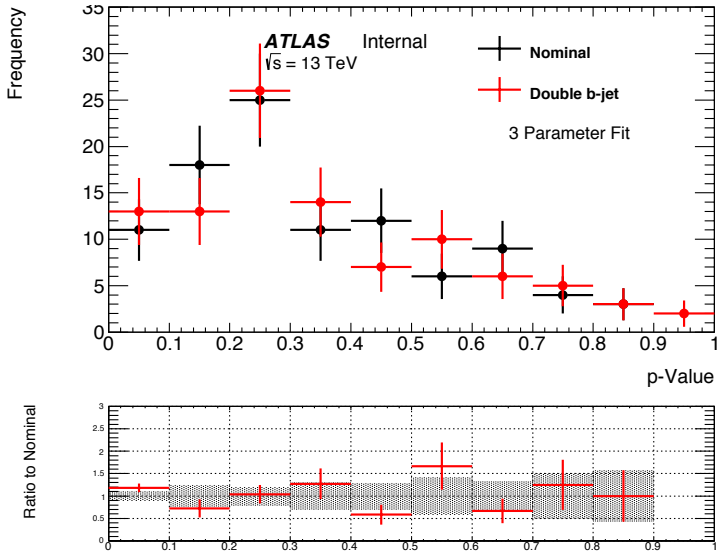
- **Proposals moving forward**
  - We want to increase the number of ‘data-like’ distributions studied.
  - 100 -> 200+
  - This will increase statistics for plots slide 8, and make plot clearer.
  - This should be turned around soon (certainly by Weds for our next meeting)

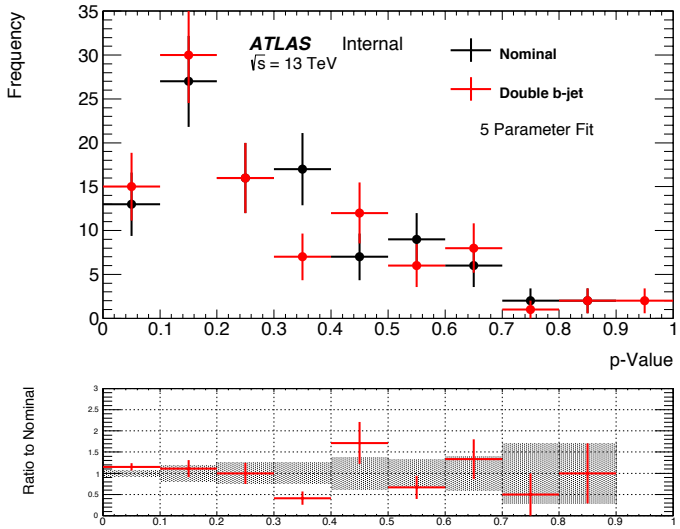
Backup:

Previous 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





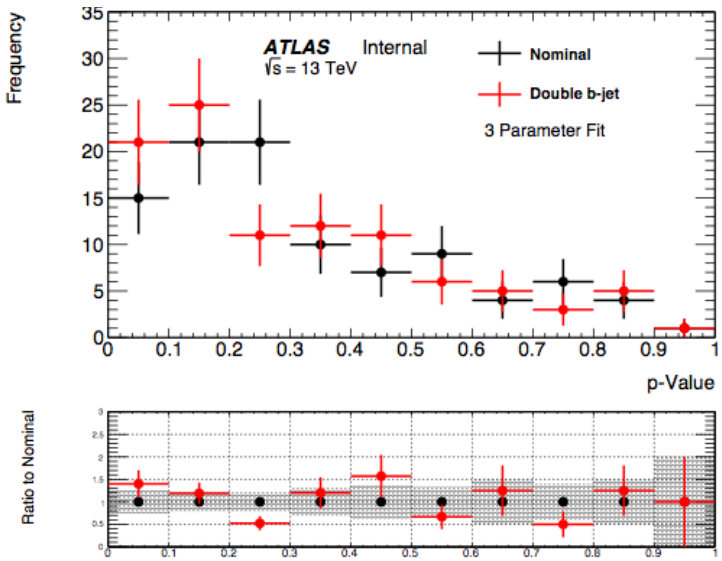


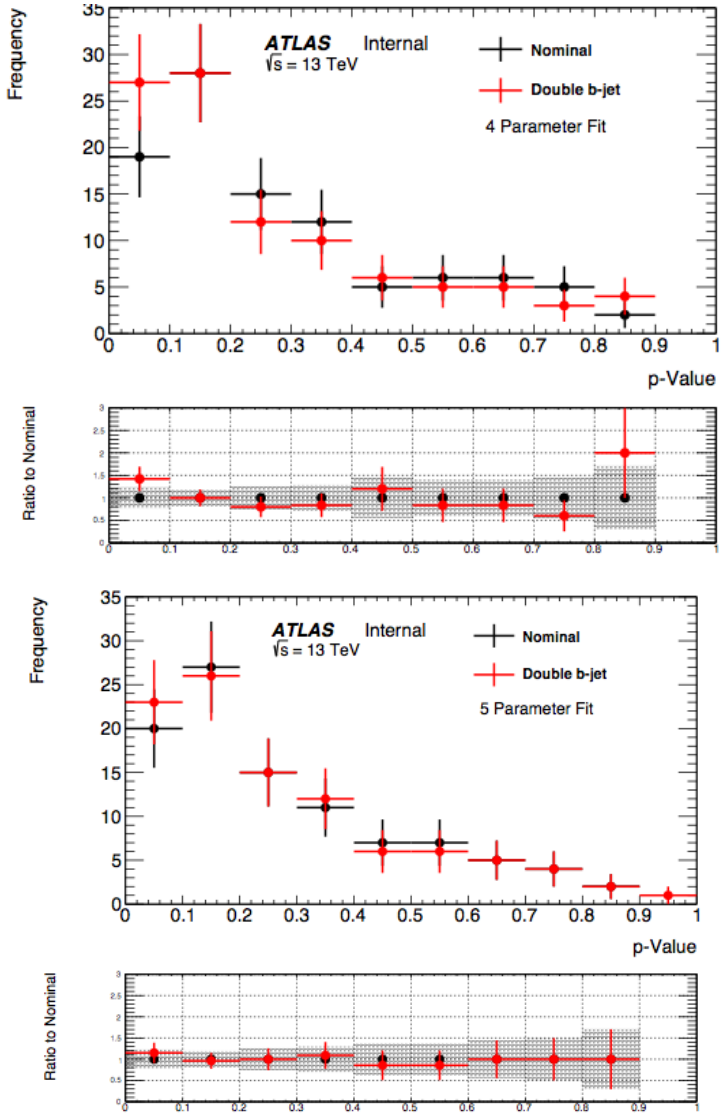
Mean p-values

	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

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