Precision Studies of Electroweak Zjj Production at ATLAS

2026 CAP Congress

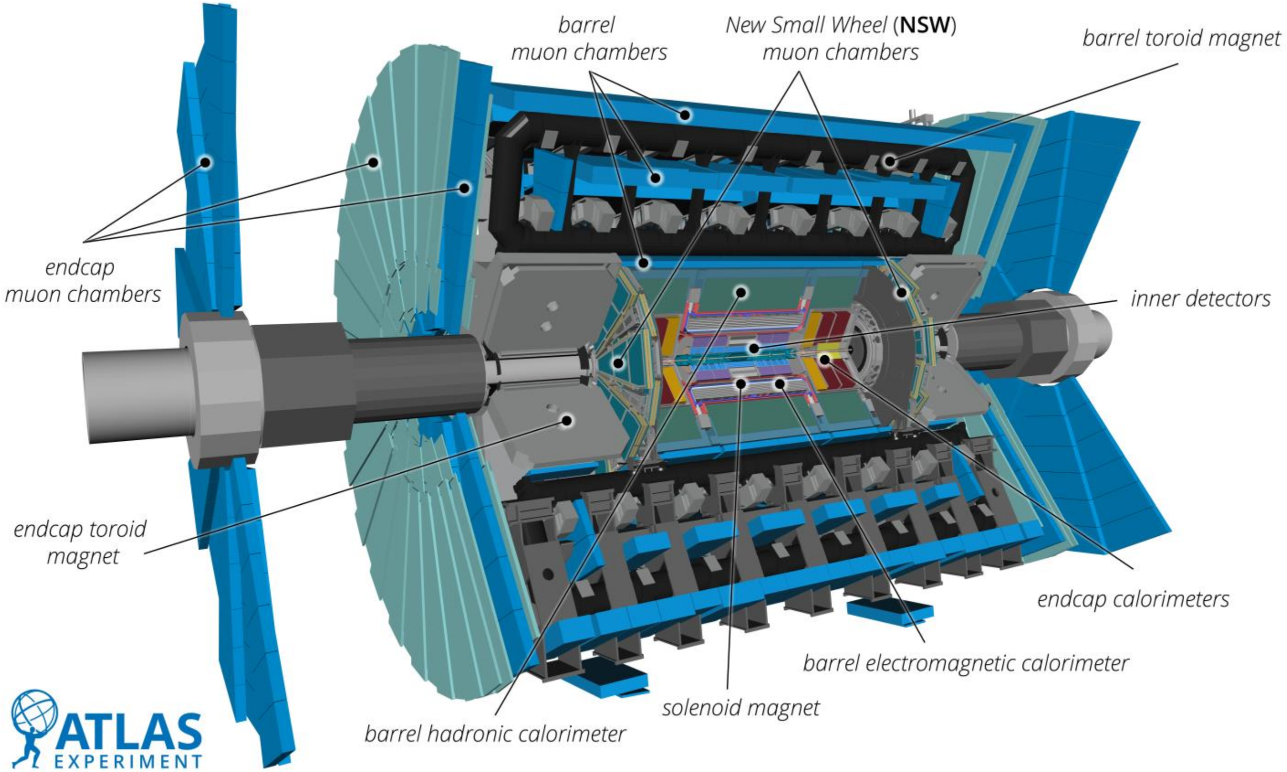
June 22-26, 2026

Ishan Vyas, PhD Candidate

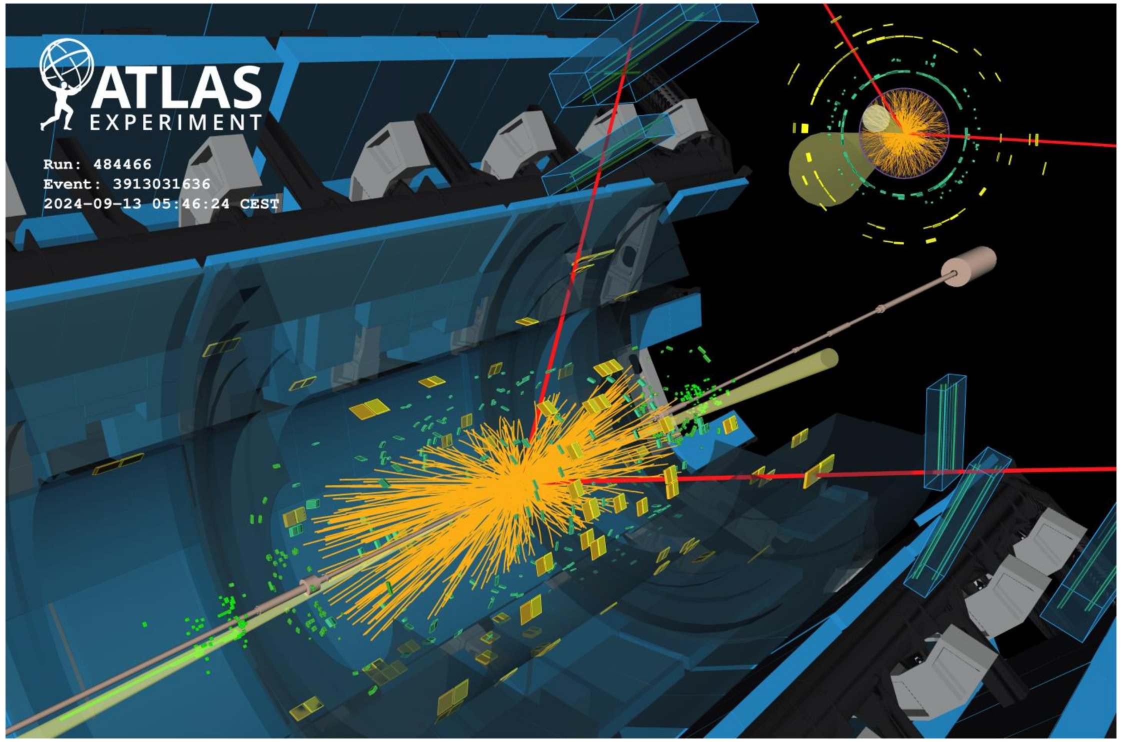
Supervisor: Prof. Alain Bellerive

The ATLAS Experiment

- Multi-layer detector at the CERN Large Hadron Collider
- Inner detector tracks particles close to the collision point.
- A set of calorimeters measure energy of particles
- Muon spectrometer is optimized for precise muon detection.

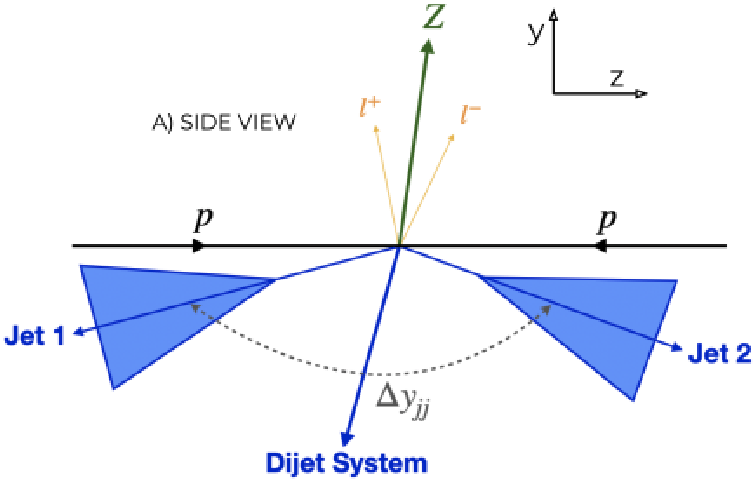


Zjj Production in ATLAS



- Z boson produced with two forward jets arising from scattered quarks
- Z boson decays into a pair of leptons
  - $Z \rightarrow e^+e^-$
  - $Z \rightarrow \mu^+\mu^-$
- EM calorimeters reconstruct electrons, muon spectrometer tracks muons
- Forward Calorimeters detect the two jets.

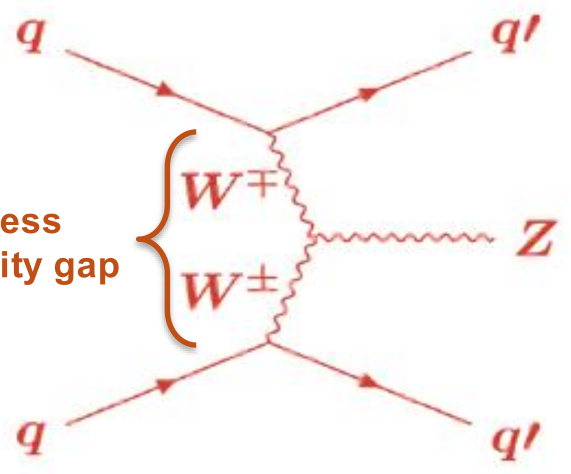
Takeaway: Probing the electroweak sector requires combined capabilities of the ATLAS detector



Zjj Production Modes

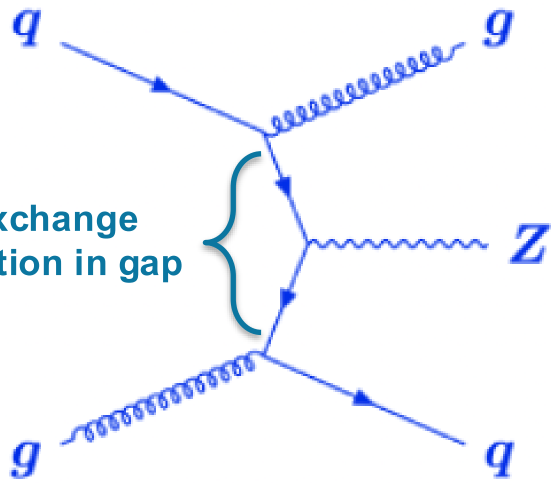
- Process: $pp \to Z + 2\ \text{Jets}$
- **Electroweak production (EW Zjj)** occurs through t-channel vector boson exchange
  - purely EW process (Signal)
  - Heavily suppressed by $\sim 10^4$
  - clean rapidity gap
- **Strong Production (Strong Zjj)** occurs via color exchange
  - QCD-induced Z production (Background)
  - Dominant mode of production
  - QCD radiation possible in the gap

EW Zjj (Signal)



Pure EW process
→ clean rapidity gap

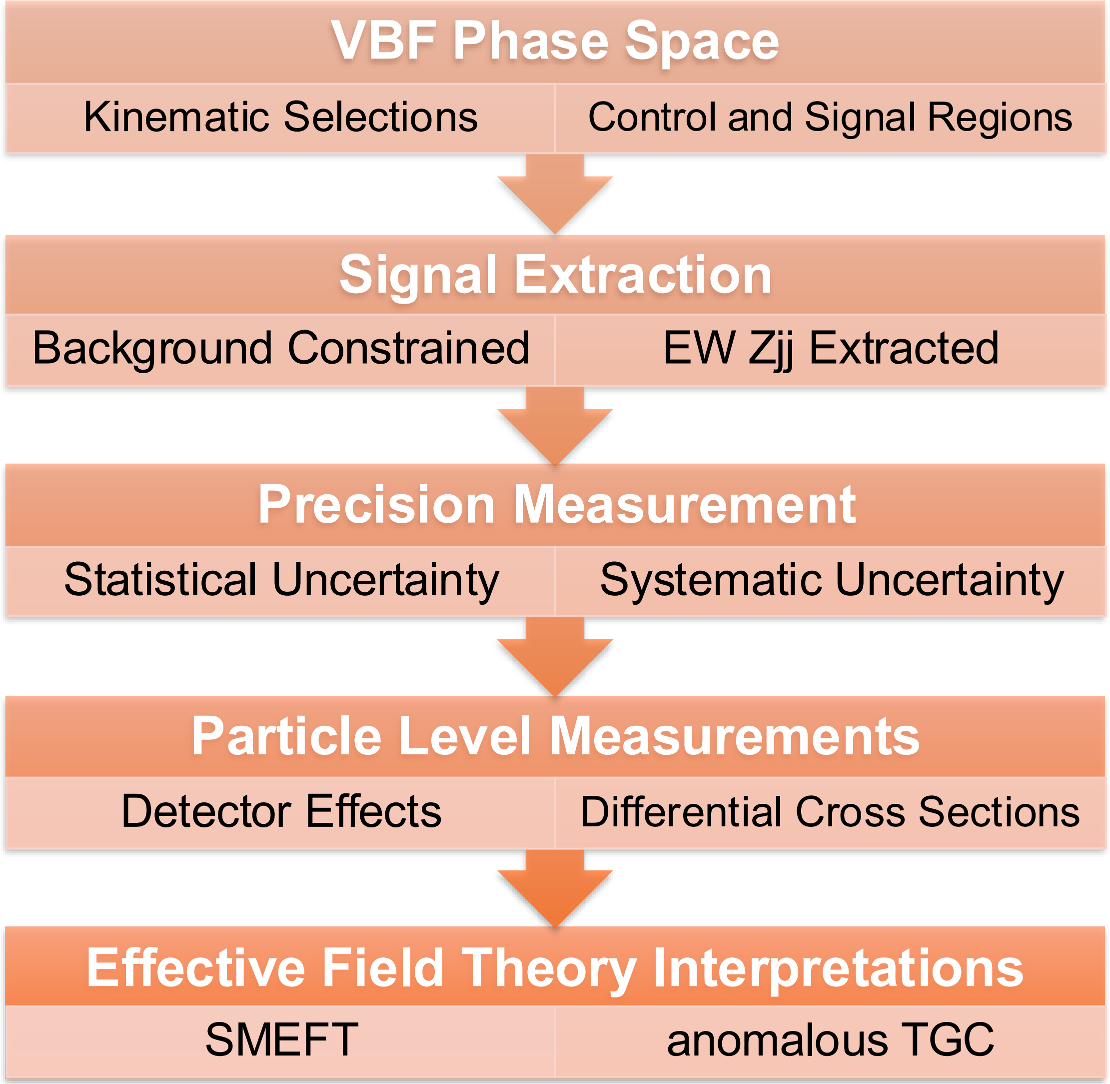
Strong Zjj (background)



Color exchange
→ radiation in gap

Analysis Strategy

- Measure electroweak Zjj production in a Vector Boson Fusion (VBF) - like topology.
- **Extract signal using a Poisson likelihood fit**
- **Evaluate uncertainties from multiple sources affecting the measurement**
- Obtain differential cross sections for key observables
- Interpret results within the Standard Model Effective Field Theory (SMEFT)



VBF Phase Space

Kinematic Selections | Control and Signal Regions

Signal Extraction

Background Constrained | EW Zjj Extracted

Precision Measurement

Statistical Uncertainty | Systematic Uncertainty

Particle Level Measurements

Detector Effects | Differential Cross Sections

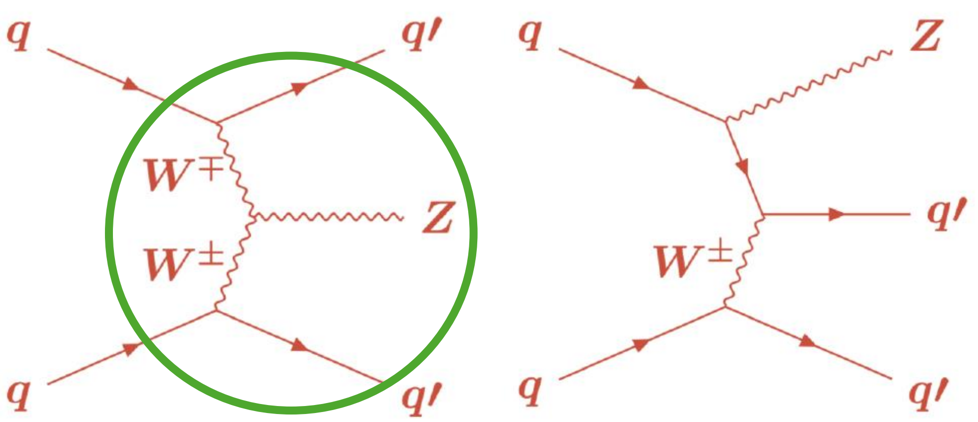
Effective Field Theory Interpretations

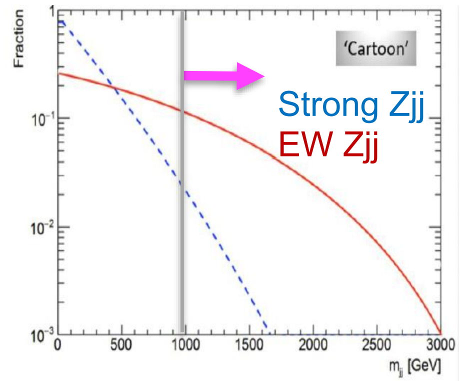
SMEFT | anomalous TGC

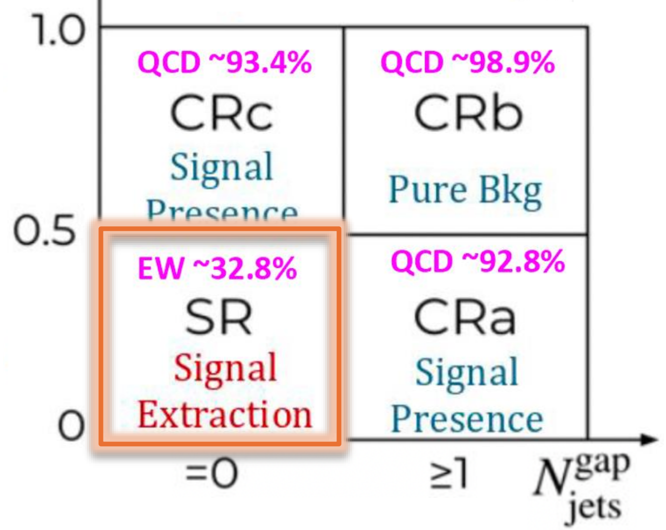
Vector Boson Fusion

- Z boson produced from WWZ vertex.
- Two forward jets with large dijet mass (m_{jj}) and rapidity separation (Δy_{jj})
- **Kinematic selections favoring VBF process are applied**
- Central jet activity $(N_{\text{jets}}^{\text{gap}})$, Centrality of Z boson (ξ_Z) defines:
  - a high-purity signal region with no GapJets
  - three control regions to constrain Strong Zjj

Takeaway: VBF-like phase space provides a clean probe of electroweak gauge-boson interactions





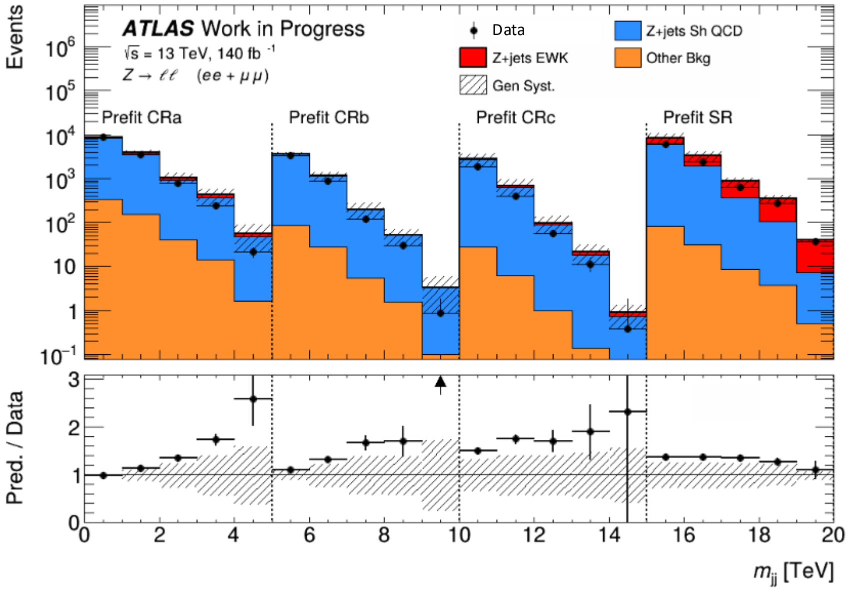


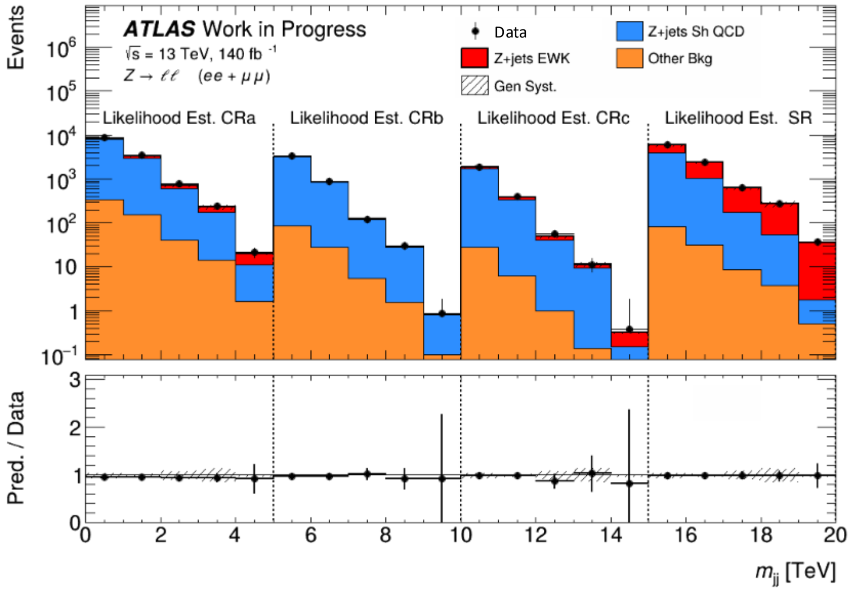
ξ_Z	$N_{\text{jets}}^{\text{gap}}$ =0	$N_{\text{jets}}^{\text{gap}}$ ≥1
0.5 – 1.0	QCD ~93.4% CRc Signal Presence	QCD ~98.9% CRb Pure Bkg
0 – 0.5	EW ~32.8% SR Signal Extraction	QCD ~92.8% CRa Signal Presence

Signal Extraction

- Data-MC mismodelling motivates a likelihood fit.
- **Simultaneous fit performed in the signal and control regions on ATLAS Run 2 dataset**
- Data in control regions constrain the dominant Strong Zjj background in the signal region
- Reduced reliance on Monte Carlo predictions

Takeaway: Precise extraction of the EW Zjj signal yield



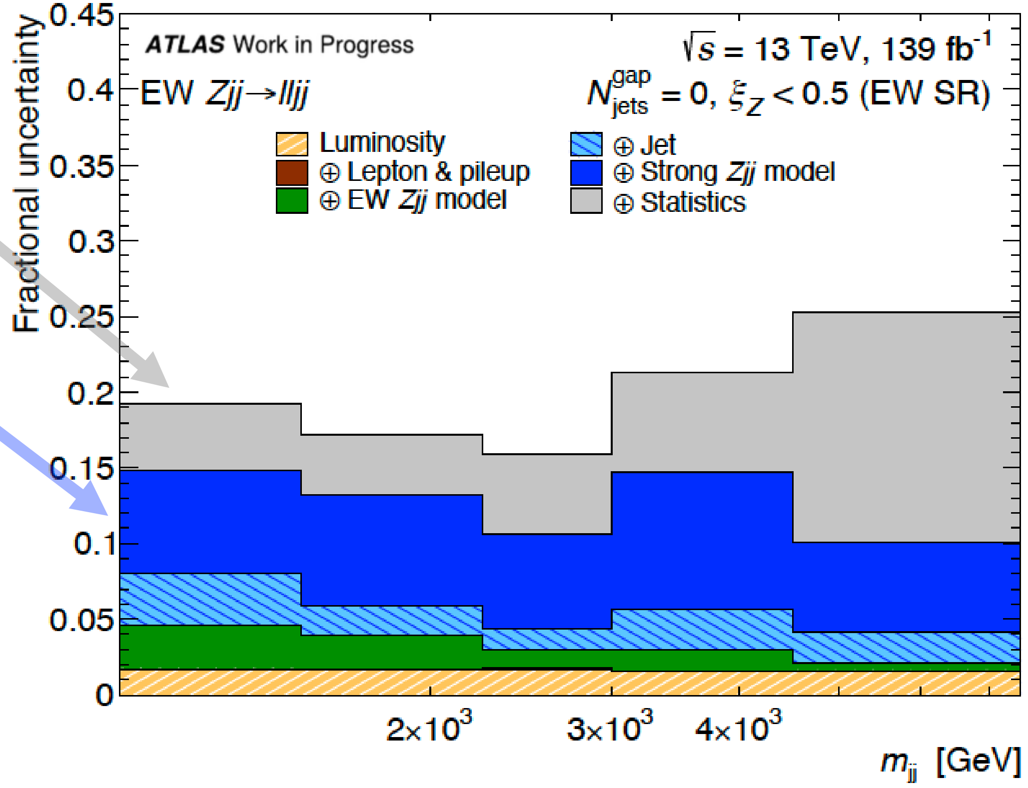


Precision Measurement Challenges

- Precision limited by both statistical and systematic uncertainties.
- Statistical uncertainty determined by the available data.
- **Systematic uncertainties arise from theory modelling and detector effects**
- Previous analysis dominated by:
  - Strong Zjj generator-modelling uncertainty
  - Jet-related detector systematic uncertainties

Takeaway: Reducing systematic uncertainties improves the measurement precision

Reference Analyses Results

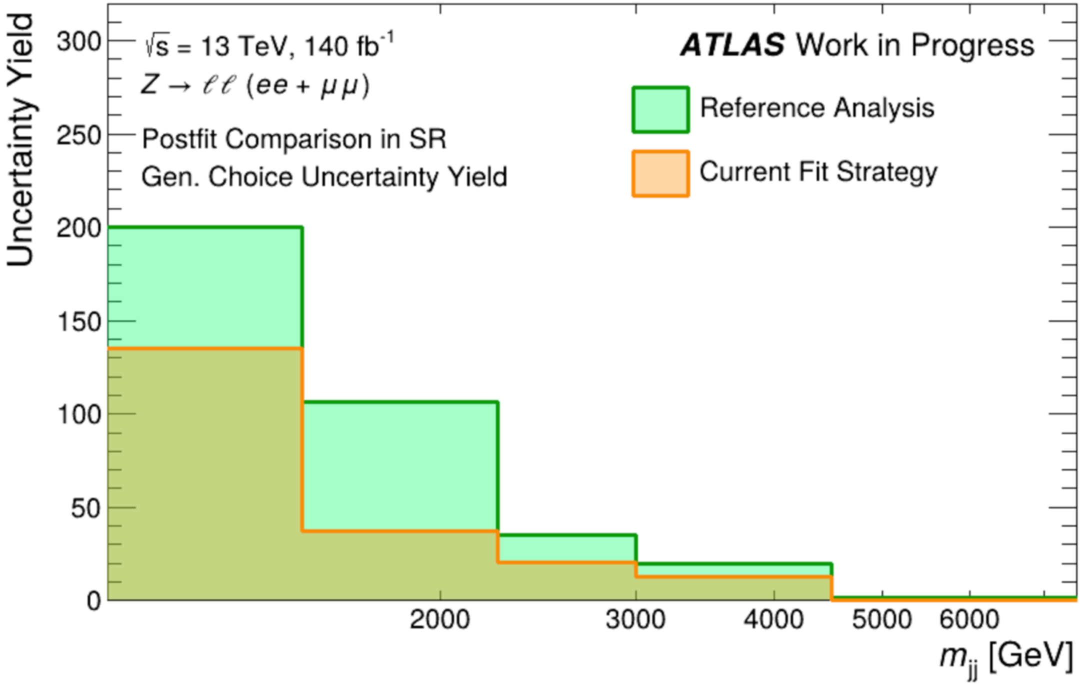


STDM-2017-27, EPJC 81 (2021) 163

Strong Zjj Modeling

- Strong Zjj predictions depend on radiation modelling
- Different MC generators predict different Zjj distributions, leading to generator choice uncertainty
- **Data-driven fit constrains the dominant Strong Zjj background**, reducing reliance on MC predictions

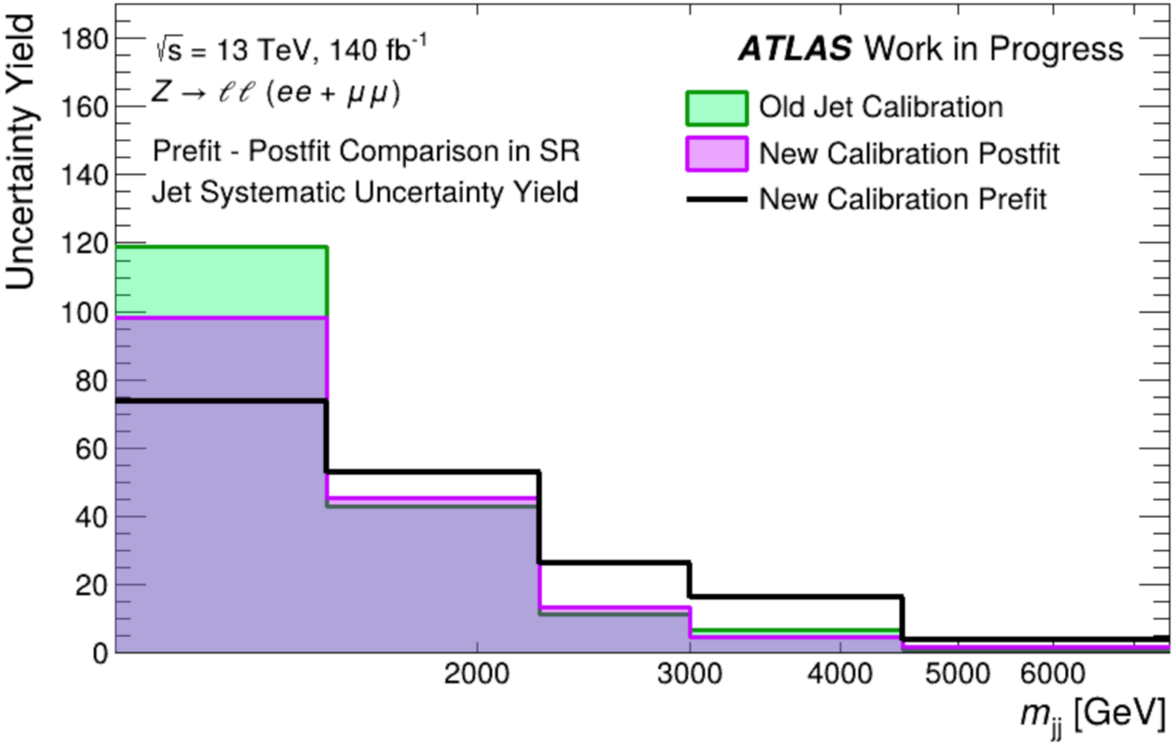
Takeaway: Significant reduction in generator-choice uncertainty relative to the previous analysis



Jet Energy Scale and Resolution

- Jets are central to the VBF Zjj event selection
- Jet uncertainties affect both event acceptance and kinematic distributions
- **Latest Jet calibrations show slight improvement → Propagated through the fit**
- Thanks to the JetEtMiss group!

Takeaway: Latest Jet systematic uncertainties have a reduced impact on the extracted EW Zjj signal

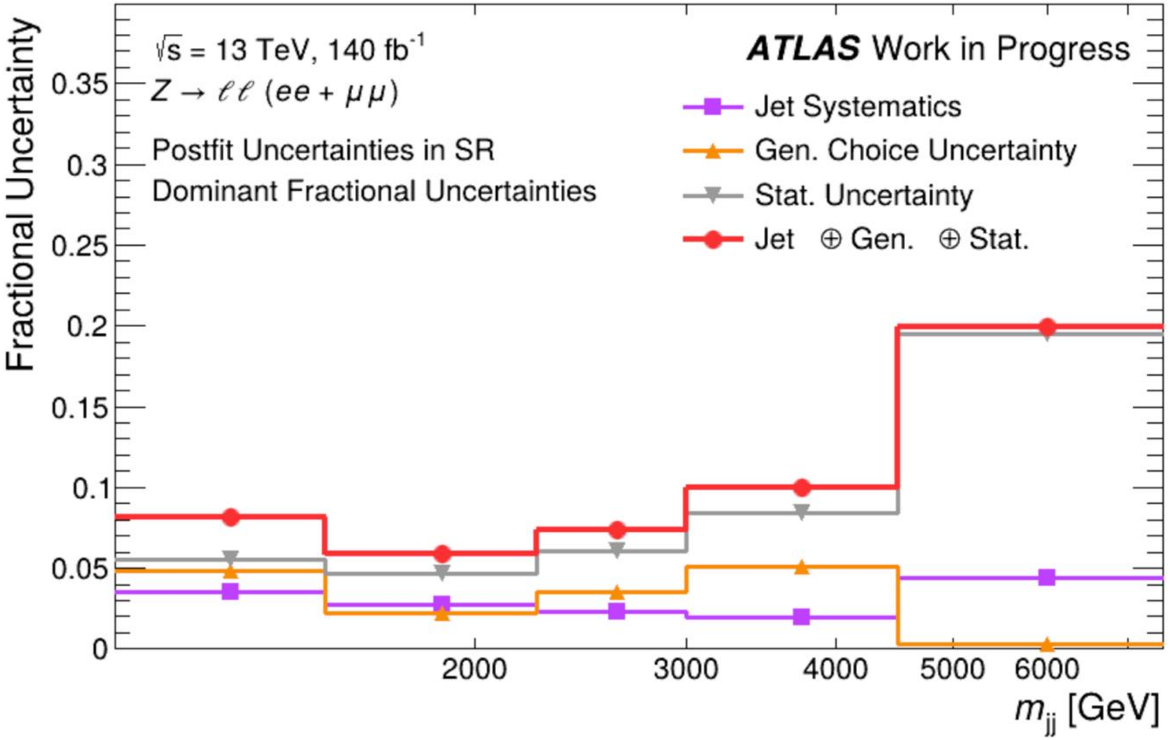


Estimated Total Uncertainty (Dominant Sources Only)

- Dominant uncertainty sources propagated through the fit, shown as fractional uncertainties
- **Generator-choice and jet-related uncertainties contribute ~ 5%,** as opposed to ~7-9% in previous analysis
- Statistical uncertainty now dominates the dominant uncertainty budget

Takeaway: Work in Progress shows potentially significant reduction in uncertainty

Current Precision Status

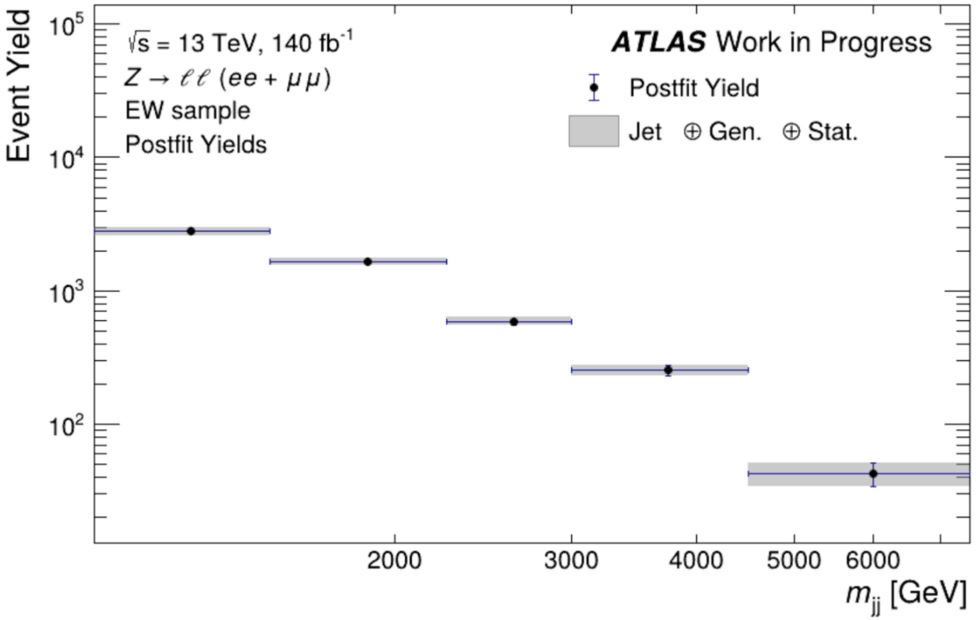


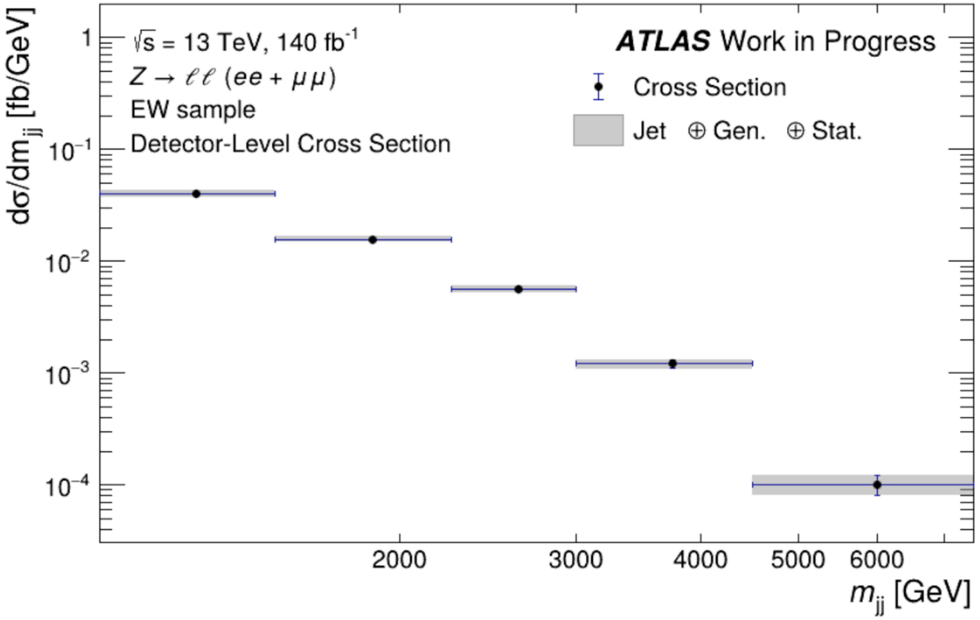
Detector-Level Cross Sections

- Extracted EW Zjj yields are shown with the uncertainty band
- Detector-level differential cross sections are evaluated as

$$\left(\frac{d\sigma}{dm_{jj}}\right)_i^{\text{det}} = \frac{N_{i,\text{EW}}^{\text{postfit}}}{\mathcal{L}\,\Delta m_{jj,i}}$$

Takeaway: Improved precision provides a strong foundation for unfolding studies

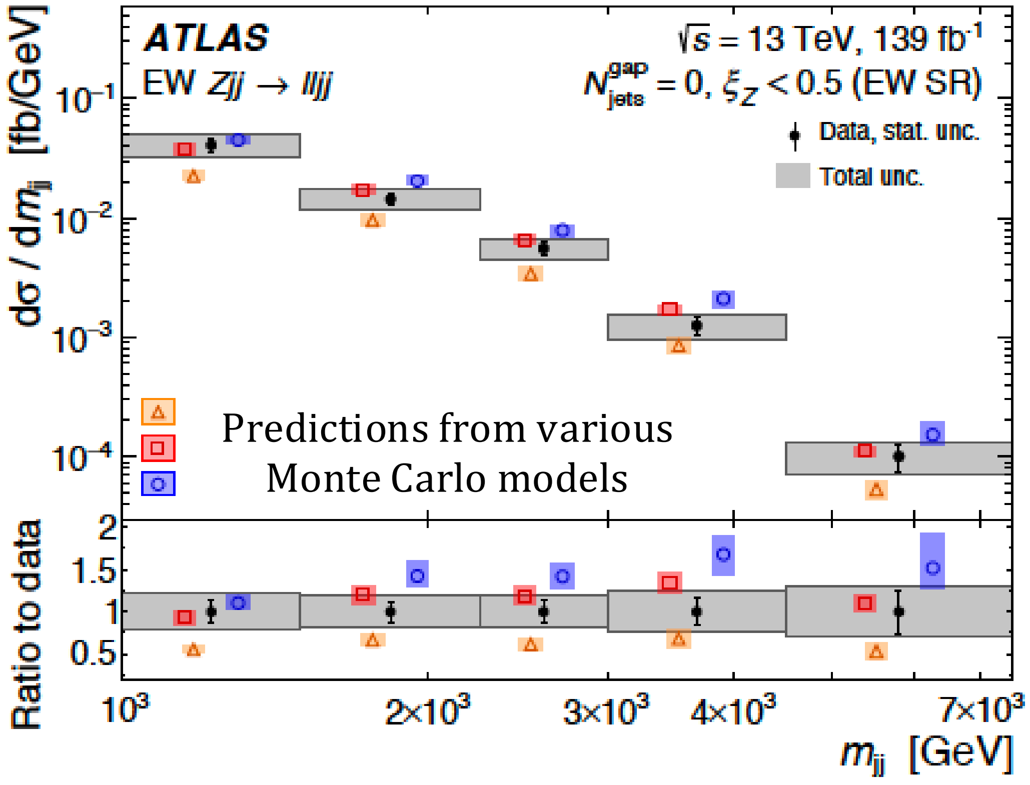




Outlook: Towards Particle-Level Measurements

- Incorporate the remaining systematic uncertainties
- Extend the analysis to the larger Run 3 ATLAS dataset
- **Unfold detector-level measurements to particle-level cross sections**
- Compare measurements with Standard Model predictions
- Set limits on anomalous Triple Gauge Couplings with SMEFT
- Publish the EW Zjj differential cross sections

Reference Analyses Results



Predictions from various Monte Carlo models

STDM-2017-27, EPJC 81 (2021) 163

Summary

Electroweak Zjj provides a clean probe of high-energy electroweak interactions

A data-constrained likelihood fit substantially reduces dominant systematic uncertainties

The improved precision will be propagated to particle-level differential cross sections through unfolding

The larger Run 3 dataset will further extend the measurement’s precision and physics reach

Thank you

Backup Slides

Generalized Likelihood Formulation

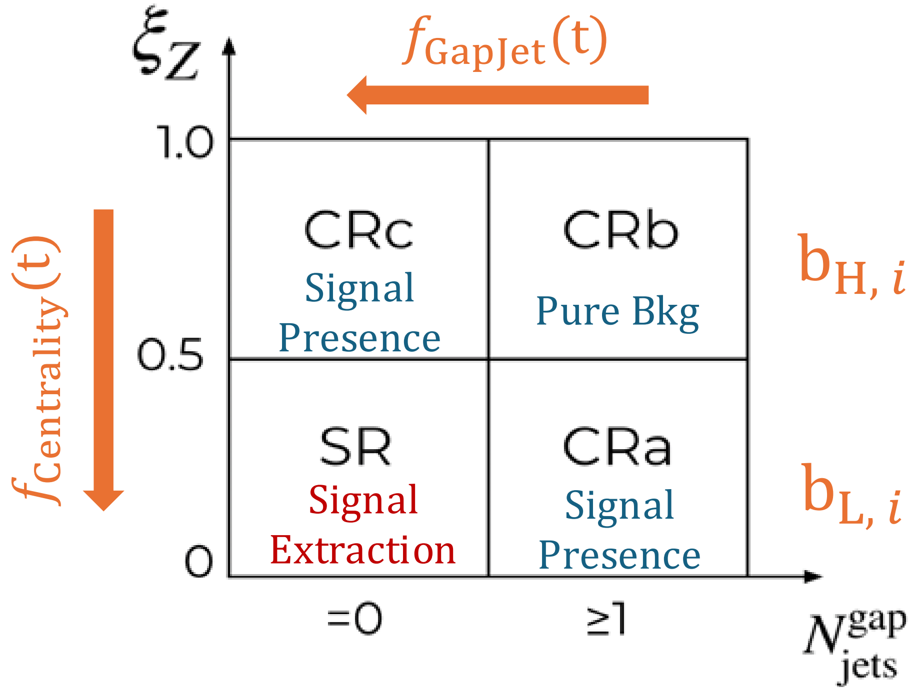
Exploit Data-Driven Scale Factors

- ★ Define the ratio $\mathrm{t} = \left[\frac{\mathrm{b}_{\mathrm{H},\,i}}{\mathrm{b}_{\mathrm{L},\,i}}\right]$.
- ★ Define functions $f(\mathrm{t})$.

Two Boundaries, Two Corrections

- ★ $f_{\mathrm{GapJet}}(\mathrm{t})$ corrects for crossing the GapJet boundary CRb→CRc and CRa→SR.
- ★ $f_{\mathrm{Cent.}}(\mathrm{t})$ corrects for crossing the ξ_{Z} boundary CRc→SR.
- ★ Since ξ_{Z} is continuous and better modeled, the correction is expressed in a simpler form.
- ★ Generalized functional form:
  - $f_{\mathrm{GapJet}}(\mathrm{t}) = ct^{n}$
  - $f_{\mathrm{Cent.}}(t) = t$

- $\lambda_i^{\mathrm{CRa}} = \mu_i \mathrm{E}_i^{\mathrm{CRa}} + \mathrm{b}_{\mathrm{L},\,i}\, \mathrm{Q}_i^{\mathrm{CRa}}$
- $\lambda_i^{\mathrm{CRb}} = \mathrm{b}_{\mathrm{H},\,i}\, \mathrm{Q}_i^{\mathrm{CRb}}$
- $\lambda_i^{\mathrm{CRc}} = \mu_i \mathrm{E}_i^{\mathrm{CRc}} + f_{\mathrm{GapJet}}(\mathrm{t})\mathrm{b}_{\mathrm{H},i}\, \mathrm{Q}_i^{\mathrm{CRc}}$
- $\lambda_i^{\mathrm{SR}} = \mu_i \mathrm{E}_i^{\mathrm{SR}} + f_{\mathrm{Cent.}}(\mathrm{t}) \,.\, f_{\mathrm{GapJet}}(\mathrm{t})\mathrm{b}_{\mathrm{L},\,i}\, \mathrm{Q}_i^{\mathrm{SR}}$



Deriving the Scaled Ratio form

First Order formulation

- Setting $n = 1$ yields,

$$f_{\mathrm{CRc}}(t) = f_{\mathrm{GapJet}}(t) = ct$$

$$f_{\mathrm{SR}}(t) = f_{\mathrm{Cent.}}(t).\, f_{\mathrm{GapJet}}(t) = ct^2$$

- **This is exactly the Scaled Ratio fit.**

Key features

- It is the simplest non-trivial functional form ($n = 1$).
- The same formulation is applied across all observables.
- Maximally leverages the data-driven scale factors.
- Minimal additional assumptions (no observable dependent corrections).

- $b^{\mathrm{CRa}} = \mathrm{b}_{\mathrm{L},\,i}\, \mathrm{Q}_i^{\mathrm{CRa}}$
- $b^{\mathrm{CRb}} = \mathrm{b}_{\mathrm{H},\,i}\, \mathrm{Q}_i^{\mathrm{CRb}}$
- $b^{\mathrm{CRc}} = c \left[\frac{\mathrm{b}_{\mathrm{H},\,i}}{\mathrm{b}_{\mathrm{L},\,i}}\right]^{\mathrm{n}} .\ \mathrm{b}_{\mathrm{H},\,i} \mathrm{Q}_i^{\mathrm{CRc}}$
- $b^{\mathrm{SR}} = c \left[\frac{\mathrm{b}_{\mathrm{H},\,i}}{\mathrm{b}_{\mathrm{L},\,i}}\right]^{\mathrm{n+1}} .\, \mathrm{b}_{\mathrm{L},\,i} \mathrm{Q}_i^{\mathrm{SR}}$

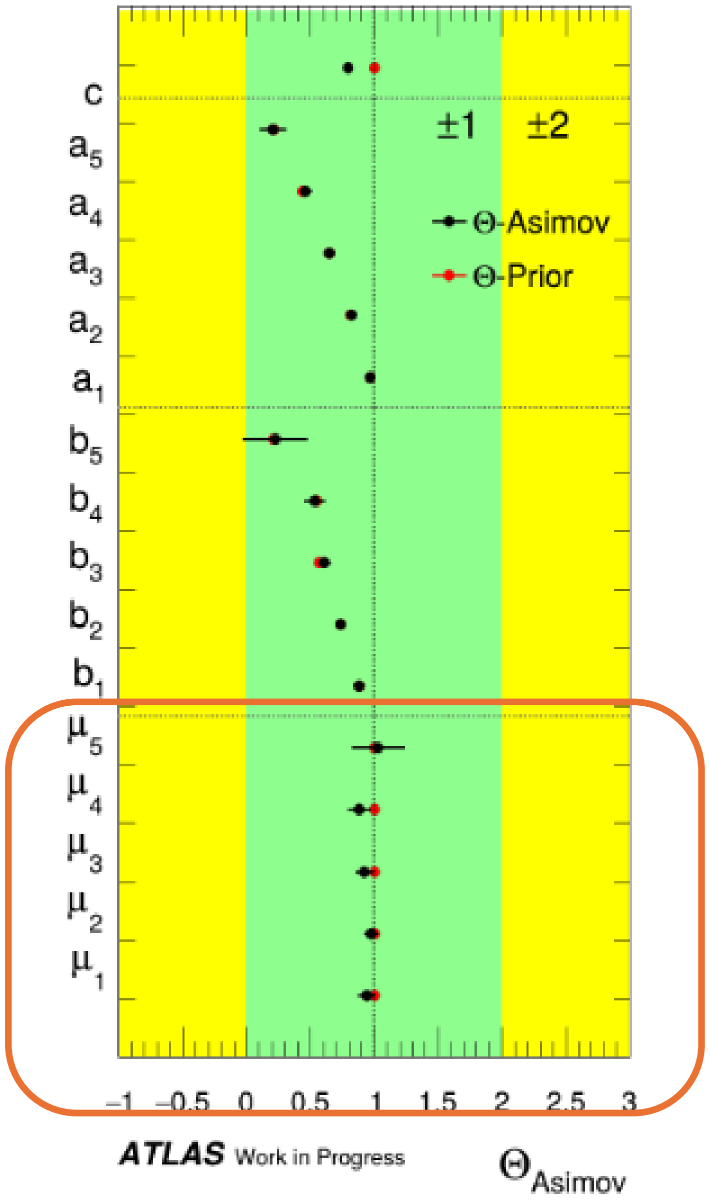
Reduces to Scaled Ratio Fit for n = 1

- $b^{\mathrm{CRa}} = \mathrm{b}_{\mathrm{L},\,i}\, \mathrm{Q}_i^{\mathrm{CRa}}$
- $b^{\mathrm{CRb}} = \mathrm{b}_{\mathrm{H},\,i}\, \mathrm{Q}_i^{\mathrm{CRb}}$
- $b^{\mathrm{CRc}} = \mathrm{c} \frac{\mathrm{b}_{\mathrm{H},\,i}^2}{\mathrm{b}_{\mathrm{L},\,i}} \mathrm{Q}_i^{\mathrm{CRc}}$
- $b^{\mathrm{SR}} = \mathrm{c} \frac{\mathrm{b}_{\mathrm{H},\,i}^2}{\mathrm{b}_{\mathrm{L},\,i}} \mathrm{Q}_i^{\mathrm{SR}}$

Asimov Fit: Fitting Sherpa to MadGraph

- **Pseudo Data:** Sherpa EW + MadGraph FxFx QCD + Others
- **Fitted MC:** Sherpa EW, Sherpa QCD and Others

- Asimov fit tells us how biased the fit is.
- Sensitivity of the fits to prior QCD modelling, observed in the distribution of μ_i.
- Not expecting a_i, b_i to be 1, but I expect μ_i close 1.
- In the previous analysis, this shift in μ_i, was quoted as a modelling error.
- EW prediction is the insensitive to QCD modelling differences.



Validation: Procedure for Ensemble test

Aim: **Check the fit bias and quantify the effect of modeling difference between Sherpa and MadGraph QCD generators.**

$$N_{r,i}^{\text{Nominal}} = N_{r,i}^{\text{EW,MC}} + N_{r,i}^{\text{MadGraph QCD, MC}} + N_{r,i}^{\text{Other,MC}}$$

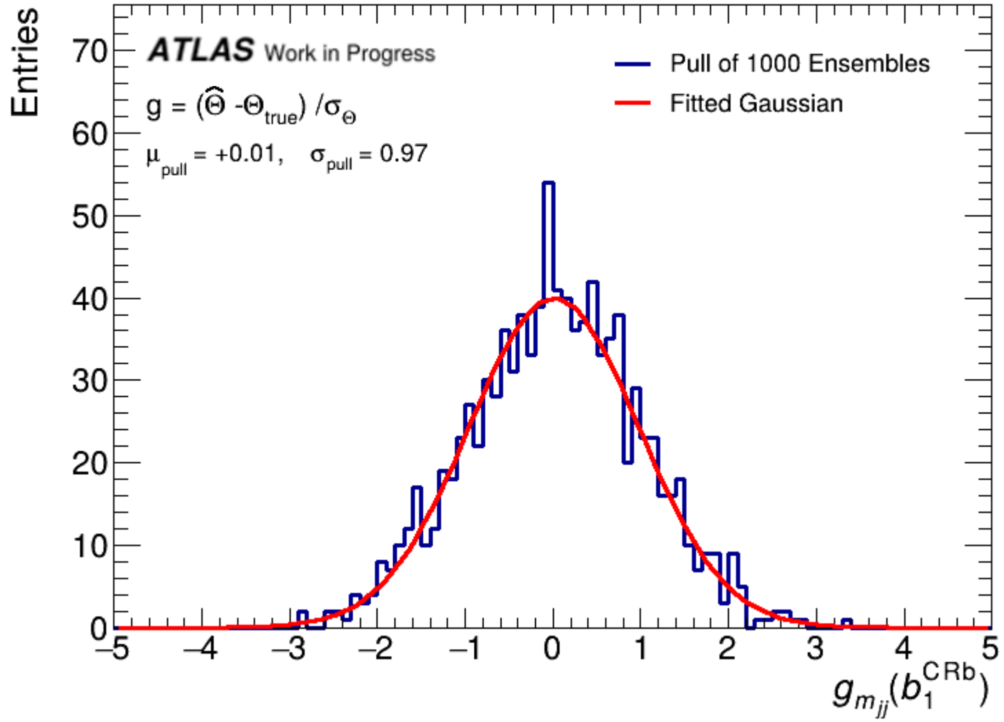
$$N_{r,i}^{\text{Ensemble}} = \text{Poisson}\left(N_{r,i}^{\text{Nominal}}\right)$$

$$\text{Pull } g = \frac{\hat{\theta} - \theta_{\text{Asimov}}}{\sigma_{\hat{\theta}}}$$

Procedure:

- Generate 1000 toy datasets adding per-bin Poisson variations of EW, QCD and other backgrounds.
- Perform likelihood fit using Sherpa QCD and EW samples on nominal and the toy datasets.
- Extract θ_{Asimov} from the nominal fit and $\hat{\theta}$ from toy fits.
- θ_{Asimov} gives information about the estimated QCD and EW yields .
- Plot the pull for $\hat{\theta}$ obtained from each of the toy fits.
- Fit a Gaussian and extract the mean and σ for each of the $\hat{\theta}$.
- Pull N(0,1) confirms that fit is bug free and toy datasets are indeed Poisson variated.

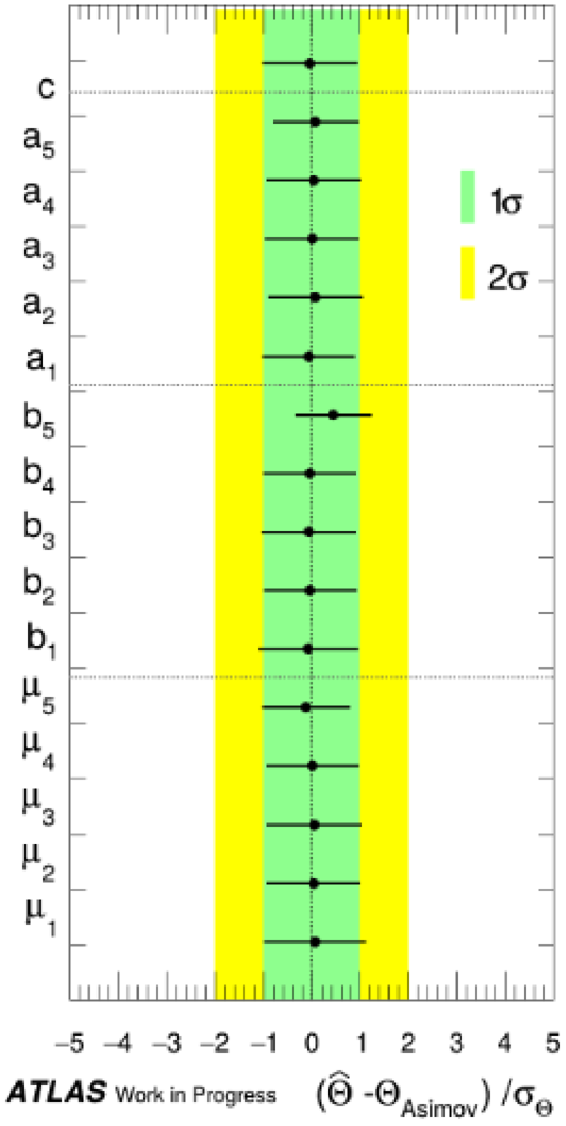
Example Pull



Pull from Ensemble Test: Fitting MadGraph to Itself

- **Pseudo Data:** Poisson (Sherpa EW + MadGraph FxFx QCD + Others)
- **Fitted MC:** Sherpa EW, MadGraph FxFx QCD and Others

- The pull N(0, 1) VALIDATES the implementation of the likelihood fit.
- Fitting MadGraph to itself is a trivial check before performing Ensemble test using different predictions.



Pull from Ensemble Test: Fitting Sherpa to MadGraph

- **Pseudo Data:** Poisson (Sherpa EW + MadGraph FxFx QCD + Others)
- **Fitted MC:** Sherpa EW, Sherpa QCD and Others

- Summary of the Ensemble test conducted in fitting Sherpa QCD to the toy datasets.
- The pull N(0, 1) VALIDATES the implementation of the likelihood fit.
- This confirms that our toy datasets are indeed Poisson distributed.
- The slight deviations in the new fits are because of systematic error introduced by ignoring the EW component in CRb.
- This was independently confirmed by creating a separate dataset with no EW contribution in CRb, where the pulls lined up to N(0, 1).
- Later this error is added as a systematic uncertainty on the final EW and QCD yields.