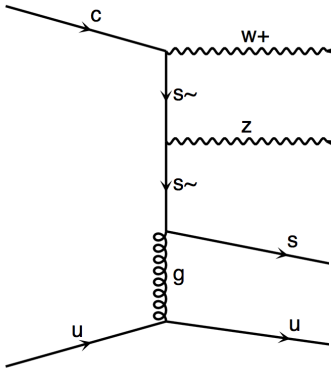
Modeling of Signal and Background Processes in Vector Boson Scattering
Experimental Perspective

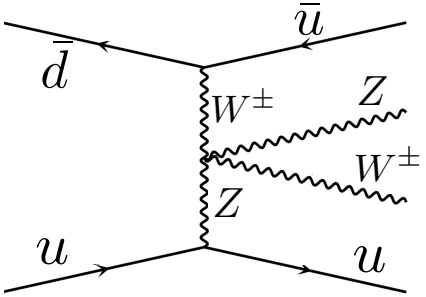
Kenneth Long
University of Wisconsin — Madison

Introduction and Motivation

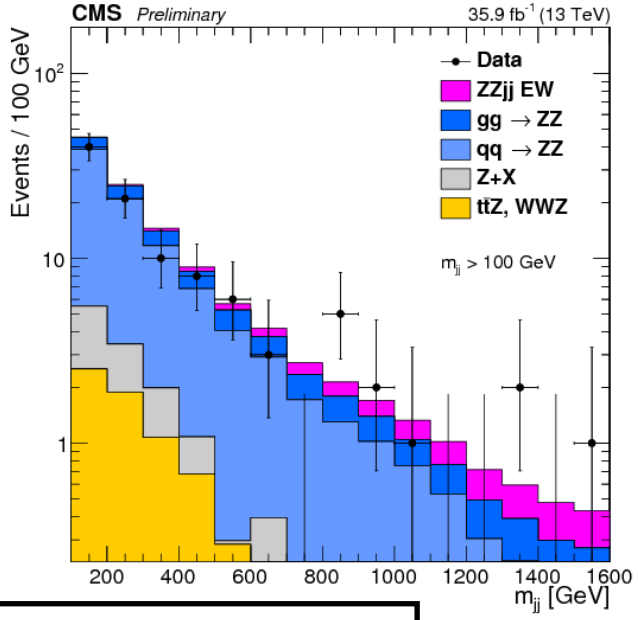
- Diboson production via vector boson scattering (VBS)
  - EWK production (α^4 at LO)
    - Distinct signature from forward jets
- Major background: VV+jets production with radiated jets
  - ➡QCD production ($\alpha^2\alpha_s^2$ at LO)
- Interference ($\alpha^3\alpha_s$)
  - Often taken as background or uncertainty on background
- Simulation is challenging … but important
  - Leveraged for signal vs. background categorization
    - fit to sensitive distribution(s) or via MVA
  - Avoid variables with poor modeling (e.g. 3rd jet)



QCD Background



EWK: VBS Signal

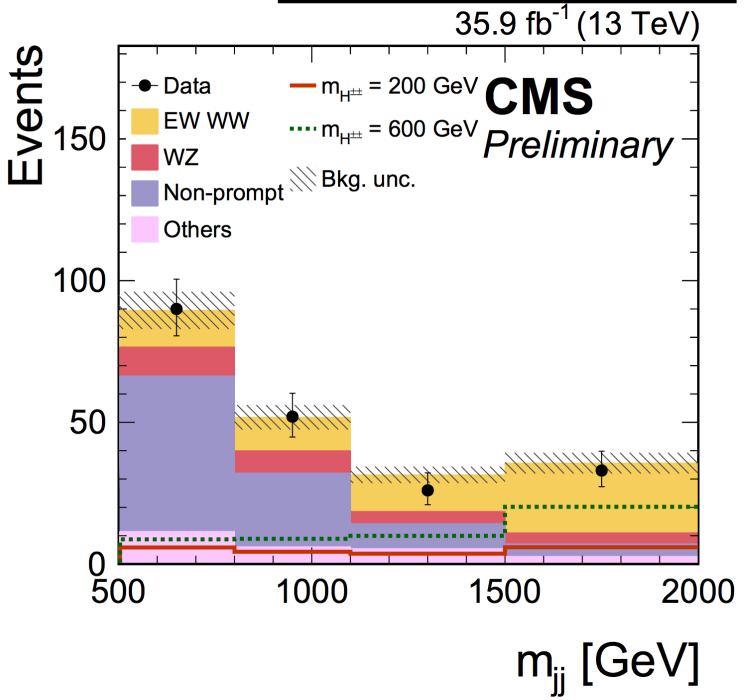


CMS-PAS-SMP-16-019

Approach for 2017 Results: $W^{\pm}W^{\pm}jj$

CMS-PAS-SMP-16-019

- First measurement with > 5σ significance
- Signal region defintion
  - Jets: anti-kt, ΔR = 0.4
    - p_T < 30 GeV, |η| < 5.0
  - $m_{jj} > 500$
  - $\Delta\eta(j_1,j_2) > 2.5$
  - $z^*_l = |\eta_l - (\eta_{j1} + \eta_{j2})/2| / \Delta\eta_{jj} < 0.75$
- Signal extraction with simultaneous fit to m_{jj} and m_{ll}
- Background composition unique from other VBS(F) analysis
  - QCD induced production is small
  - Dominated by Non-prompt (ttbar with jet faking lepton) and leptonic WZ with one lepton lost
    - ★ Non-prompt fully data driven,
    - ★ WZ (shape) normalized to data in bins of m_{jj} in WZ control region
- ➡ Much smaller dependence on simulation than other channels



Approach for 2017 Results: ZZjj

CMS-PAS-SMP-16-019

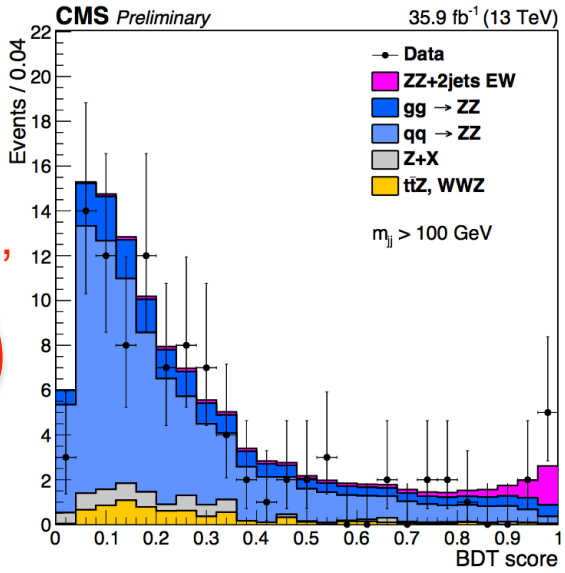
- Very low background from non-ZZ processes, but S/B ~ 1/20
  - BDT training to optimize sensitivity
    - 7 Inputs: m_{jj}, $\Delta\eta_{jj}$, $z_{\ell 1}^*$, $z_{\ell 2}^*$, $R(p_T)$, dijet p_T balance, m_{4l}

➡ Observed significance 2.7σ (expected 1.6σ)

$$\mu = 1.39^{+0.72}_{-0.57}(\text{stat})^{+0.46}_{-0.31}(\text{syst})$$

Dominated by JES/JER, background modeling

Reduced by ~√3 with full Run II



- Background almost entirely QCD-induced ZZ
  - qq/qg processes modeled using MadGraph5_aMC@NLO at LO and NLO
    - LO: ZZjj+Pythia8 (no merging)
    - NLO: ZZ+≤2@NLO with FxFx merging
      - Z bosons generated on-shell and decayed with MadSpin
  - gg initiated from MCFM and and MG5_aMC
- Background uncertainties
  - Compare variations in BDT output for different generators and scale variations

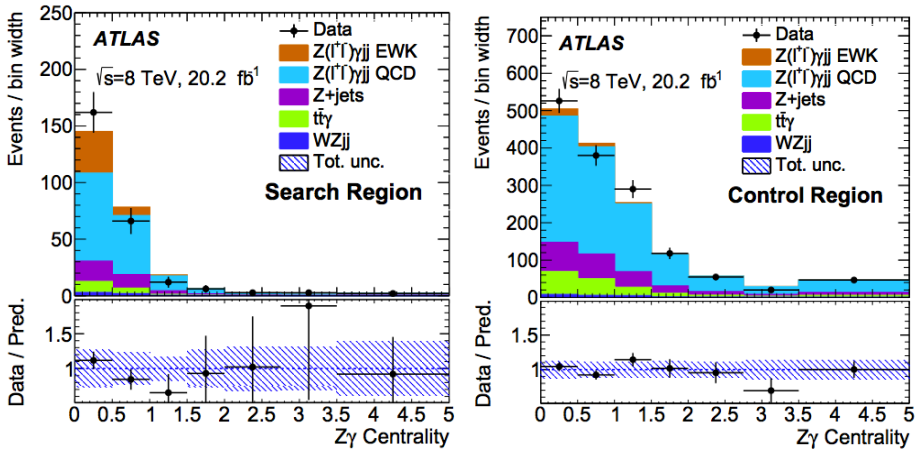
ATLAS Zγjj and Summary

- 8 TeV Zγjj analysis at ATLAS
  - Interpretation with and without EWK extraction
  - For EWK extraction, QCD shape and interference from Sherpa
  - Normalization from control region

arXiv:1705.01966

From N. Martinez, LHCP

Source of uncertainty	EWK [%]	Total (EWK+QCD) [%] SR	CR
Statistical	40	9	4
Jet energy scale	36	9	4
Theory	10	5	4
All other	8	5	6
Total systematic	38	11	8



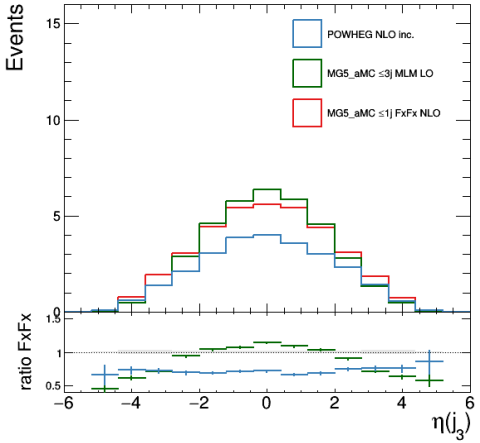
- **Inclusive region :** for checks
- **Control region**: 150<m_{jj}<500 GeV (constrain QCD norm, <5% of signal)
- **Search region**: m_{jj} > 500 GeV (VBS enhanced, >20%) ; N_{exp}=22.8 +/- 1.5

- Summary
  - Background composition in W±W±jj is unique
  - For other channels, signal vs. background from shapes of sensitive distributions
    - Constraints (e.g. normalization) in control regions from data, but rely on MC for categorization

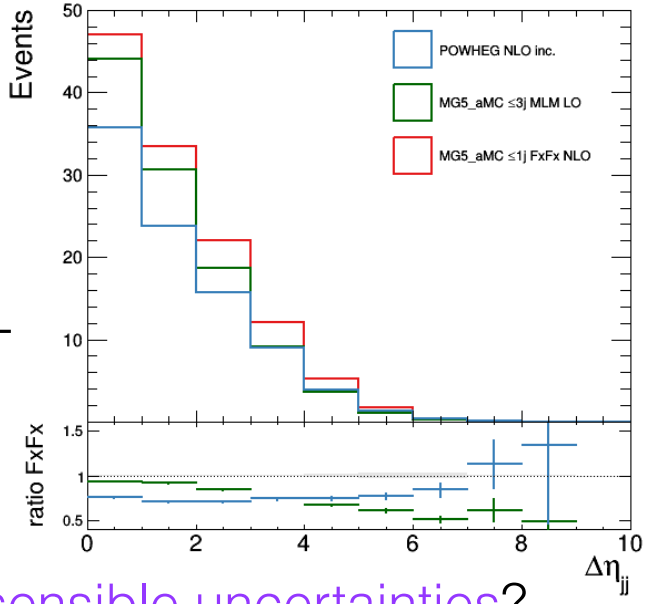
Comparisons of QCD Background

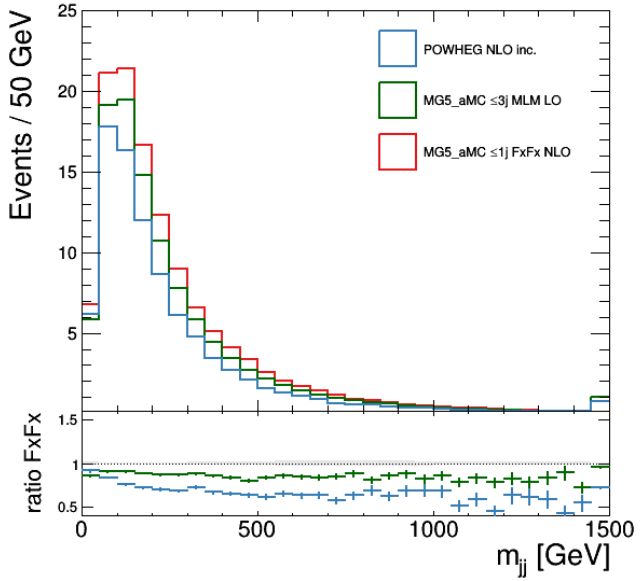
- Example: background comparisons for WZ
  - All showered with Pythia 8 (CUETM1 tune)

Generator	Order (merging)	Cards
MG5_aMC@NLO LO	LO (≤3j MLM)	cards dir
MG5_aMC@NLO NLO	NLO (≤1j FxFx)	cards dir
POWHEG Box	NLO (None)	powheg.input



- Expect differences: fundamental differences between generators
- Process availability + resources mean we can't study everything





- How can we derive sensible uncertainties?

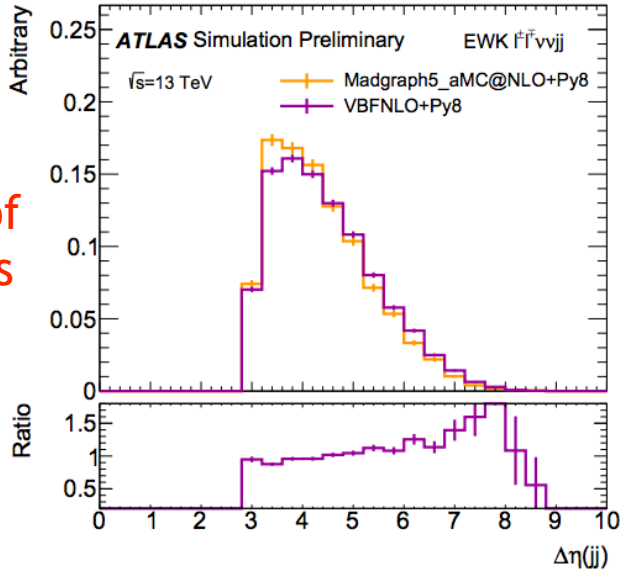
EWK Comparisons

- For showered+hadronized events, differences in EWK processes aren't always within published (fixed order) uncertainties
  - Extensive comparisons published by ATLAS
- What does this tell us on how we should derive uncertainties?

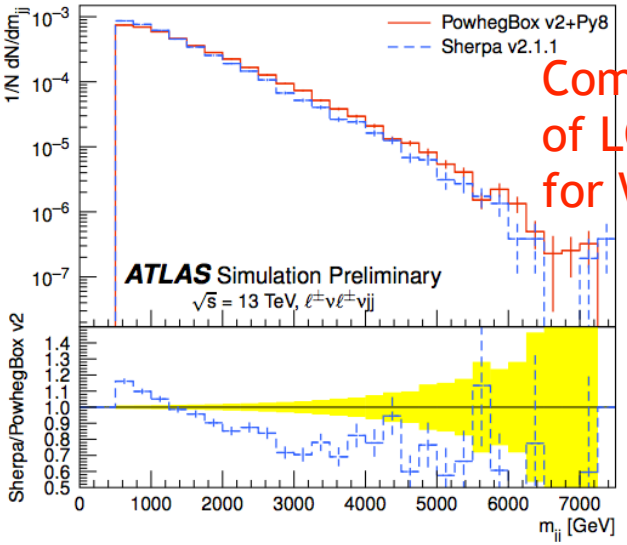
ATL-PHYS-PUB-2017-005

		$VV+2j$	$VV+3j$	$VV+\geq 4j$
$VVjj = \ell^{\pm}\ell^{\mp}2\nu jj$	VBFNLO+PYTHIA8	LO	PS	PS
	MadGraph5_aMC@NLO+PYTHIA8	LO	PS	PS
$VVjj = \ell^{\pm}\ell^{\pm}2\nu jj$	Sherpa	LO	PS	PS
	PowhegBox+PYTHIA8	NLO	LO	PS

Comparison of LO generators for $W^{\pm}W^{\mp}jj$



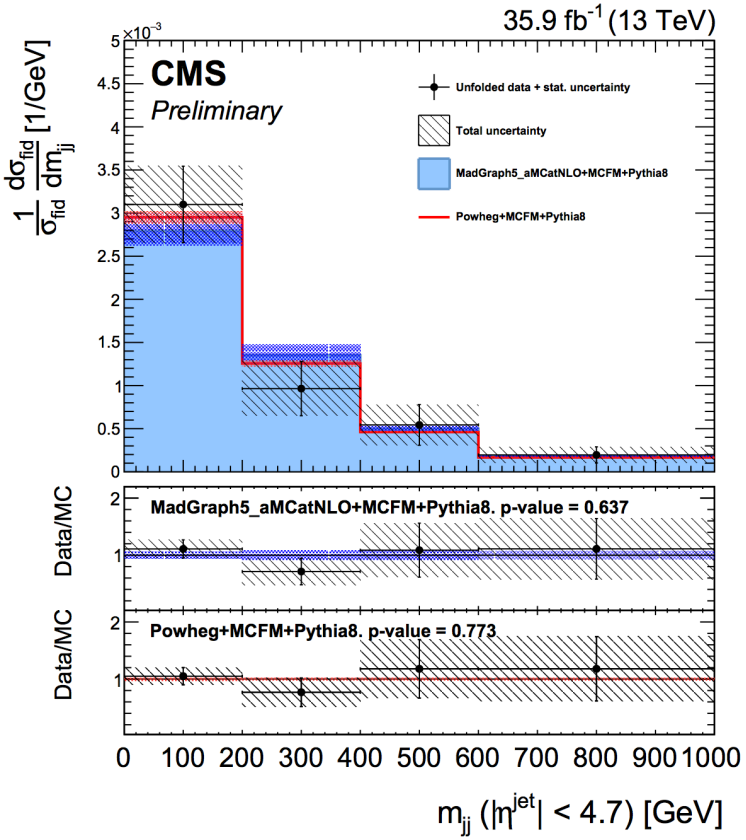
Comparison of LO vs NLO for $W^{\pm}W^{\pm}jj$

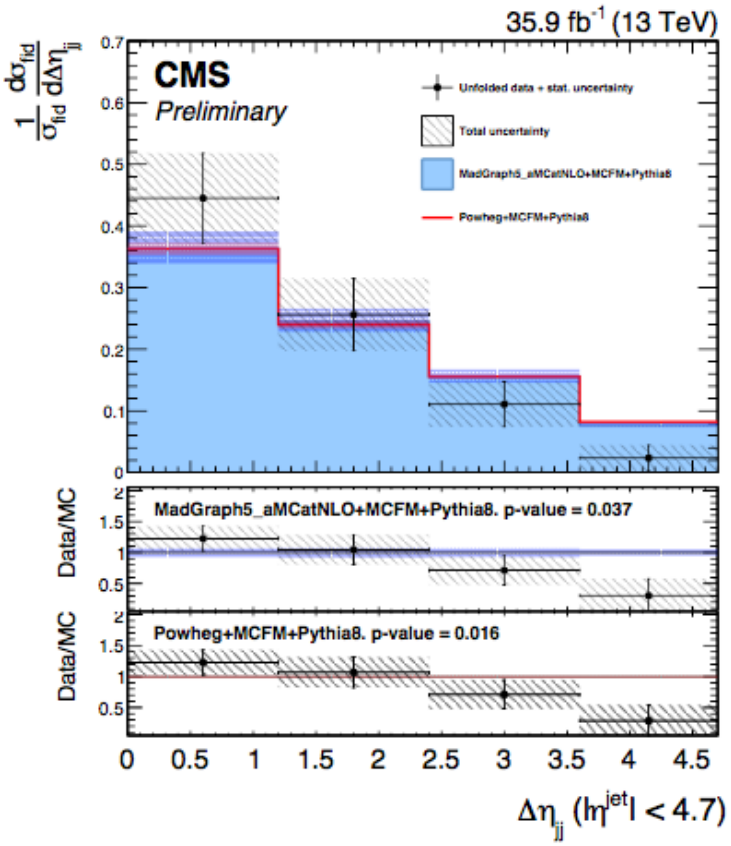


(d) m_{jj}

Background Constraints from Data?

- Following V+jets treatment in mono-jet
  - Is it theoretically well motivated to use measurements in one VV+jets channel to constrain background in other channels?
  - Example: ZZ+jets to contrain QCD induced WW/WZ





Conclusions and Moving forward

- Experimental measurements of VBS (often) rely on modeling to separate QCD induced backgrounds and VBS signal
  - Stat uncertainties becoming subdominant with 2016 (and beyond)
  - Modeling uncertainties similar to experimental ones (e.g. JES/JER) in some cases
  - Attributing sensible modeling uncertainties is a challenge when options are limited and give varying results
    - Being too conservative directly hurts sensitivity

- We would be interested in a broad study of the modeling options and performance for background processes
  - How applicable are studies of background modeling in one channel to another?
- Would also benefit from studies of signal modeling
- Open to suggestions on the signal and background treatment