b-Tagged Dijet Analysis:
Fit Function Studies

Laurie McClymont

UCL Meet

29 January 2016

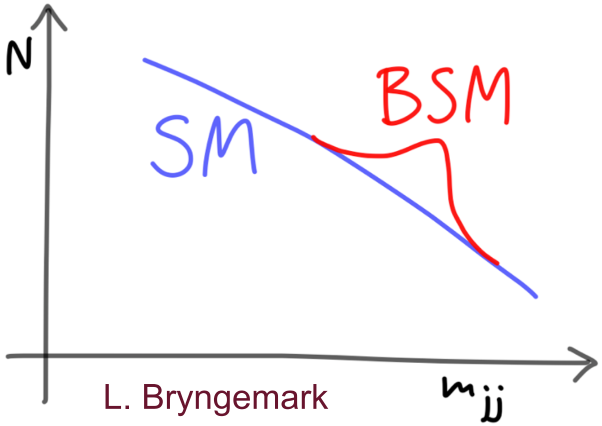
Introduction

- **Overview of Analysis**
  - On course for Moriond!
  - Passed JDM group approval last friday
  - Supporting documentation: CDS Entry
  - Paper draft in place

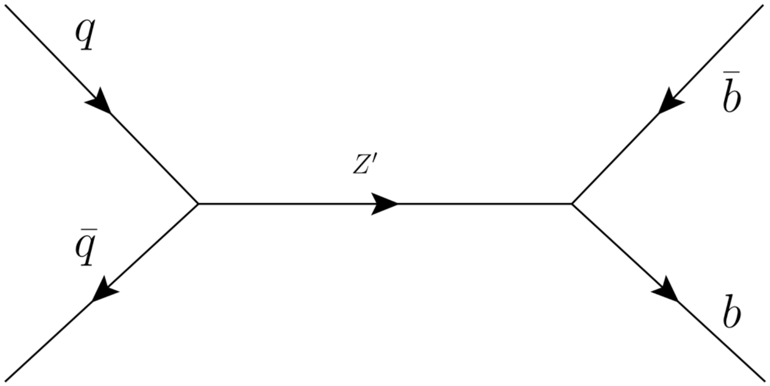
- **Fit function studies**
  - Are we robust to changes in flavour fraction?
  - Do we need an additional systematic for flavour fractions.

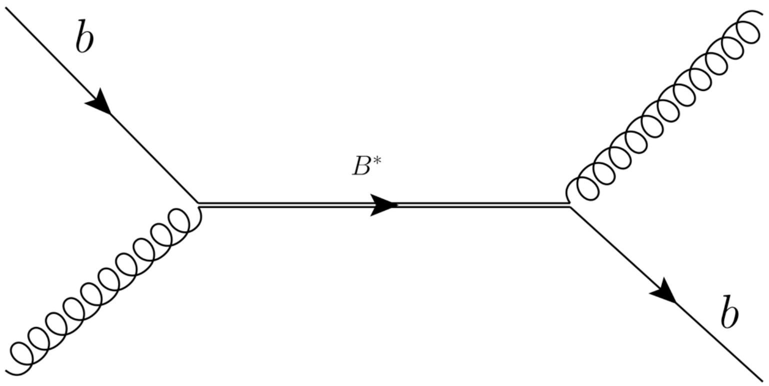
Analysis Intro

- **Search for resonance in invariant mass spectrum of b-tagged jets**
  - Fit QCD background using smoothly falling function.
  - Follows a similar path to dijet analysis



- **In addition, b-tagging is applied.**
  - Three categories - 0, 1 and 2 b-tags
- **Search for generic di-jet resonance**
  - Gaussian with width similar to benchmark models.
  - Two benchmark models, which we will set limits on.
    1) **Z' => bb**, double b-jet final state. (Sequential SM Z')
    2) **b* => bg**, single b-jet final state.



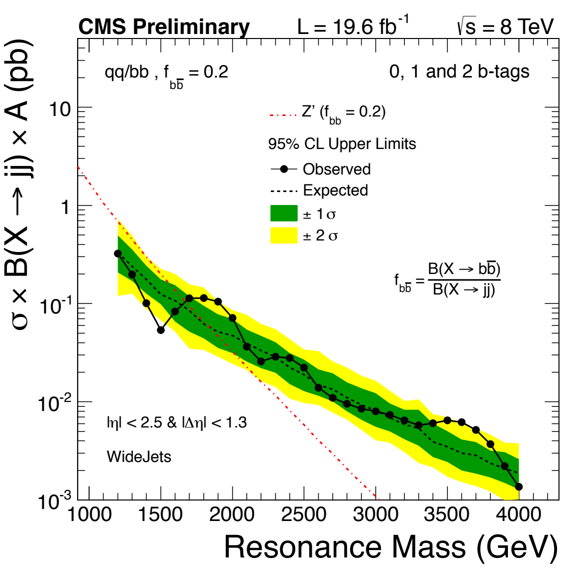


Motivation

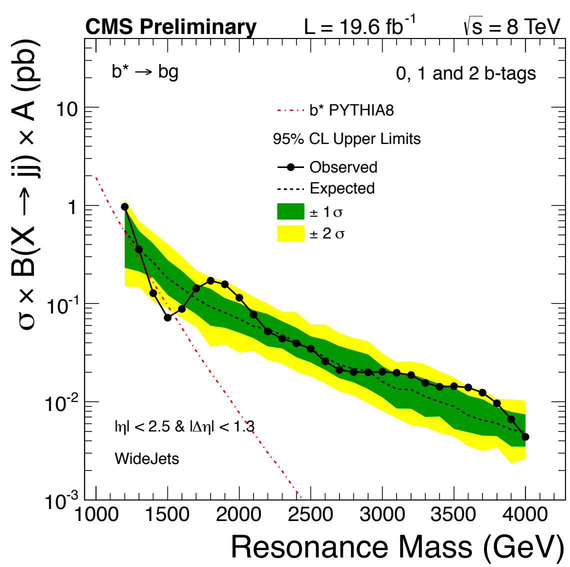
- Many BSM models predict heavy particles that decay into bb or bg.
  - Z', b*, Randall-Sundrum graviton...
  - QCD background is dominated by light jets (u, d, s and gluon)
  - b-tagging can be used to increase sensitivity to these models.

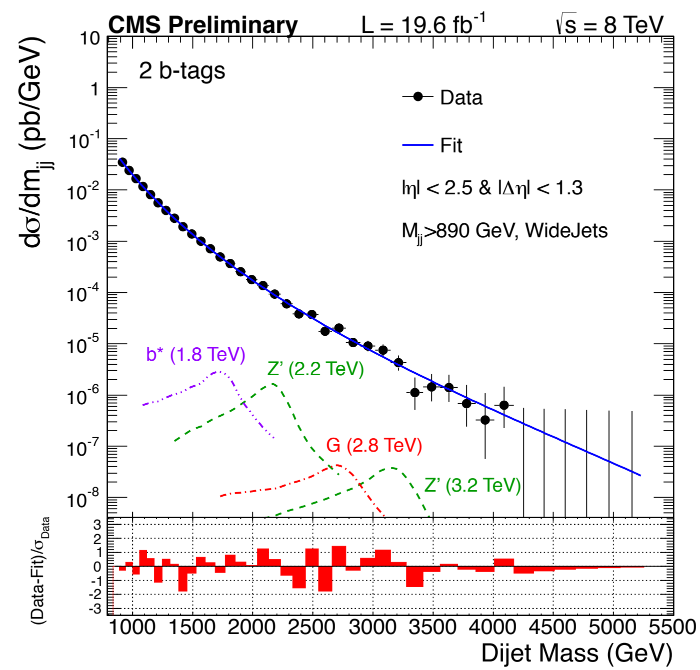
- Generic search performed searching for high mass resonance decaying to b-tagged jets.
  - Performed at CDF and CMS - ([CMS-PAS-EXO-12-023]())
  - No ATLAS result from Run-1

Z' excluded at 1.7 TeV



b excluded at 1.5 TeV*





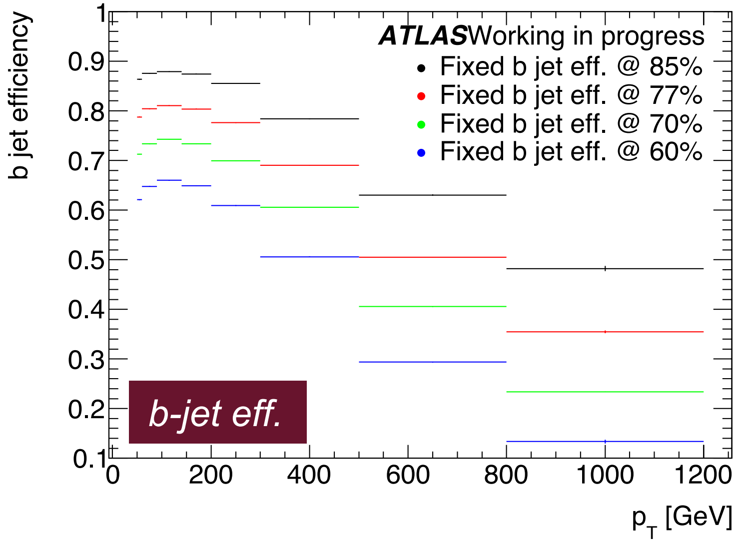
Analysis Details

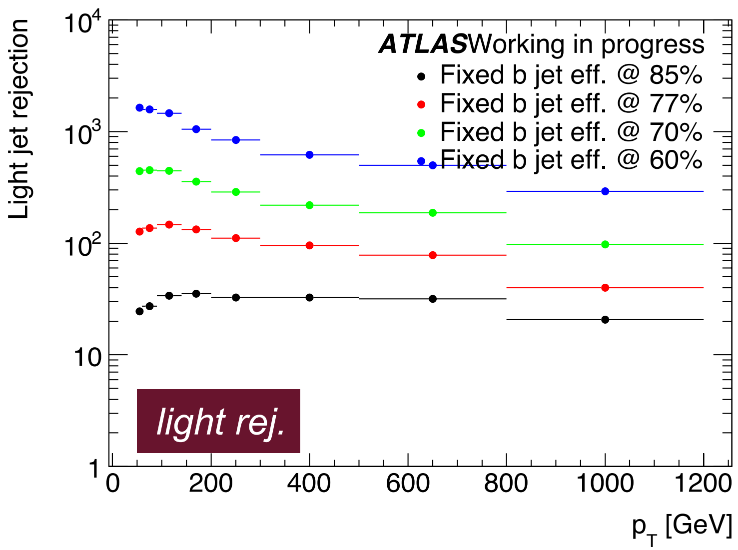
- **Trigger = HLT_j360**
  - Lowest unprescaled single jet trigger

- **Event Selection** (Full list in backup)
  - Same as dijet analysis
  - **m_{jj} > 1100 GeV**, on the trigger plateau.
  - **|y*| < 0.6**, where $y^* = 0.5*(y_1 - y_2)$
    - Central region more sensitive to BSM physics.
  - **Jet |η| < 2.4**
    - In tracking volume for *b*-tagging.

- **MV2c20**
  - **Fixed cut 85% efficiency working point**
  - Loose WP provides best sensitivity compared to others.

- b-jet efficiency ~ 50% at jet-p_T ~ 1 TeV
- Light-jet rejection ~ 30
  - Approx. flat, good for background modelling.

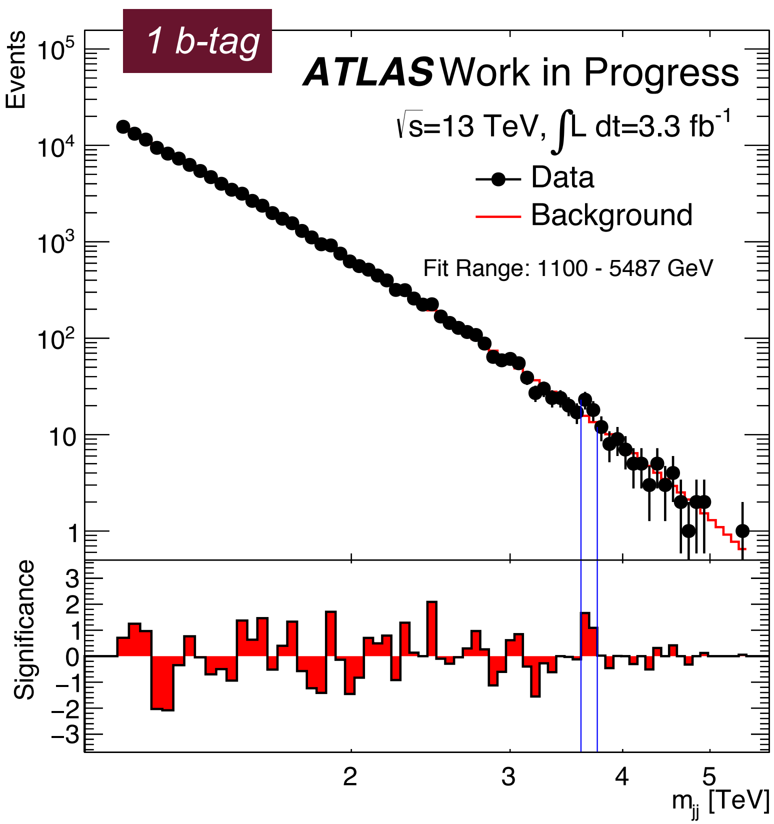


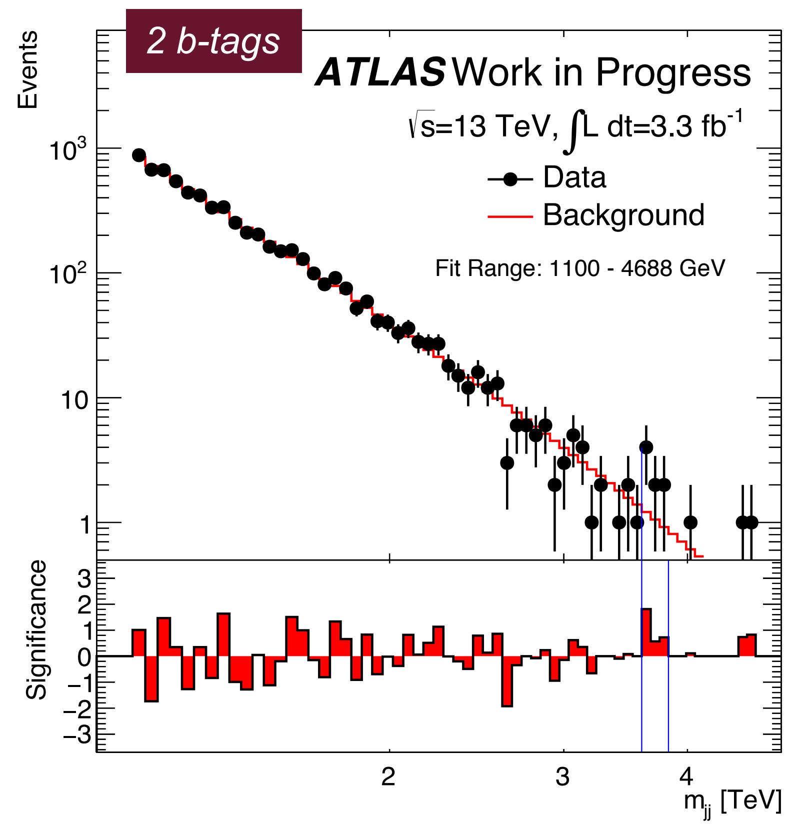


Data - Bump Hunter

- Mass spectra in three tag categories
  - 3.2 fb^{-1} of data, full data set
  - Background fitted with **smoothly falling function**
- Bump Hunter searches for resonances using Gaussian signal
  - Searches for statistically significant deviations

- **No excess found more significant than 2σ**

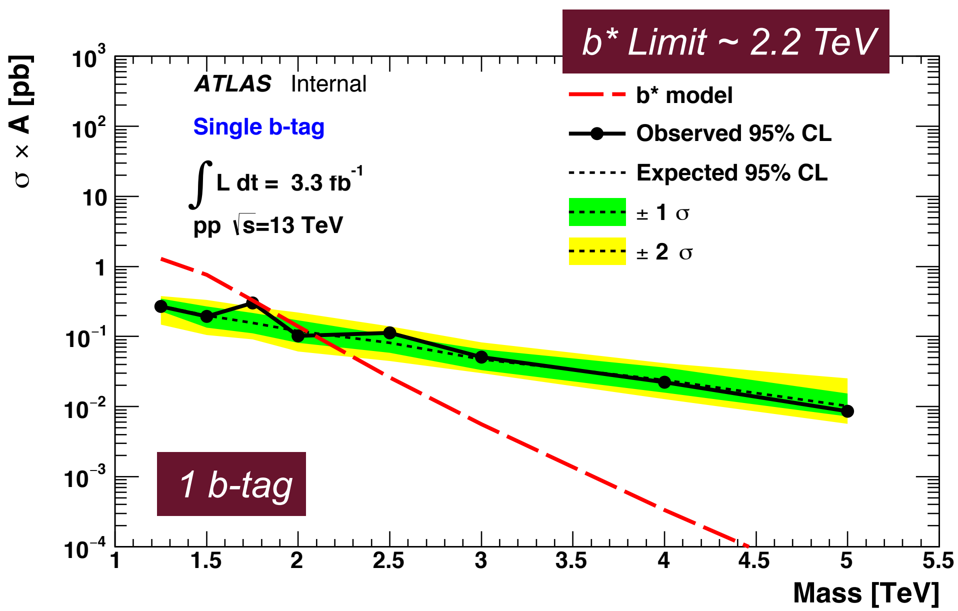


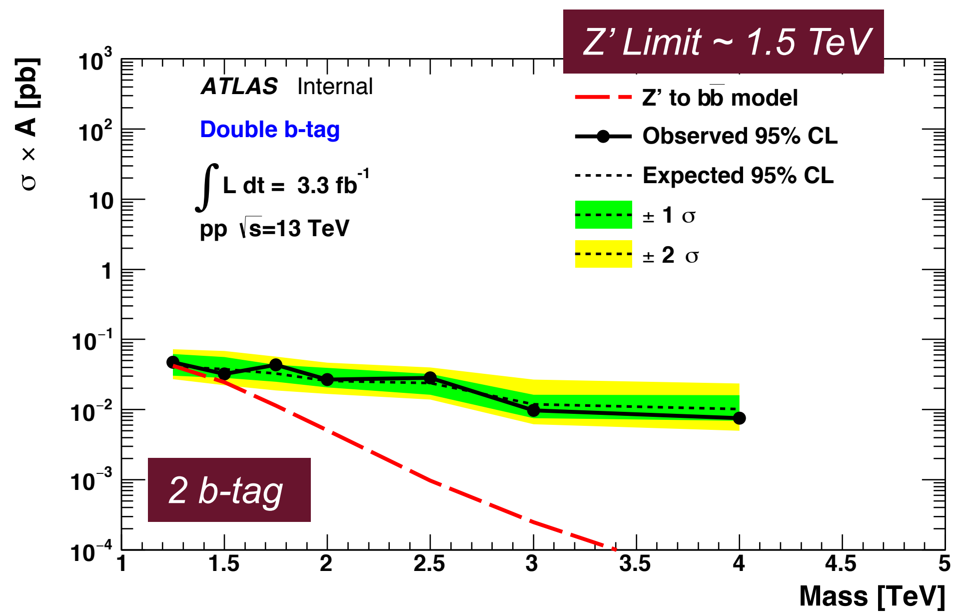


7 Limit Setting

- 95% C.L. upper limits set for *b** and *Z'*
  - Use bayesian approach for limit setting.
  - No correction for acceptance.

- Systematics:
  - Luminosity uncertainty - 5% - From luminosity group
  - Background uncertainty
    - **Uncertainties on fit function choice and fit parameters**
  - Signal uncertainty
    - JES and JER uncertainties
    - An additional BJES uncertainty
    - *b*-tagged scale factor uncertainty



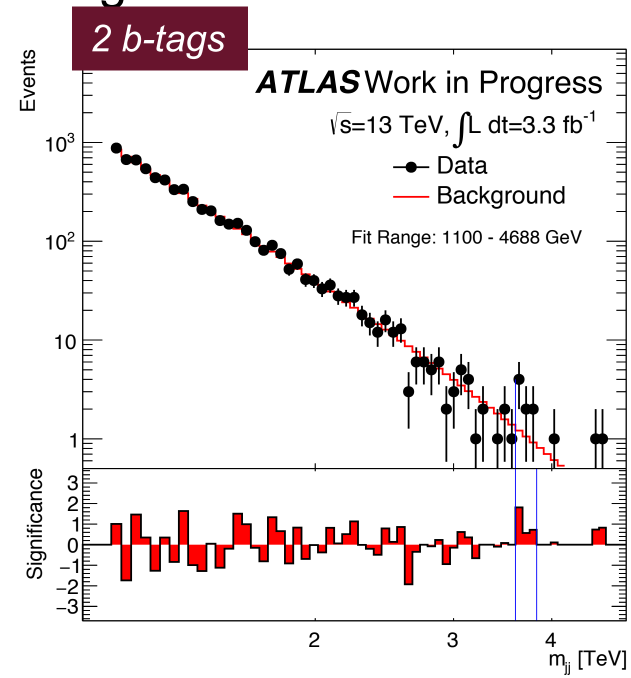


Fitting Function Studies

- **Fitting Function**
  - We use a smoothly falling fitting function to fit to background

$$f(x) = p_1(1-x)^{p_2}(x)^{p_3+p_4 \ln x + p_5 \ln x^2}$$

where, $x = m_{jj}/\sqrt{s}$

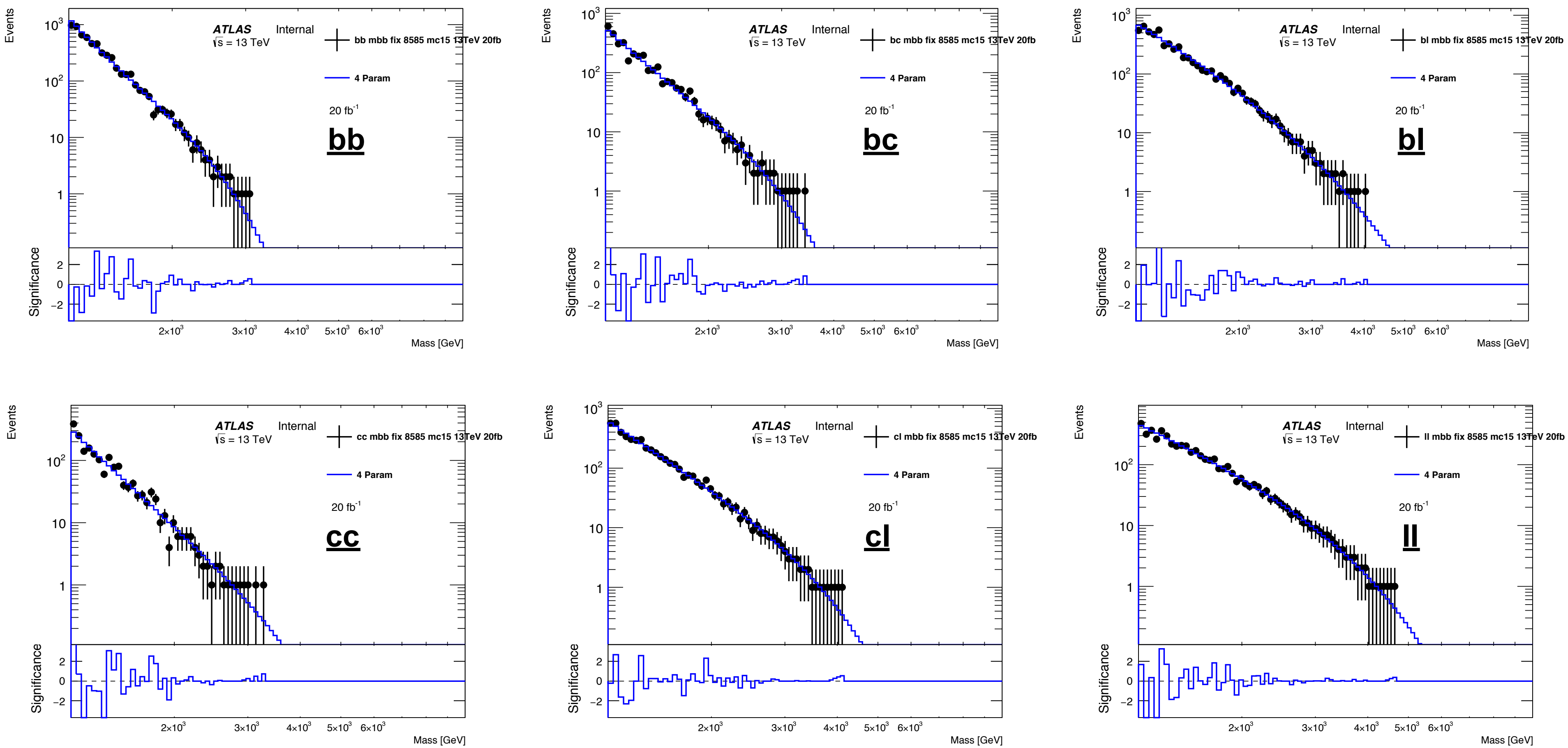


  - This comes in a 3, 4 or 5 parameter versions
    - Setting p_4 and p_5 to zero

- **Varying Flavour Composition**
  - It is known that the fitting function can fit to the Monte-Carlo
  - However MC is not a perfect prediction of reality
  - What if there are more *b*-jets in the data than in the MC
    - Can we still fit to data in this case? Is our fitting function robust.

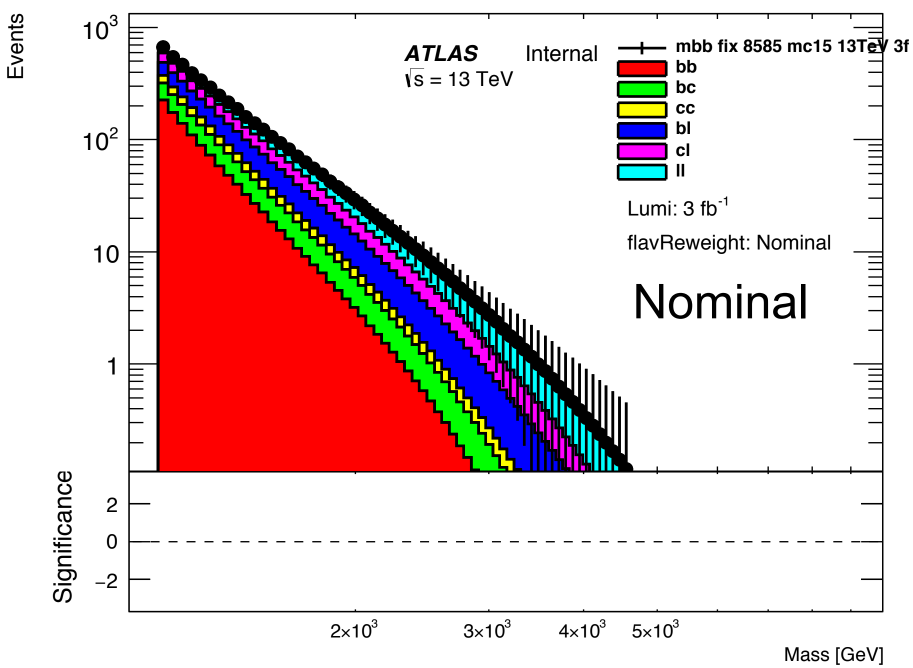
Getting the Flavour Fractions

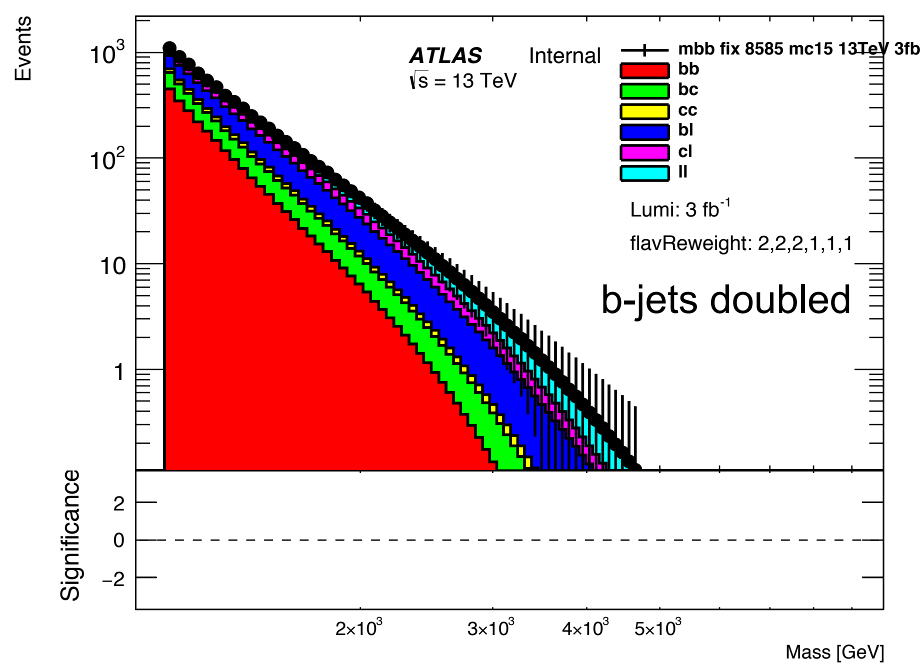
- Flavour fractions are extracted from MC using truth information.
- The dijet mass spectrums for these flavour fractions are then scaled to $20fb^{-1}$
- The dijet mass spectrums are fitted to using the 4-parameter fit function.



Stacking the Flavour Fractions

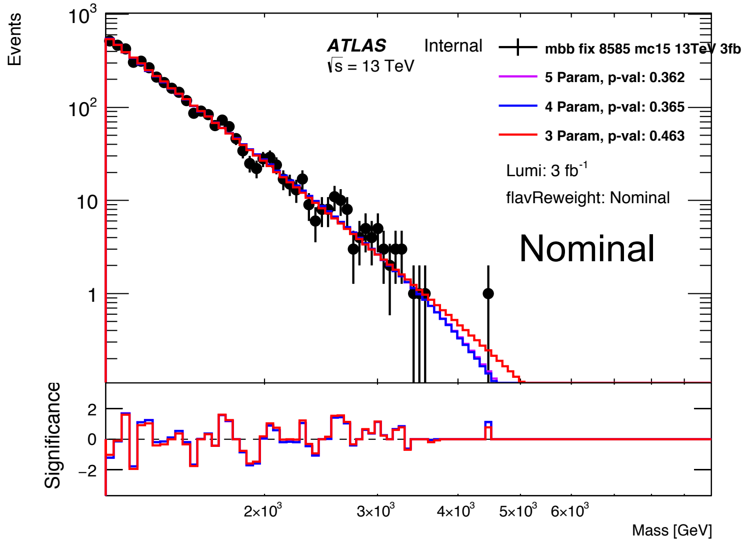
- We take the fits to the flavour fractions from previous slide
- These are then scaled to $3fb^{-1}$ and stacked to create a dijet mass
- I also create the case where b-jet content is doubled (bb, bc and bl)

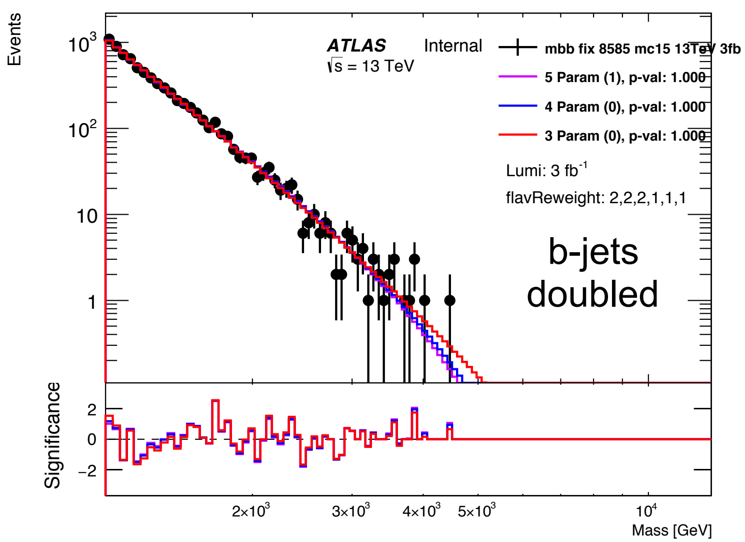




Making Data-Like

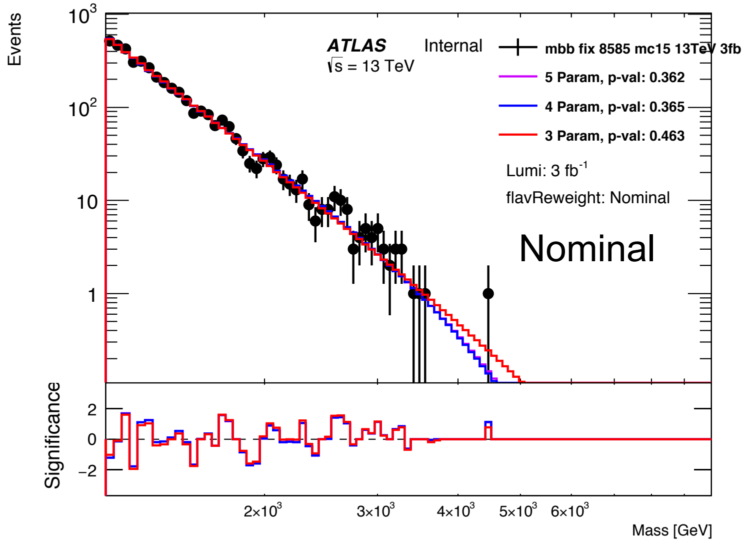
- By applying poisson fluctuations we can create ‘data-like’ distribution
- These are fitted using the 3, 4 and 5 parameter fit function

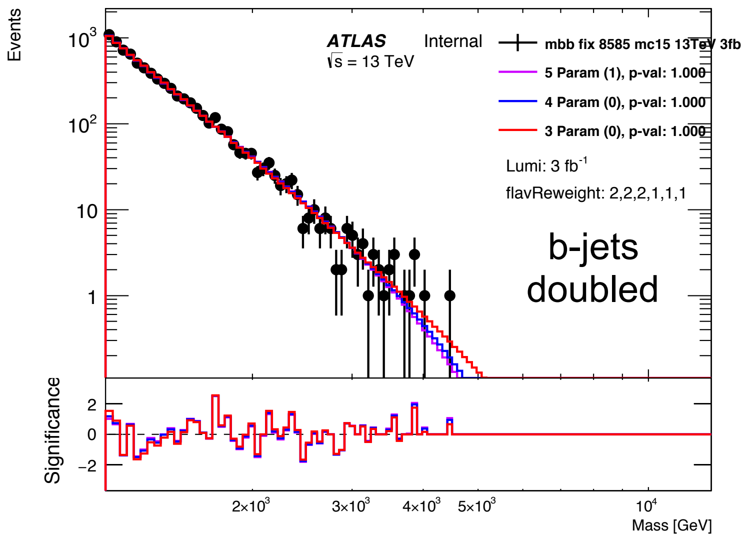




Making Data-Like

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



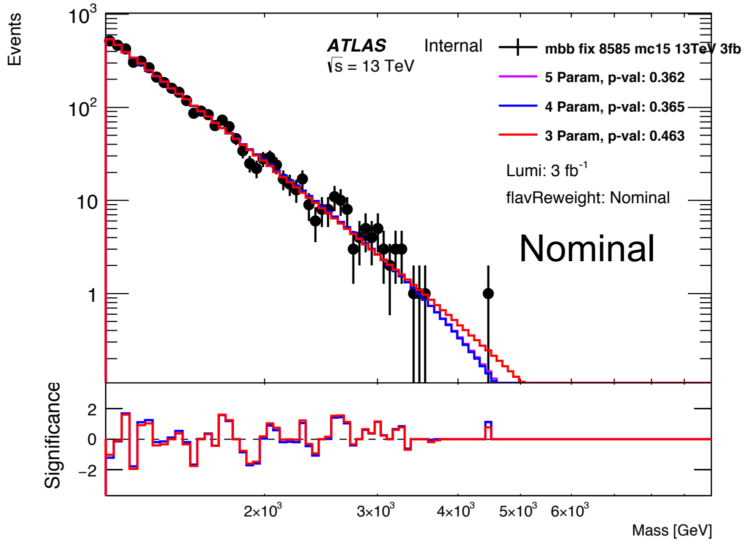


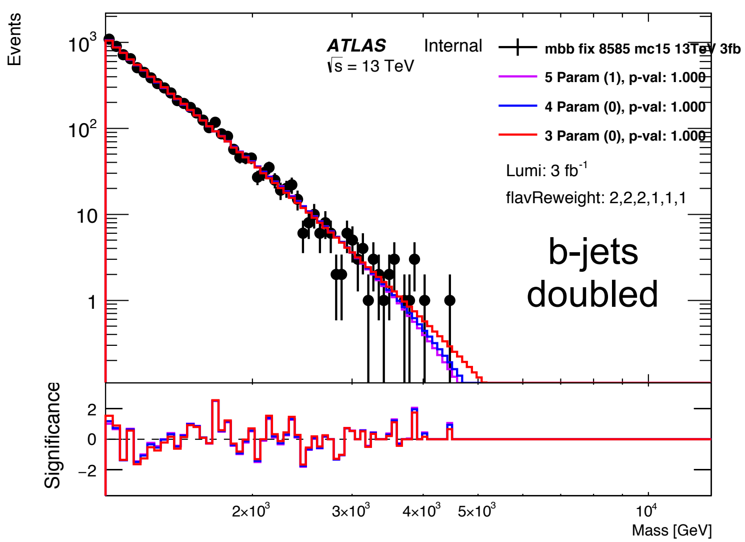
To calculate p-value of a fit:

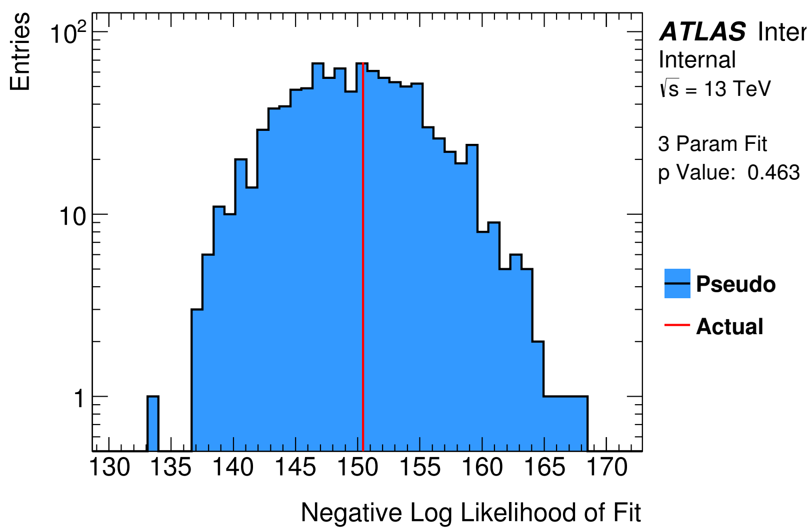
1. Take the fit function and apply poisson fluctuations. (Pseudo-experiment)
2. Re-fit to the pseudo-data using the same fit function.
3. Compare quality of fit to pseudo-experiment to that of the original fit.
   - For a measure quality of fit I use negative log likelihood
4. Repeat 1000 times and count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

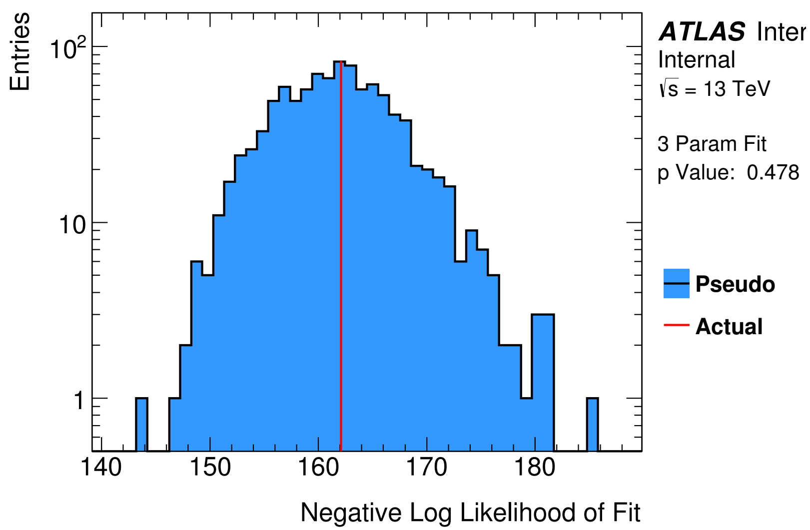
Making Data-Like

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



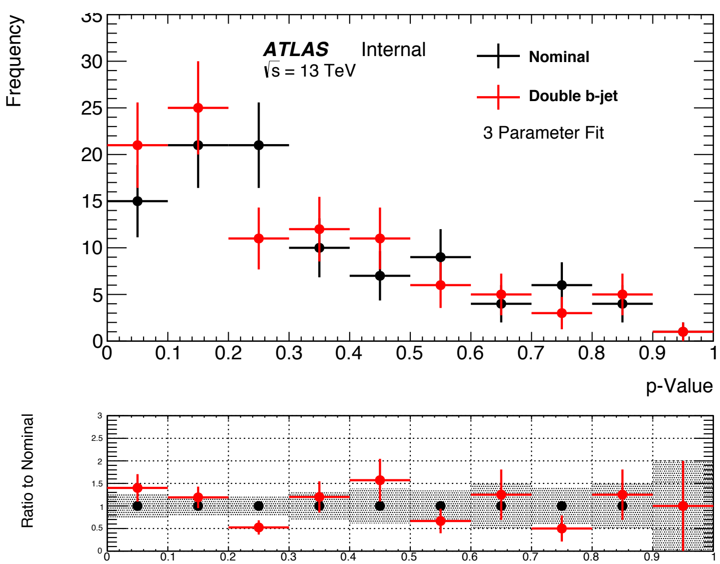


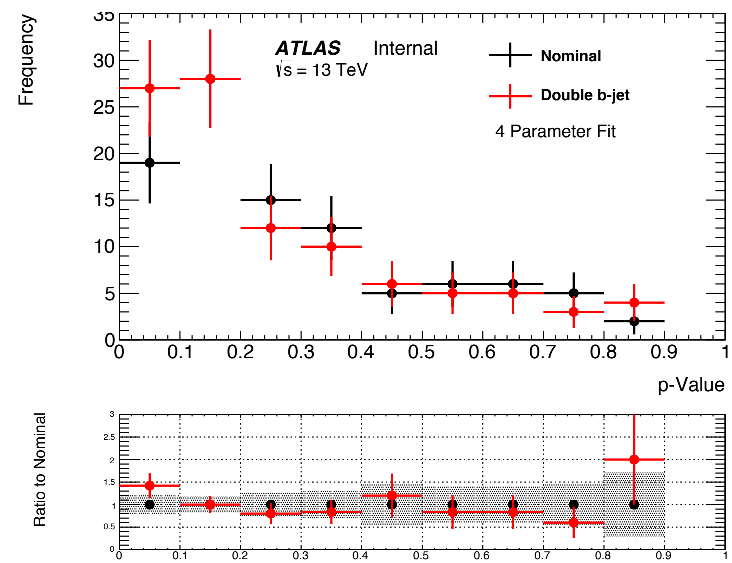


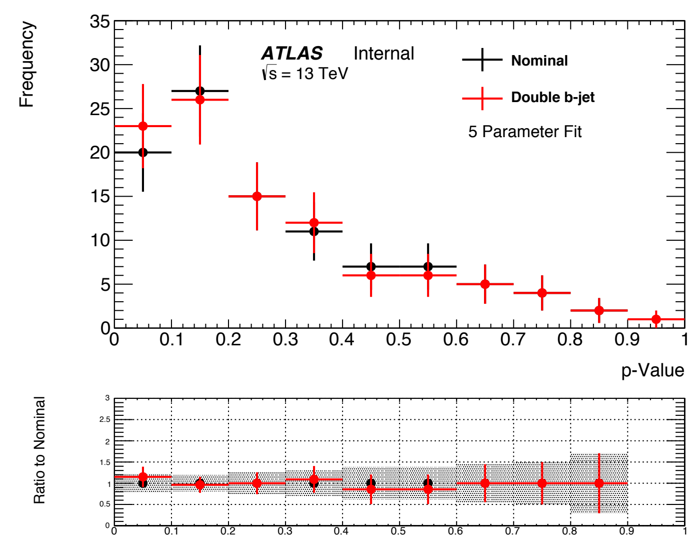


Many More Data-Like

- 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





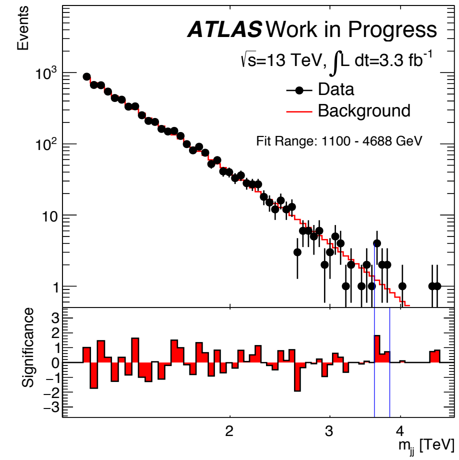


Mean p-values

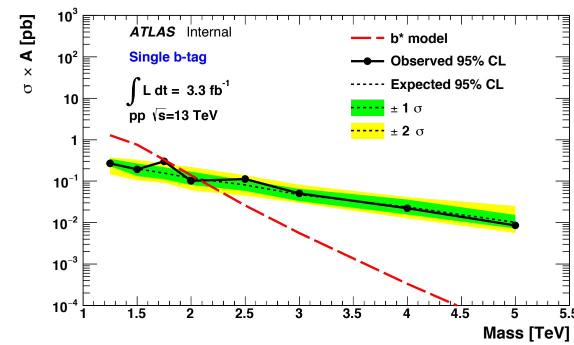
	3-Para. Fit	4-Para. Fit	5-Para. Fit
Nominal	0.325 +/- 0.024	0.280 +/- 0.023	0.283 +/- 0.022
b-jet Doubled	0.308 +/- 0.024	0.267 +/- 0.022	0.276 +/- 0.022

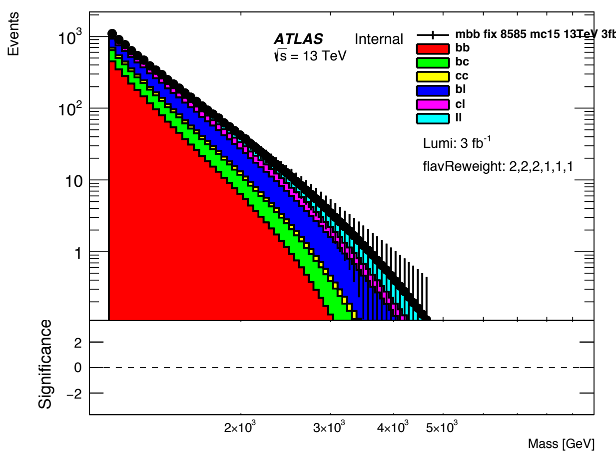
16 Conclusions

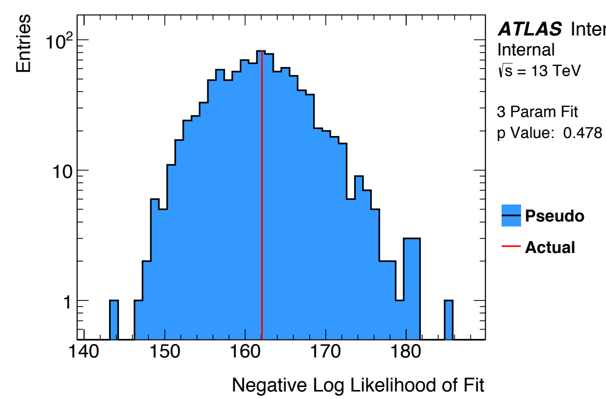
- **Search performed resonance in invariant mass spectrum of b-tagged jets**
  - Dijet mass spectra fitted to using fit function
  - No excesses more significant than 2σ
  - Limits set on Z' and b* models



- **Fit Robust to Changes in Flavour Fractions**
  - 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.







Backup:
b-Tagged Dijet Analysis

- **Aiming for a paper for Moriond**

- Documentation in Place
  - SVN Area

- Ed Board Assembled
  - A. Glazov (Chair), F. Parodi, L. Tompkins
  - First Ed Board Meet - 11th November
  - Agenda can be found here

- Exotic Plenary Talk here

- **Data Used**
  - 25ns data with luminosity of **3.2 fb^{-1}** (Periods D-J)
  - Exclude runs with IBL Off - Due to huge drop in b-tagging performance.
  - GRL: *data15_13TeV.periodAllYear_DetStatus-v70-pro19-04_ DQDefects-00-01-02_PHYS_StandardGRL_All_Good_25ns.xml*

- **Trigger**
  - **HLT_j360**, lowest unprescaled single jet trigger

- **Event Selection**
  - Reject events with problematic calo. reconstruction (LAr, Tile and Core Errors)
  - At least two jets.
  - **Leading-jet $p_T > 440$ GeV**, Subleading jet $p_T > 50$ GeV
  - **$m_{jj} > 1100$ GeV**, such that we are on the trigger plateau.
  - **$|y^*| < 0.6$**, where $y^*=0.5*(y_1 - y_2)$
    - Central region more sensitive to BSM physics.

- **Jet Selection**
  - Standard jet calibration (with JES correction applied)
  - 2015 loose jet quality cuts applied.

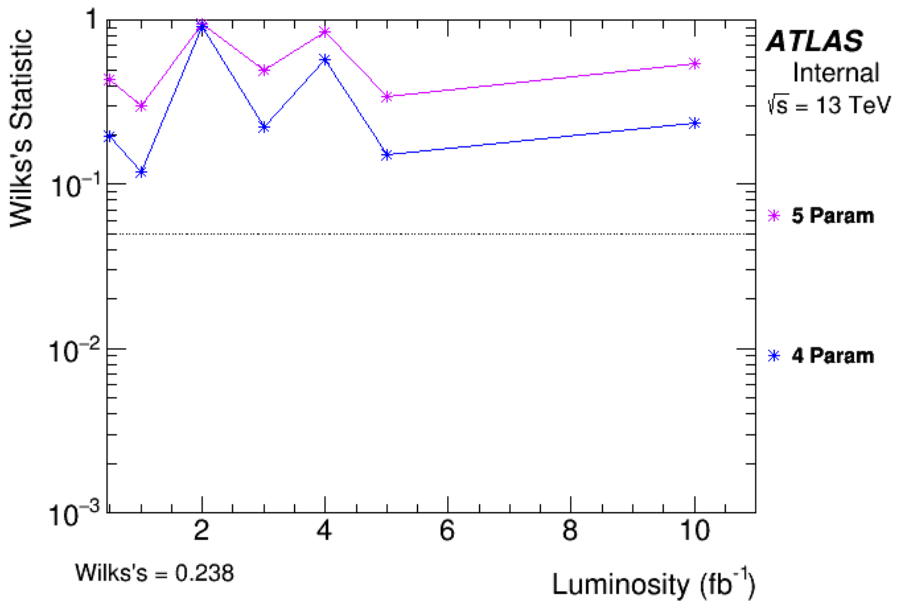
Background Modelling — *b-tagged di-jet*

- Fit to background using smoothly falling function:

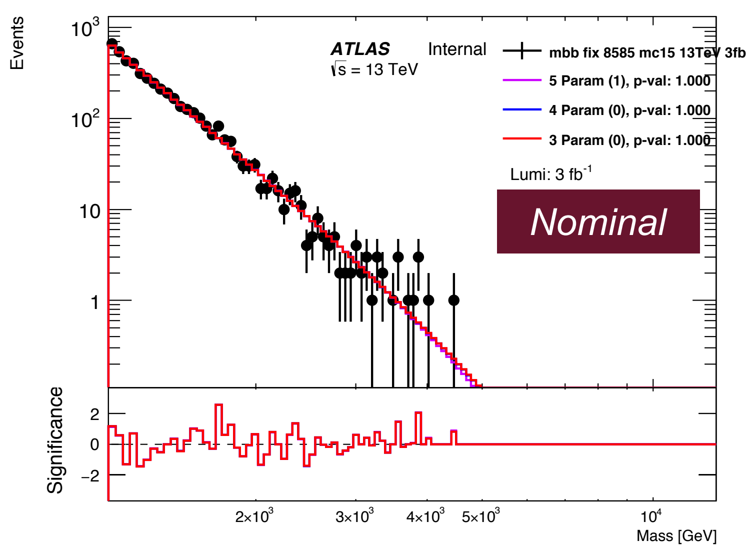
$$f(x) = p_1(1-x)^{p_2}(x)^{p_3+p_4\ln x+p_5\ln x^2}$$

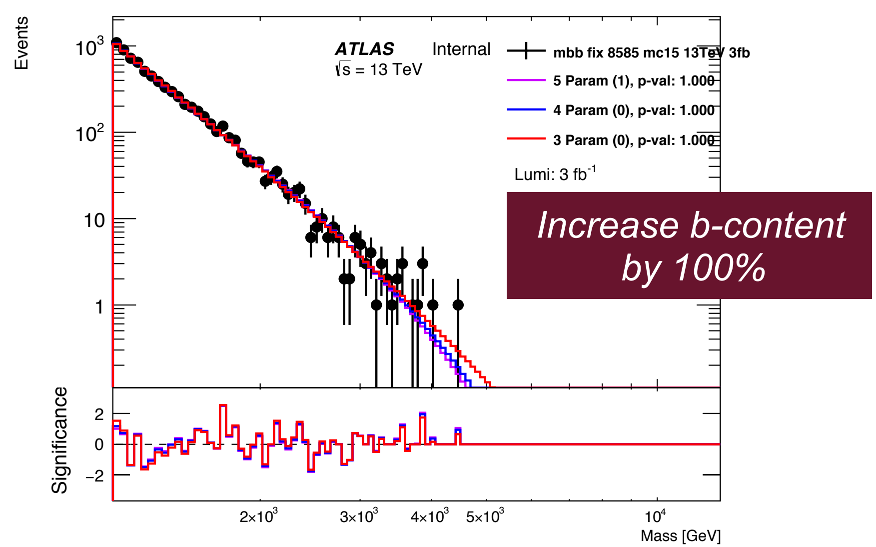
where, $x = m_{jj}/\sqrt{s}$

- This comes in 3, 4 and 5 parameter functions.
  - 3 and 4 parameter set $p_5 = 0$ and then $p_4 = 0$.
- Use Wilks' statistic to choose fit function
  - Default option is 3 parameter fit function.
  - Use Wilks' to test if we need to change function
  - MC tests show we expect to be able to use 3 parameter fit up to 10 fb^{-1}



- Performing cross-checks confirming that we are robust to changes in flavour fraction





List of Systematics

- Luminosity - 9% uncertainty

- Background
  - Fit function and fit parameters

- Signal
  - JES Uncertainty
    - Branches available in analysis nTuple
    - < 4%
  - JER Uncertainty
    - Assume to be negligible
  - BJES Uncertainty
    - Studies performed
  - B-tagging scale factor uncertainty

- Studies to be carried out
  - Then will be added to limit setting procedure.

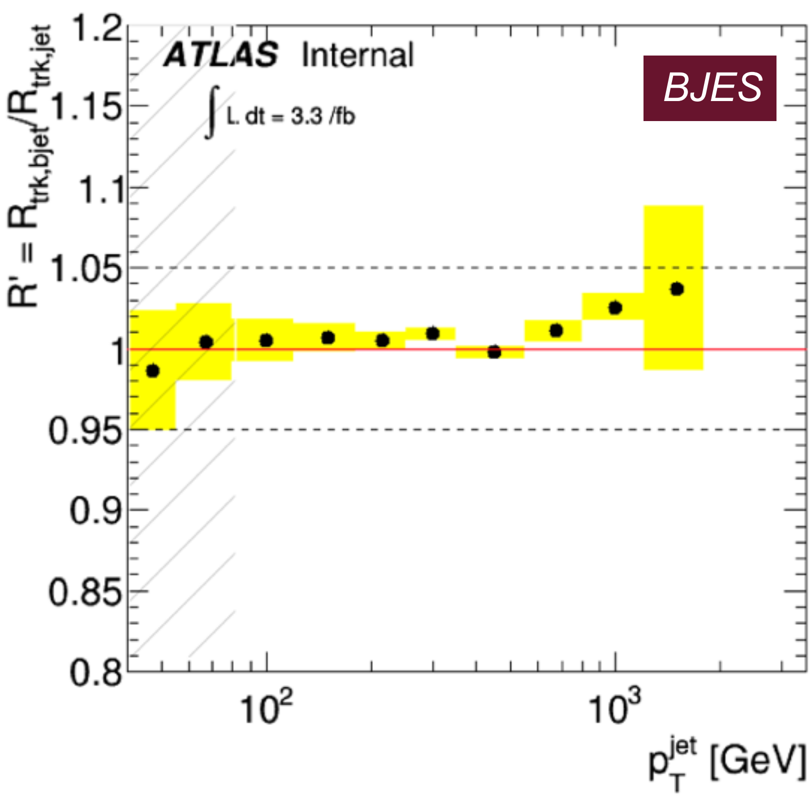
b-tagged di-jet

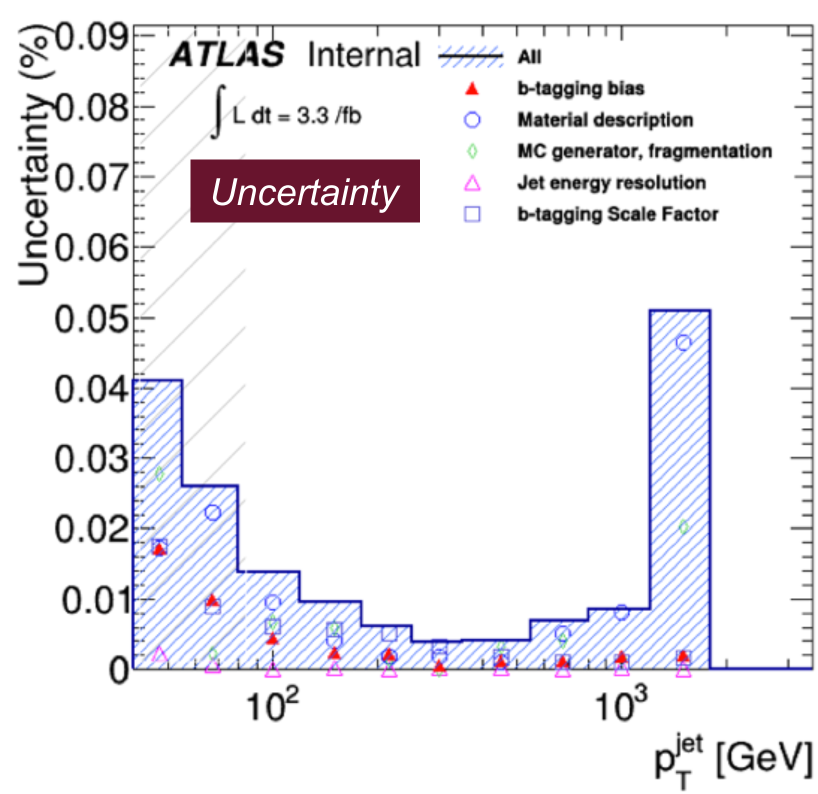
- Calculate using ratio of tracks within jet cone to reconstructed calo jet.
  - Use a double ratio between b-tagged jets and inclusive jets

$$R' = \frac{< r^{trk}_{bjet} >_{Data} / < r^{trk}_{bjet} >_{MC}}{< r^{trk}_{inc} >_{Data} / < r^{trk}_{inc} >_{MC}}$$

where $r^{trk} = \frac{\sum \vec{p_T}^{trk}}{p_T^{jet}}$

- Ongoing study
  - Further work required
  - Regularly presented in JES/JER Meetings

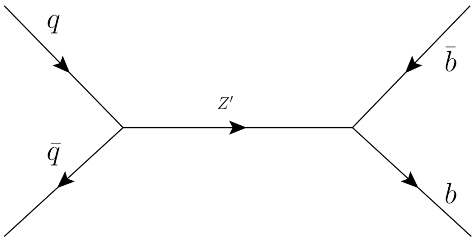


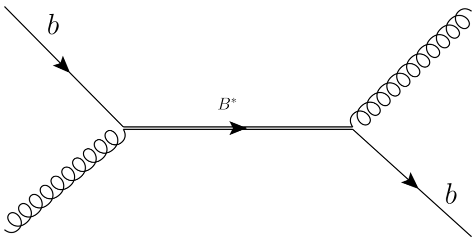


Signal Models

b-tagged di-jet

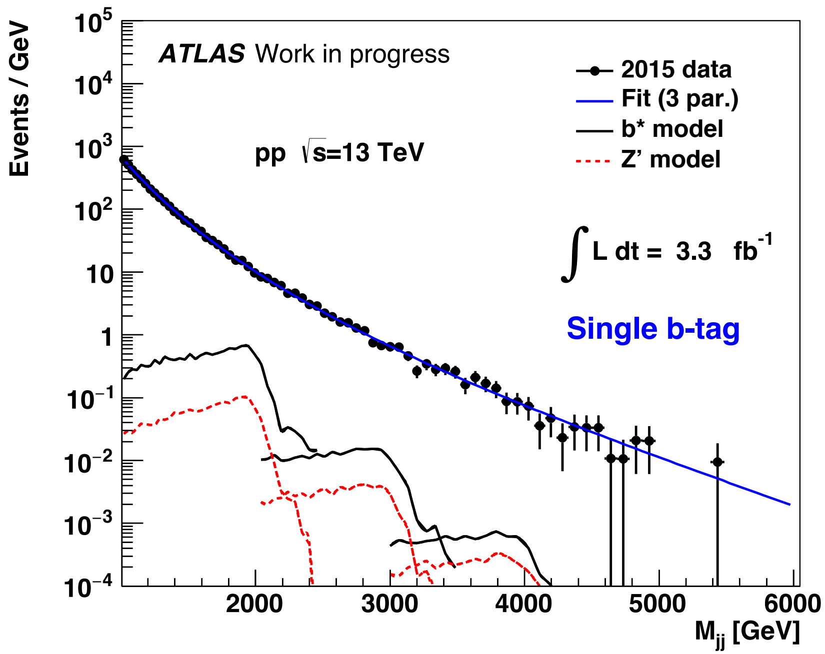
- Two benchmark models - We can set limits here.
  - **Z' => bb** - 1.25, 2, 3 and 4 TeV
  - **b* => b+X** - 1.25, 2, 3, 4 and 5 TeV
  - Templates taken from MC samples

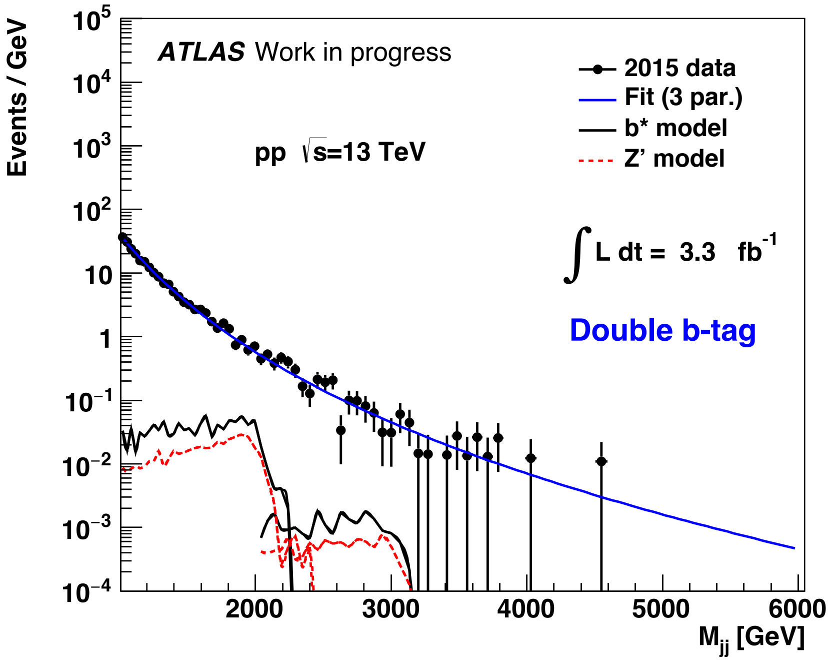




- One cross-check channel
  - **q*** - 2.5, 3, 3.5 and 4.5 TeV

- Generic search performed for a Gaussian signal.
  - Resonance width taken from the benchmark





Data/MC Jet pT Comparison

b-tagged di-jet