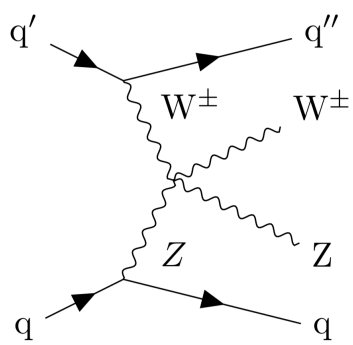
Vector boson fusion and scattering measurements at ATLAS and CMS

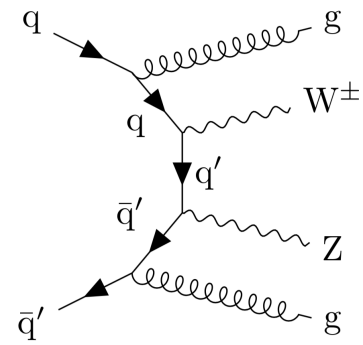
Kenneth Long
University of Wisconsin — Madison
for the ATLAS and CMS Collaborations

Introduction and motivation

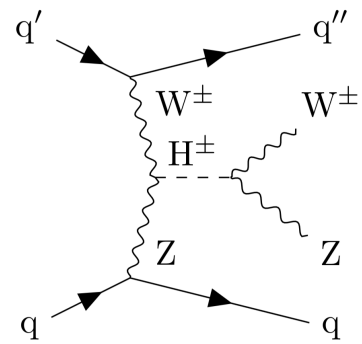
- V/VV production via vector boson fusion/scattering
  - Important component of VVjj (Vjj) production proceeding entirely via EW interactions at tree level
  - Given SM Higgs, interactions with vector bosons, and V self-interactions precisely predicted
    - Deviations from predictions signal new physics in EW sector



VBS WZjj



Non-VBS WZjj

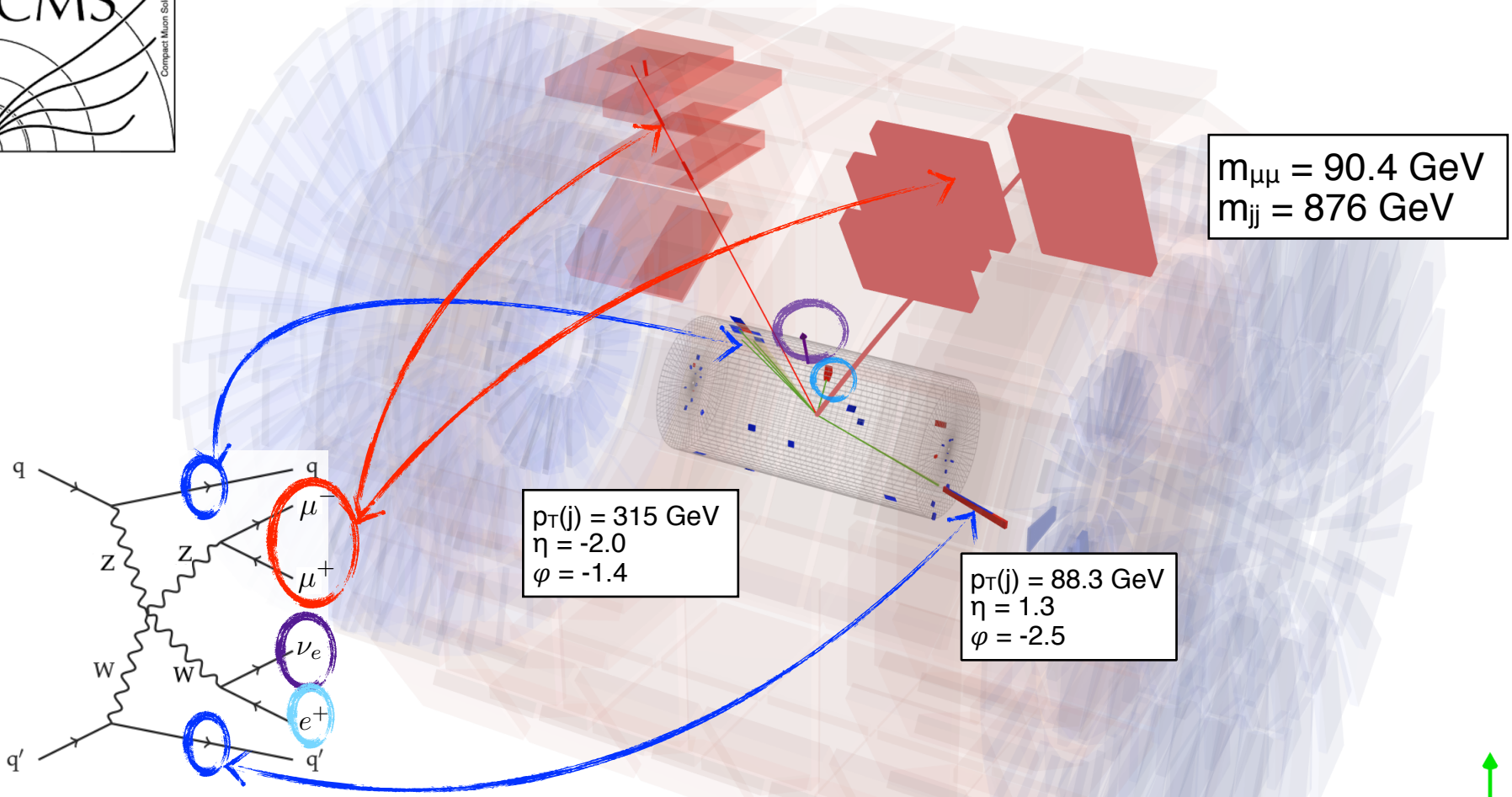


BSM H± Production

- VBF: standard candle for VBF H, VBS
- Low cross sections for VBS just becoming accessible
  - Does VBS production occur with the rate predicted by the SM?
  - Do distributions show any signs of BSM physics?

Characteristics of VBS/VBF events

- Radiation of vector bosons, lack of color flow between jets
  - ➡ Distinct kinematic signature for VVjj EW component



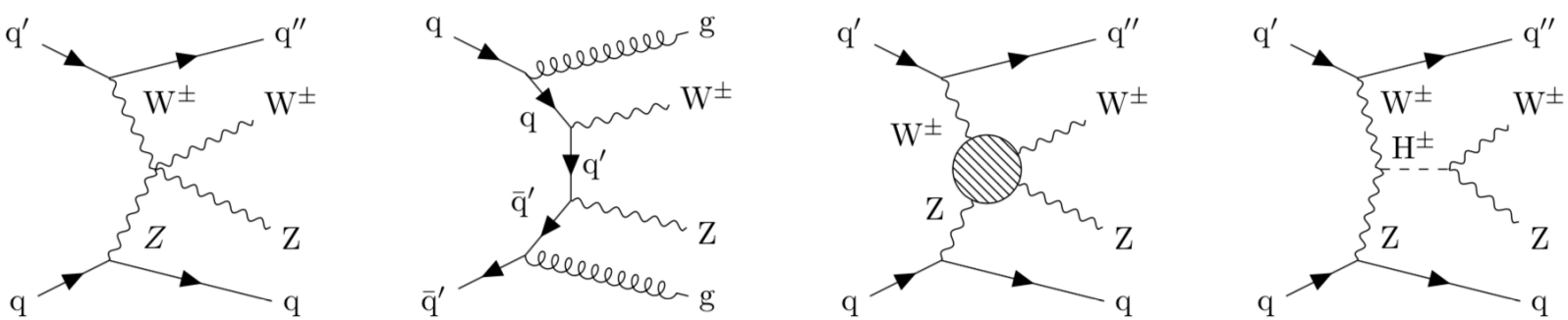
CMS Experiment at LHC, CERN
Data recorded: Wed Oct 12 18:07:34 2016 CDT
Run/Event: 283043 / 94262902

- Forward and high momentum jets
- Leptons central wrt jets

Overview of VBF (VBS) analyses

- Select V (VV) events with VBF-like jets
  - Dominant experimental uncertainty: jet energy scale
- Estimate non-V (VV) backgrounds — usually data driven

1. Measure Vjj (VVjj) cross section (treat (a) + (b) as signal)
   - Theoretical dependence minimal for cut-and-count analysis
2. Distinguish EW and QCD production mechanisms through kinematics variables (e.g., of two highest p_T jets)
   - Treat (a) as signal, (b) as background
   - Modeling uncertainties important for MC-driven backgrounds
   - Multi-variate — best sensitivity, less explicit theoretical assumptions
3. Look for new physics modifying the VVV (VVVV) interaction (**Talk by C. Lee**)
   - Interpret in terms of generic (EFT) (c) or explicit models (d)

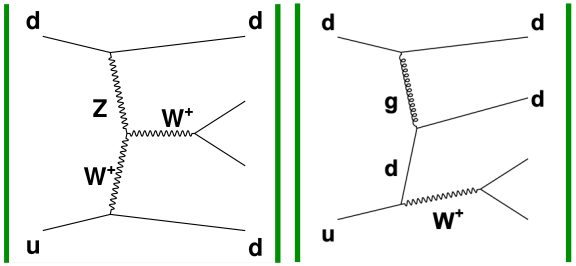


(a) $O(\alpha^4)$ (b) $O(\alpha_s^2\alpha^2)$ (c) (d)

Electroweak VBF W production

- Important "standard candle" for VBF Higgs
- Very high statistics ⟹ precision measurement
- VBF contribution not uniquely separable
  - All fully EW vertices $O(\alpha^4)$ as signal
  - QCD-induced $O(\alpha_s^2\alpha^2)$ as background
  - Interference fully included (as background)

$\sigma_{int} \sim$

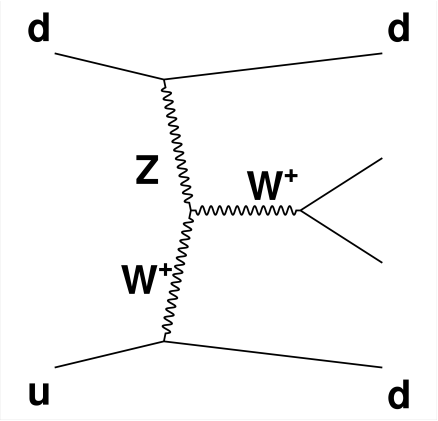


- Selection:
  - 1 lepton > 30 (25) GeV for e (μ),
  - p_T^{miss} > 40 GeV
  - m_{jj} > 200 GeV
- Backgrounds:
  - W+jets (QCD Wjj) very dominant
    - Modeled with MG5_aMC ≤2j@NLO (FxFx) and MG5_aMC ≤4j@LO (MLM) +Pythia8

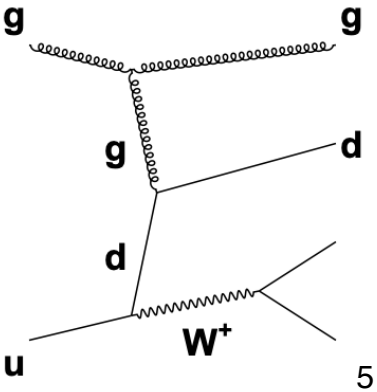
arXiv:1903.04040

Submitted to EPJC

VBF production



QCD production



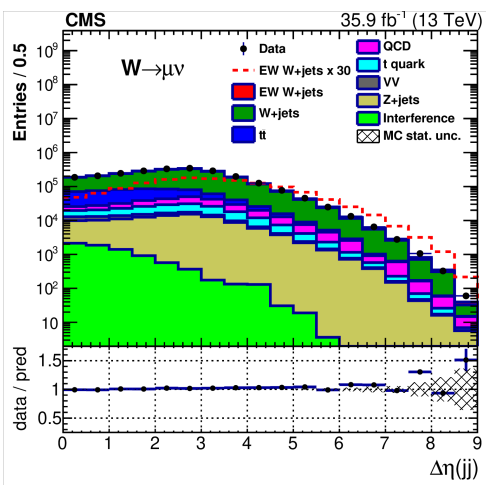
VBF W at CMS: Procedure

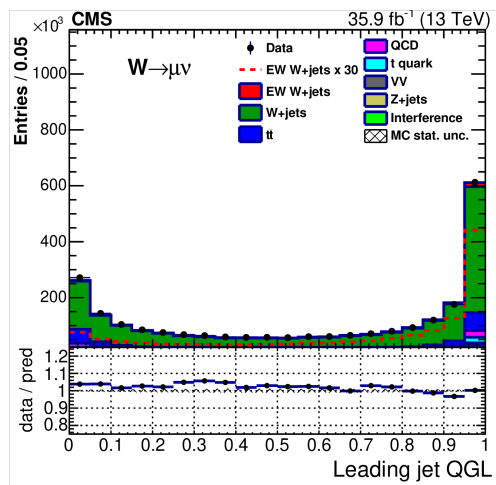
- Train BDT with 4 discriminating variables
  - m_{jj}, $\Delta\eta_{jj}$, $z^*(W)$, quark/gluon likelihood (QGL)
  - BDT trained and fit separately for ee and μμ
- LO MG5_aMC+Pythia8 used for training, NLO+Pythia8 used for background in fit
  - Data generally well-modeled by NLO MC
  - Shape uncertainty from NLO scale+PDF
- σ via fit to transformed BDT output (BDT')

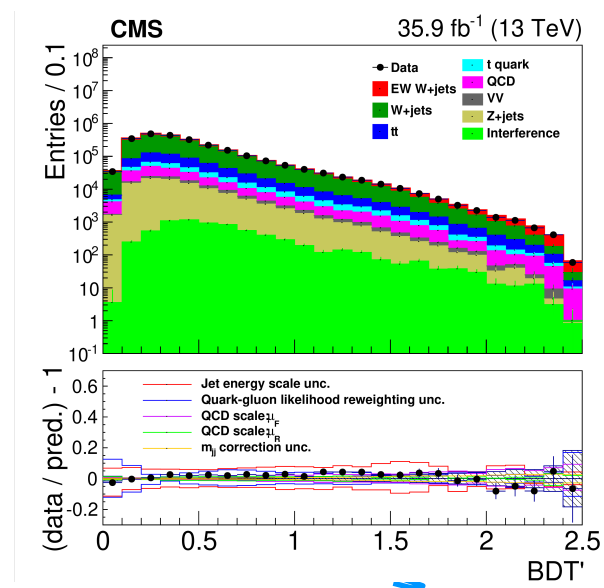
arXiv:1903.04040

Submitted to EPJC

$$\eta^*(W) = \eta(W) - 1/2(\eta_{j1}+\eta_{j2})$$
$$z^* = \eta^*(W)/\Delta\eta_{jj}$$







VBF W at CMS: Results

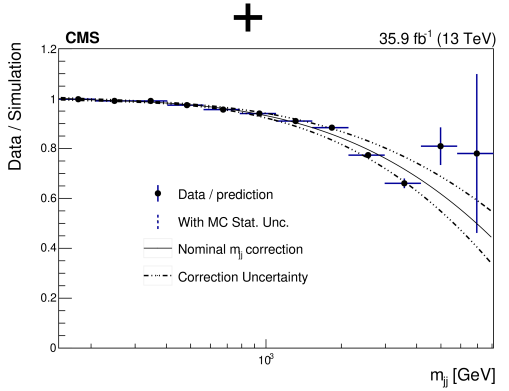
- Correct m_{jj} distribution for lower triggering efficiency in data than in MC
  - Energy shape in forward (|η| > ~2.5) ECAL different from expectation
    - Forward jets associated with earlier bunch crossing
  - Objects from correct bunch crossing lost
- Derive correction to m_{jj} in signal depleted region $R(p_T) = \frac{|\vec{p}_{Tj_1} + \vec{p}_{Tj_2} + \vec{p}_{TW}|}{|\vec{p}_{Tj_1}| + |\vec{p}_{Tj_2}| + |\vec{p}_{TW}|} < 0.2$
  - uncertainty assessed by varying the fit within uncertainties —> dominant uncertainty

$$\sigma(\mathrm{EW}\ \ell\nu jj) = 6.23 \pm 0.12\,(\mathrm{stat}) \pm 0.61\,(\mathrm{syst})\,\mathrm{pb}$$

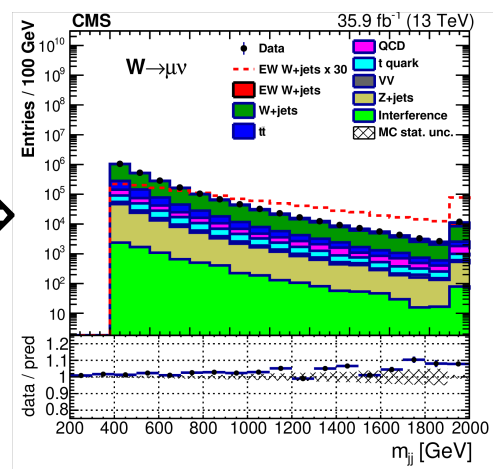
$$\sigma_{\mathrm{LO}}(\mathrm{EW}\ \ell\nu jj) = 6.81^{+0.03}_{-0.06}\,(\mathrm{scale}) \pm 0.26\,(\mathrm{PDF})\,\mathrm{pb}$$

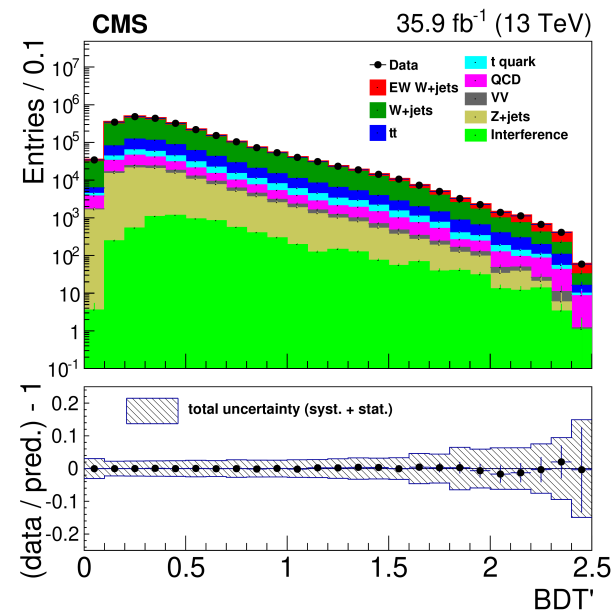
uncorrected m_{jj}

+



⟹



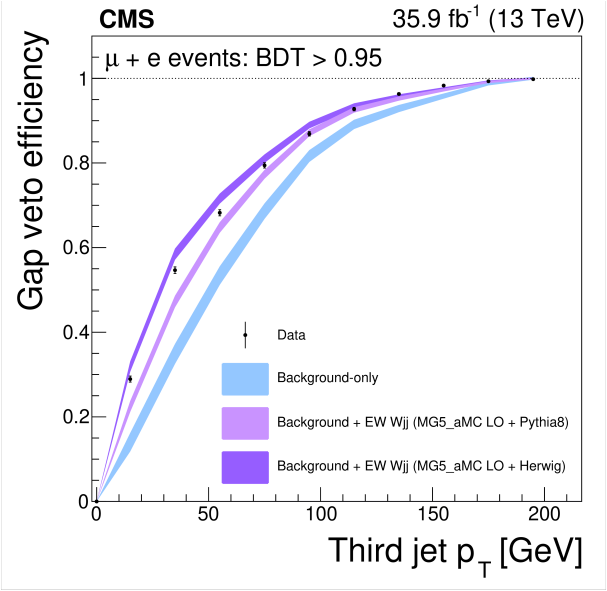


VBF W at CMS: rapidity gap

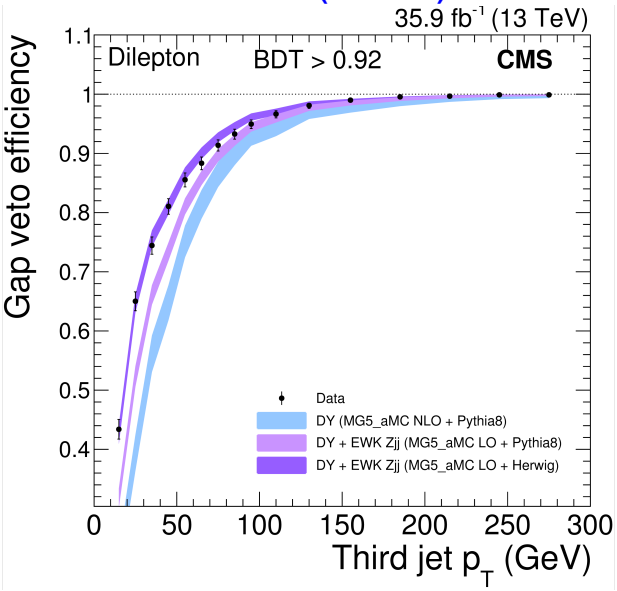
- Study of hadronic activity in VBF-enhanced region
  - Compare EW Wjj signal at LO + Herwig++ and Pythia8 shower
  - Select events with BDT' > 0.95
- Measure fraction of events with no central jet vs. $p_T(j)$ threshold (other "veto" conditions also considered)
- Results in agreement with analogous study for EW Zjj
  - Excess in central radiation observed in other VBS/VBF processes
  - Likely linked to treatment of color connections of ingoing/outgoing quarks

- Statistical uncertainties only in simulation
- Slight preference for Herwig++, but syst. uncertainties in simulation likely large

EW Wjj at CMS



EW Zjj at CMS
EPJC 78 (2018) 589



Electroweak VBS W±W± production

PRL 120, 081801 (2018)

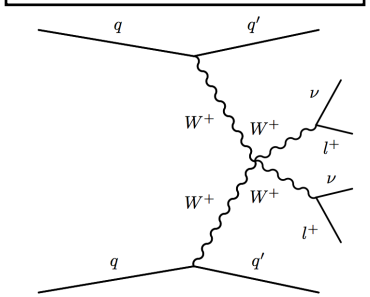
ATLAS-CONF-18-030

- Why W±W±jj ⟶ℓ±ℓ±jj?
  - Golden channel of VBS measurements
    - EW production dominant over QCD-induced
    - Distinct same-sign (SS) lepton state

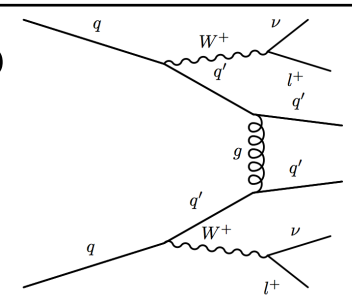
- Results from ATLAS and CMS with similar procedures and selections

Requirement	CMS	ATLAS
Third lepton veto	Applied	Applied
B-tagging veto	Applied	Applied
E_T^{miss} [GeV]	40	30
$m_{\ell\ell} >$ [GeV]	20	20
$p_T^j >$ [GeV]	30/30	65/35
$m_{jj} >$ [GeV]	500	500
$\lvert\Delta\eta_{jj}\rvert >$	2.5	2.0
$max(z_l^*) <$	0.75	N/A

VBS production



QCD production



- Backgrounds
- ≥ 2 prompt SS leptons (WZ, QCD WW) ⟹ from Monte Carlo
  - Correct WZ using data in 3ℓ control region
- Non-prompt backgrounds ⟹ data driven
- Charge mis-ID
  - Simulation corrected with data

Electroweak W±W±: CMS Results

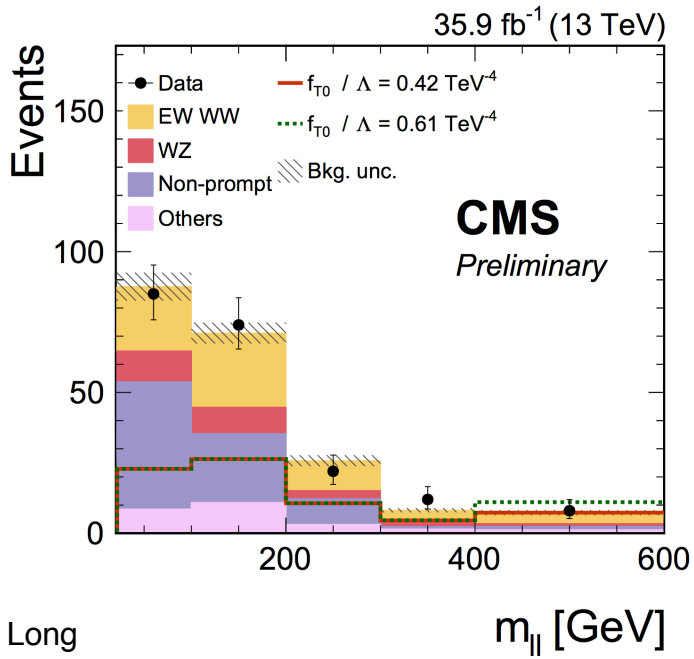
- EW significance and cross section measurement via fit to 2D distribution of m_{jj} and m_{ll}
  - Simultaneous fit to 3ℓ control region to correct WZ m_{jj} shape per bin
- Observed (expected) significance of 5.5σ (5.7σ)

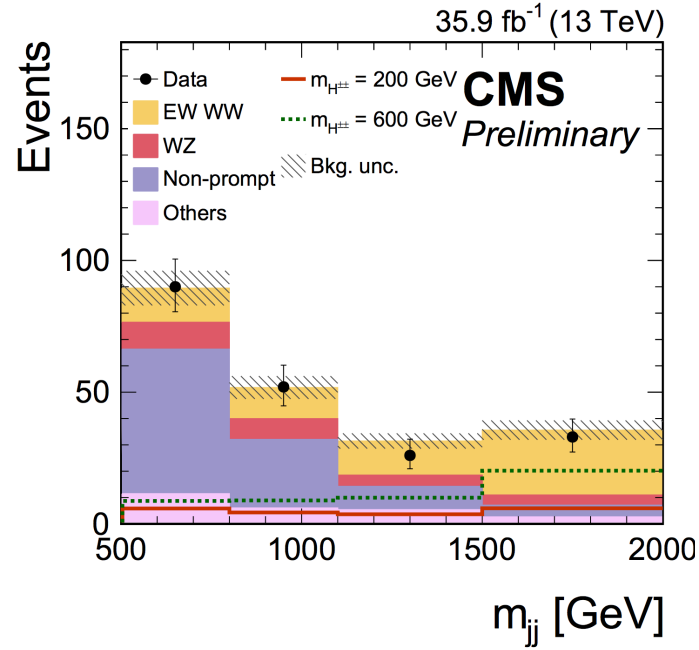
σ_{fid} = 3.83 ± 0.66 (stat) ± 0.35 (syst) fb

PRL 120, 081801 (2018)

Requirement	CMS	ATLAS
$p_T^j > [\mathrm{GeV}]$	30/30	65/35
$\lvert\eta^j\rvert <$	5.0	4.5
$\lvert\Delta\eta_{jj}\rvert >$	2.5	2.0
E_T^{miss} [GeV]	N/A	30
$m_{\ell\ell} > [\mathrm{GeV}]$	N/A	20

- Agrees with MG5_aMC prediction, σ_{LO} = 4.25 ± 0.27





Electroweak W±W±: ATLAS Results

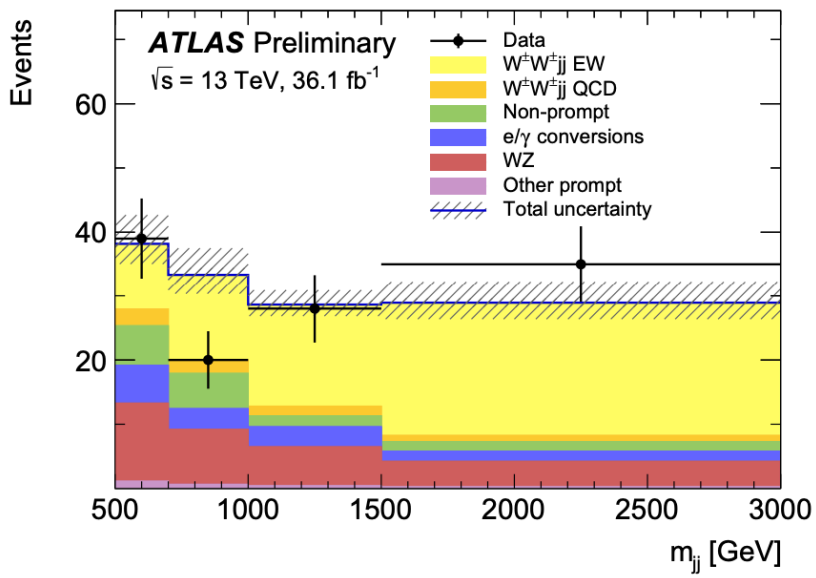
- EW significance and cross section measurement via fit to distribution of m_{jj}
  - Simultaneous fit to yield in 3ℓ control region, WZ normalization free
- Observed (expected) significance of 6.9σ (4.6σ)
  - Expected wrt sherpa

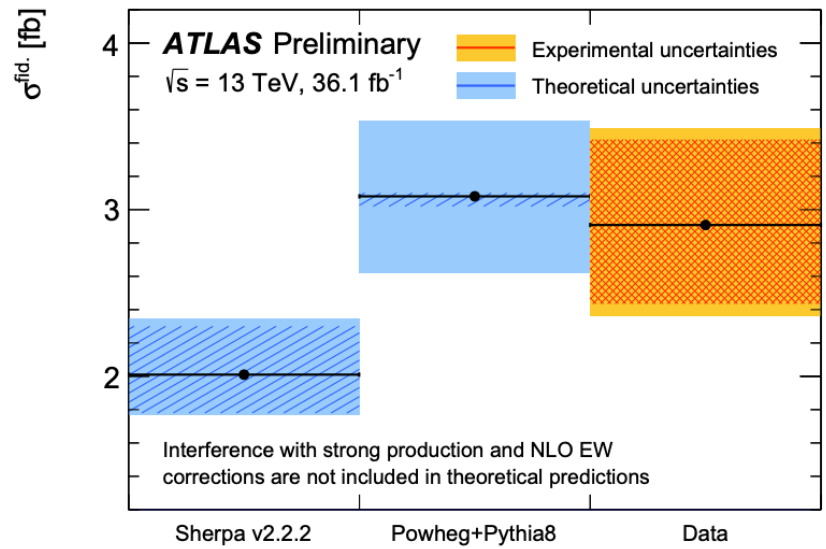
$$\sigma_{fid} = 2.91^{+0.16}_{-0.14}\ (stat) \pm 0.35\ (syst)\ fb$$

ATLAS-CONF-18-030

Requirement	CMS	ATLAS
$p_T^j >$ [GeV]	30/30	65/35
$\lvert\eta^j\rvert <$	5.0	4.5
$\lvert\Delta\eta_{jj}\rvert >$	2.5	2.0
E_T^{miss} [GeV]	N/A	30
$m_{\ell\ell} >$ [GeV]	N/A	20

- Small tension wrt Sherpa, good agreement with POWHEG





Low cross section in Sherpa linked to color flow modeling

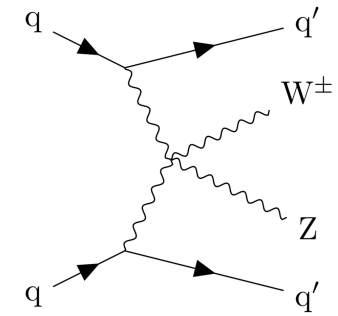
Electroweak WZ production

CMS-SMP-18-001

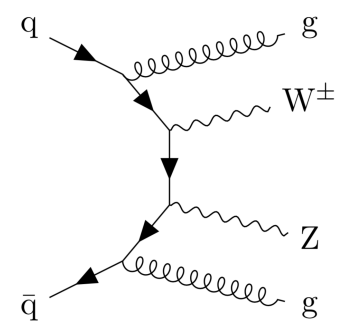
Submitted to PLB

- Why WZjj ⟶3ℓνjj?
  - Sensitive to charged resonances or couplings
  - Less clean signature than ZZ, $W^{\pm}W^{\pm}$, but cross section accessible with large dataset
- Results from ATLAS and CMS with ~similar selections, significant differences in procedure
  - Exactly 3 leptons, moderate p_T + p_T^{miss} or $m_T(W)$
  - Two jets with p_T > 50 (40) GeV
  - CMS: Tight cuts on selected jets
    - $m_{jj} > 500$ && $|\Delta\eta_{jj}| > 2.5$ && $|\eta^*(3\ell)| < 2.5$
  - ATLAS: $m_{jj} > 500$
- Backgrounds
- ≥ 3 prompt leptons ⟹ from Monte Carlo
  - Constrain QCD WZ in control regions
    - ATLAS: $m_{jj} \in [150, 500]$ GeV
    - CMS: $m_{jj} > 100$ GeV + failing signal dijet cut
- Non-prompt backgrounds ⟹ data driven

VBS production

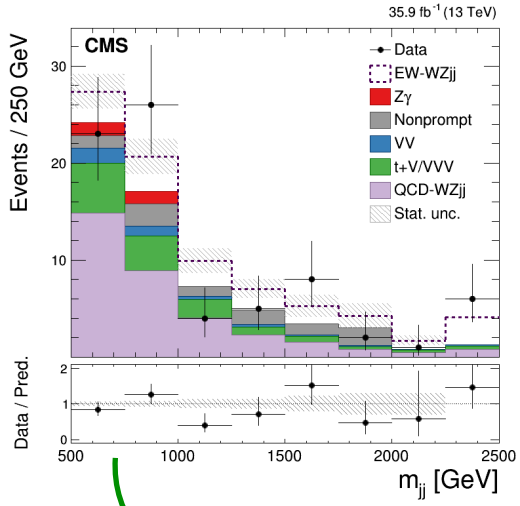


QCD production



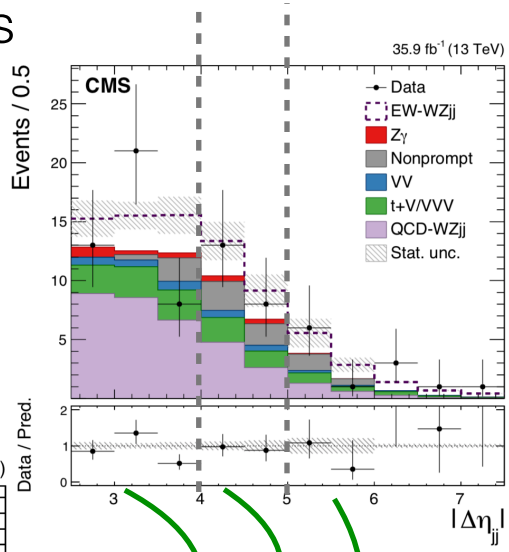
Electroweak VBS WZ: CMS results

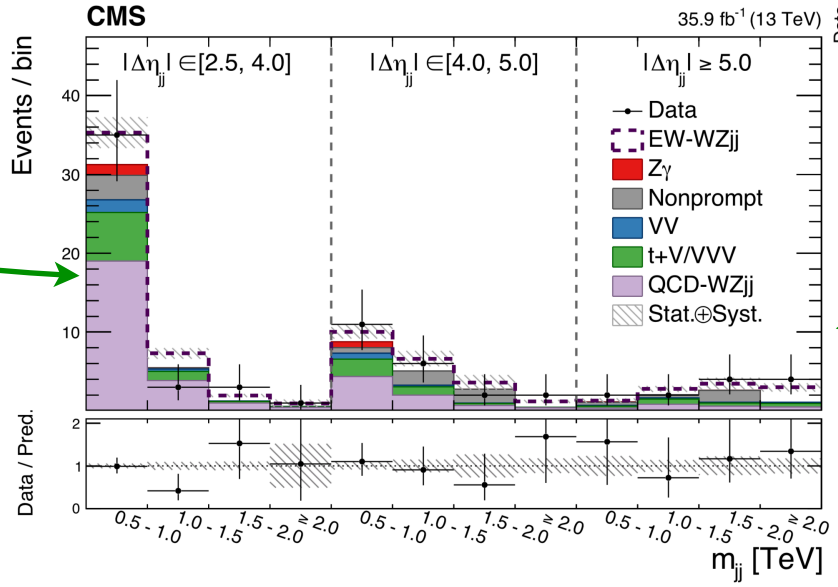
- Simultaneously fit to yield in control region and 2D dist. of m_{jj} and $\Delta\eta(j_1,j_2)$
  - Validate QCD WZ modeling in CR and with MC comparisons
    - Norm. constrained around input + uncertainties
  - Well studied variables, explicit modeling assumptions



$$\mu_{\mathrm{EW}} = \sigma_{obs}/\sigma_{LO,\,MG} = 0.82^{+0.51}_{-0.43}$$

- Observed (expected) significance of 2.2 (2.5)
- Normalization in CR from fit ~1





EW contribution (purple dashed, stacked) rises with increasing $m_{jj}/\Delta\eta_{jj}$

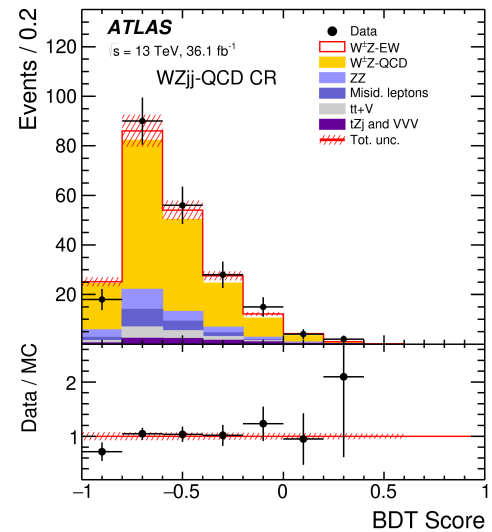
Electroweak VBS WZ: ATLAS results

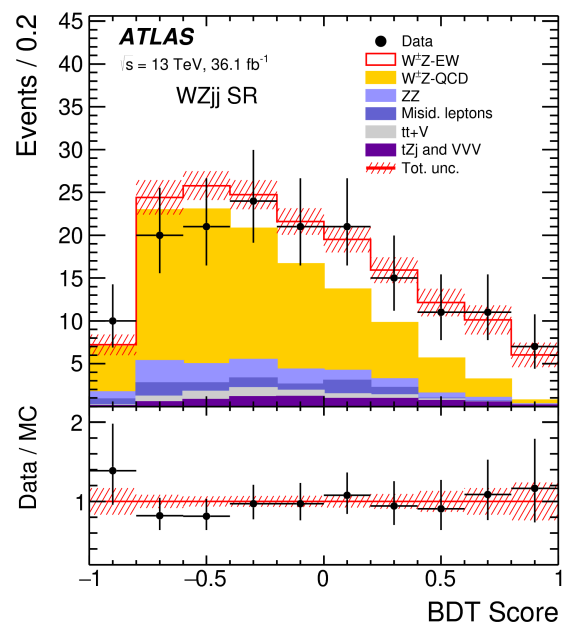
- Train BDT for EW vs QCD discrimination with 15 vars
  - Jet, jet/lepton, vector boson kinematics
    - m_{jj}, N_{jets}, p_T:j_1, j_2 ,W, Z; η: j_1, j_2 ,, $\Delta\eta_{jj}$, $\Delta\phi_{jj}$ …
- Validated MC w/ distributions, low m_{jj} for BDT
- Fit BDT distribution simultaneously with
  - m_{jj} in QCD-WZjj background region
  - #b jets in $n_b > 0$ CR (ttV/tV)
  - m_{jj} in 4ℓ CR (ZZ)
  - ★Normalizations unconstrained in fit

$$\sigma_{\mathrm{EW}}^{\mathrm{fid}} = 0.57^{+0.16}_{-0.14}\ \mathrm{fb}$$

$$\sigma_{WZjj-\mathrm{EW}}^{\mathrm{fid.,\ Sherpa}} = 0.321 \pm 0.002\,(\mathrm{stat.}) \pm 0.005\,(\mathrm{PDF})^{+0.027}_{-0.023}\,(\mathrm{scale})\ \mathrm{fb}$$

  - Observed (expected) significance of 5.3 (3.2)
- Best fit results
  - μ_{QCD} = 0.6
  - μ_{EW} = 1.77





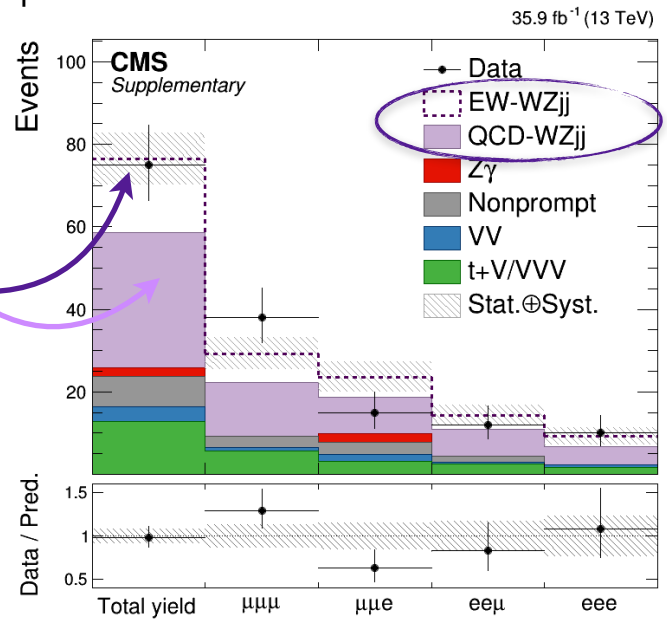
WZjj EW+QCD: ATLAS and CMS

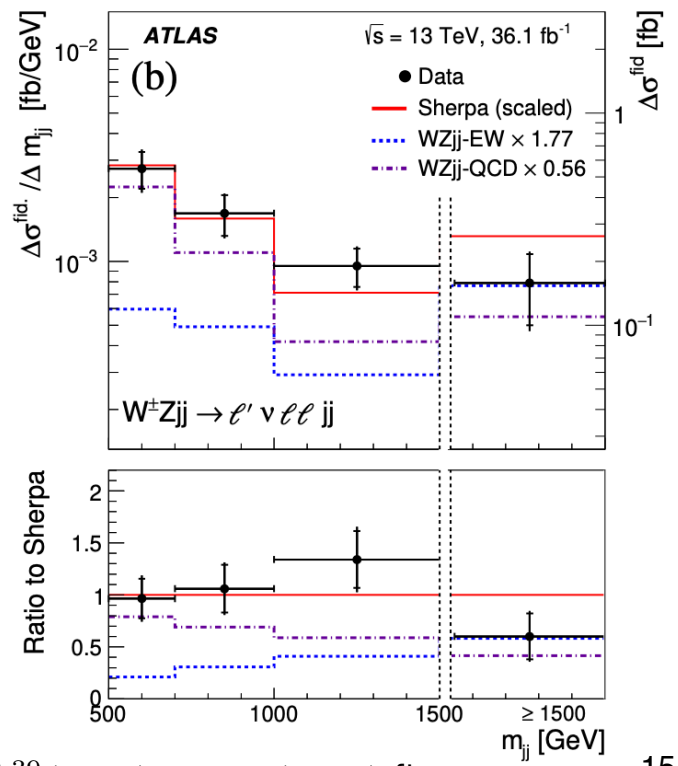
- Also publish combined EW+QCD results
  - CMS: separate fit to the yields by channel with EW+QCD as signal
    - Loss in sensitivity, but nearly no theoretical dependence
  - ATLAS: combine best-fit rate for EW and QCD WZjj from fit to BDT
    - Also present unfolded distributions

$$\sigma^{\text{fid}}_{\text{WZjj}} = 1.68 \pm 0.16\,(\text{stat}) \pm 0.18\,(\text{syst})\ \text{fb}$$

$$\sigma^{\text{fid., Sherpa}}_{W^{\pm}Zjj} = 2.15 \pm 0.01\,(\text{stat.}) \pm 0.05\,(\text{PDF})^{+0.65}_{-0.44}\,(\text{scale})\ \text{fb}$$

EW+QCD treated together as signal!





Single signal strength, loose via acceptance

Tight $\sigma^{\text{fid}}_{\text{WZjj}} = 3.18^{+0.57}_{-0.52}\,(\text{stat})^{+0.43}_{-0.36}\,(\text{syst})\ \text{fb}$

Loose $\sigma^{\text{fid,loose}}_{\text{WZjj}} = 4.39^{+0.78}_{-0.72}\,(\text{stat})^{+0.60}_{-0.50}\,(\text{syst})\ \text{fb}$

Compare tight fid: $\sigma_{\text{fid,MG}} = 3.27\,^{+0.39}_{-0.32}(\text{scale}) \pm 0.15\,(\text{PDF})\ \text{fb}$

Interpretation of results

- Fiducial regions aren't trivially comparable (definitions in backup)
  - Expected: experimental selections are different
  - ➡ Similar $p_T(\ell)$, $\eta(\ell)$, $\eta(j)$, differences in cuts + …
  - CMS: no $p_T(\nu)$ cut, $p_T(j) > 50$, $m_{jj} > 500$ GeV, $|\Delta\eta_{jj}| > 2.5$, $|\eta^*(3\ell)| < 2.5$
  - ATLAS: $p_T(j) > 40$, $m_T(\ell, \nu) > 40$, $m_{jj} > 500$ GeV
  - is there a middle ground?
- In addition, MC predictions may differ significantly

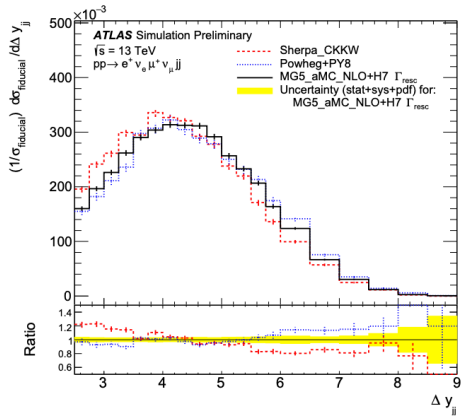
➡ Difficult to conclude if data/MC ratio is a difference in data or MC

Process	Experiment	Obs. (fb)	Pred. (fb)	Obs. ratio	Region
EW WZjj	ATLAS	$0.57^{+0.16}_{-0.14}$	$0.321^{+0.13}_{-0.11}$	$1.77^{+0.49}_{-0.43}$	ATLAS SR
	CMS	—	$1.25^{+0.13}_{-0.11}$	$0.82^{+0.51}_{-0.43}$	CMS tight SR
WZjj (EW+QCD)	ATLAS	$1.68^{+0.25}_{-0.25}$	$2.15^{+0.65}_{-0.44}$	0.78	ATLAS SR
	CMS	$3.18^{+0.71}_{-0.63}$	$3.27^{+0.42}_{-0.35}$	$0.98^{+0.22}_{-0.20}$	CMS tight SR
QCD WZjj	ATLAS	—	—	$0.56^{+0.16}_{-0.16}$	ATLAS CR
	CMS	—	$18.6^{+0.31}_{-0.25}$	~1.02	CMS tight CR

Monte Carlo for VBF/VBS

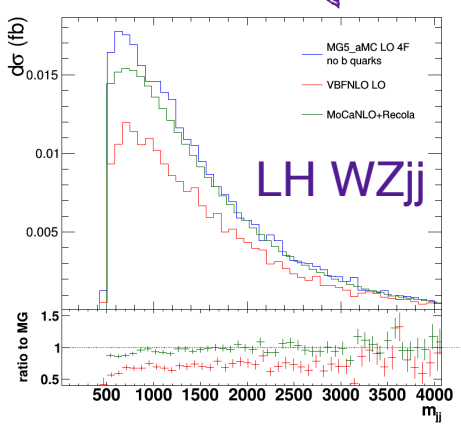
- Monte Carlo simulations are critical to VBF/VBS measurements
  - Rely on differential predictions for EW/QCD separation
- Major theoretical efforts made for state-of-the art predictions
  - e.g., NLO EW: arxiv:1904.00882, arxiv:1611.02951
- In addition to pushing state-of-the-art, critical to validate "workhorse" tools
  - ATLAS study of SS WW
  - Les Houches (LH) study of WZ
  - VBSCan study of SS WW

➡ Theory+experiment efforts: excellent agreement can be obtained, but not for free



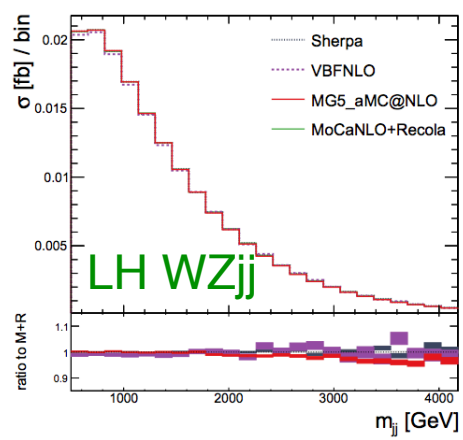
ATLAS W±W±jj

Untuned



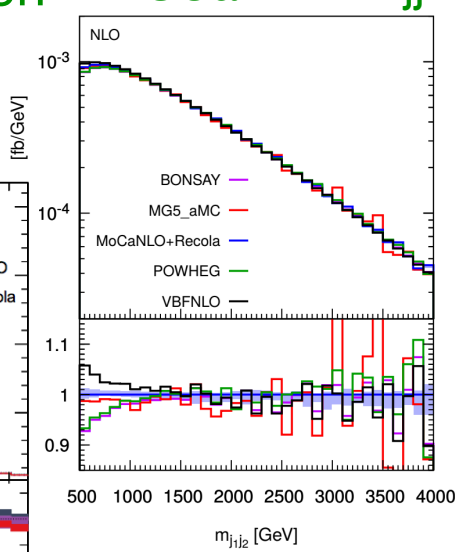
LH WZjj

Tuned configuration



LH WZjj

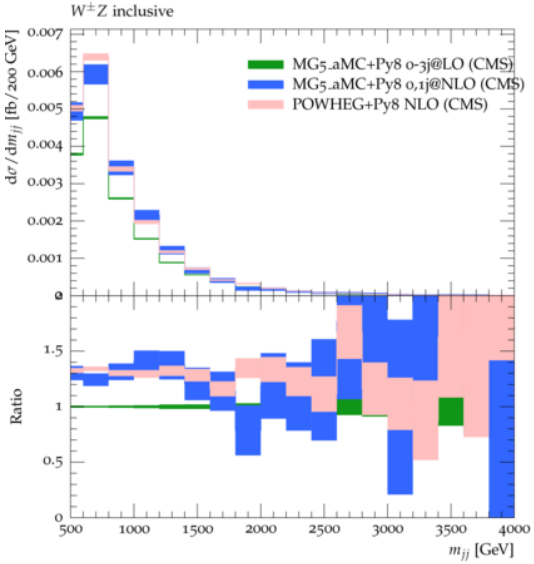
VBSCan W±W±jj



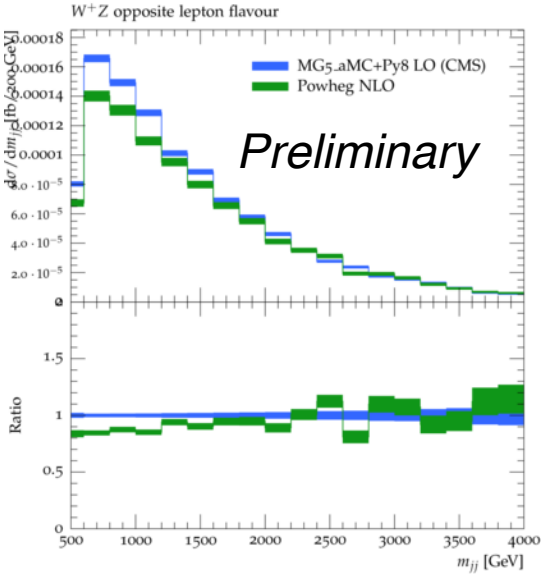
MC and interpretation of results

- Valuable insight to MC developers and experimentalists from such studies
  - Not trivial to claim a "unique" σ_{LO}
    - Scale choice? Showering? Other input params?
  - There is likewise no one $\sigma_{MG5\_aMC@NLO}$, σ_{Sherpa}, etc…
- Continue to work to fully validate simulations
  - When are differences uncertainties vs. misconfigurations?
- Experimental studies must acknowledge this
- Can we better share info about our simulations?
  - Rivet obvious choice
  - For QCD WZ, is $0.56\sigma_{Sherpa}(ATLAS) = 1.02\sigma_{MG}(CMS)$?
  - For WW, is Sherpa the outlier (seems yes)?
- Interest is there, work is ongoing in EWWG, VBSCan

Predictions for QCD WZjj used in CMS analysis



CMS MG simulation vs. new NLO EW WZjj in POWHEG
Jager, Karlberg, Sheller

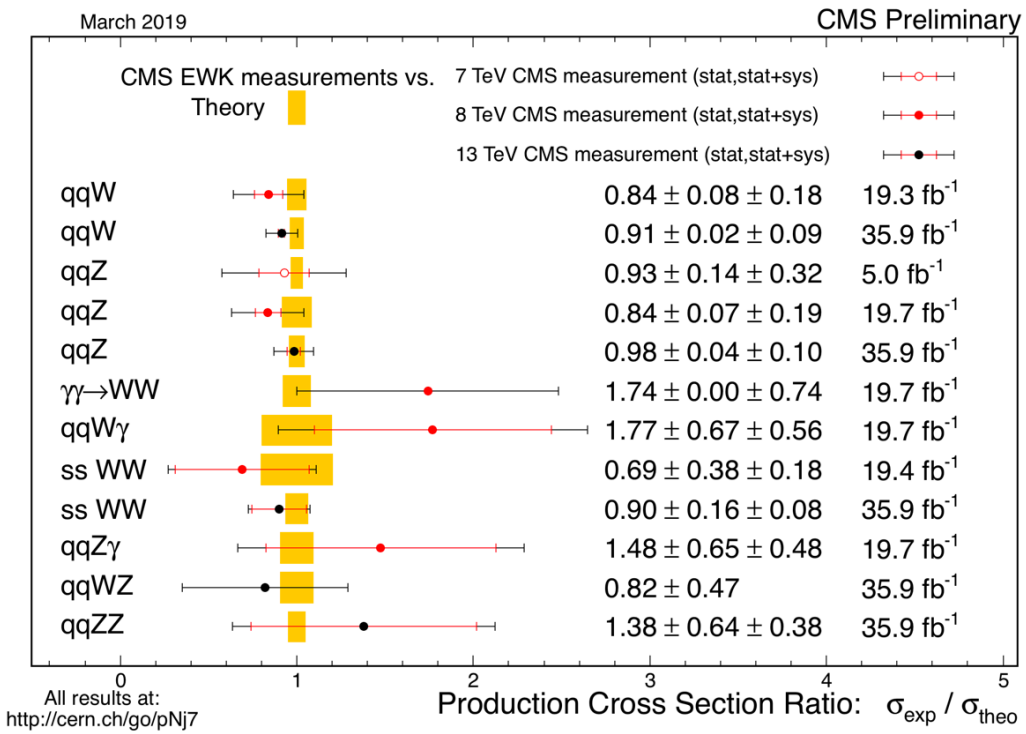


★ Implemented in Rivet — publicly available

Conclusions

- New, exciting measurements from ATLAS and CMS at 13 TeV
  - VBF V measurements with high precision
  - Moving from observation to measurement phase in VBS
- Run II data is being analyzed
  - High-precision measurements ⟹ theory/cross-experiment comparisons interesting
- Improvements in simulation, and how we exploit it, will be valuable

- New tests of the SM… which the SM is withstanding
  - Deviations could be subtle
  - More data and improved experimental and theoretical techniques help look for cracks with increased precision and at higher energy scales



Conclusions

Backup

Overview of VBS WZ fiducial regions

CMS

	Tight fiducial	Loose fiducial
$p_T^{\ell'_1}$ [GeV]	> 25	> 20
$p_T^{\ell'_2}$ [GeV]	> 15	> 20
p_T^{ℓ} [GeV]	> 20	> 20
$\lvert\eta^{\mu}\rvert$	< 2.5	< 2.5
$\lvert\eta^{e}\rvert$	< 2.5	< 2.5
$\lvert m_{\ell'\ell'} - m_Z\rvert$ [GeV]	< 15	< 15
$m_{3\ell}$ [GeV]	> 100	> 100
$m_{\ell\ell}$ [GeV]	> 4	> 4
p_T^{miss} [GeV]	-	-
$\lvert\eta^{j}\rvert$	< 4.7	< 4.7
p_T^{j} [GeV]	> 50	> 30
$\lvert\Delta R(j,\ell)\rvert$	> 0.4	> 0.4
n_j	≥ 2	≥ 2
p_T^{b} [GeV]	-	-
n_b	-	-
m_{jj}	> 500	> 500
$\lvert\Delta\eta_{jj}\rvert$	> 2.5	> 2.5
$\left\lvert\eta^{3\ell} - \frac{1}{2}(\eta^{j_1} + \eta^{j_2})\right\rvert$	< 2.5	-

ATLAS

- $p_T(\ell_{Z1}) > 15$ GeV
- $p_T(\ell_{Z2}) > 15$ GeV
- $p_T(\ell_W) > 20$ GeV
- $\eta(\mu, e) < 2.5, 2.5$
- $m_T(W) > 30$ GeV
- $p_T(j) > 40$ GeV
- $\eta(j) < 4.5$
- $\Delta R(\ell, j) > 0.3$
- $m_{jj} > 500$ GeV (signal)
  - No other dijet cuts to define signal region

Overview of MC generators for WZ

- **CMS**: All samples use Pythia 8 for shower+had., with exception of GEN studies for signal modeling
  - Signal: MG5_aMC LO
  - QCD Background: MG5_aMC ≤3j@LO, MG5_aMC ≤1j@NLO
  - tZq: NLO MG5_aMC@NLO 4f, scaled to 5f NLO
  - ttV: NLO MG5_aMC@NLO
  - Others: Generally NLO MG

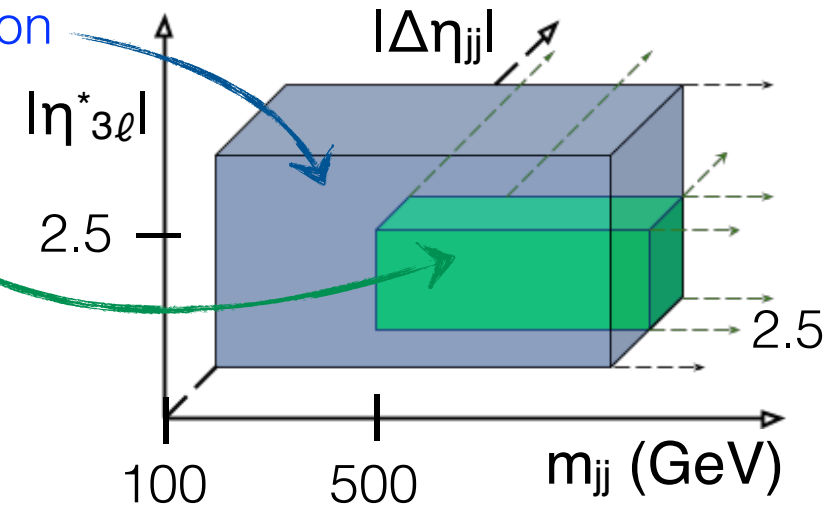
- **ATLAS**: Generally use Sherpa for ME+shower+had., otherwise shower from Pythia
  - Signal: Sherpa LO
  - QCD Background: Sherpa v2.2 with ≤1j@NLO+≤3j@LO,
    - POWHEG+Pythia,Herwig++
  - tZq: LO MG5_aMC@NLO
  - ttV: LO MG5_aMC@NLO
  - Others: Generally Sherpa LO

Overview of MC generators for $W^{\pm}W^{\pm}$

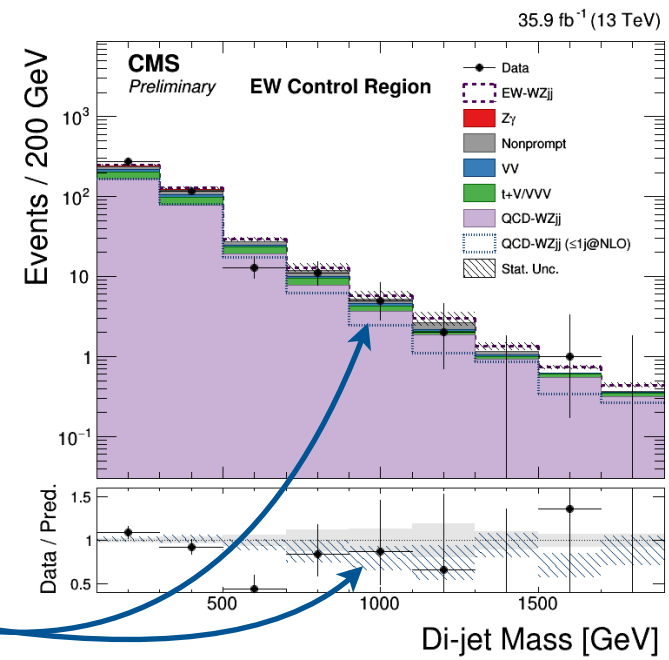
Process	CMS	ATLAS
EWK same-sign WW	MADGRAPH	SHERPA
QCD same-sign WW	MADGRAPH	SHERPA
$V\gamma$	MADGRAPH	SHERPA
VVV	aMC@NLO	SHERPA
$t\bar{t}V$	aMC@NLO	aMC@NLO
QCD WZ	MADGRAPH	SHERPA
EWK WZ	MADGRAPH	SHERPA
QCD ZZ	MADGRAPH	SHERPA
EWK ZZ	MADGRAPH	SHERPA

CMS: Constraining QCD WZ modeling

- Constrain normalization in sideband region
  - m_{jj} > 100 GeV, fail dijet signal cuts
  - Good agreement in CR



- Check that scale+PDF uncertainty covers uncertainty
- Nominal sample: simulated with MG5_aMC+Py8 WZ+≤3j@LO
  - Compare to predictions from MG5_aMC+Py8 WZ+≤1j@NLO,
    - each normalized to data in sideband region
  - Good agreement in sideband and signal regions
  - Uncertainty: LO scale+PDF+10% normalization covers largest differences



Alternative QCD-WZjj