



Flavour Composition Studies

Laurie McClymont, Andreas Korn

13 January 2016



- 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.
- Comment on CDS
 - => A lot of p-values are 1, is this plausible
 - => This is what I want to address

Figure 17: Same comment about the fit legend text—why do they all have a p-value of 1? Also, does doubling the b-content just happen to improve the fits such that each of the 3,4,5 parameter fits are identical or are the higher order contributions just close to 0 for these fits?

Details

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

Standard Dijet Resonance Cuts

- Leading Jet $p_T > 410$ GeV
- Sublead Jet $p_T > 50$ GeV
- $|y^*| < 0.6$
- $m_{jj} > 1100$ GeV

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

Use DijetHelpersPackage:

Create Histograms and merge slices

- *makeStandardHistograms.py*
- *plotStandardPlots.py*
- Fit 20ifb smooth histos using *singleFit.py*
- Create data-like using from these fits

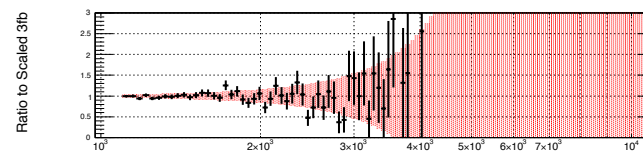
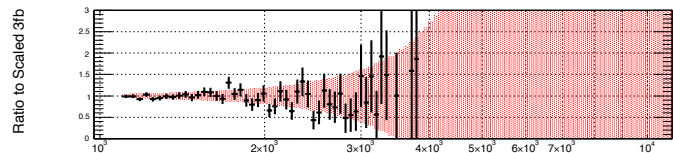
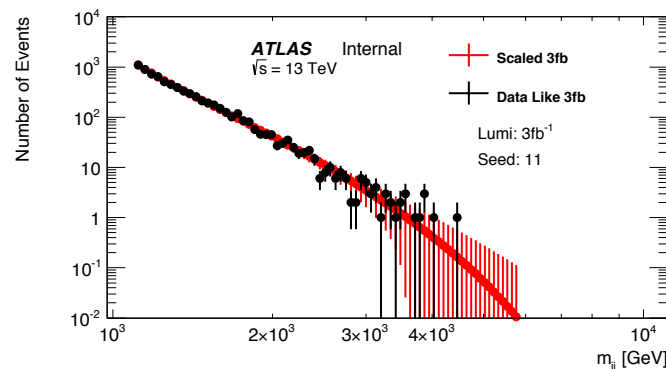
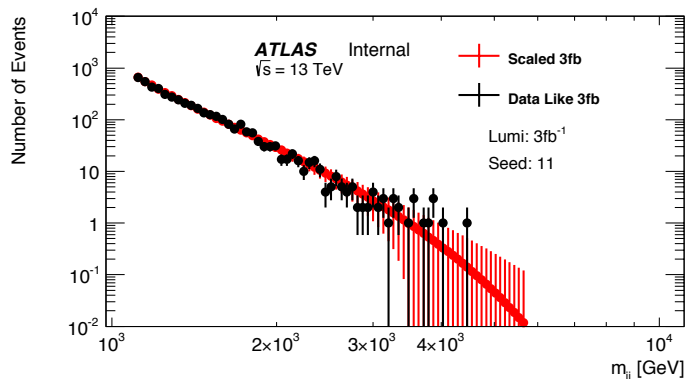
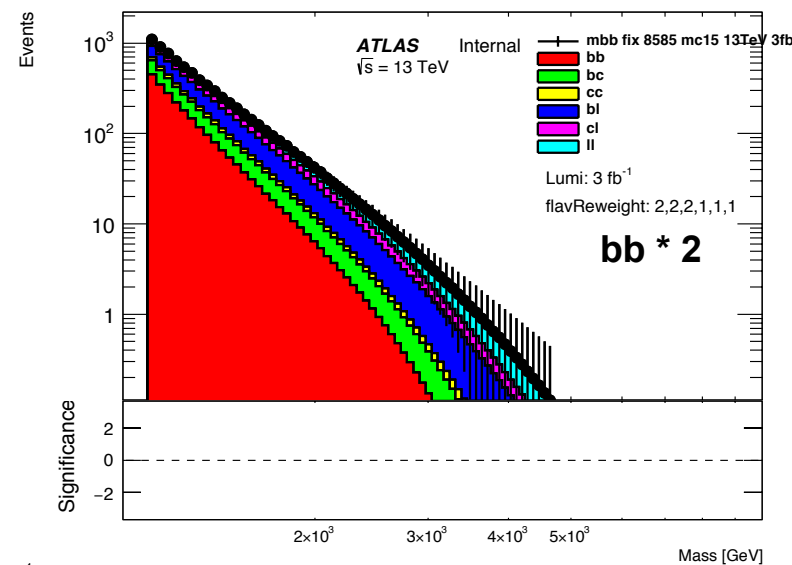
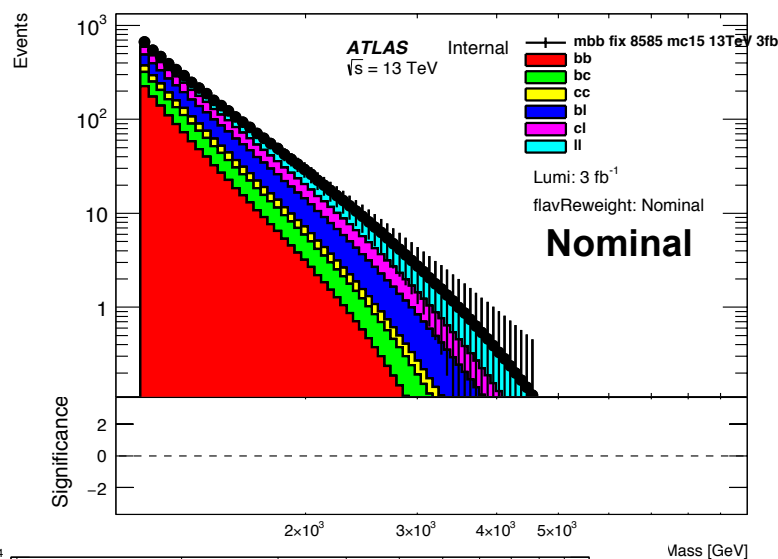
Fit spectrums and make some plots

- *singleFit.py*
- *plotSingleFit.py*



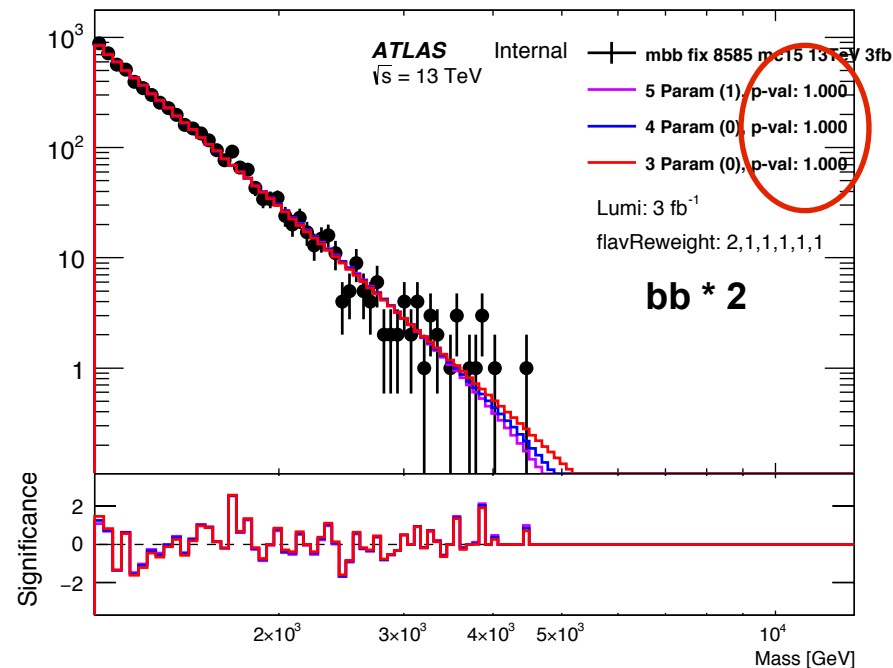
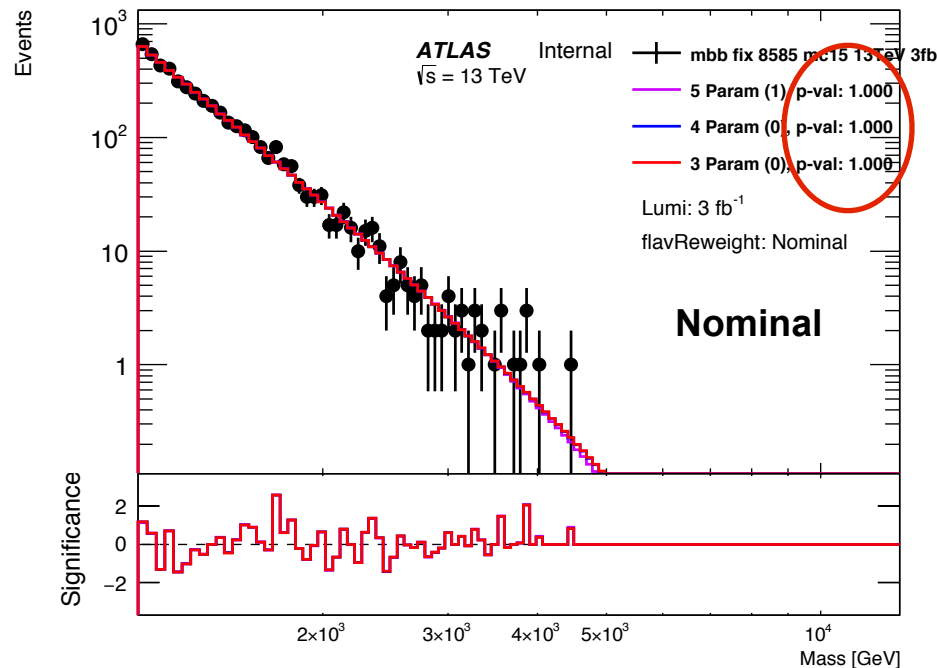
3 Combining Flavour Composition

- Create flavour fractions
 - => Creating histograms from fits to 20fb scaled to 3fb.
 - => Adding the fractions in different ways to produce various spectra.
- This creates new scaled distributions from which we can create data-like MC.





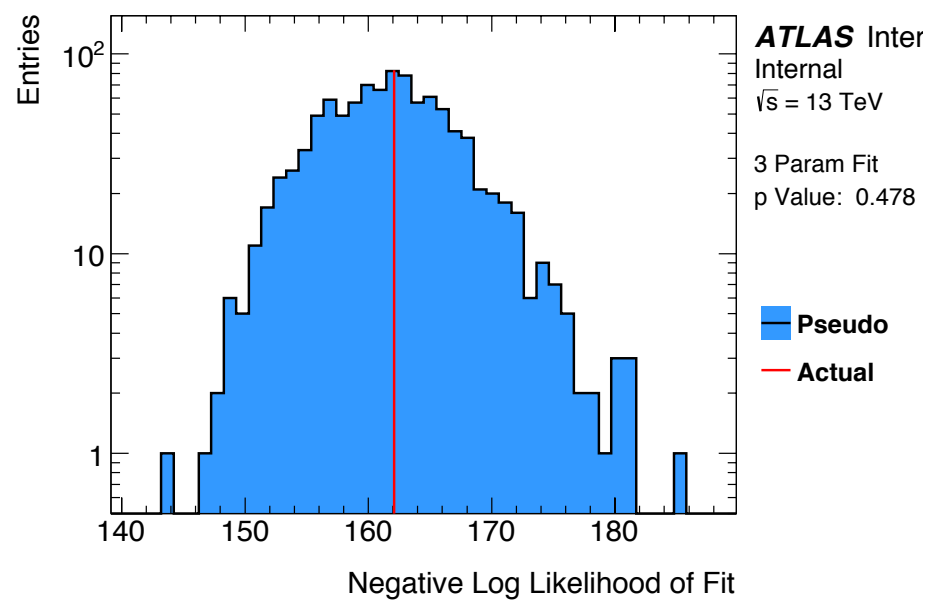
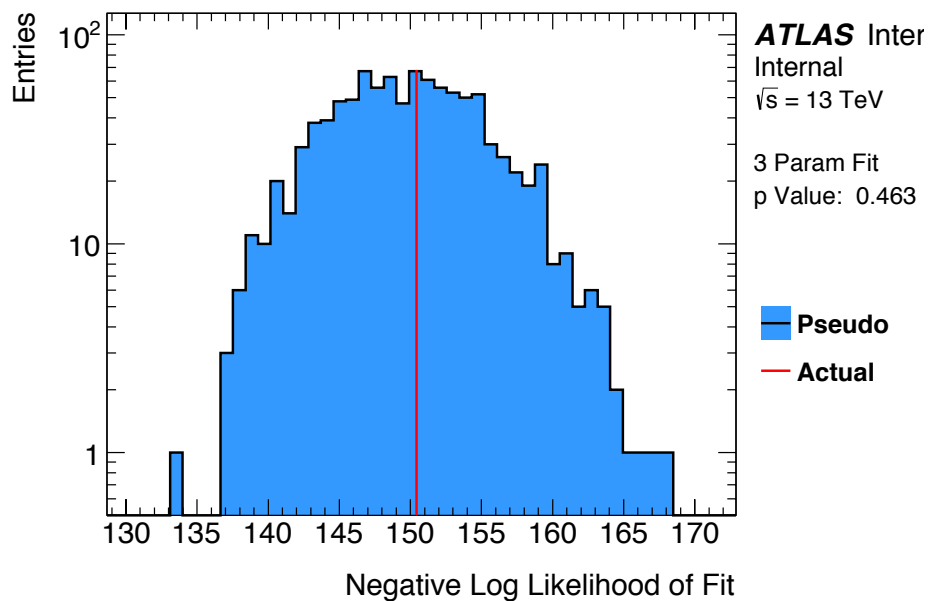
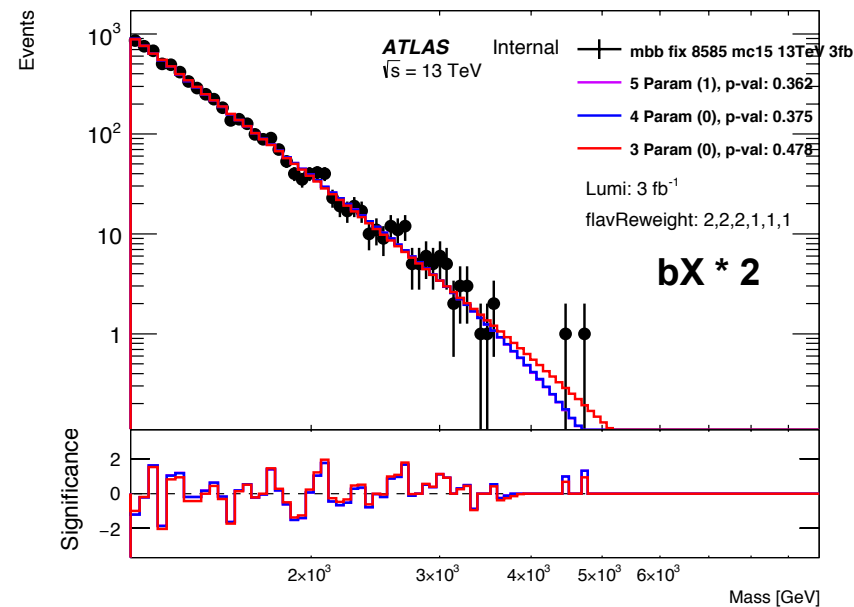
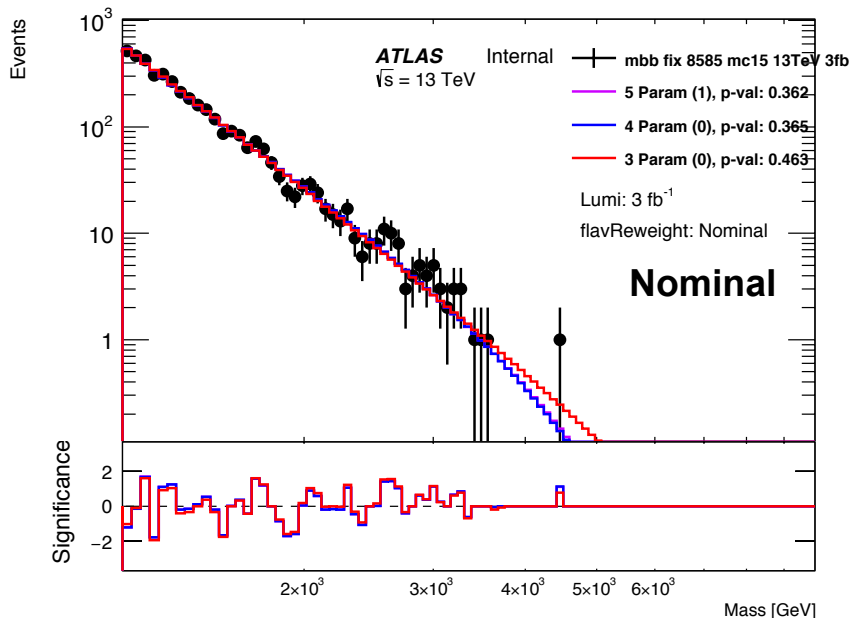
- We then fit to data-like distributions



- Very high p-value ($p \sim 1.000$, really??)
 - => These values taken from Chi^2/Ndof
 - => Question on this on cds
- Try using toys to create p-Values (Advised by Jeff)
 - => Produce pseudo-data distributions based on the fit function parameters.
 - => Fit to these distributions.
 - => Ask number of times that fit to the pseudo is better than the original fit
 - => I use 1000 toys here.



5 Fitting Procedure with toys for p-value





- Each seed represents a different initial fluctuation, a different data-like distribution
=> For each seed a p-value can be created
- P-values are quite low, even in nominal case.
=> However no notice-able drop between nominal and bX^2

Seed	Nominal			bX*2		
	3	4	5	3	4	5
11	0.182	0.170	0.150	0.187	0.165	0.137
12	0.463	0.365	0.362	0.478	0.375	0.362
13	0.459	0.524	0.546	0.258	0.340	0.365
14	0.687	0.633	0.635	0.623	0.524	0.497
15	0.007	0.001	0.007	0.053		0.033
16	0.071	0.061	0.066	0.428		0.404
17	0.194	0.170	0.190	0.268	0.297	0.324
18	0.287	0.197	0.214	0.182		0.117
19	0.344		0.248	0.153	0.108	0.107
20	0.065	0.059	0.105	0.095	0.089	0.123
21	0.440	0.563	0.556	0.436	0.647	0.638
22	0.203	0.125	0.136	0.361	0.298	0.297
23	0.237	0.155	0.174	0.161	0.119	0.105
24	0.120	0.043	0.069	0.242	0.145	0.142
25	0.130	0.106	0.124	0.013	0.013	0.016
26	0.232	0.161	0.194	0.203	0.160	0.163
27	0.582	0.519	0.541	0.413	0.349	0.361
28	0.010	0.008	0.010	0.053	0.036	0.017
29	0.389	0.361	0.339	0.093	0.092	0.098
30	0.092	0.061	0.068	0.025	0.021	0.012
Average	0.260	0.225	0.237	0.236	0.222	0.216



Conclusions

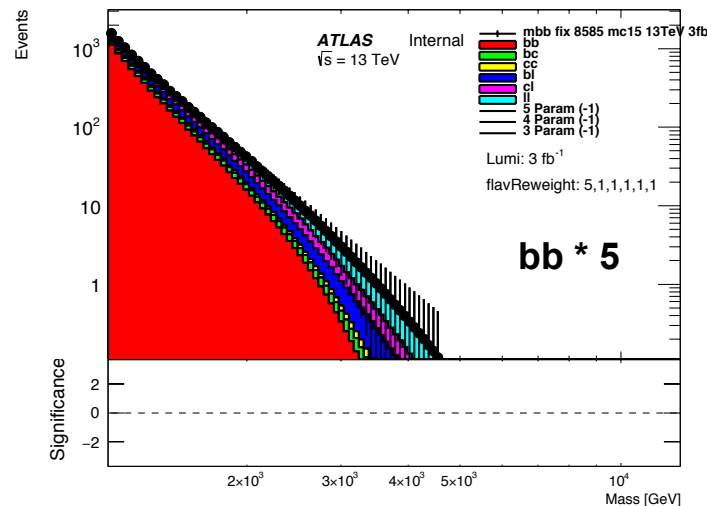
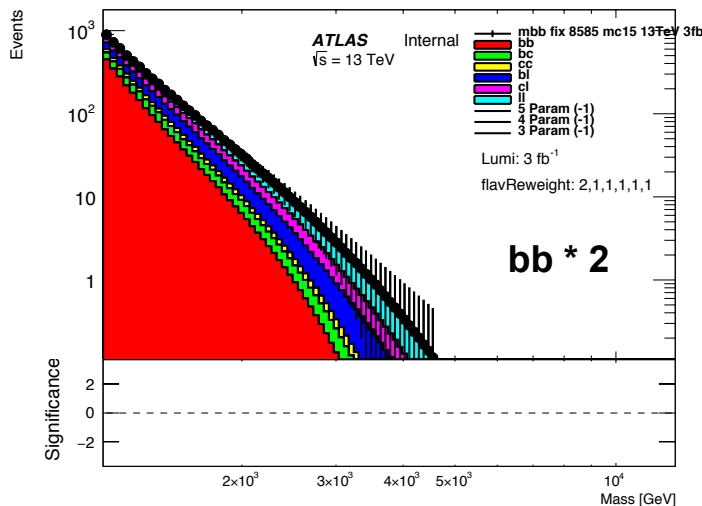
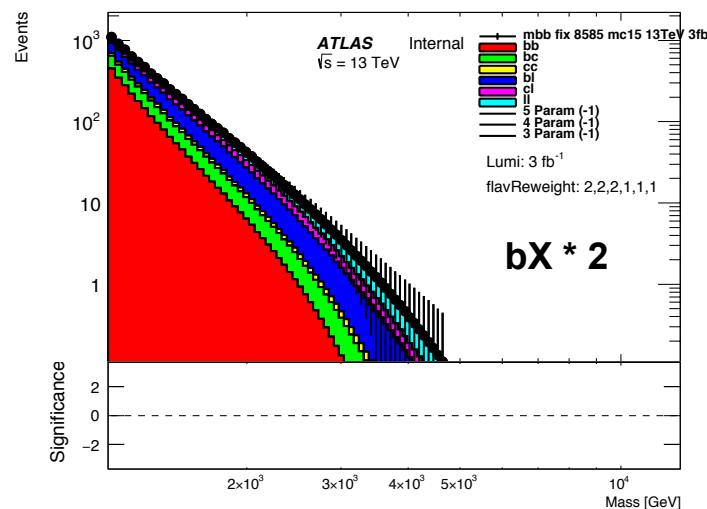
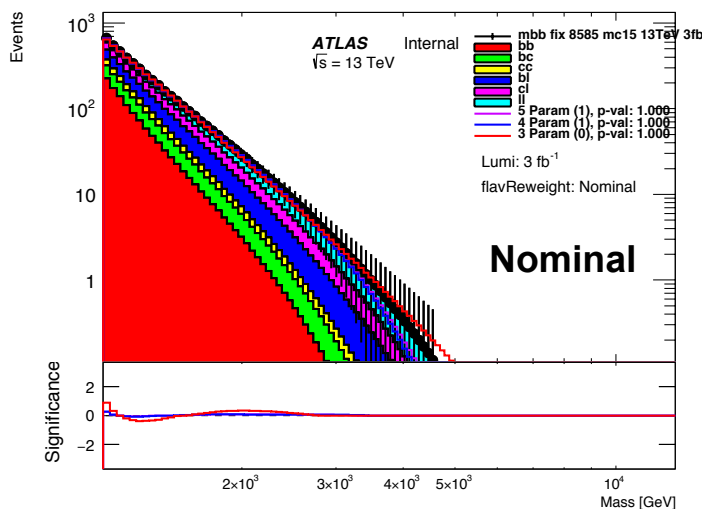
- Creating p-values by using pseudo-experiments.
- Low p-values when using toys.
=> Really want to try and understand why...
- What is the best way of displaying results
=> Choose one typical example
=> Show a grid of results
=> An average of several p-values

-

Need to look at errors, here they are poisson. => Maybe I should fit scaled that is truncated at 0.5

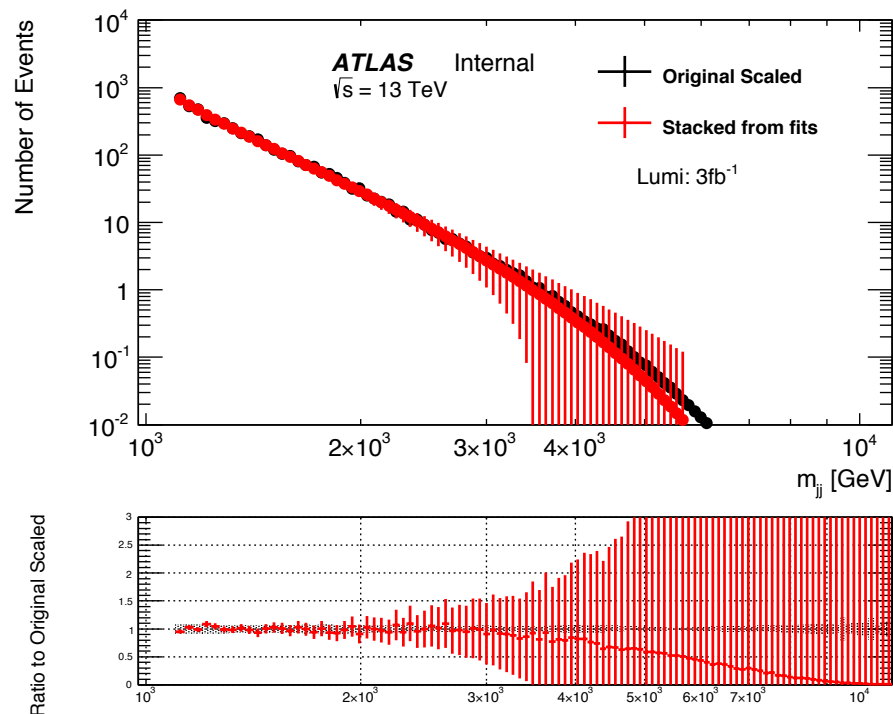


- Then combine histograms by,
 - => Creating histograms from fits scaled to 3fb.
 - => Adding the fractions in different ways to produce various spectra.
- This creates new scaled distributions from which we can create data-like MC.





- Compare nominal produced from stacking to original MC sample
=> To validate that our procedure for producing distributions is valid.



Smooth means

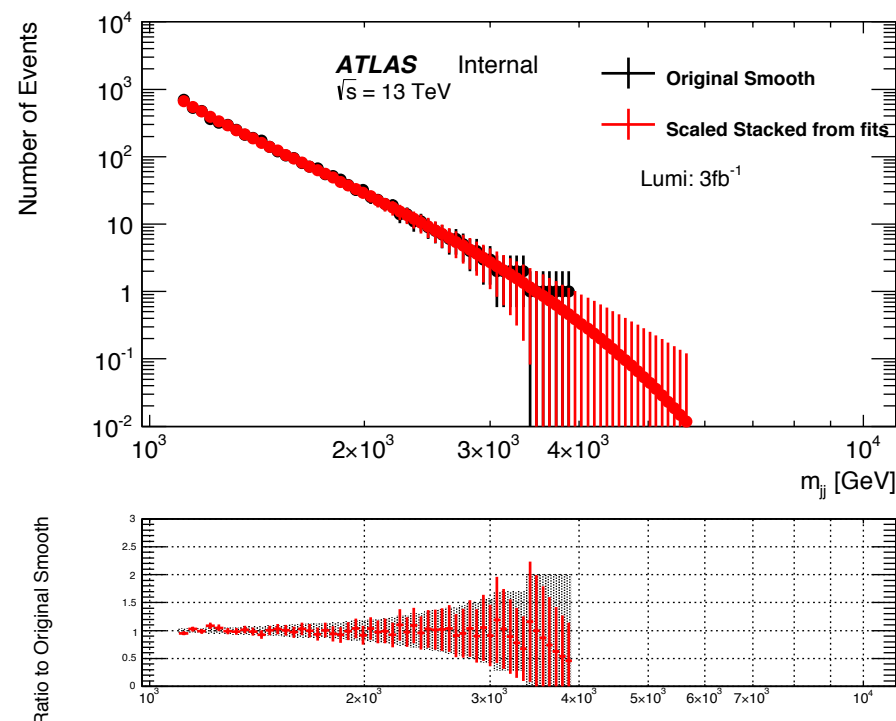
- rounded to one
- poisson errors

Smooth distributions was what was originally fitted to
=> Comparison is good.

Matches well in region where nEvents > 1

Some differences in tail of distribution.

=> All within errors.



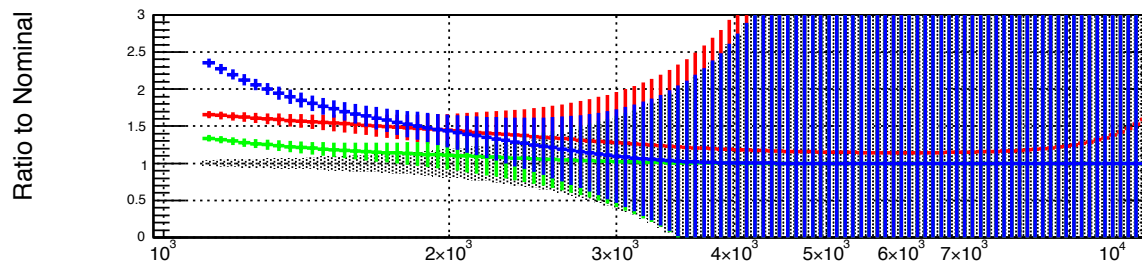
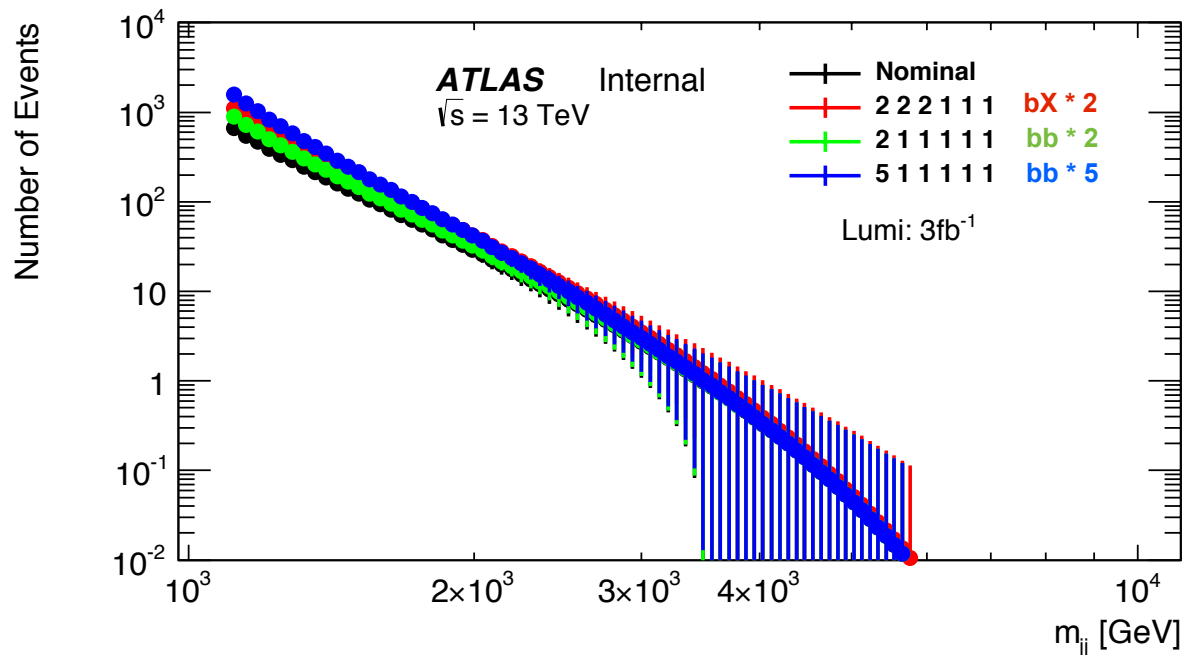
- Compare how different these spectra actually are...

=> 1 1 1 1 1 1 = Nominal, no reweighing.

=> 2 2 2 1 1 1 = double contribution of bb, bc, bl

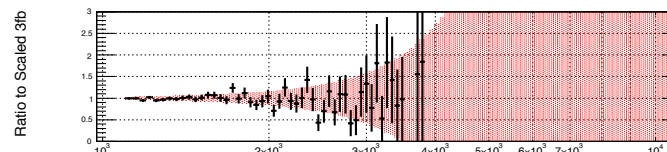
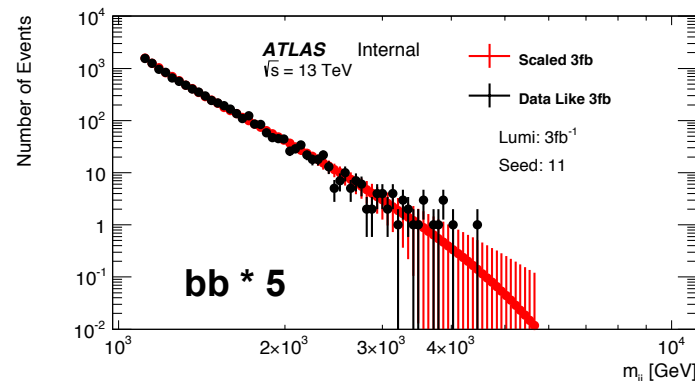
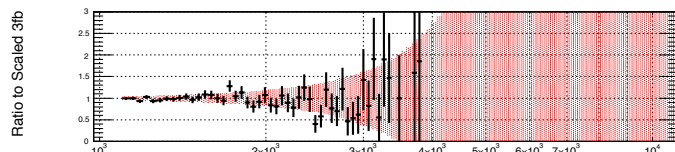
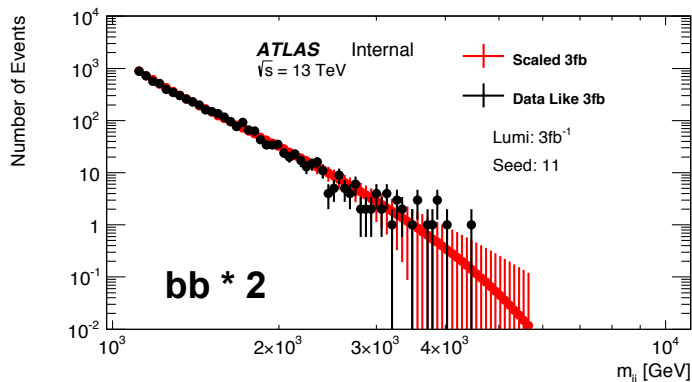
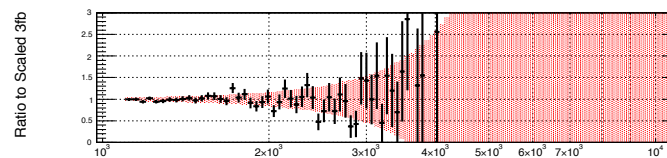
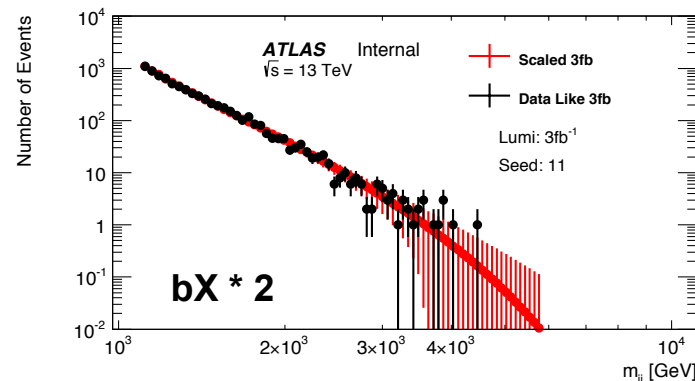
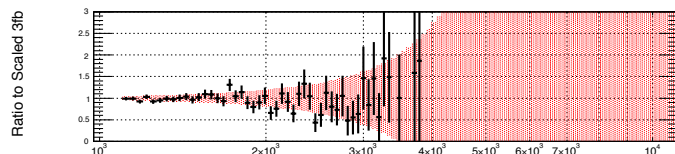
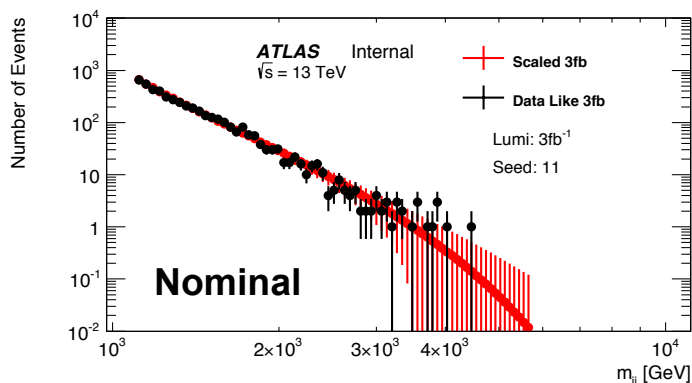
=> 2 1 1 1 1 1 = double contribution of bb

=> 5 1 1 1 1 1 = quintuple contribution of bb



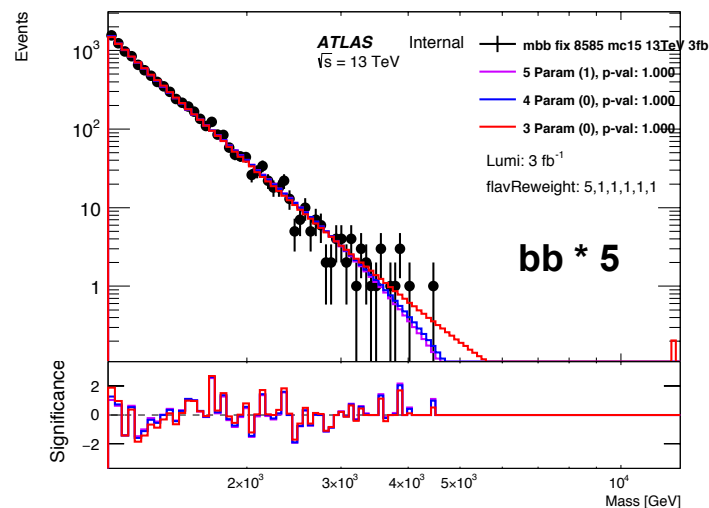
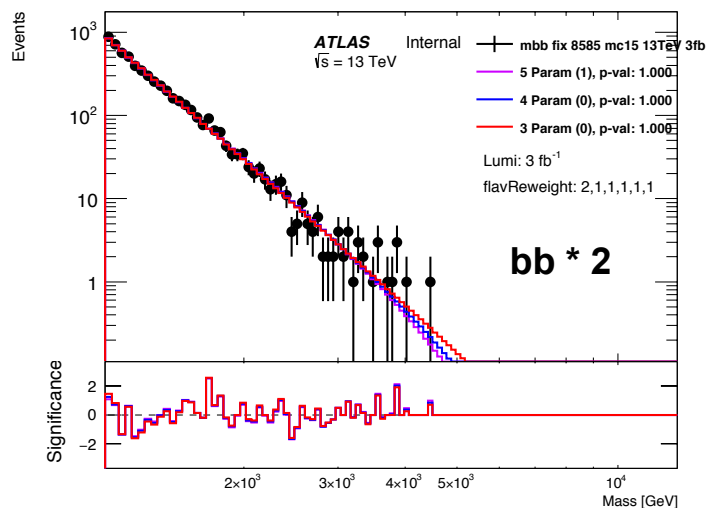
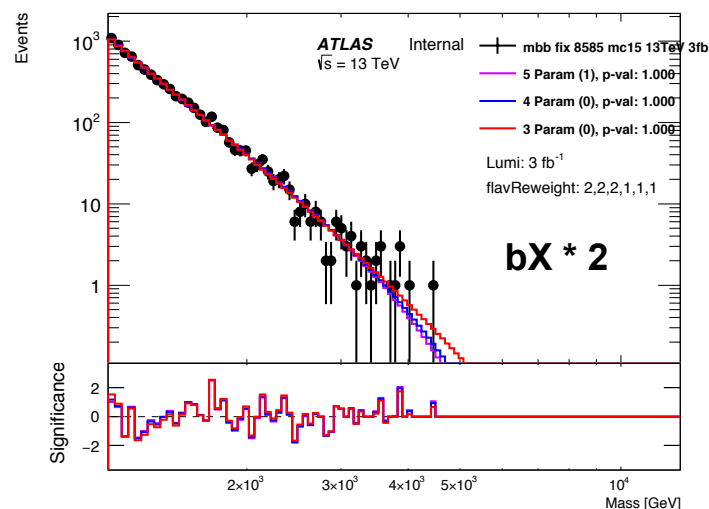
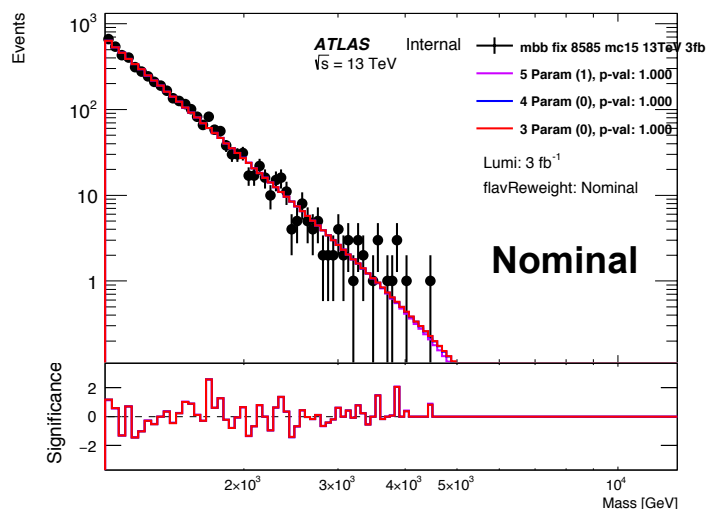


- Make data-like
 - => Poisson fluctuating using scaled distribution as mean entry.
- This creates new scaled distributions from which we can create data-like MC.





- We then fit to each of our different combinations using the all our fitting functions
- Strong performance in all cases
=> However 3 parameter fit splits from others as weighting increases.





1.4 Nominal, All Seeds

