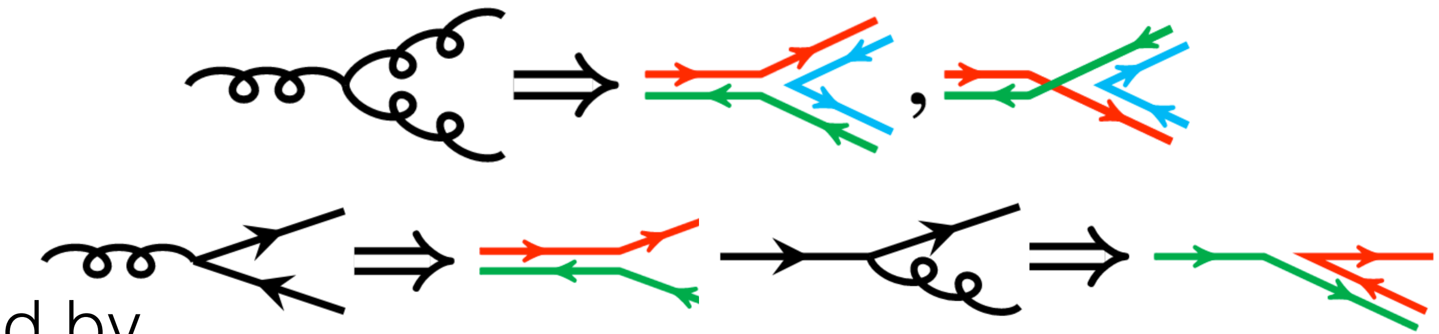
Measuring Jet Substructure Observables at the ATLAS Experiment

Jennifer Roloff, on behalf of the ATLAS collaboration

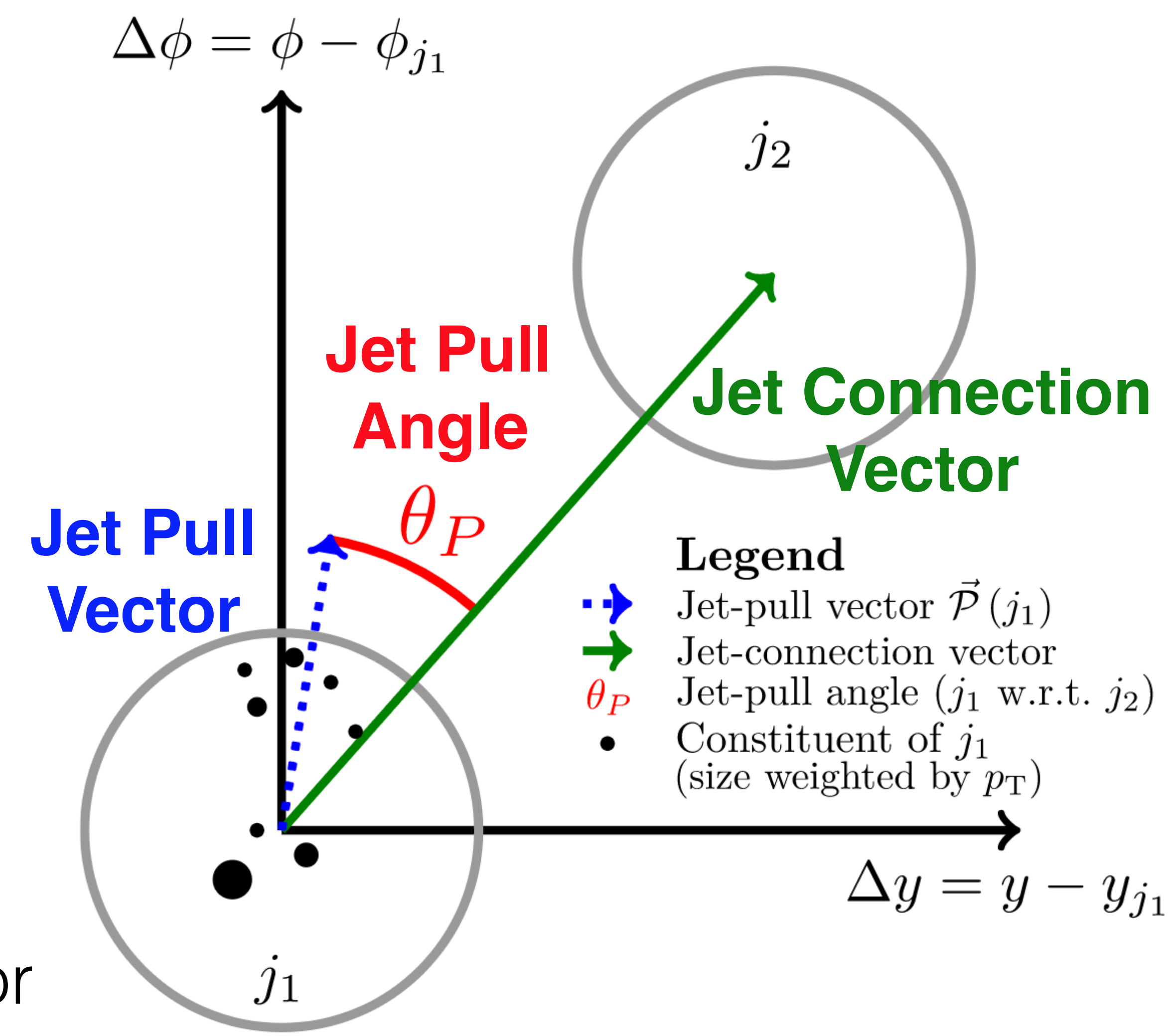
July 17, 2018

CERN-EP-2018-041

Color Flow



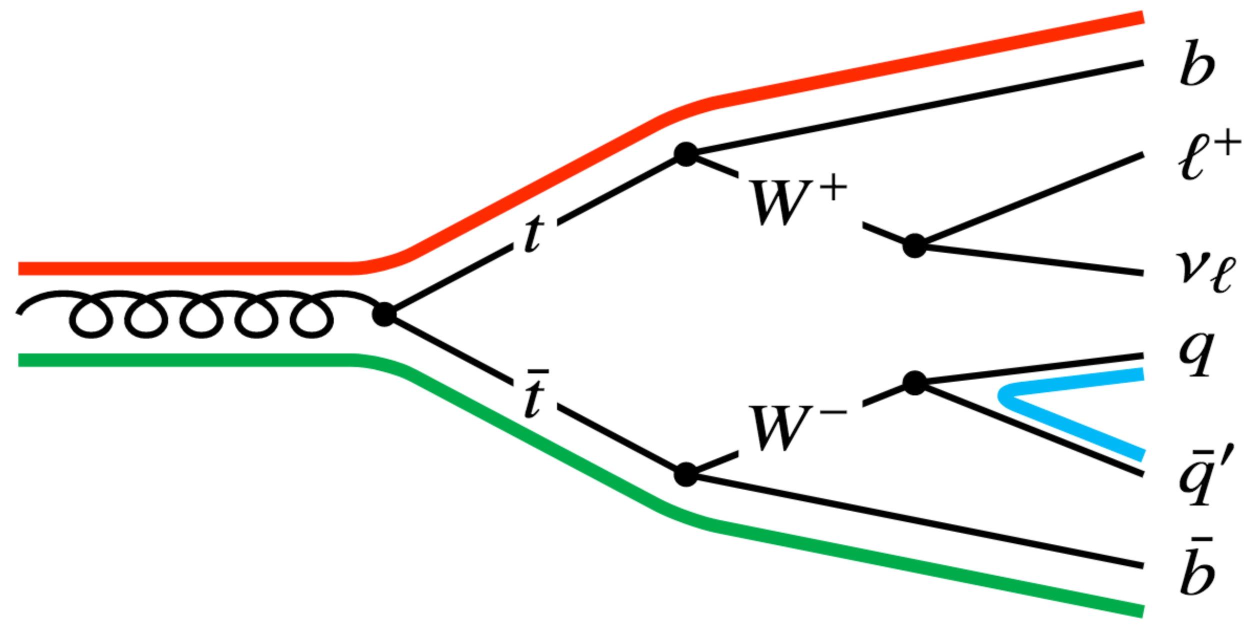
- QCD radiation can be approximated by **color strings connecting quarks and gluons**
  - Color charge flows from initial state particles to stable particles
  - Connections exist between initial particles and color-neutral hadrons
- Color connection predicts more radiation where these color strings are
  - Can use the **jet pull** to measure the color flow
- Color connections are poorly constrained in QCD

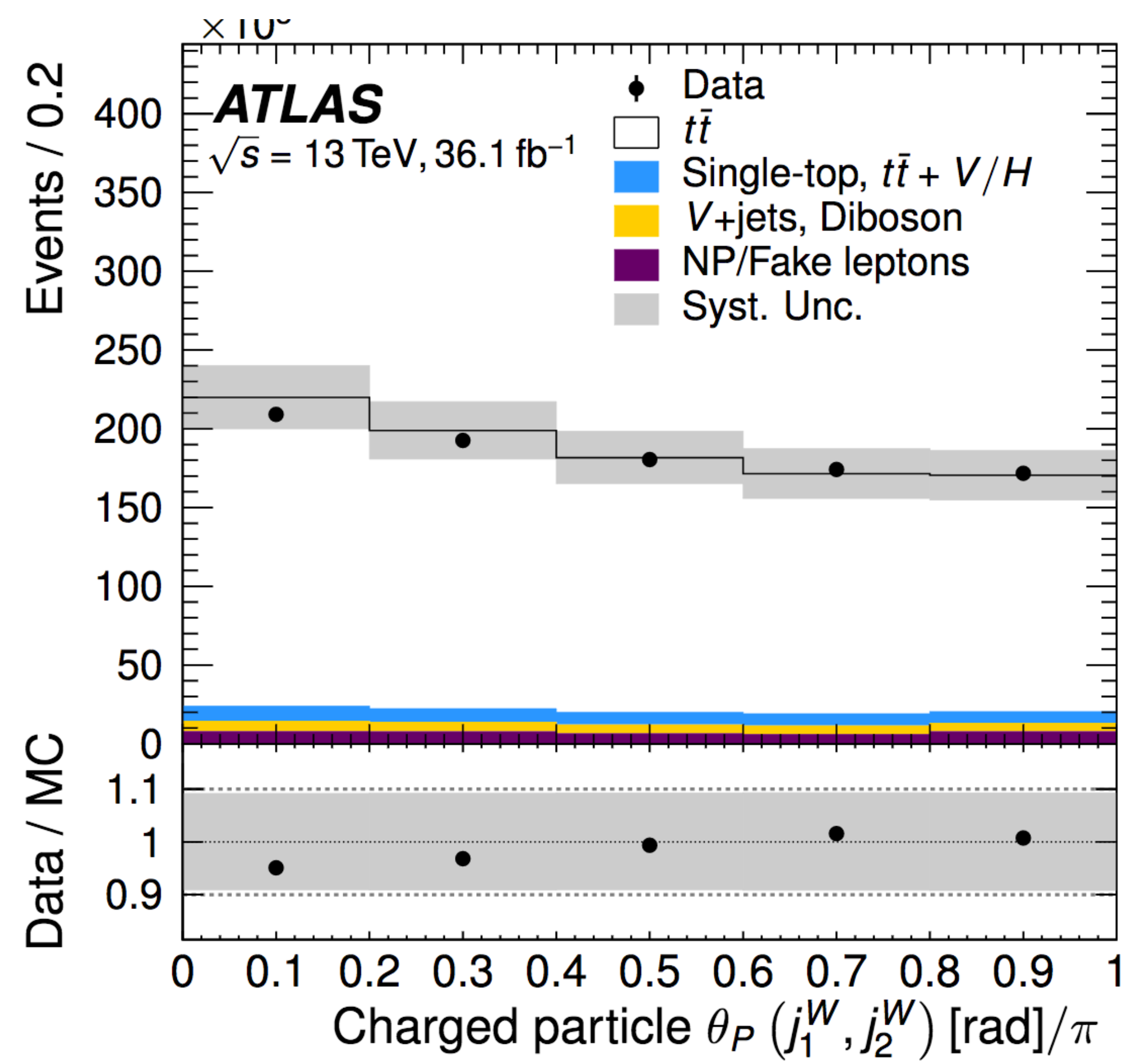


$$\vec{\mathcal{P}}(j) = \sum_{i \in j} \frac{\left|\vec{\Delta r}_i\right| \cdot p_\mathrm{T}^i}{p_\mathrm{T}^j} \vec{\Delta r}_i$$

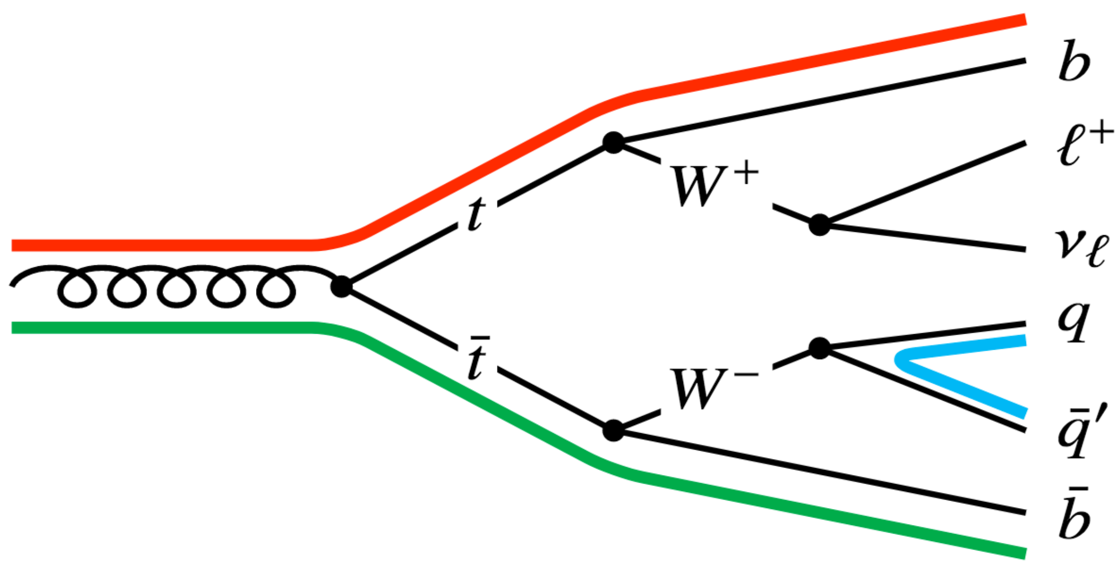
Measurement of Color Flow in ttbar events

- ttbar events can contain both color connected and non-connected jets
  - Can select these events with high purity
  - W is neutral, so its **decay products are color connected**
  - The **b-quarks** from each top decay are **not connected** to each other
- Calculating the jet pull using **tracks**
  - Provides better spatial resolution than from clusters
  - Ghost associate tracks to R=0.4 jets made from clusters
  - Recalculate the jet axis from the ghost-associated tracks
- Measuring both the **pull angle** and the **magnitude of the jet pull**





Event Selection and Unfolding



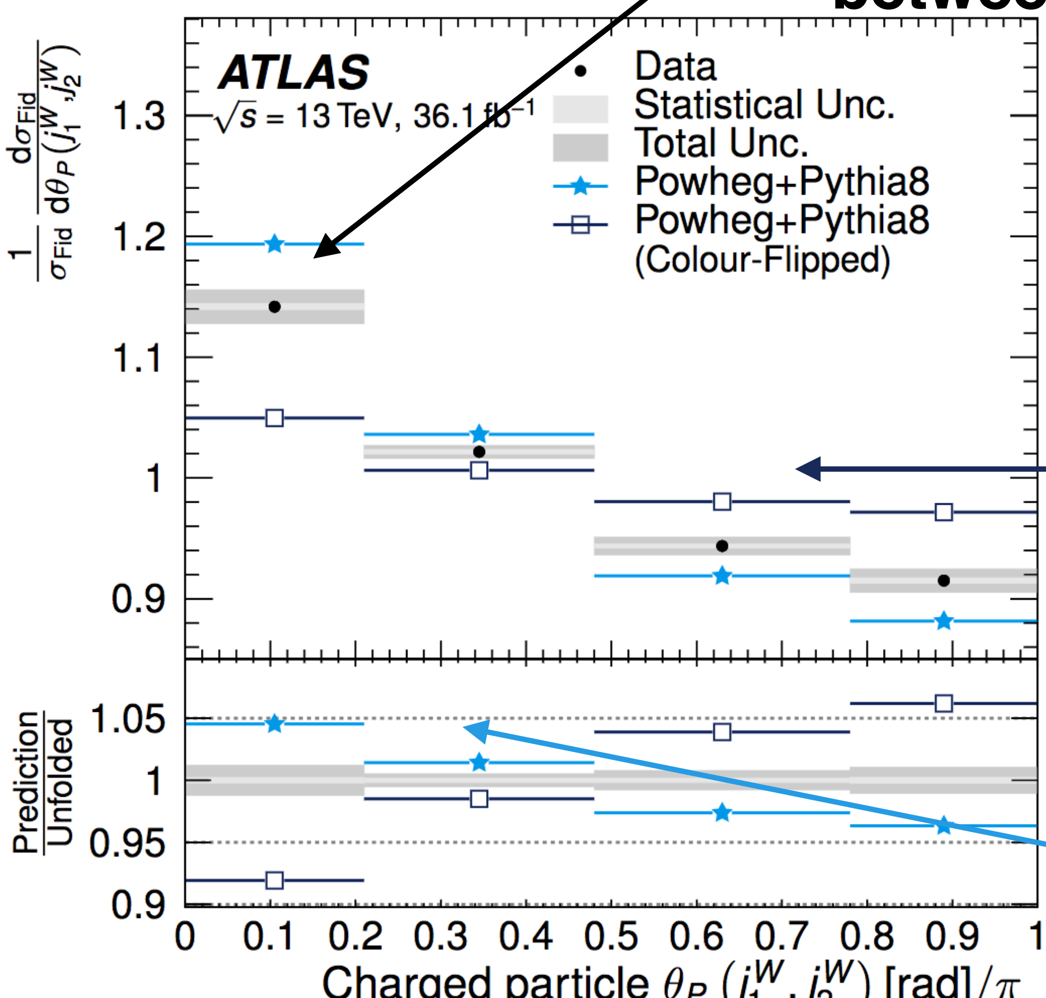
Sample	Yield		
$t\bar{t}$	1 026 000	±	95 000
$t\bar{t}V$	3270	±	250
$t\bar{t}H$	1700	±	100
Single-top	48 400	±	5500
Diboson	1440	±	220
W + jets	27 700	±	4700
Z + jets	8300	±	1400
NP/Fake leptons	53 000	±	30 000
Total expected	1 170 000	±	100 000
Observed	1 153 003		

- Triggering on a **single-electron** or **single-muon**
- Event must contain 4 jets, two of which are b-tagged
- Require MET of at least 20 GeV
- Background determination using MC, except for the NP and fake lepton background
  - Signal purity of around 88%
- **Background contributions are subtracted** bin-by-bin, then the **distribution is unfolded**
  - Using bayesian unfolding to particle level
  - Particle level jet pull calculated in similar way as reconstructed jets, using charged particles

Measurement of Color Flow in ttbar events

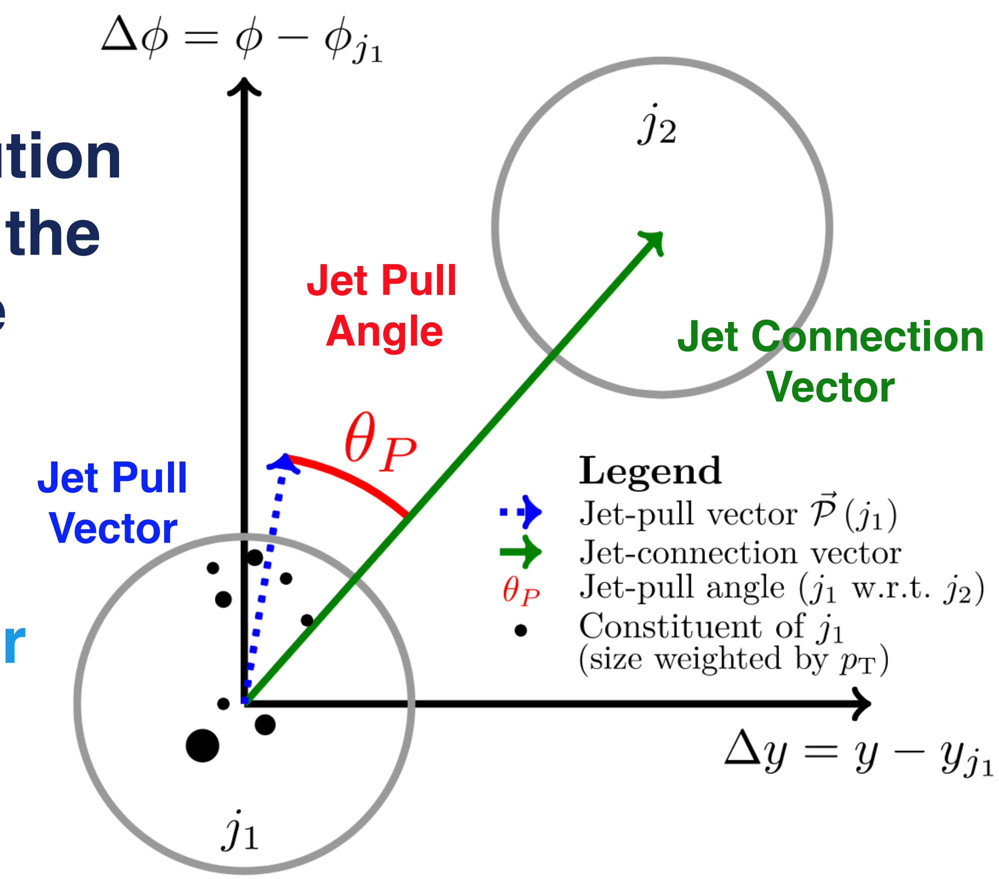
- Comparing against a **color-flipped model** to see sensitivity to color flow
- Color-connected model models the data better than the color-flipped model for both color-connected and unconnected jets, though neither are perfect

Data closer to the singlet model overall, but falls somewhere between the two



Fairly flat distribution for model where the daughters are unconnected

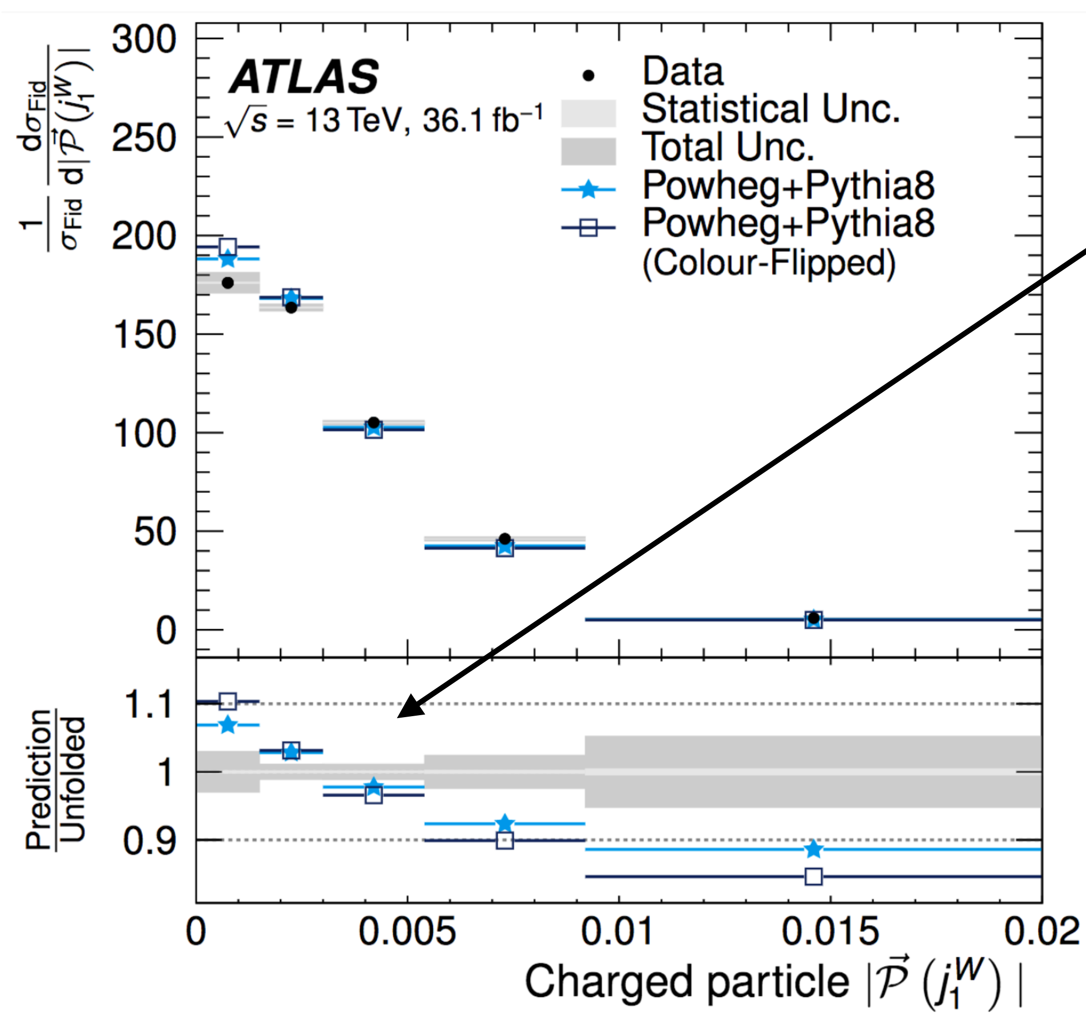
Powheg+Pythia8 predicts more color flow than is measured in data



Forward Jet Pull Angle

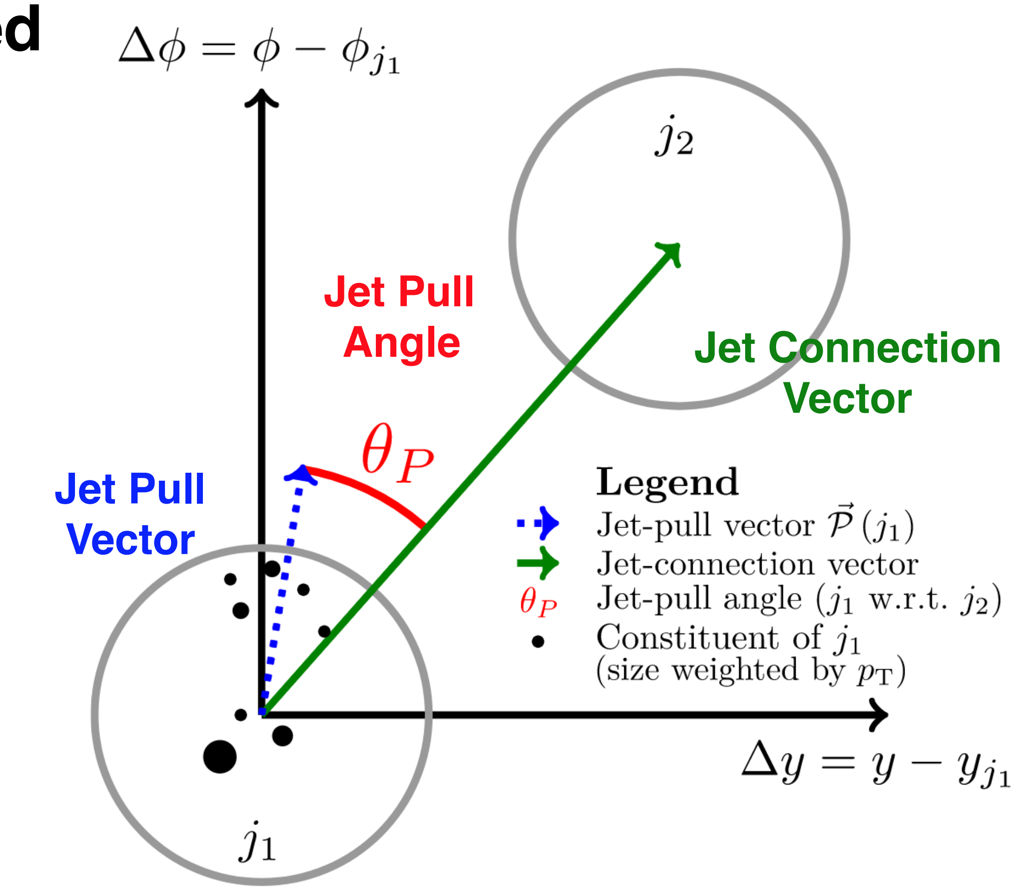
Measurement of Color Flow in ttbar events

- Comparing against a **color-flipped model** to see sensitivity to color flow
- Color-connected model models the data better than the color-flipped model for both color-connected and unconnected jets, though neither are perfect



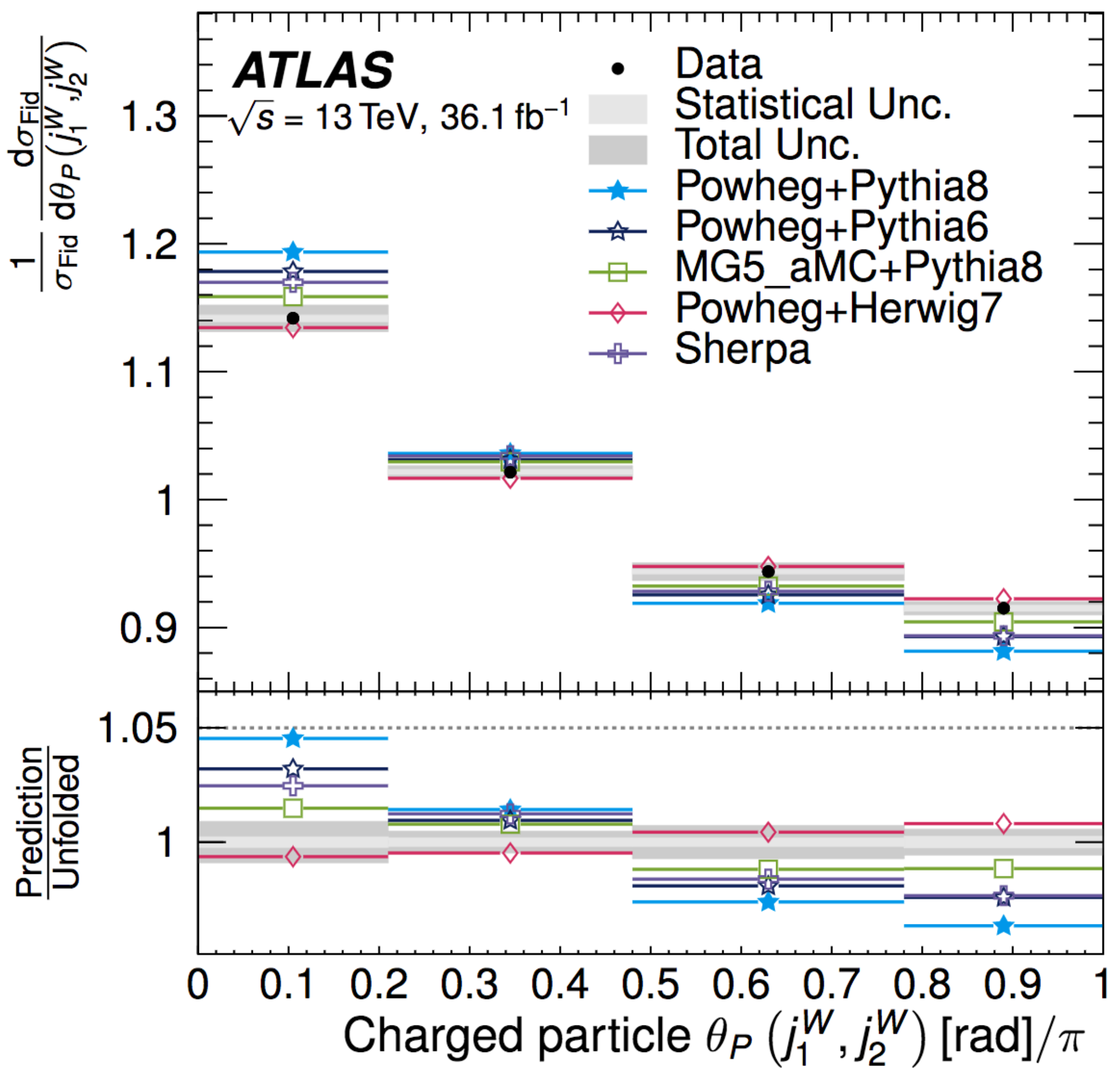
Less discrimination between standard and color-flipped models

Jet Pull Magnitude

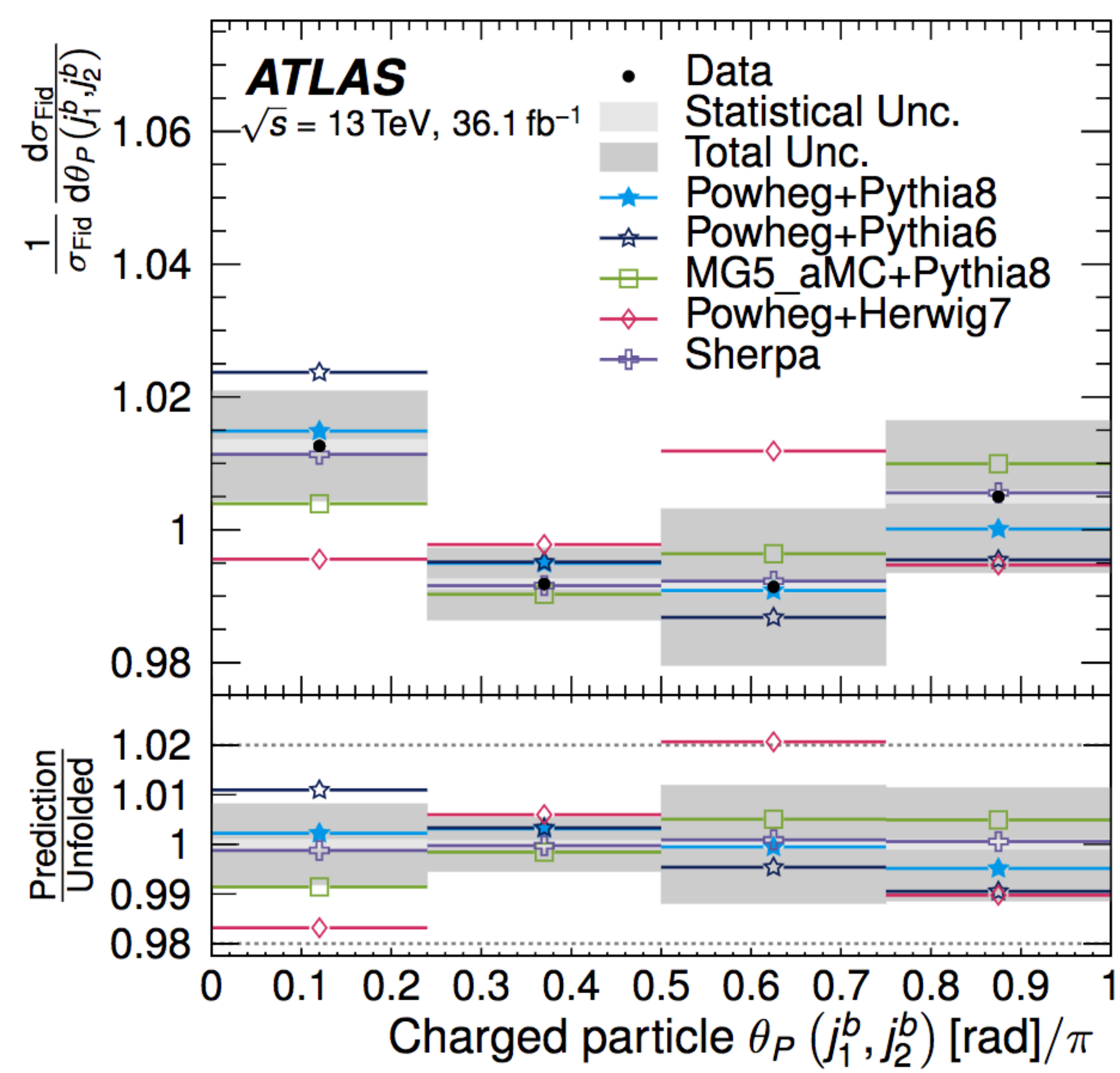


Measurement of Color Flow in ttbar events

- Comparing to several SM predictions, with combinations of different hard-scatter predictions and different hadronization models
  - Generally poor agreement from Powheg+Pythia8, and better agreement from Powheg+Pythia6 and Powheg+Herwig7
- Almost all of them **predict more color flow** than is measured in data



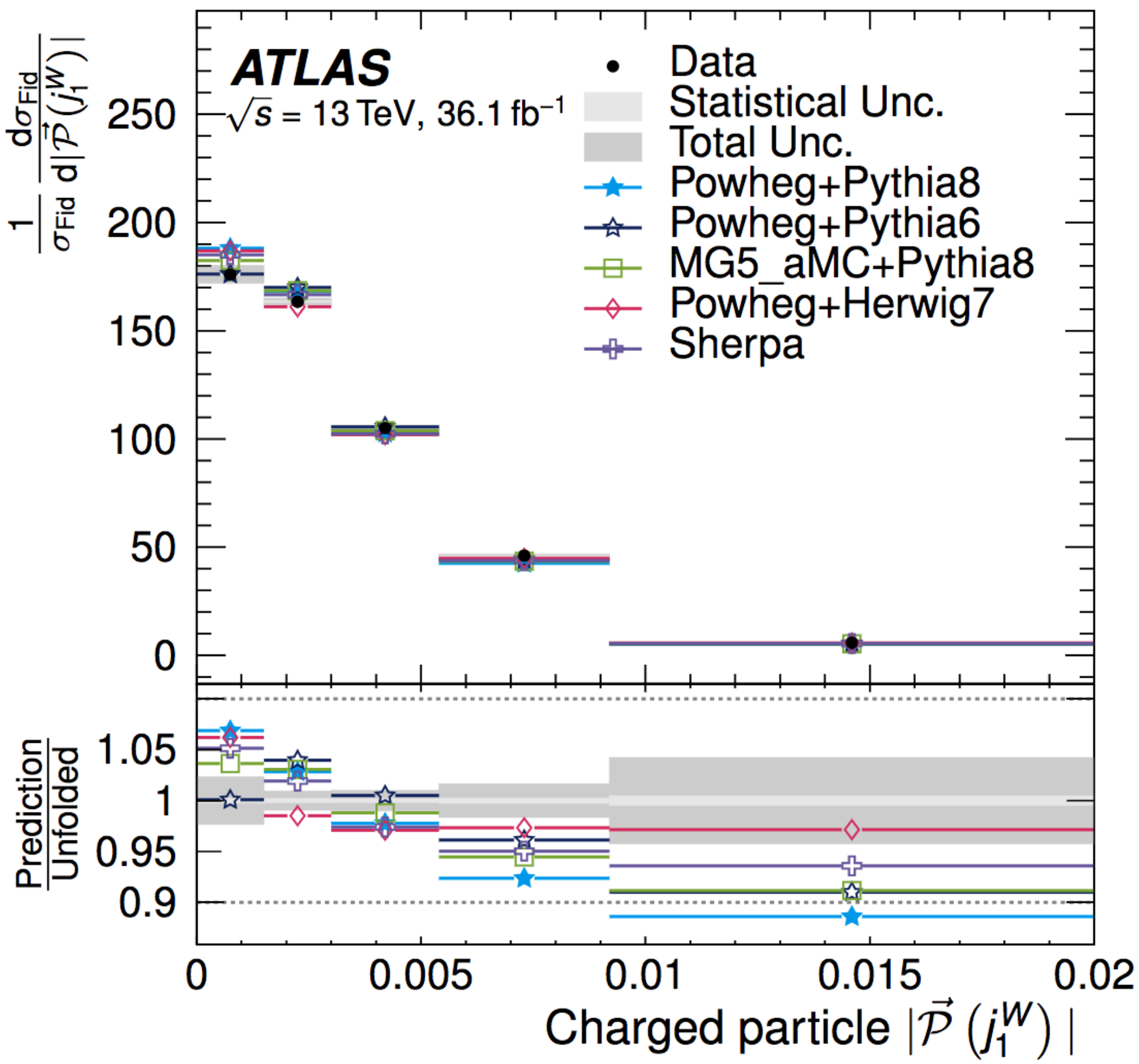
W-decay products (Color-connected)
Forward Pull Angle



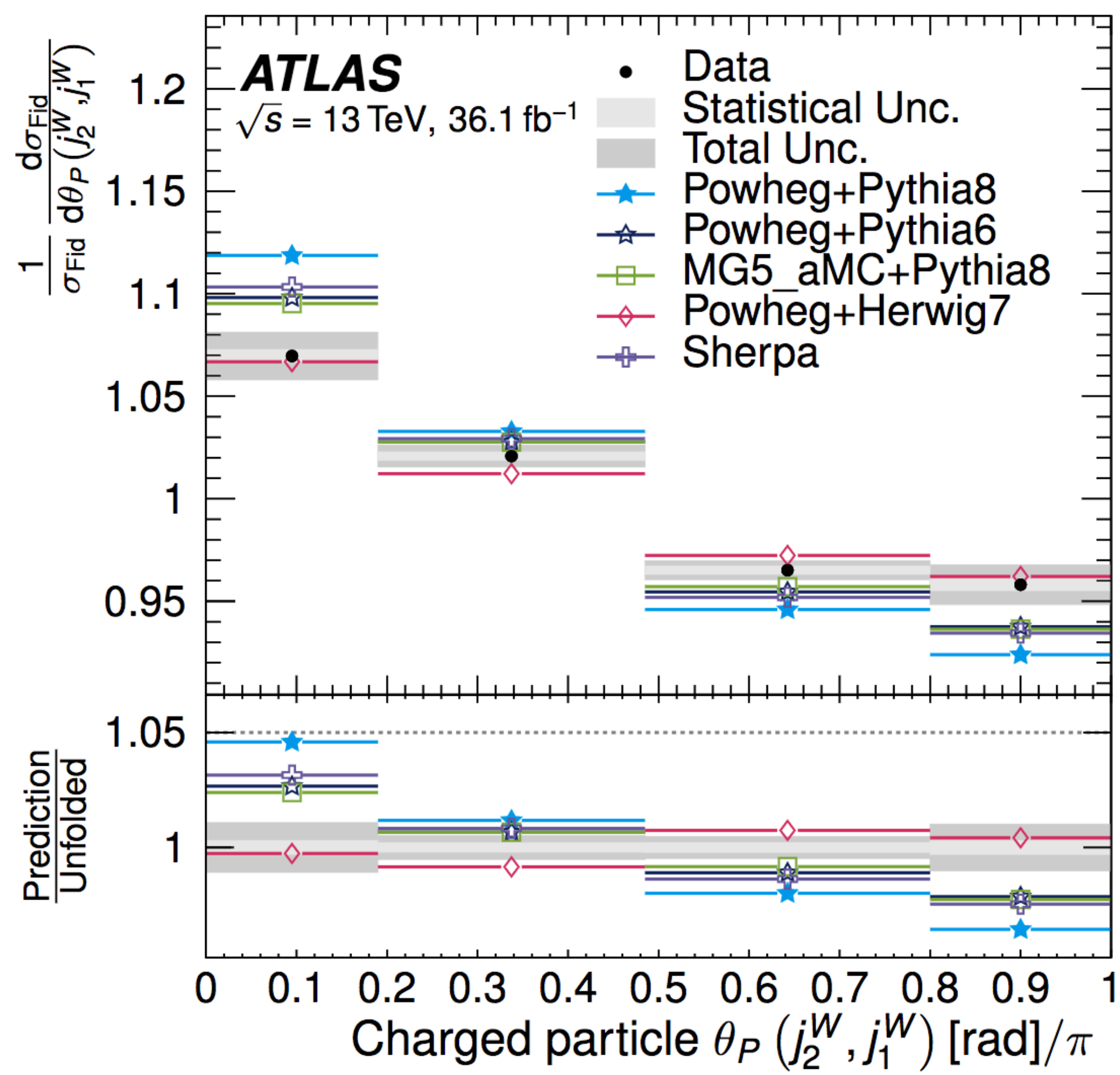
b-quarks (not color connected)
Forward Pull Angle

Measurement of Color Flow in ttbar events

- Magnitude of the pull vector is not modeled well by most of the predictions
- Backward pull angle is slightly smaller than the forward pull angle, but accuracy of modeling is similar



Magnitude of the pull vector

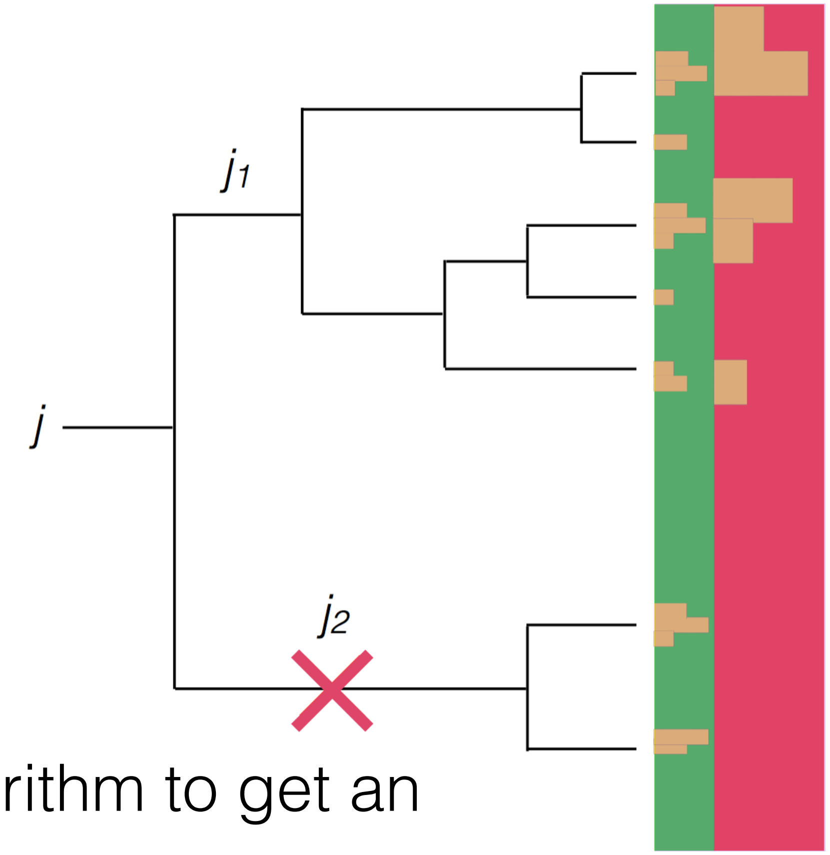


Backward pull angle

A Measurement of the Soft Drop Jet Mass

Soft Drop and the Jet Mass

CERN-EP-2017-231



- The jet mass is defined as $(m_{jet})^2 = (\Sigma E)^2 - (\Sigma p)^2$
    - Dominated by **resummation** and **not fixed-order**
- Run jet finding using the **anti-k_t algorithm**
- Recluster its constituents with the Cambridge/Aachen algorithm to get an **angular-ordered shower history**
- Starting from the last branch of the clustering history, check if

$$\frac{min(p_{T,j1}, p_{T,j2})}{(p_{T,j1} + p_{T,j2})} > z_{cut}(\frac{\Delta R_{j1,j2}}{R})^{\beta}$$

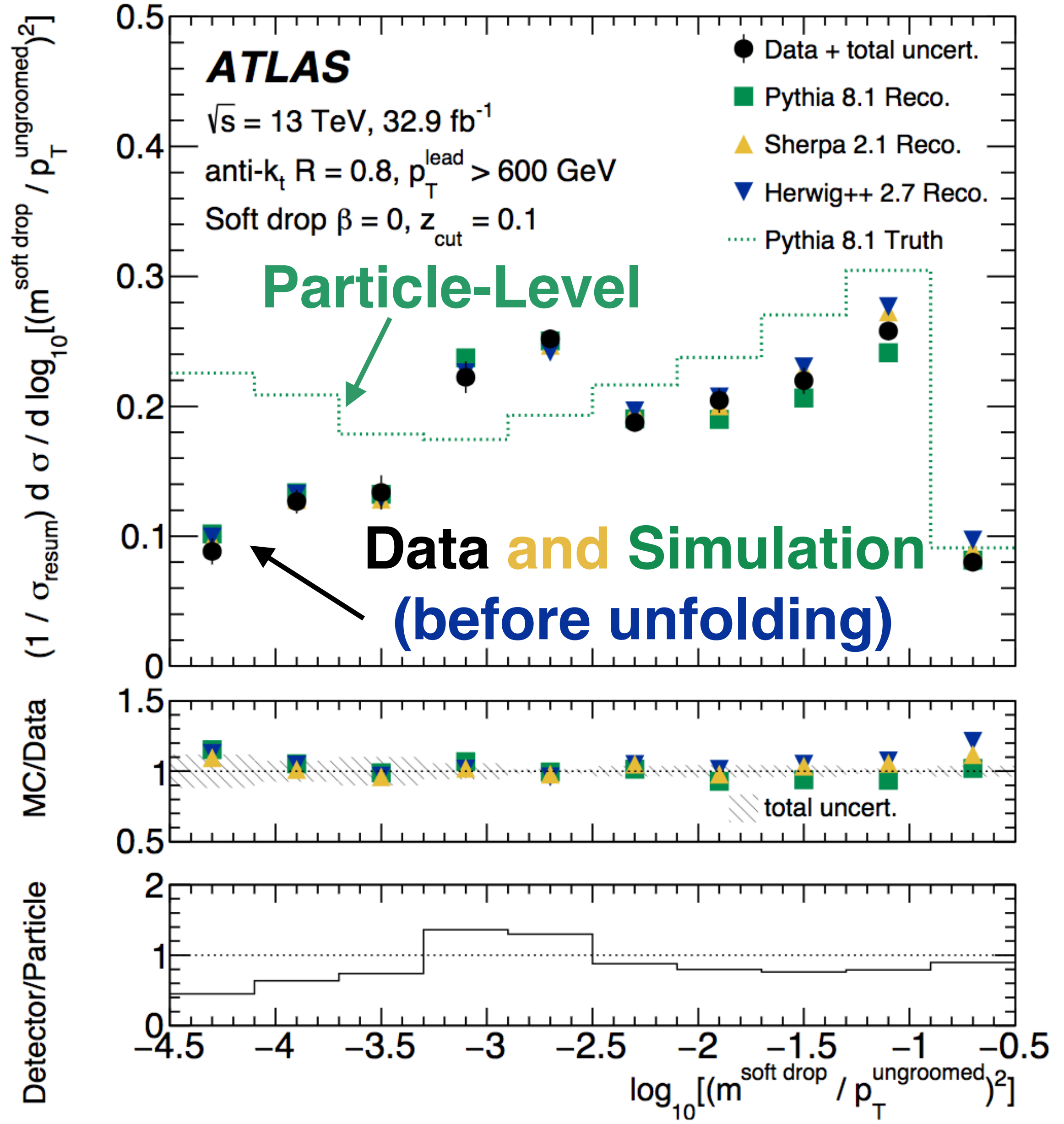
    - If so, the algorithm ends, and the jet has been groomed
    - If not, the softer branch (*j2*) is removed, and this process is repeated with the harder branch (*j1*)
- z_{cut} sets the **scale of energy removal**, only measuring $z_{cut} = 0.1$
- β determines the **sensitivity to wide-angle radiation**, measuring β = 0,1, and 2

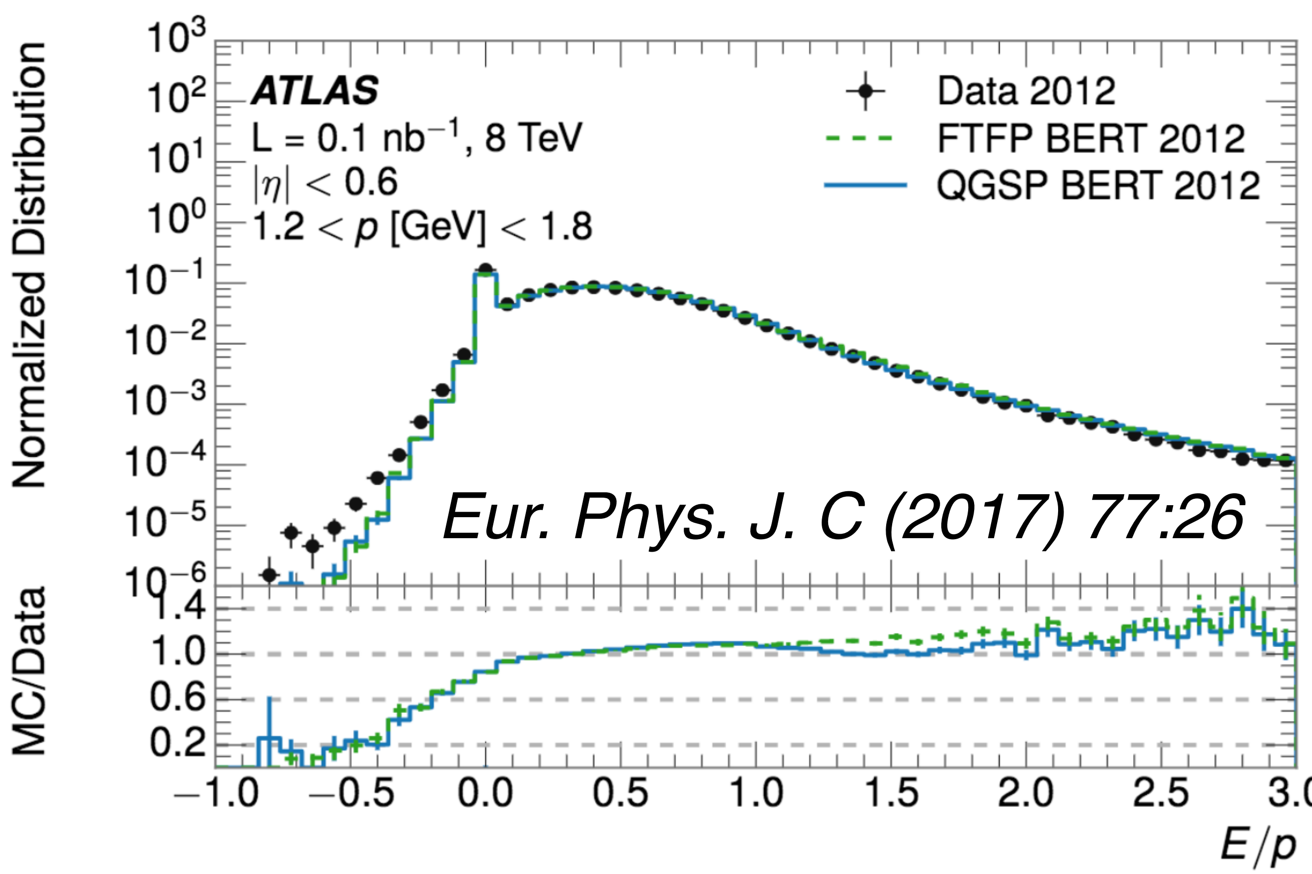
Event Selection and Unfolding

- Using the lowest unprescaled jet trigger
- Apply a **dijet selection** of $p_T^{lead} < 1.5 * p_T^{sublead}$, and require $p_T^{lead} > 600$ GeV
- Apply **Soft Drop grooming** to R=0.8 jets made from locally calibrated calorimeter-cell clusters
- Measuring $\boldsymbol{\rho} = \log[(m^{\text{Soft Drop}} / p_T^{\text{Ungroomed}})^2]$ instead of the mass directly
  - This depends logarithmically on pT, so the shape is less dependent on the pT binning
  - Using a log-scale binning to better understand the region of interest (the resumption region)
  - Using the ungroomed jet p_T, since it is collinear safe for β=0
- **Simultaneously unfold $\boldsymbol{\rho}$ and $\mathbf{p_T}$** using Bayesian unfolding
  - Only showing final results binned inclusively in p_T, individual p_T bins available on hepdata
- **Normalize** between $-3.7 < \rho < -1.7$ (the resummation region)

Cluster-Based Uncertainties

- The calorimeter-based jet mass is affected by **non-trivial detector corrections**
  - Need to understand modeling of clusters to derive systematic uncertainties
- Use ratio of the cluster energy to the track momentum (**E/p**) to characterize the clusters
  - Cluster efficiency determined by the percentage of clusters with E/p = 0 (no cluster matched to the track)

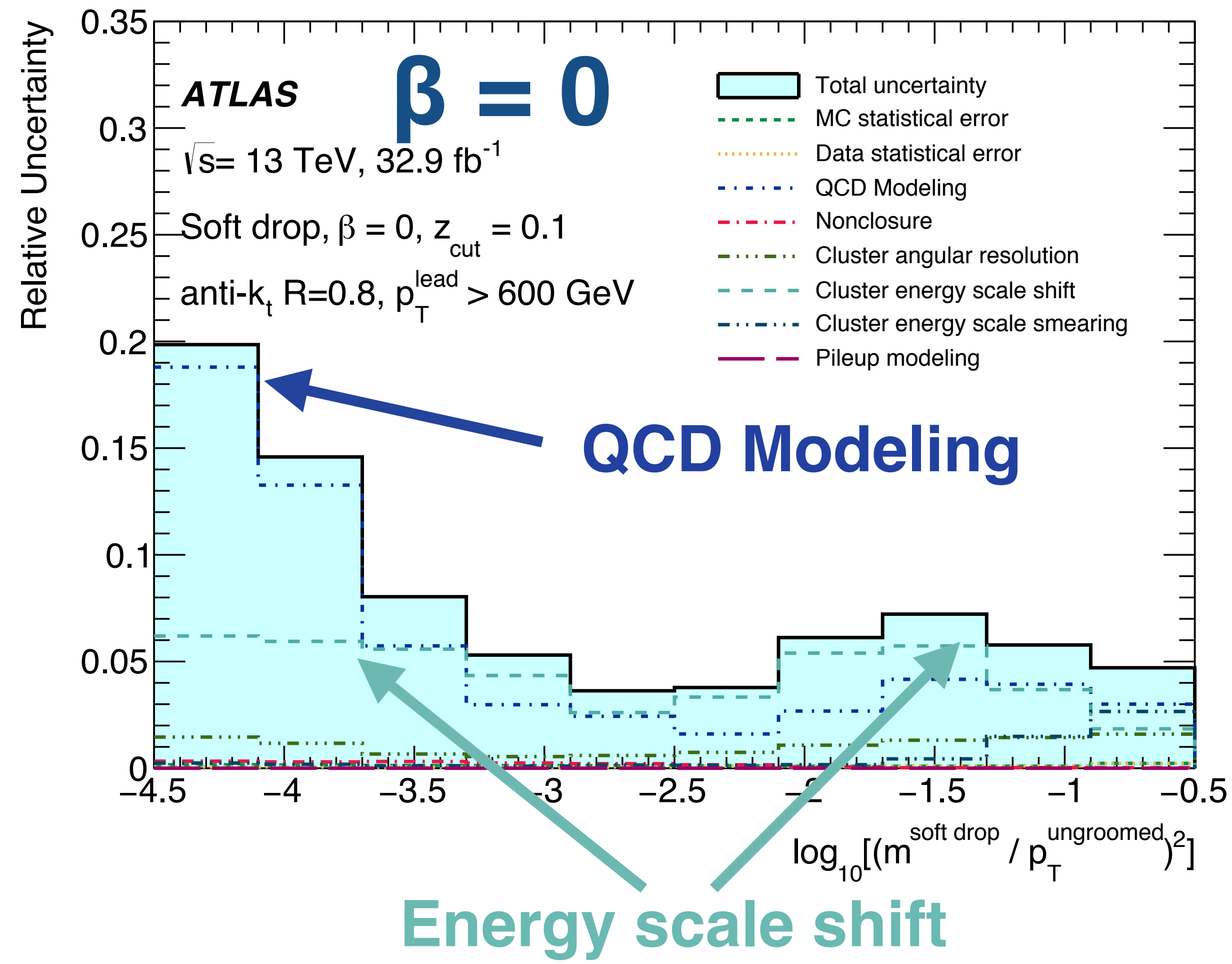


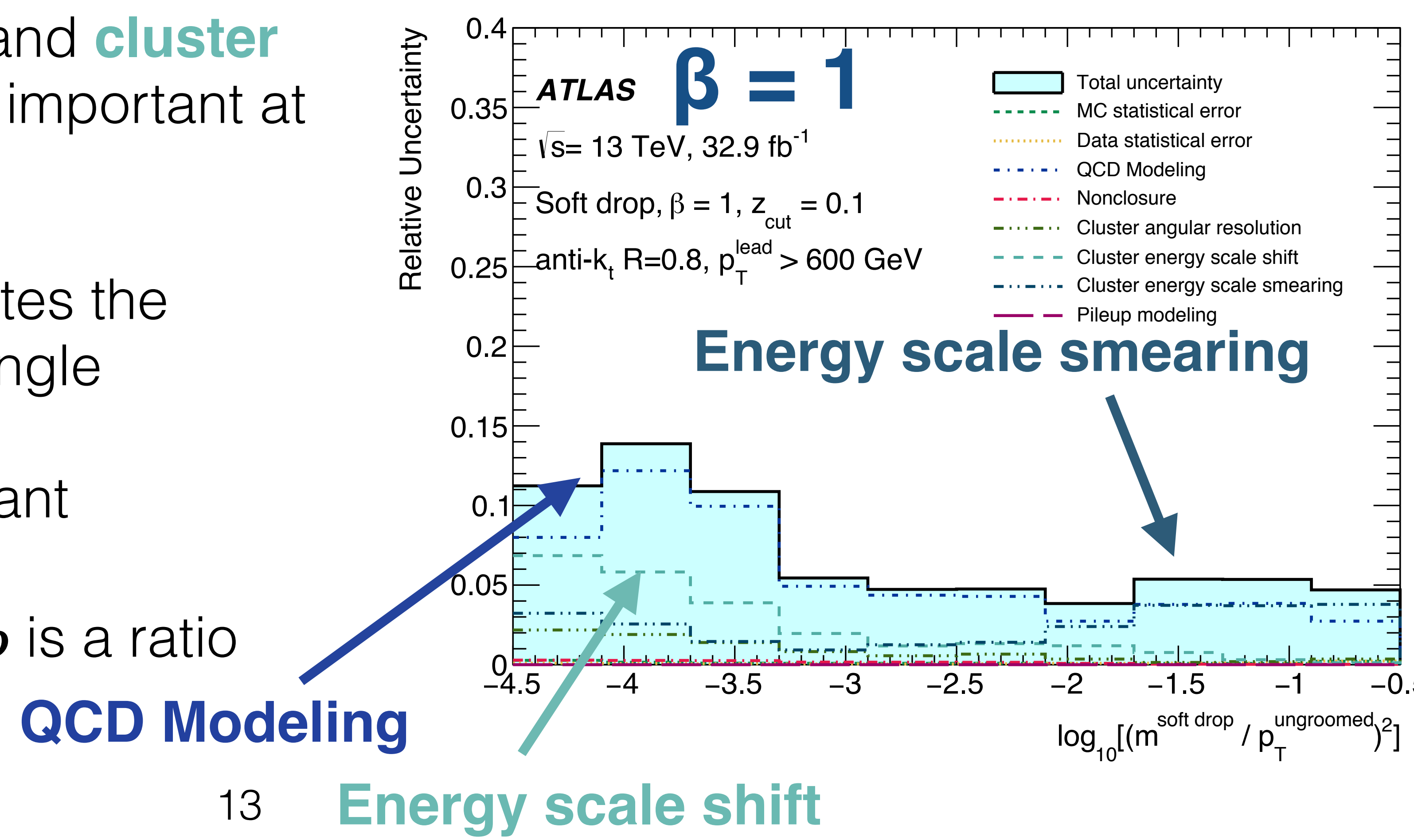


- Mean and standard deviation used for the cluster energy scale and resolution
- Cluster angular resolution determined by radial distance between tracks and clusters

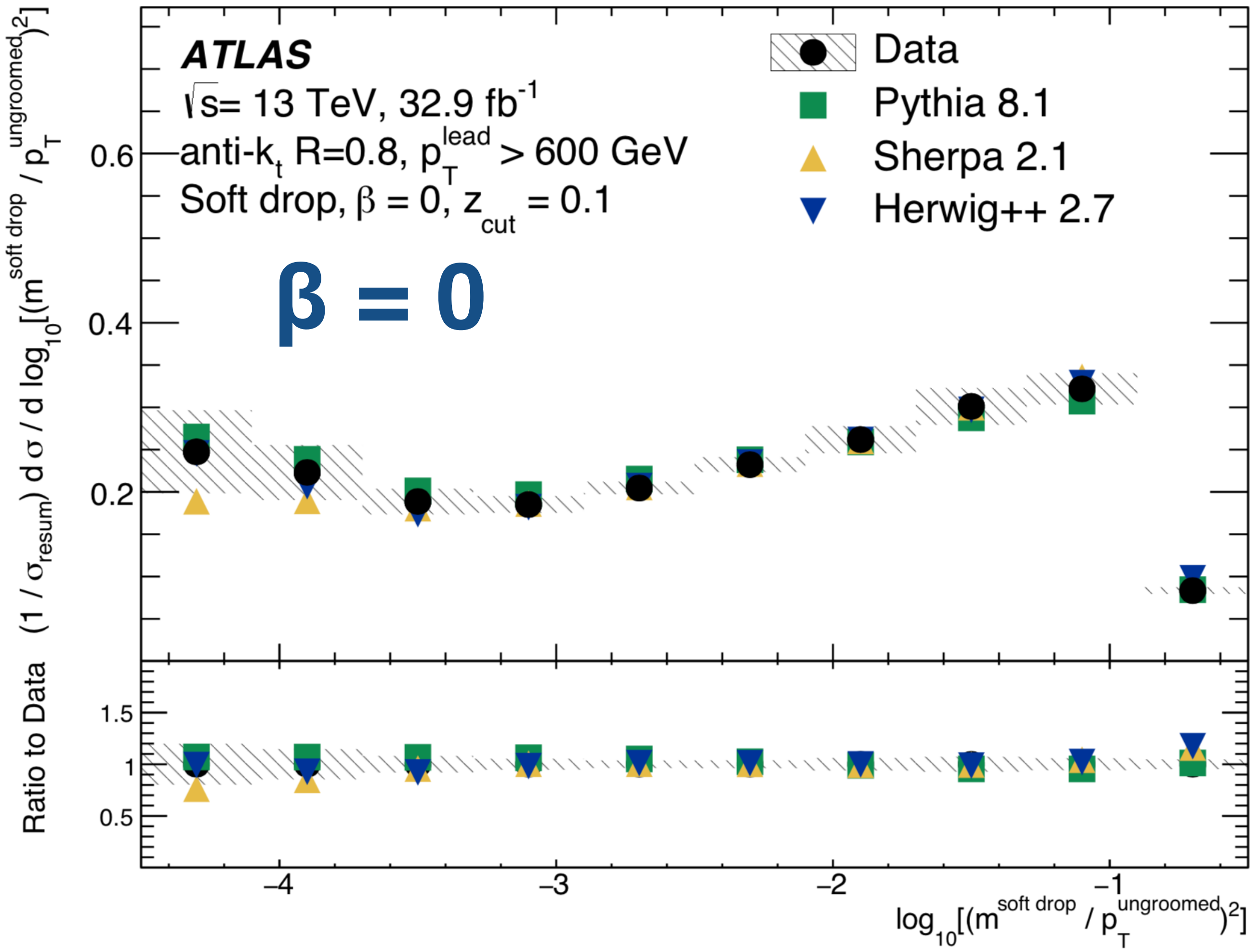
Soft Drop Mass: Uncertainties

- **QCD modeling** uncertainties dominate over the entire mass range
  - Particularly large at low mass where non-perturbative effects are largest
- **Cluster energy scale shift** uncertainty large at lower masses
  - Low cluster multiplicity
- **Cluster energy scale smearing** and **cluster energy scale shift** become more important at higher masses
  - Energy of hard prongs dominates the mass instead of the opening angle
- Other uncertainties are subdominant
- Many uncertainties cancel since ρ is a ratio



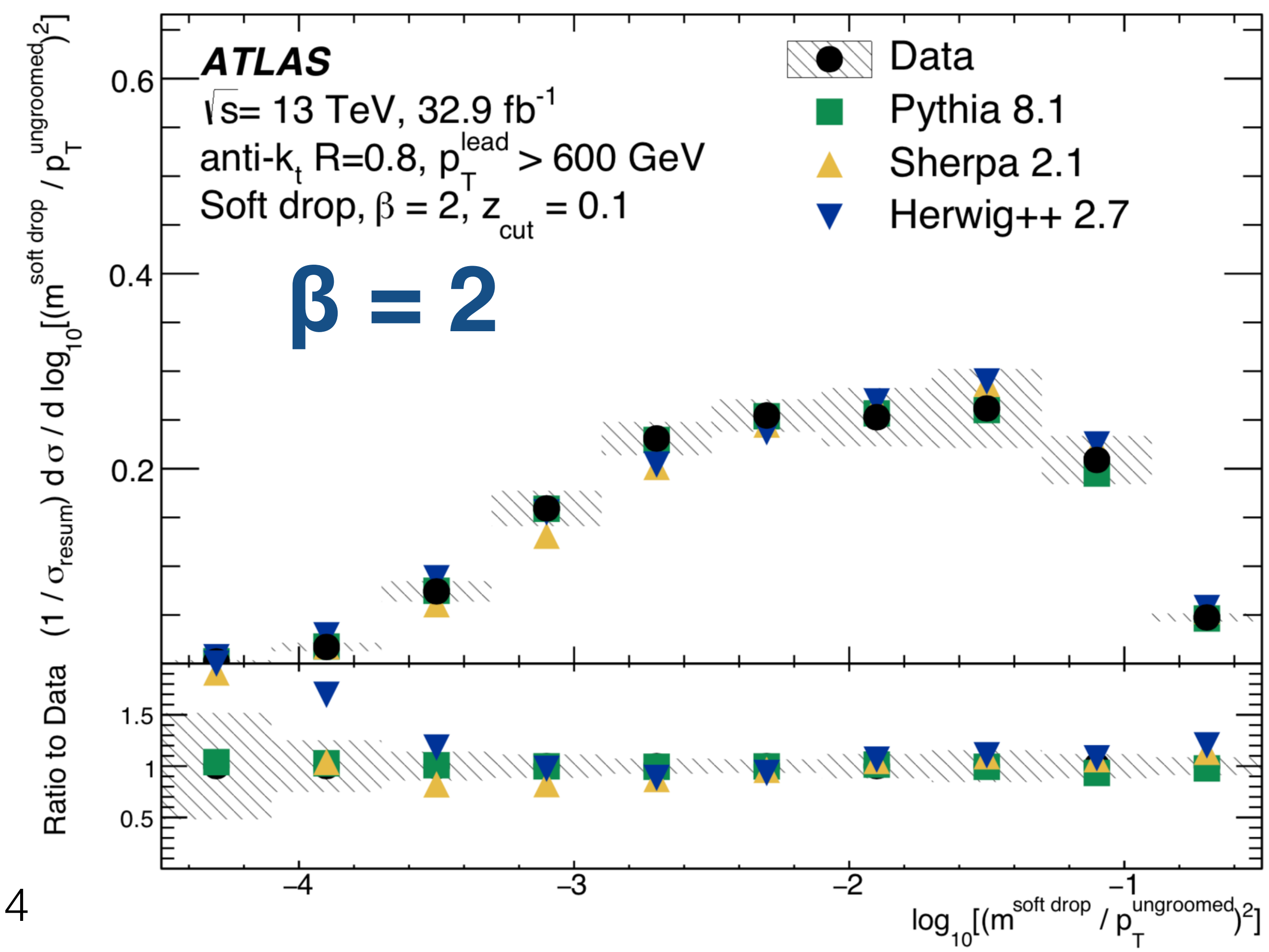


Soft Drop Jet Mass

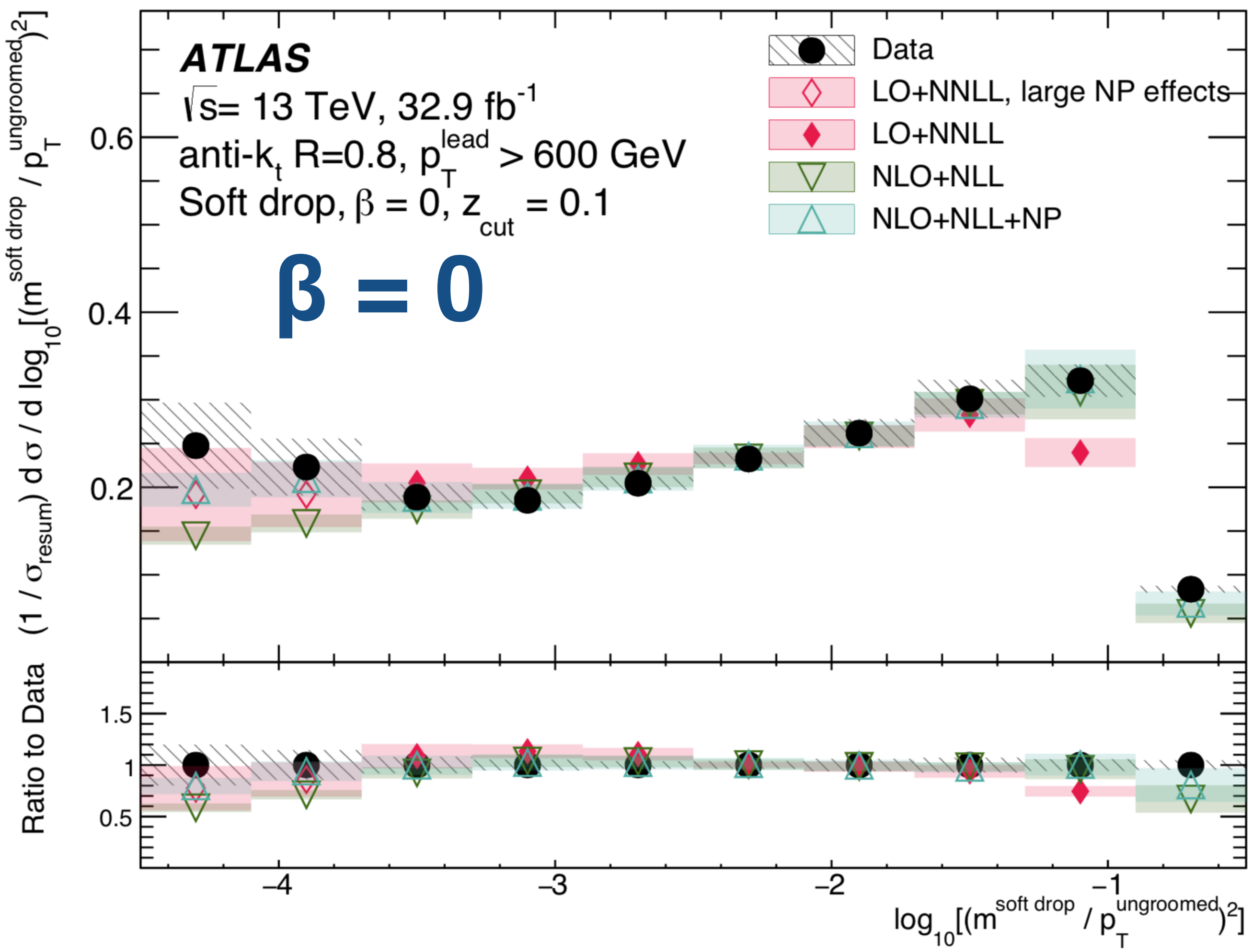


- Three main regimes in the mass distribution:
  - $\rho < -3$: Non-perturbative regime
  - $-3 < \rho < -1$: Resummation regime
  - $\rho > -1$: Fixed order regime

- Very good data/MC agreement for β=0
- Some disagreement at very low mass for β=2
  - Unsurprising since more non-perturbative effects are relevant for higher values of β
  - Pythia still models it very well



Soft Drop Jet Mass



- Three main regimes in the mass distribution:

  ρ < -3: Non-perturbative regime

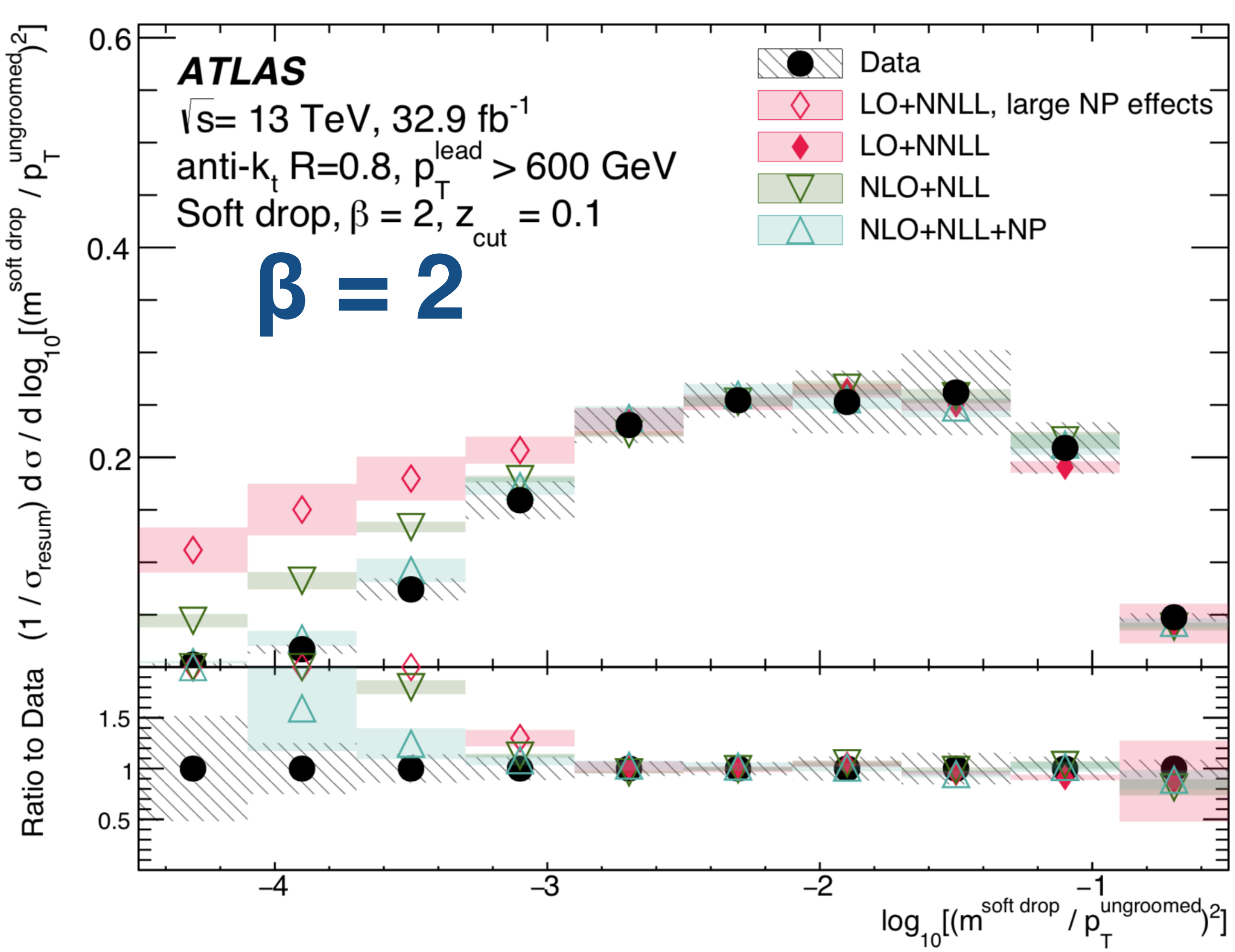
  -3 < ρ < -1: Resummation regime

  ρ > -1: Fixed order regime

- See this paper on NLO+NLL and this paper on LO+NNLL for more information about the predictions

- Both predictions agree very well with data in the resummation regime

- More disagreement in both predictions at low masses
  - Including non-perturbative effects improves the accuracy of the NLO+NLL prediction



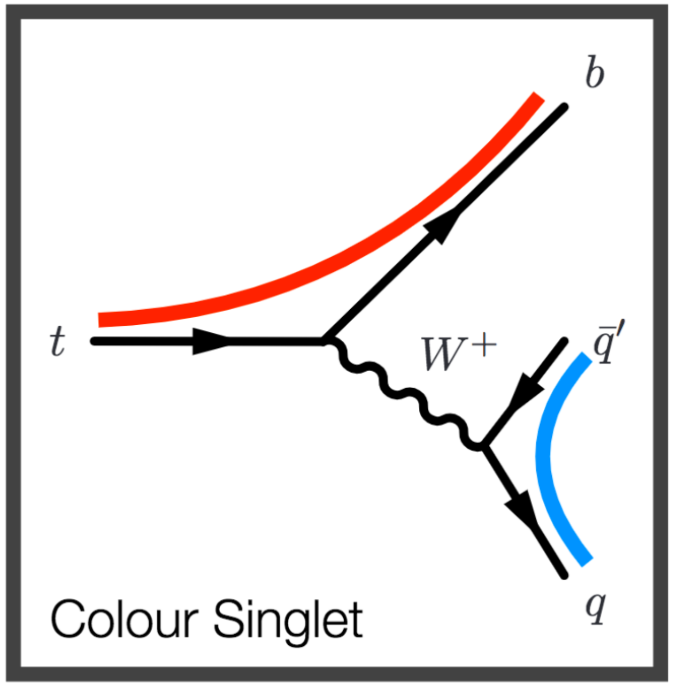
Summary

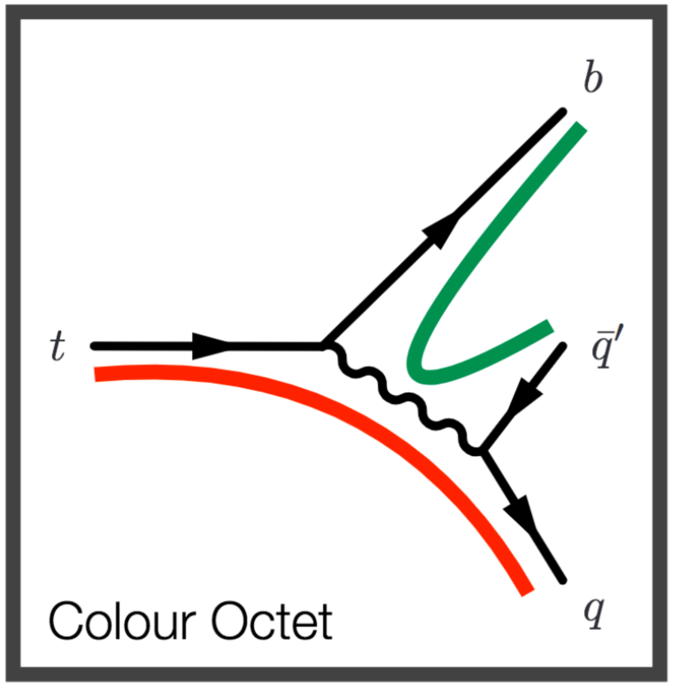
- New measurements of jet substructure are solidifying our understanding of the internal structure of jets and the theory of QCD
- The **color flow** measurements give us a better understanding of color connections in QCD
  - Comparing against a range of predictions with different hard-scatter calculations and hadronization models to understand which are modeling these effects well
  - See Andrew's talk about the first calculation of jet pull
- The **Soft Drop jet mass measurement** tests the limits of predictions of the jet mass
  - Overall good agreement with Monte Carlo simulations
  - Good agreement with predictions in the resummation region, demonstrating the capabilities of these techniques
  - Opens up the possibility of more advanced measurements using jet substructure

Backup

Color Flow

- The color flow models for the W as a color singlet and as a color octet





Color Flow

$\Delta\theta_P\left(j_1^W, j_2^W\right)$ [%]	$\theta_P\left(j_1^W, j_2^W\right)$			
	0.0 – 0.21	0.21 – 0.48	0.48 – 0.78	0.78 – 1.0
Hadronisation	0.55	0.13	0.24	0.14
Generator	0.32	0.25	0.50	0.01
b-tagging	0.35	0.13	0.20	0.31
Background model	0.30	0.16	0.16	0.27
Colour reconnection	0.22	0.16	0.16	0.18
JER	0.11	0.12	0.23	0.02
Pile-up	0.19	0.16	0.00	0.01
Non-closure	0.14	0.07	0.07	0.18
JES	0.12	0.06	0.14	0.06
ISR / FSR	0.15	0.02	0.12	0.02
Tracks	0.05	0.04	0.03	0.06
Other	0.02	0.01	0.01	0.02
Syst.	0.88	0.44	0.71	0.51
Stat.	0.23	0.19	0.19	0.25
Total	0.91	0.48	0.73	0.57

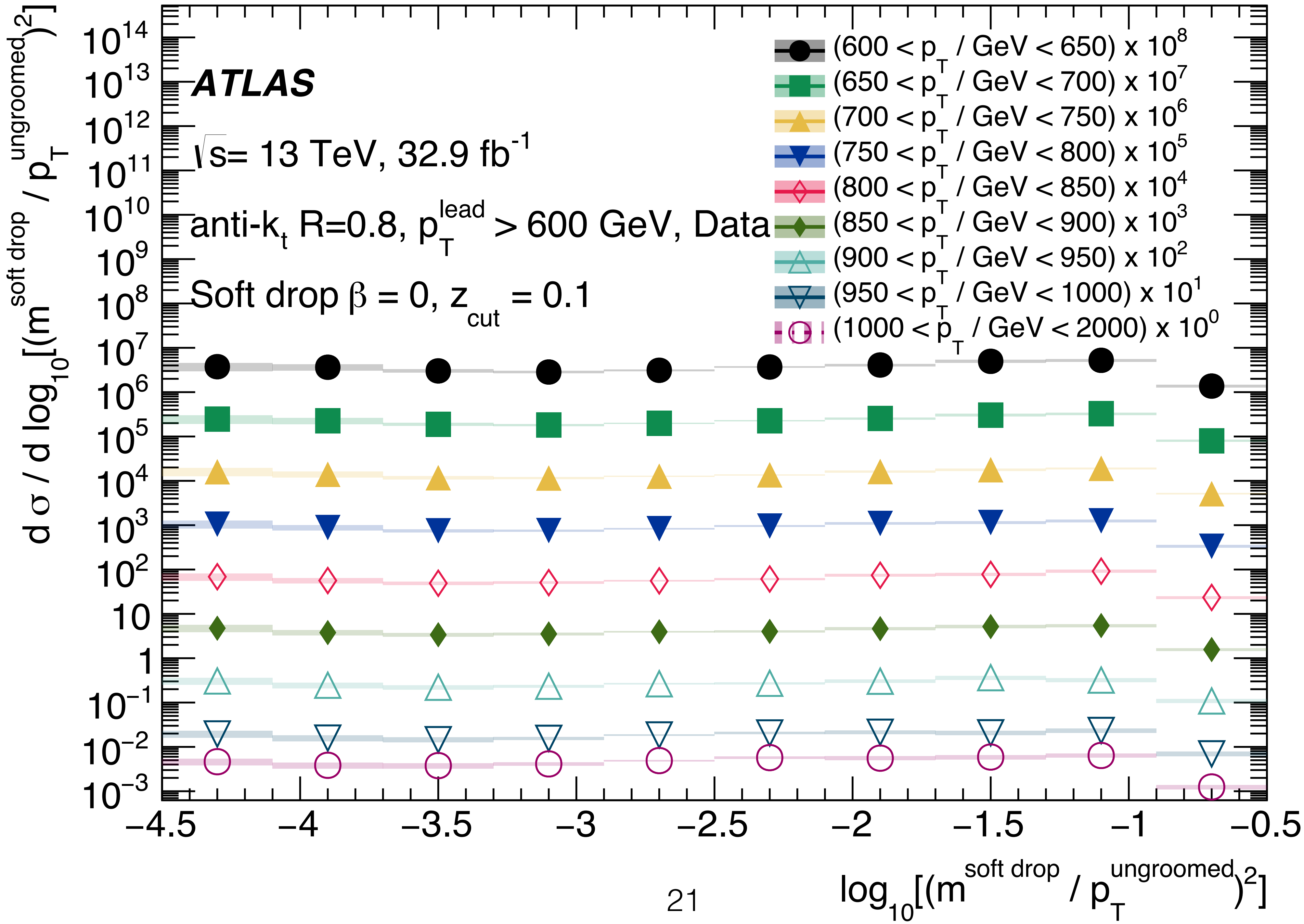
Color Flow

- Goodness of fit for each model and observable
- Powheg+Herwig 7 tends to perform well overall, with the exception of the pull angle for the b-jets

Sample	$\theta_{\mathcal{P}}\left(j_1^W, j_2^W\right)$		$\theta_{\mathcal{P}}\left(j_2^W, j_1^W\right)$		$\theta_{\mathcal{P}}\left(j_1^b, j_2^b\right)$		$\lvert\vec{\mathcal{P}}\left(j_1^W\right)\rvert$	
	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value	χ^2/NDF	p-value
Powheg+Pythia8	50.9 / 3	< 0.001	25.1 / 3	< 0.001	0.7 / 3	0.867	24.8 / 4	< 0.001
Powheg+Pythia6	23.2 / 3	< 0.001	8.2 / 3	0.042	4.2 / 3	0.240	21.1 / 4	< 0.001
MG5_aMC+Pythia8	6.8 / 3	0.077	6.7 / 3	0.082	2.0 / 3	0.563	17.6 / 4	0.001
Powheg+Herwig7	2.7 / 3	0.446	3.4 / 3	0.328	4.8 / 3	0.190	11.3 / 4	0.023
Sherpa	22.0 / 3	< 0.001	11.9 / 3	0.008	0.0 / 3	0.998	14.1 / 4	0.007
Powheg+Pythia8*	17.1 / 3	< 0.001	25.0 / 3	< 0.001	0.3 / 3	0.958	11.1 / 4	0.026
Flipped Powheg+Pythia8*	45.3 / 3	< 0.001	45.9 / 3	< 0.001	2.6 / 3	0.457	17.2 / 4	0.002

Soft Drop

- pT dependence of the Soft Drop mass distribution
  - Note how the distribution is fairly stable across pT



Soft Drop

- pT dependence of the Soft Drop mass distribution
  - Note how the distribution is fairly stable across pT