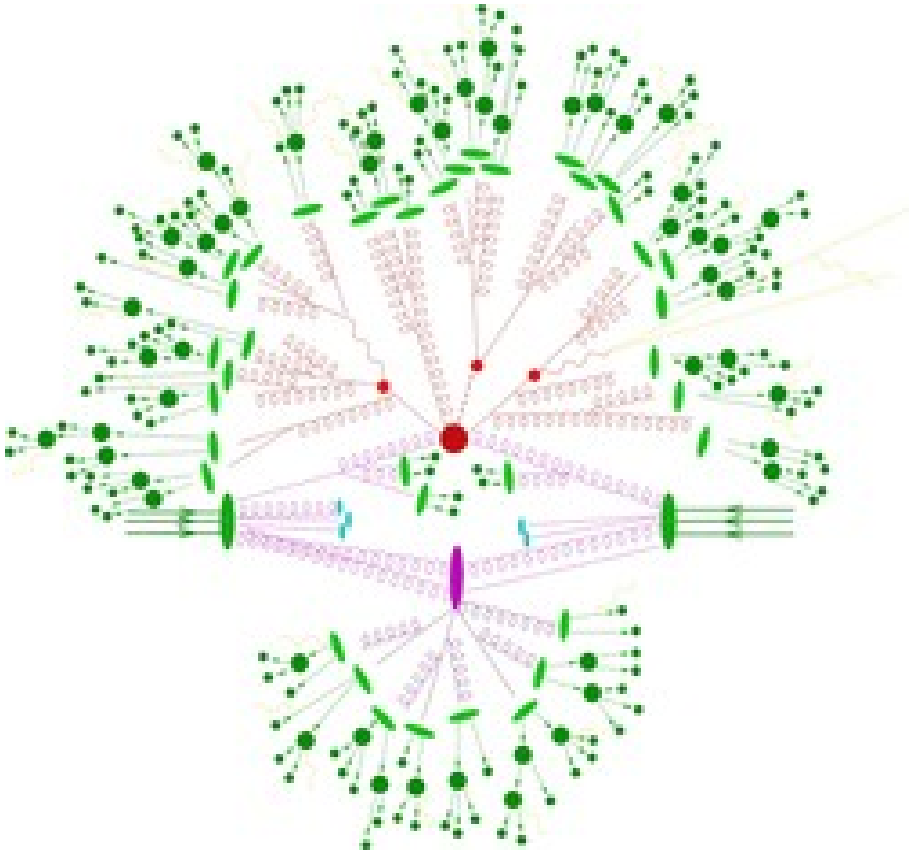
MC issues

Thorsten Kuhl

- Intro
- Few test cases:
  - Ttbar
  - W+jets
  - Heavy flavours
- Comments



Overview

- Now we have a lot of Data → we can constrain various variations
  - Scales are sometimes already constrained by data
  - Need proper heavy flavour treatment
  - We can constrain shower at low (MinBias) and higher scale (di-jet and ttbar gap fraction)
- We have a lot of very fancy Monte Carlo generators
- Producing MC is expensive (CPU time):
  - 500 events per 40 hours Sherpa ttbar+3jets (9 legs) wo phase space integration
  - Alpgen similar but only at ttbar+Np5 (7 legs)
  - Simulation: additional 1-5 Minutes

  → We have to be sure that we produce good Monte Carlo events
- Point of talk is not to blame someone or show computational issues,
  - it is about physics features and essential ingredients for good analysis

Example 1: ttbar

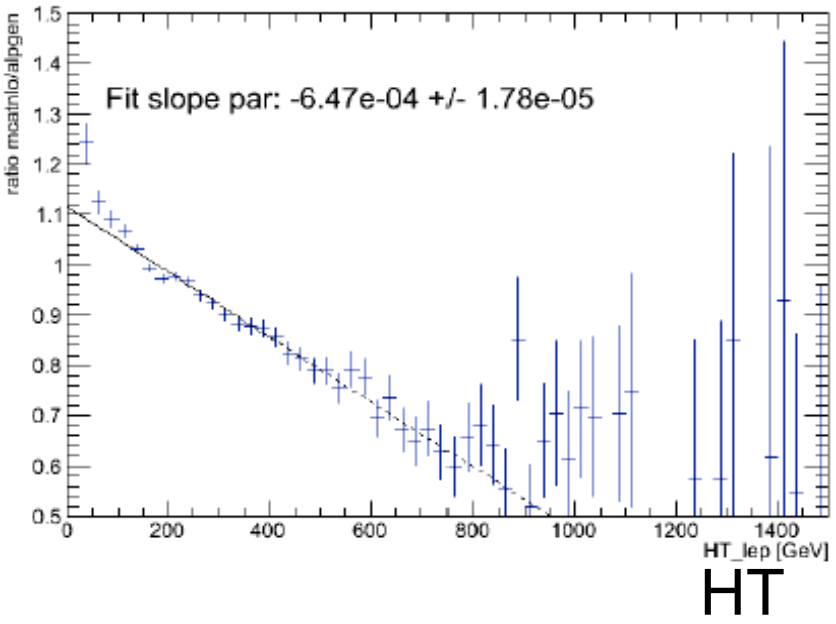
- Alpgen
  - Version: 2.13
  - Showering: Herwig 6.520, Jimmy 4.31
  - PDF: CTEQ6L1
- MC@NLO
  - Version: 4.01
  - Showering: Herwig 6.520, Jimmy 4.31
  - PDF: CT10
- Powheg-Box
  - Version: 1.0
  - Showering: Pythia 6.425
  - PDF: CTEQ6L1, CT10
- Underlying event tune: AUET2(B)

Example 1: ttbar

- Selection
  - One electron with $p_T > 25$ GeV and $|\eta| < 2.47$
  - At least two jets with $p_T > 25$ GeV and $|\eta| < 2.5$
  - Missing transverse energy > 40 GeV

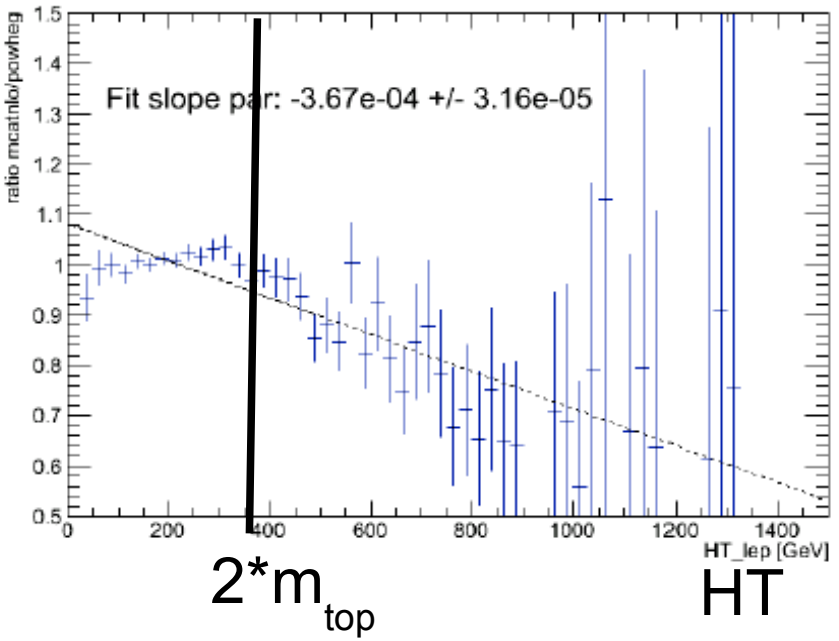
Scalar sum of up to four leading jets

MC@NLO / Alpgen



HT

MC@NLO / Powheg



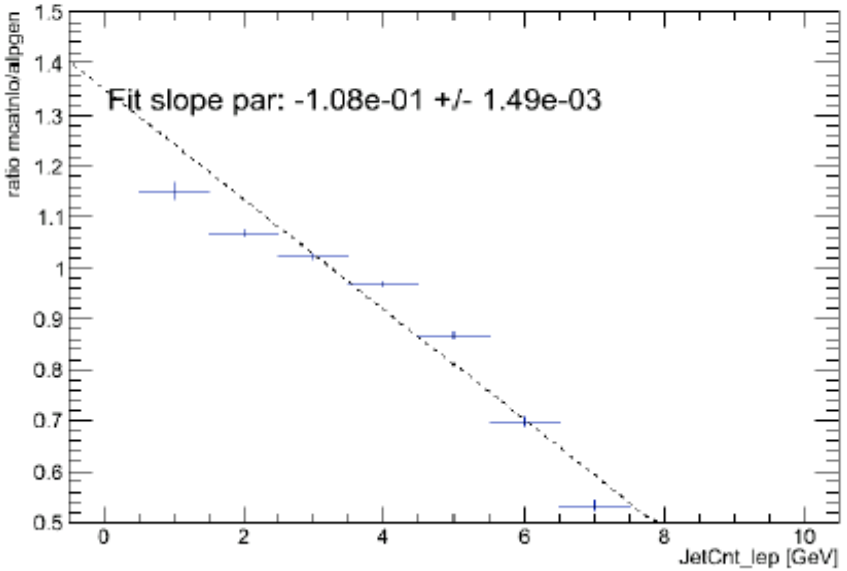
$2*m_{top}$ HT

Example 1: ttbar

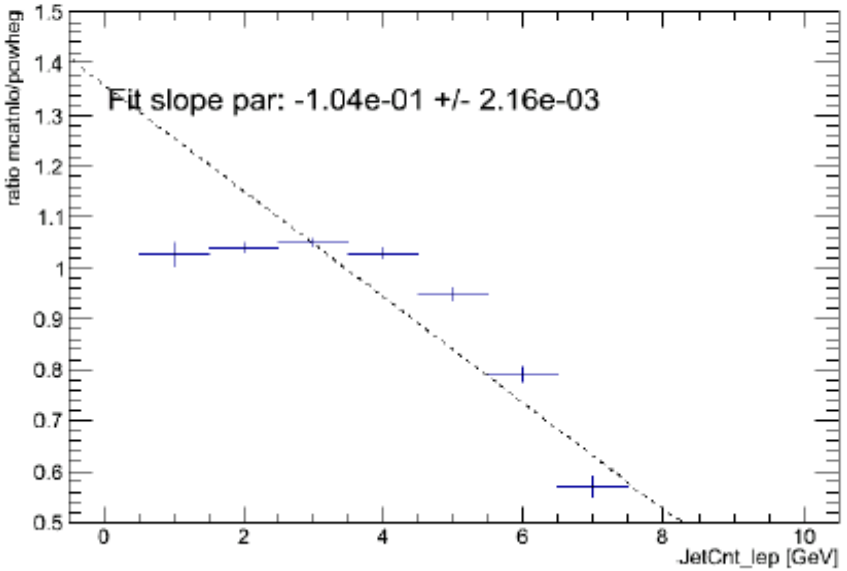
- Selection
  - One electron with $p_T > 25$ GeV and $|\eta| < 2.47$
  - At least one jet with $p_T > 25$ GeV and $|\eta| < 2.5$
  - Missing transverse energy > 40 GeV

Number of jets with $p_T > 25$ GeV

MC@NLO / Alpgen



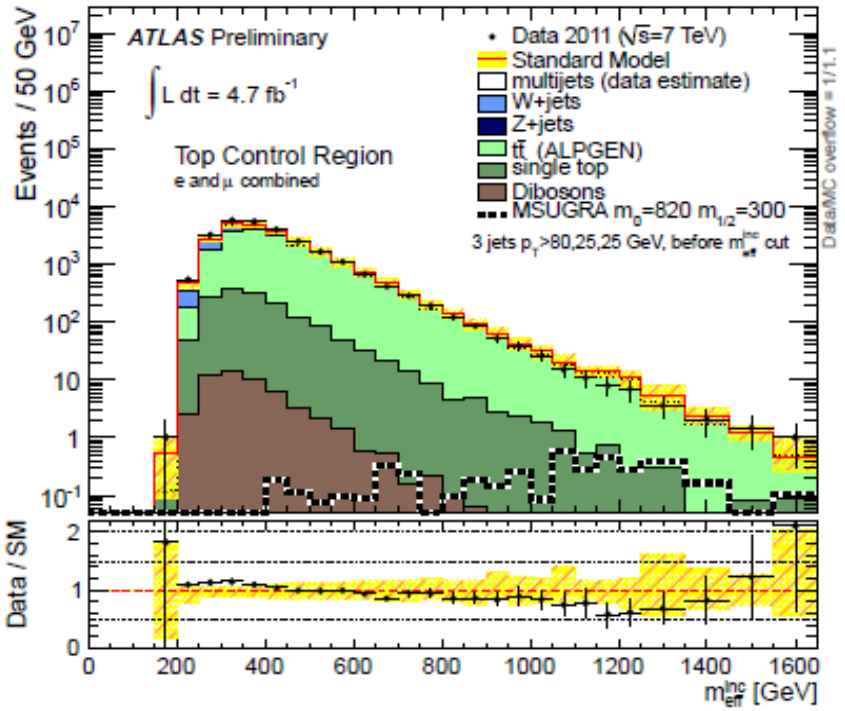
MC@NLO / Powheg



* Powheg describes data really very good, but why does it behave like a "Multi Leg generator" and not drops off after 4 jets?

Example 1: ttbar

- Selection
  - One electron/muon with $p_T > 25/20$ GeV and $|\eta| < 2.47/2.4$
  - At least three jets with $p_T > 80, 25, 25$ GeV, $|\eta| < 2.5$, at least one b-tag
  - Missing transverse energy $\in [30, 120]$ GeV
  - Transverse mass $\in [40, 80]$ GeV



ATLAS-CONF-2012-041

ALPGEN

Difference in xsection due to scale variation

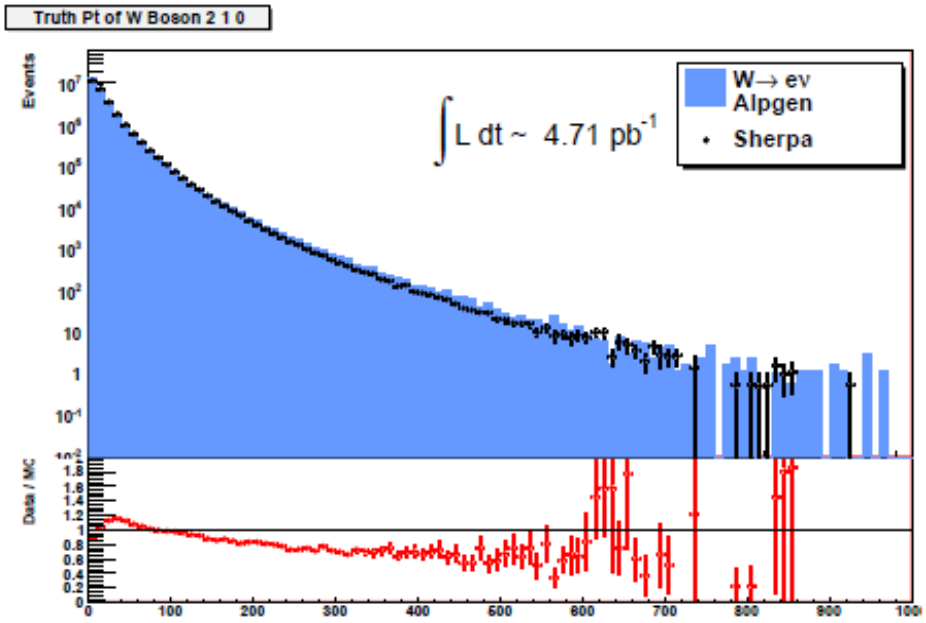
	NP0	NP1	NP2	NP3	NP4	NP5
			semileptonic $t\bar{t}$			
up	1.13	1.01	0.92	0.81	0.71	0.59
down	0.85	0.97	1.12	1.26	1.39	1.44
			dileptonic $t\bar{t}$			
up	1.09	1.00	0.96	0.82	0.73	0.65
down	0.82	0.98	1.14	1.28	1.43	1.65

We want these variations and we would like to have them consistent for all generators, experiments

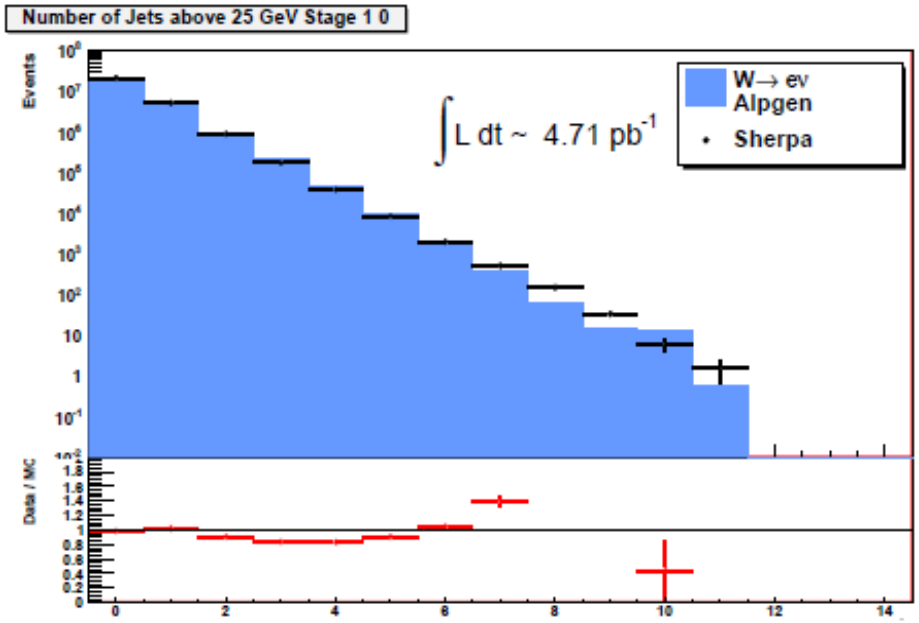
Example 2: W+jet

- Comparison of Alpgen to Sherpa
- Selection
  - One electron with $p_T > 8$ GeV and $|\eta| < 2.5$

Truth W transverse momentum



Number of jets with $p_T > 25$ GeV

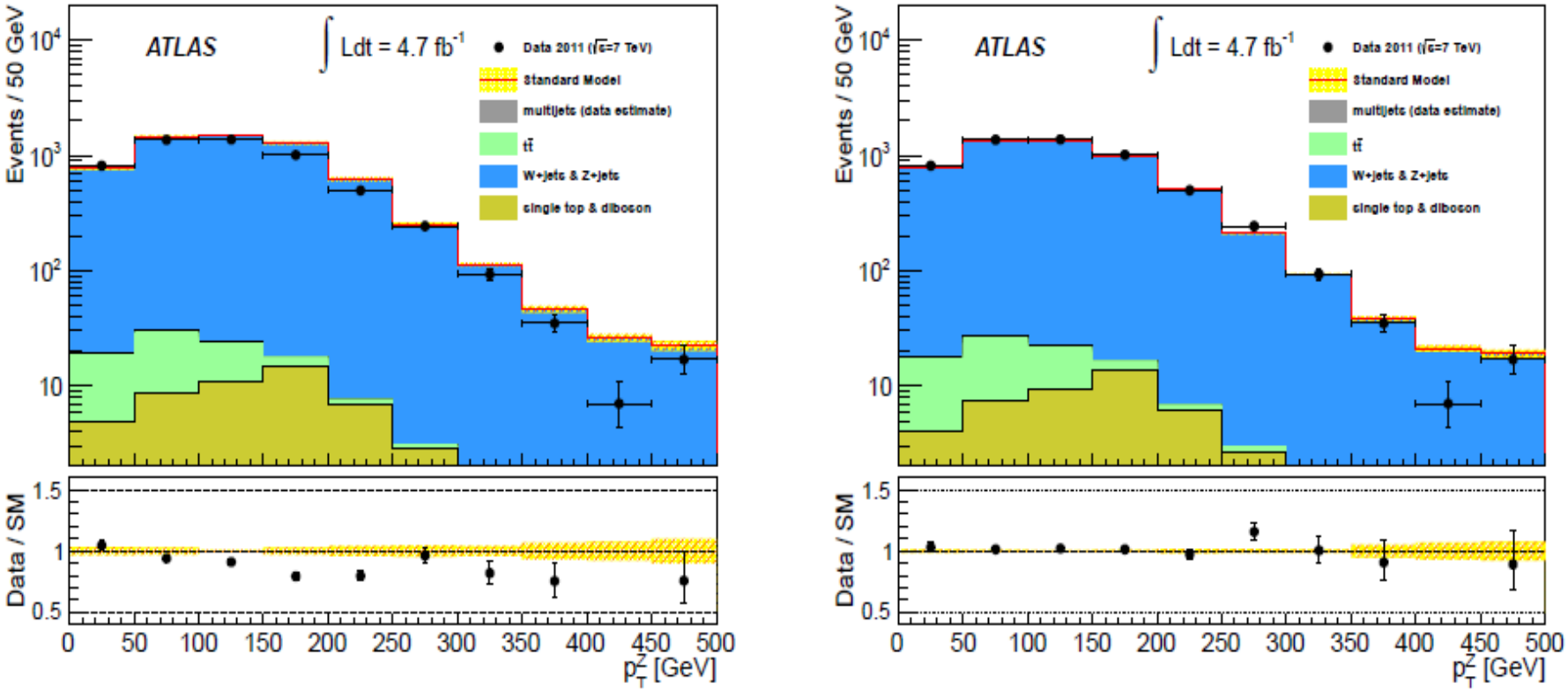


- Alpgen has more pT(W) than sherpa and more jets
- data prefer Sherpa pT(W) but Alpgen njet distribution

Example 2:W+jet

- Last time Mark Hohlfeld told that best fit is Alpgen pt(W) scaled to Sherpa
  - Instead of Alpgen/Sherpa reweighting extract correction form fit to data
    - Fit p_T^Z in data to extract reweighting for p_T^W (using Alpgen MC)
    - Use several bins of the true p_T^Z (p_T^W)

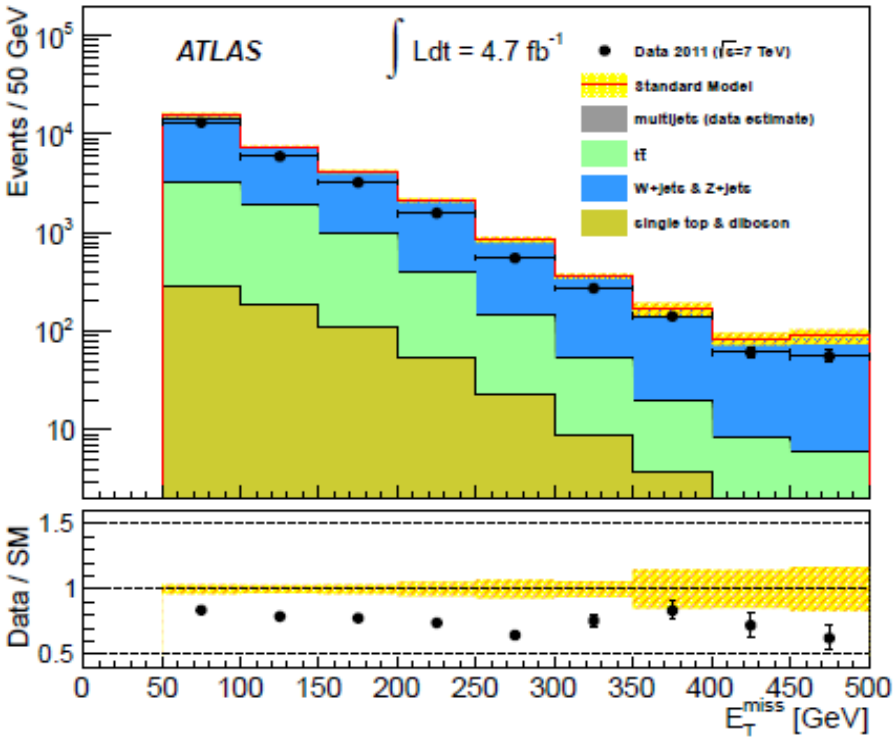
Before fit p_T^Z in Z control region After fit

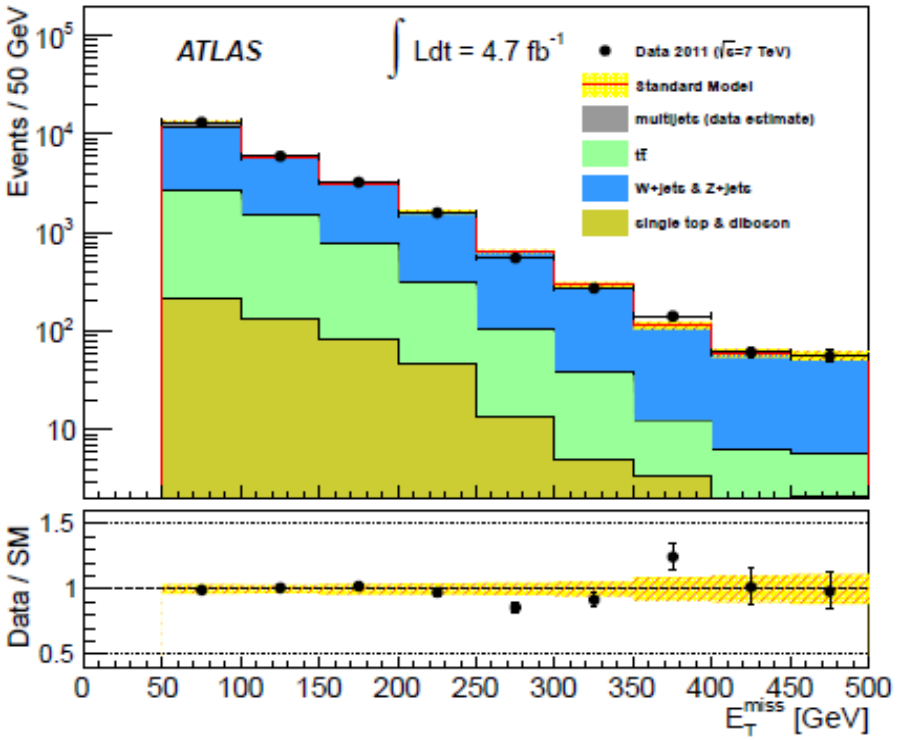


Example 2: W+jet

- Instead of Alpgen/Sherpa reweighting extract correction form fit to data
  - ▲ Fit p_T^Z in data to extract reweighting for p_T^W (using Alpgen MC)
  - ▲ Use several bins of the true p_T^Z (p_T^W)

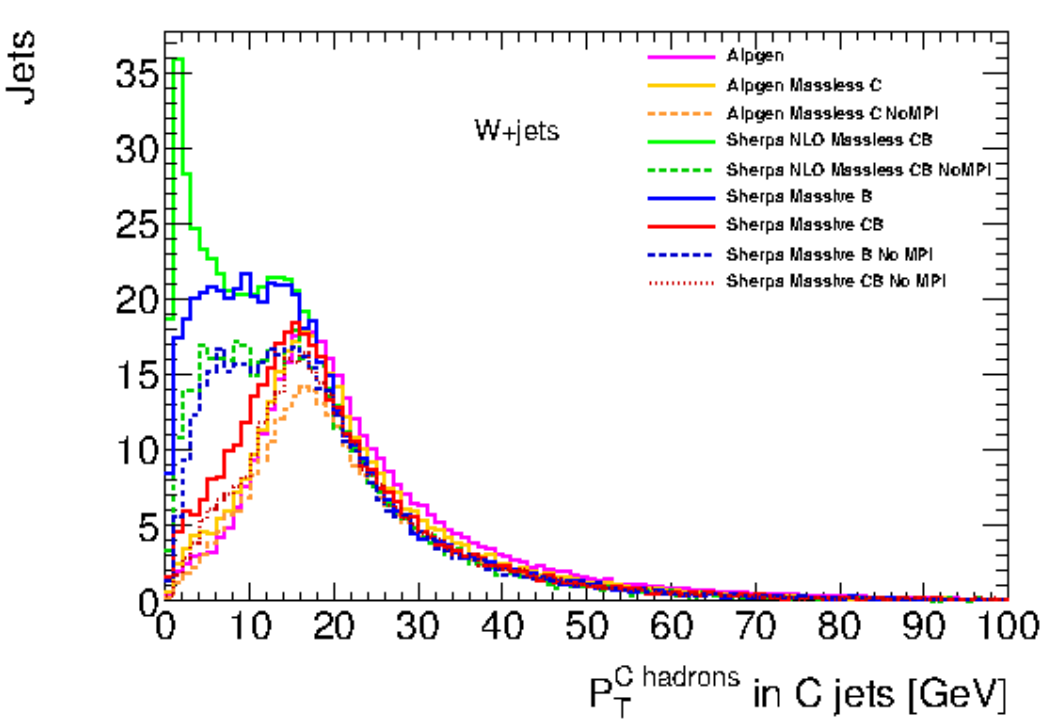
Before fit $\not{E}_T$ in W+jets control region After fit

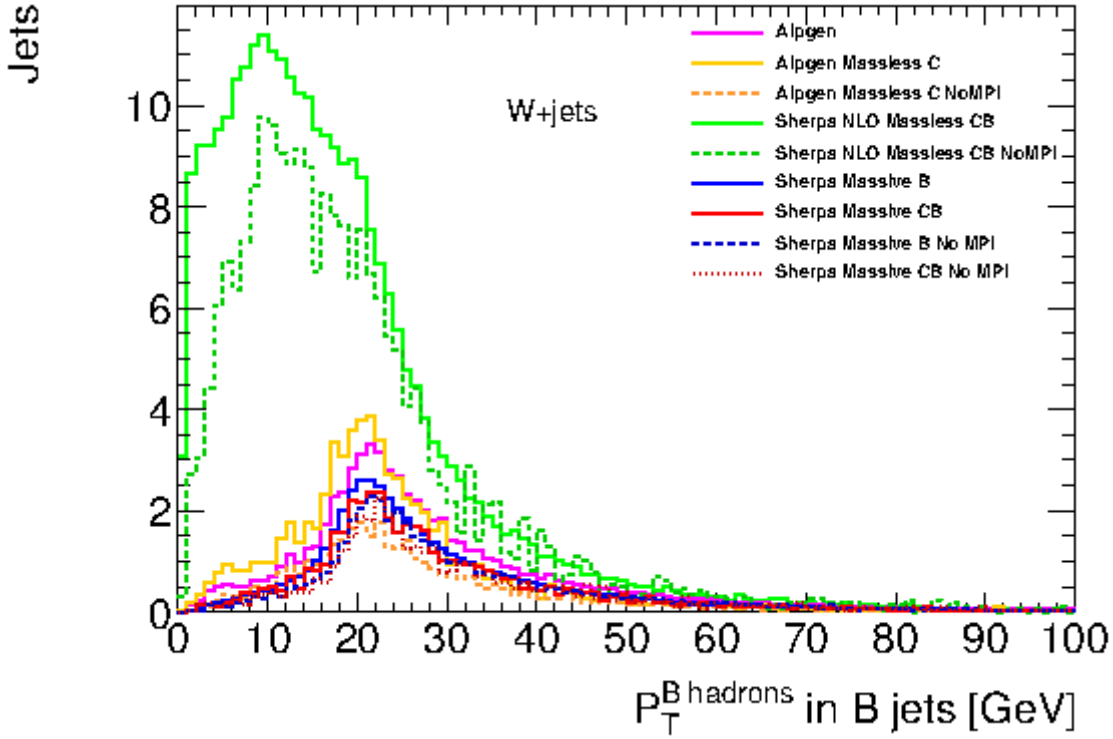




- Need to use data to tune MC (Rivet) where possible
- Cost a lot of manpower, combined effort?

Example 3: Heavy flavor treatment





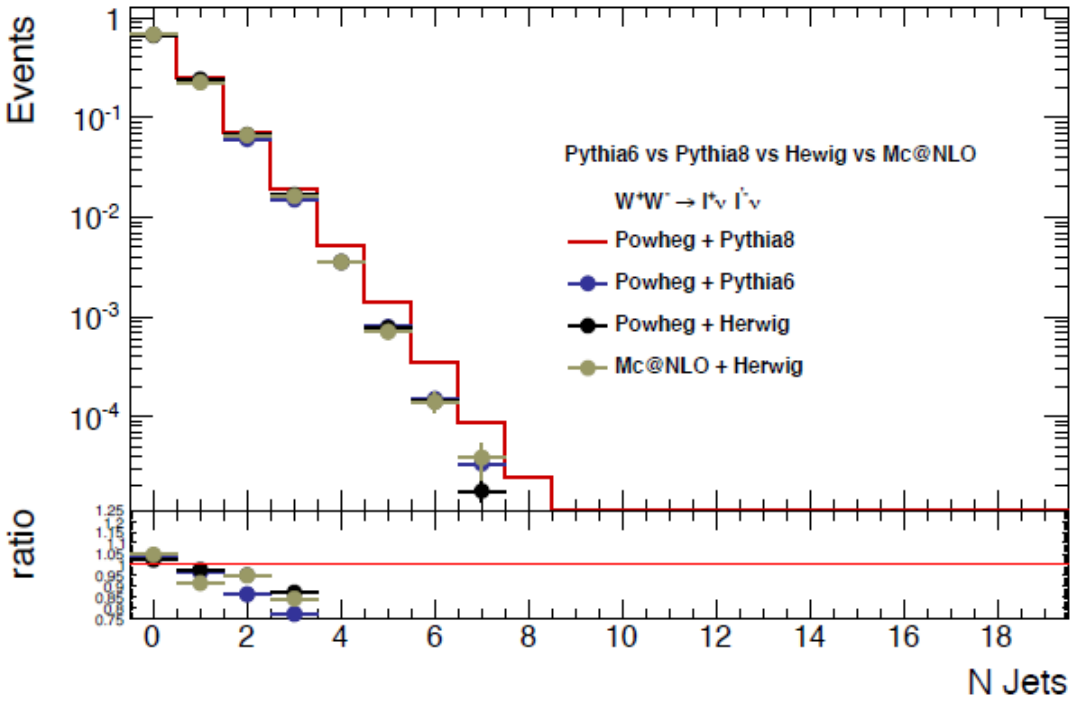
- Heavy flavour content (Hadron-Level) for different Alpgen and Sherpa Monte Carlo
- Too many heavy flavours in Sherpa NLO (massless treatment)
- Big different in treatment of massless c in Alpgen and Sherpa
- Massive heavy flavours looking very similar

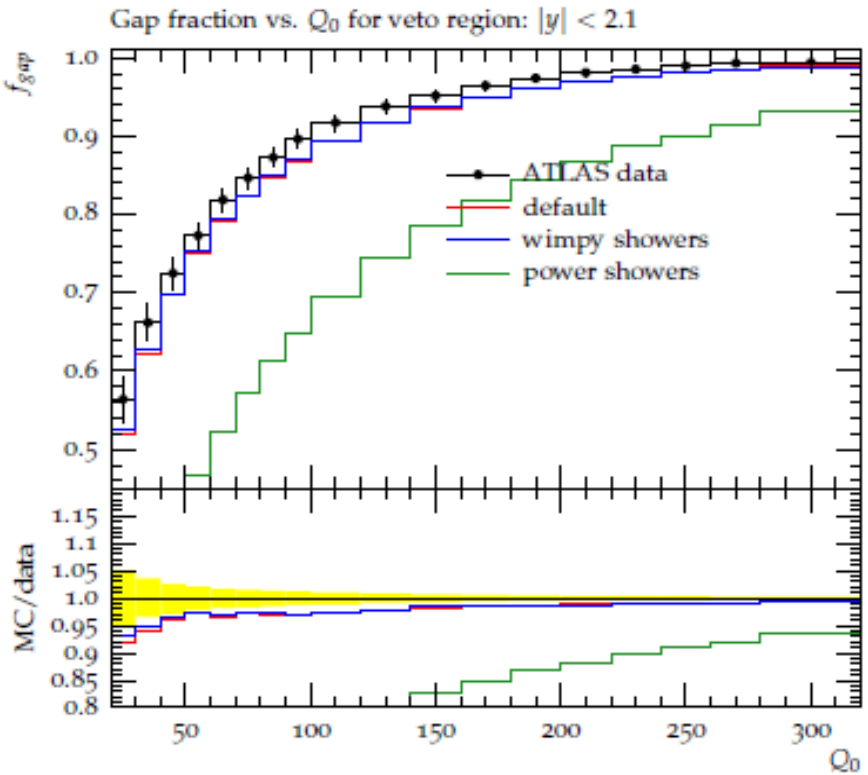
Example 3: Heavy flavor treatment

- Heavy flavour tagging is a important feature at LHC:
  - A lot of searches/measurements use it for background suppression
- Heavy flavour fractions/spectrum(fragmentation functions) and decays should be handled properly
- Very important: overlap-removal from b/c produced in shower and ME
  - Some generator have internal tool (sherpa)
  - some we developed them by our self (Alpgen), ad-hoc, not that cleanly solved but good in most use cases
  - preferable would be a possibility of variation
- Some times it is important to have tool to split sample in different flavour contents and components (from ME, shower, MPI) → data driven reweighting, tuning
- Features of the generator should be communicated (why massless c different in alpgen and sherpa)

Example 4: Pythia8 Wimpy/power shower

- Pythia8 has different default of shower treatment for different powheg lhe_f input processes
- Examples: WW (left), pythia8 behaviour very different vs others. Power shower because no real gluon in lhe_f?
- Right side: default is overwritten and power shower forced (ttbar)





Gap fraction analysis in rivet important tool to constrain shower

Summary

- We have a lot of excellent generator tools
  - newest and best calculations
  - Excellent matching between ME and shower
- We are working on a lot of new generator+shower setup (→ herwig++ and pythia8) and new NLO generator setups (sherpa NLO, aMC@NLO):
  - It would be good to have example setups compared to data or other commonly used generator setups
  - Would be good to compare to rivet if setups agree with data
- For precise data analysis the smallish things are important:
  - Heavy flavour treatment, avoidance of double counting
  - Helicities in decays (taus)
  - Use data to constrain properties (shower, W/Z pt , gap fraction)
  - PDF treatment (4 vs 5 flavor schema)
- If one of the ingredient does not work than you should not wonder that we still use ACER for single Top t-channel unfolding