Observation of vector boson scattering ZZjj process with the ATLAS detector

Jing Chen

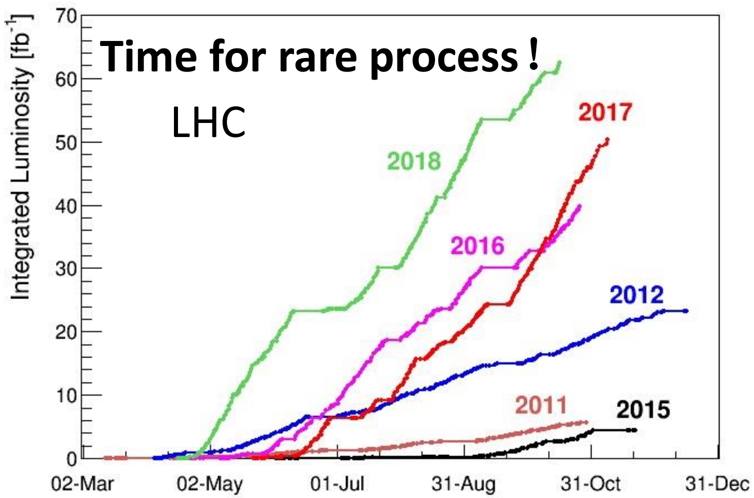
The 5th China LHC Physics Workshop

University of Science & Technology of China

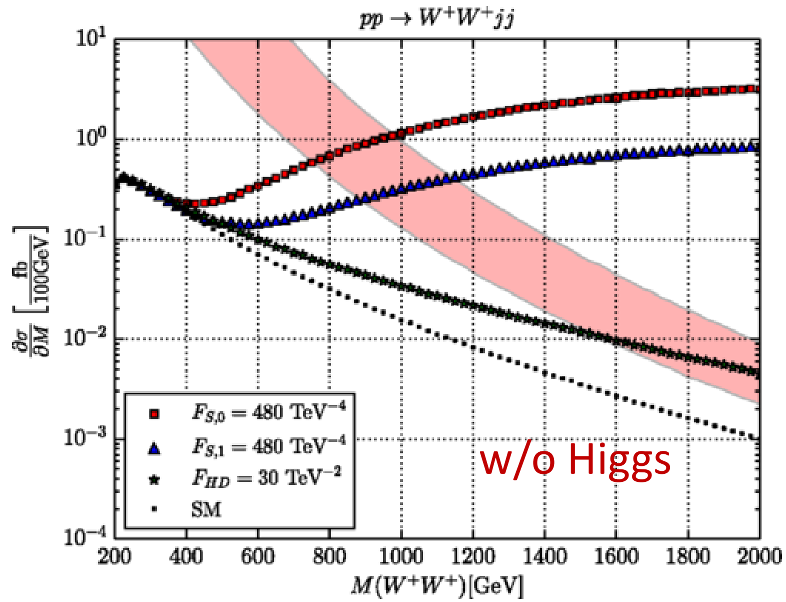
Oct. 24th, 2019

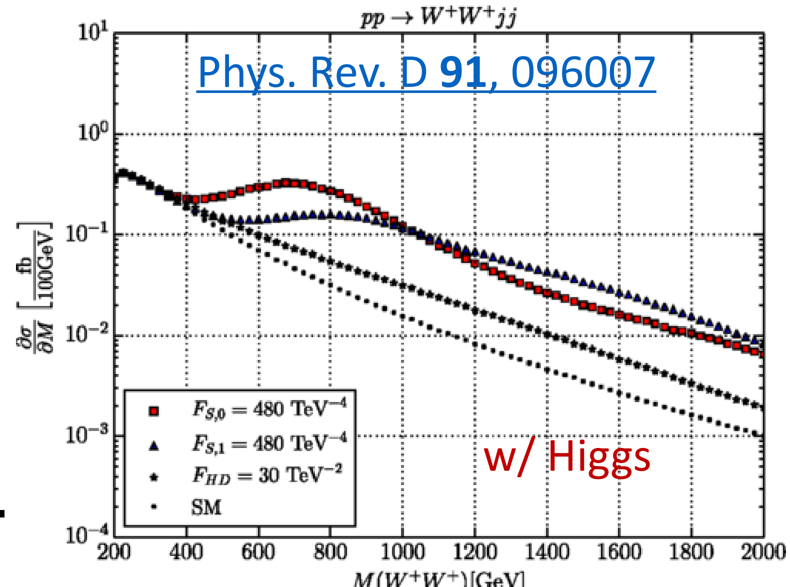
Report Number：ATLAS-CONF-2019-033

Introduction



https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.luminosity





- The presence of the Higgs boson prevents the VBS amplitudes from violating unitarity at the TeV scale!

- **Vector boson scattering is a key process to probe the nature of electroweak symmetry breaking.**

VBS VVjj measurements in ATLAS and CMS

13TeV	Observed (expected) significance		Challenges
	ATLAS	CMS	
ssWW	$6.5(4.4)\sigma$ Phys. Rev. Lett. 123 (2019) 161801	$5.5(5.7)\sigma$ PRL 120 (2018) 081801	First observation of VBS: large ratio of EW to strong production cross-sections.
WZ	$5.3(3.2)\sigma$ Phys. Lett. B 793 (2019) 469	$2.2(2.5)\sigma$ PLB 795 (2019) 281	Similar cross-section as ssWW, but larger QCD background.
ZZ	$llll jj + ll\nu\nu jj$ 5.5(4.3) ATLAS-CONF-2019-033	$llll jj$ $2.7(1.6)\sigma$ *Phys. Lett. B 774 (2017) 682*	$llll jj$ channel: small cross-section, low background, fully reconstructed final state. $ll\nu\nu jj$ channel: relatively large cross-section, complex background components, large uncertainties from jet/E_T^{miss} reconstruction.

Frst observation of EW ZZjj production

Analysis overview

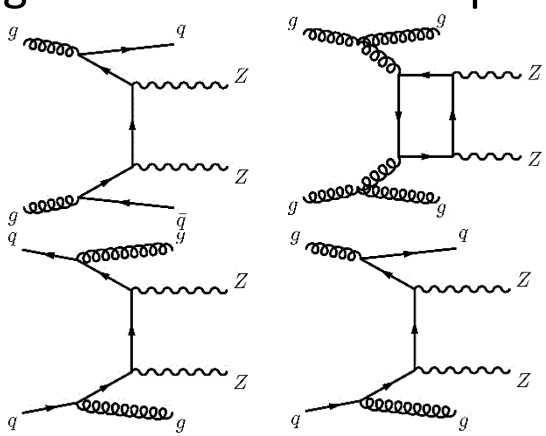
- **Physics goals**

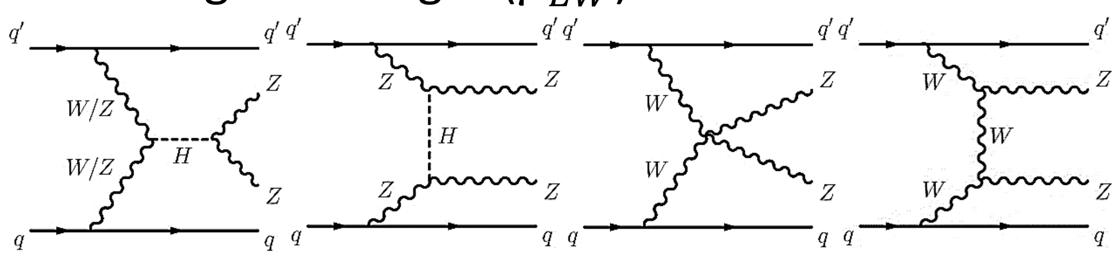
  - Measure the ZZjj cross-section (EW + QCD):

    Measure the inclusive cross-section for $llll jj$ and $ll\nu\nu jj$ separately in their corresponding fiducial volume.

  - Evidence on EW VBS ZZjj

    Combine $llll jj$ and $ll\nu\nu jj$, fit the MVA output to extract the significance of EW component and signal strength (μ_{EW}).





- **Data and MC samples**

  - Full Run2 datasets (139fb^{-1})

  - MC samples

EW ZZjj	MadGraph
QCD qqZZjj	Sherpa222
QCD ggZZjj	$llll jj$:Sherpa222; $ll\nu jj$:gg2vv
Triboson and other diboson	Sherpa222, Powheg
Z+jets	Sherpa221
Top	Powheg, MadGraph, Sherpa221

Selections

	$\ell\ell\ell\ell jj$	$\ell\ell\nu\nu jj$
Electrons	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.47$ $\lvert d_0/\sigma_{d_0}\rvert < 5$ and $\lvert z_0 \times \sin\theta\rvert < 0.5$ mm	
Muons	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.7$ $\lvert d_0/\sigma_{d_0}\rvert < 3$ and $\lvert z_0 \times \sin\theta\rvert < 0.5$ mm	$p_{\mathrm{T}} > 7$ GeV, $\lvert\eta\rvert < 2.5$
Jets	$p_{\mathrm{T}} > 30$ (40) GeV for $\lvert\eta\rvert < 2.4$ ($2.4 < \lvert\eta\rvert < 4.5$)	$p_{\mathrm{T}} > 60$ (40) GeV for the leading (sub-leading) jet
ZZ selection	$p_{\mathrm{T}} > 20, 20, 10$ GeV for the leading, sub-leading and third leptons Two OSSF lepton pairs with smallest $\lvert m_{\ell^+\ell^-} - m_Z\rvert + \lvert m_{\ell'^+\ell'^-} - m_Z\rvert$ $m_{\ell^+\ell^-} > 10$ GeV for lepton pairs ★ $\Delta R(\ell,\ell') > 0.2$ $66 < m_{\ell^+\ell^-} < 116$ GeV	$p_{\mathrm{T}} > 30$ (20) GeV for the leading (sub-leading) lepton One OSSF lepton pair and no third leptons ★ $80 < m_{\ell^+\ell^-} < 100$ GeV ★ ★ No b-tagged jets ★ $E_{\mathrm{T}}^{\mathrm{miss}}$ significance > 12 ★
Dijet selection	Two most energetic jets with $y_{j_1} \times y_{j_2} < 0$ ★ $m_{jj} > 300$ GeV and $\Delta y(jj) > 2$ ★ ★	$m_{jj} > 400$ GeV and $\Delta y(jj) > 2$ ★

- ★ To reject events from low mass resonances
- ★ Relatively loose m_{jj} cut to keep more events for further MVA studies
- ★ Reduce top background
- ★ Suppress W background
- ★ Suppress Z+jets background
- ★ Back-to-back topology, enhance S/B ratio

Background estimation

- ***llll jj* QCD background ($qq \rightarrow ZZjj$ and $gg \rightarrow ZZjj$):**

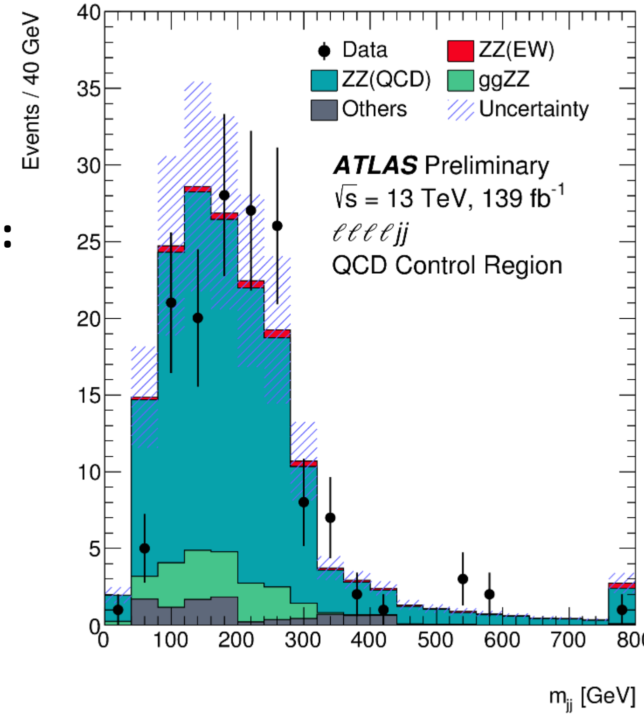
  -- QCD CR: $|\Delta Y_{jj}| < 2$ or $m_{jj} < 300$ GeV

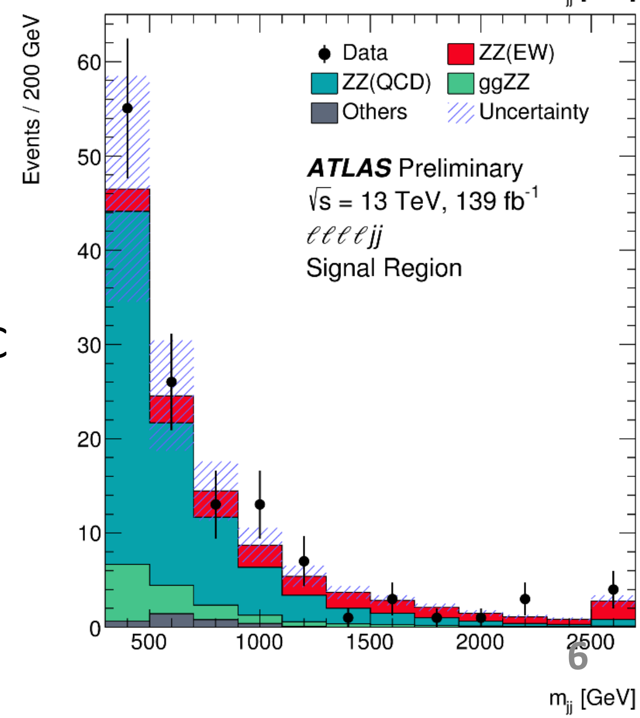
  -- Simultaneous fit SR & QCD CR.

  -- Theoretical uncertainty mainly from generator modelling uncertainty (**Sherpa vs. MG**).

  -- Jet pile-up uncertainty: high-mu vs. low-mu comparison as additional systematic.

  -- The modelling of QCD ZZjj has been cross checked in another high centrality CR.





- ***llll jj* Others background:**

  -- Fake lepton background

  - Fake factor method is used.
  - Systematics: varying "poor" lepton definition, MC contamination, use one bin fake factor instead of p_T/η dependent ones, fake factor difference from data and MC

  -- WWZ…: MC

Background estimation

- **ll$\nu\nu$ Non-Resonant ($t\bar{t}$, WW, Wt, Z→ $\tau\tau$, single top) background:**

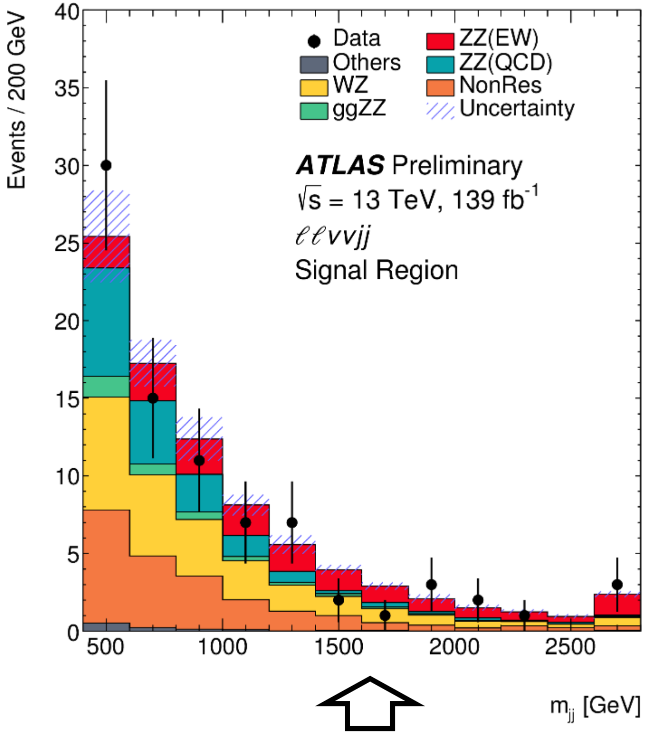
-- Non-Resonant background: $e\mu$ CR

- Exploit ratio of decays ee:$\mu\mu$:eμ = 1:1:2
- In kinematic region $Q(p_T, \eta)$:

$$\epsilon^Q = \sqrt{\frac{N_{ee}^Q}{N_{\mu\mu}^Q}}$$

$$N_{SR\ ee}^{Q,e\mu} = \frac{1}{2} \times \epsilon^Q \times N_{e\mu\ CR}^{sub,bkg}$$

$$N_{SR\ \mu\mu}^{Q,e\mu} = \frac{1}{2} \times \frac{1}{\epsilon^Q} \times N_{e\mu\ CR}^{sub,bkg}$$



Main backgrounds: **Non-Resonant** and **WZ.**

-- Systematics:

- ϵ factor's dependency on different binning method.
- ϵ factor's uncertainty due to data stat. uncertainty.
- Shape difference between MC and data driven based methods.

Background estimation

- **ll$\nu\nu$ WZ background:**

-- WZ background: 3l CR (eee, eeμ, μμe, μμμ)

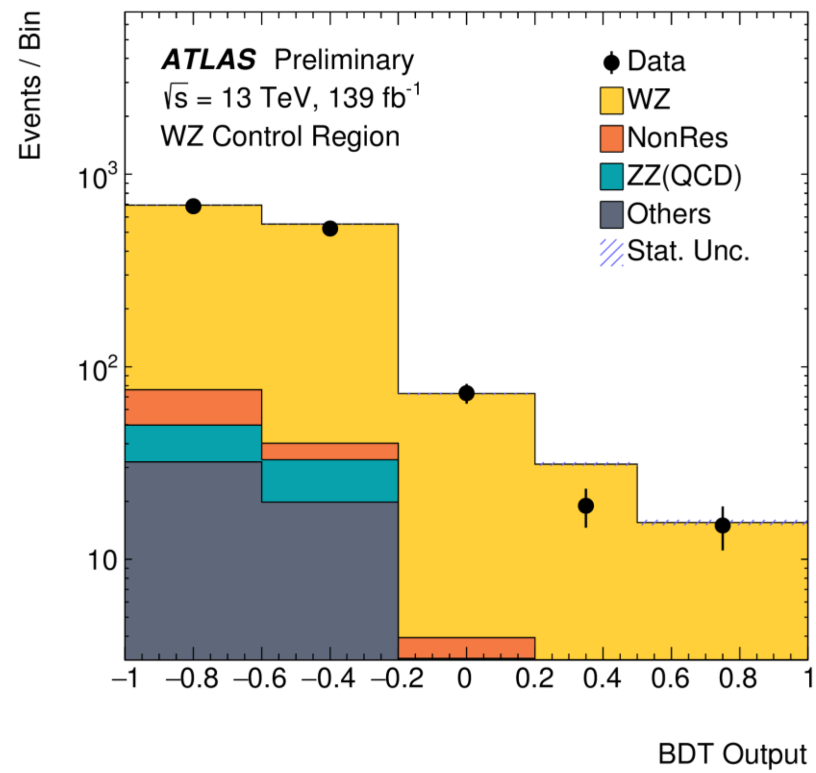
$$N_{WZ}^{est} = N_{MC}^{2l\,SR} \times sf_{WZ} = N_{MC}^{2l\,SR} \times \frac{N_{data-nonWZMC}^{3l\,CR}}{N_{MC}^{3l\,CR}} = N_{data-nonWZMC}^{3l\,CR} \times \frac{N_{MC}^{2l\,SR}}{N_{MC}^{3l\,CR}}$$

-- Scale factor: 0.85

-- Systematics:

- Statistical uncertainty on sf_{WZ} due to CR data.
- Experiment and theory uncertainties on the WZ transfer factor and signal region WZ shape.
- EW WZ cross section: Use the SM cross section for EW WZ. Treat the prediction v.s. measurement ($\mu_{EWKWZ} = 1.77$) difference as the cross-section uncertainty for EW WZ.

[Phys. Lett. B 793 (2019) 469](#)



Background estimation

- **ll$\nu\nu$ Others background:**

  -- Z+jets background:

  - Use SR MC to estimate Z+jets shape.
  - Choose low E_T^{miss} significance region as CR.
  - Fit CR data – nonZjMC. Extrapolate fit result to SR to derive SR Z+jets event yield.
  - Systematics: variations in the fitting functions, differences between estimated and simulated yields and distributions.

  -- ZZ → $llll$,VVV, ttV, ttVV backgrounds: MC

Process	$\ell\ell\ell\ell jj$	$\ell\ell\nu\nu jj$
EW $ZZjj$	20.6 ± 2.5	12.3 ± 0.7
QCD $ZZjj$	77.4 ± 25.0	17.2 ± 3.5
QCD $ggZZjj$	13.1 ± 4.4	3.5 ± 1.1
Non-resonant-$\ell\ell$	-	21.4 ± 4.8
WZ	-	22.8 ± 1.1
Others	3.2 ± 2.1	1.2 ± 0.9
Total	114.3 ± 25.6	78.4 ± 6.2
Data	127	82

- **Event yields:**

  -- uncertainties: stat.+syst.

  -- Minor backgrounds are summed together as 'Others'.

Cross-section

- The definition of fiducial regions are very similar with detector-level selections by using particle-level physics objects.
  -- $llllјj$ channel: loose $m_{l^+l^-}$ to [60, 120]GeV to reduce migration effect.
  -- $ll\nu\nu jj$ channel: to simplify the lepton selections, loose electron eta cut to 2.5 . Use truth E_T^{miss}>130GeV to instead of E_T^{miss} significance>12.

- Fiducial cross-sections for the inclusive production of the EW and QCD processes are measured separately in individual channels.

$$C = \frac{N_{detector-level}}{N_{FV-truth}} \qquad \sigma = \frac{N_{data} - N_{background}}{\mathcal{L} \times C}$$

llllјj **C factor**	0.699 $\pm$ 0.003 (stat) $\pm$ 0.012 (theo) $\pm$ 0.028 (exp)
llννjj **C factor**	0.216 $\pm$ 0.003 (stat) $\pm$ 0.008 (theo) $\pm$ 0.008 (exp)

	Measured fiducial σ [fb]	Predicted fiducial σ [fb]
$\ell\ell\ell\ell jj$	$1.27 \pm 0.12(\text{stat}) \pm 0.02(\text{theo}) \pm 0.07(\text{exp}) \pm 0.01(\text{bkg}) \pm 0.03(\text{lumi})$	$1.14 \pm 0.04(\text{stat}) \pm 0.20(\text{theo})$
$\ell\ell\nu\nu jj$	$1.22 \pm 0.30(\text{stat}) \pm 0.04(\text{theo}) \pm 0.06(\text{exp}) \pm 0.16(\text{bkg}) \pm 0.03(\text{lumi})$	$1.07 \pm 0.01(\text{stat}) \pm 0.12(\text{theo})$

Search for EW ZZjj

- **MVA**

  - Gradient Boosted Decision Tree **(BDTG)** method is used.
  - In $lllljj$ channel, training is performed based on EW and QCD samples.
  - In $ll\nu\nu jj$ channel, the training signal is EW, all other backgrounds are used except Z+jets.

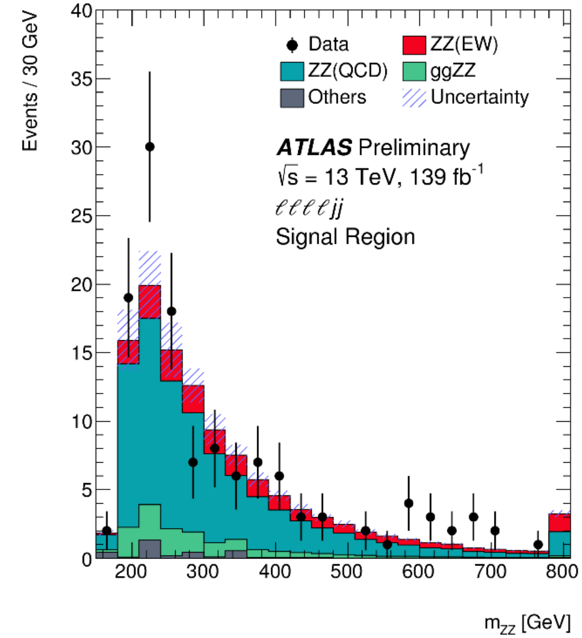
- **Fit procedure**

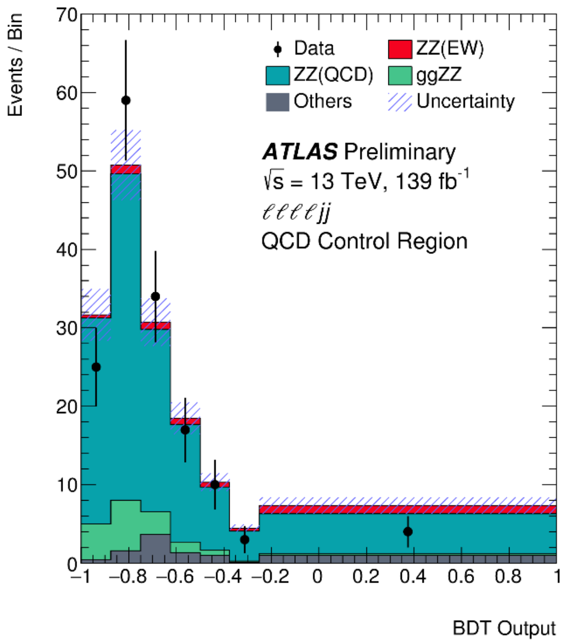
  - To extract EW process, a profile likelihood fit is performed on BDTG response.
  - Simultaneous fit in SR and QCD $lllljj$ CR. The measured fiducial cross-section over the SM prediction for EW ZZjj production **($\boldsymbol{\mu_{EW}}$)** is taken as the **parameter of interest**. μ_{QCD}^{lllljj} represents the normalization factor of QCD ZZjj production.
  - Systematics enter the likelihood as nuisance parameters with Gaussian constrains.
  - For **signal process**, only **shapes uncertainties** are considered.
  - An additional 1.7 k-factor is applied to gg sample.Phys. Rev. D 92 (9 2015) 094028
  - Experiment systematics are treated as correlated.
  - Theoretical uncertainties for the ZZjj production are uncorrelated between the two channels (different fiducial volume).
  - QCD scale uncertainty in $lllljj$ SR, CR and generator modelling uncertainty are uncorrelated (large phase-space difference).

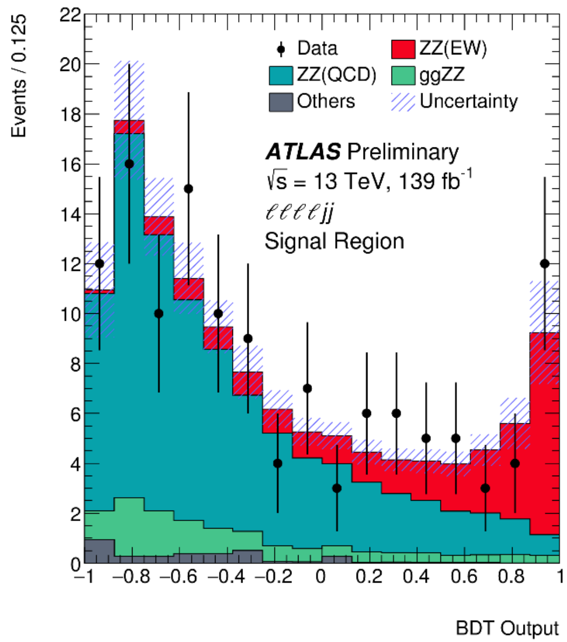
Search for EW ZZjj

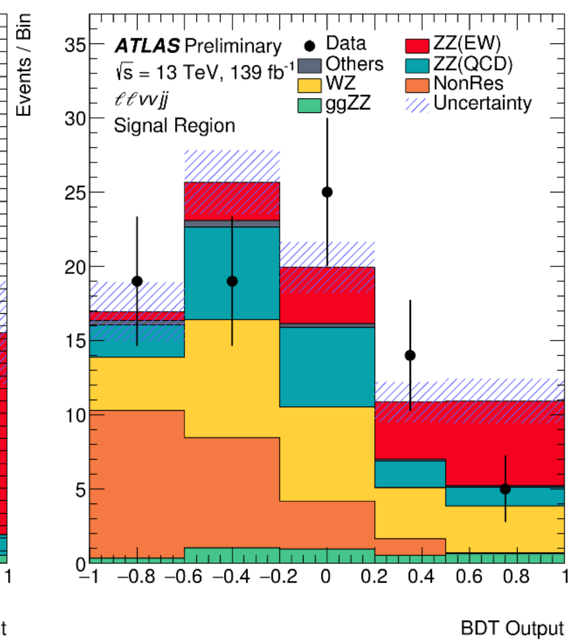
- **Observed and expected distributions:**

  -- M_{ZZ} is scaled with μ_{EW} and μ_{QCD} form final fit.
  -- BDT output plots are post-fit results.
  -- Data distributions are consistent with predict ones.









Search for EW ZZjj

- Normalizations and shapes uncertainties of background processes are considered, while **theoretical uncertainties associated to the EW signal normalization are dropped.**

	μ_{EW}	$\mu_{\mathrm{QCD}}^{\ell\ell\ell\ell jj}$	Significance Obs. (Exp.)
$\ell\ell\ell\ell jj$	1.54 ± 0.42	0.95 ± 0.22	5.48 (3.90) σ
$\ell\ell\nu\nu jj$	0.73 ± 0.65	-	1.15 (1.80) σ
Combined	1.35 ± 0.34	0.96 ± 0.22	5.52 (4.30) σ

- The EW ZZjj cross-section (combing the two channels) in the fiducial volume is **0.82 $\pm$ 0.21 fb**, calculated as $\mu_{EW} \times \sigma_{SM}$ ($\sigma_{SM} = 0.61 \pm 0.03\text{fb}$).

Summary

- First observation of VBS ZZjj process with full Run2 datasets ($139fb^{-1}$).

  -- **Inclusive cross sections** for $llll jj$ and $ll\nu\nu jj$ channels **are measured** in dedicated fiducial volume and found to be **consistent** with the SM predictions.

  -- Observed(expected) significance of EW production is **5.5σ(4.3σ)**.

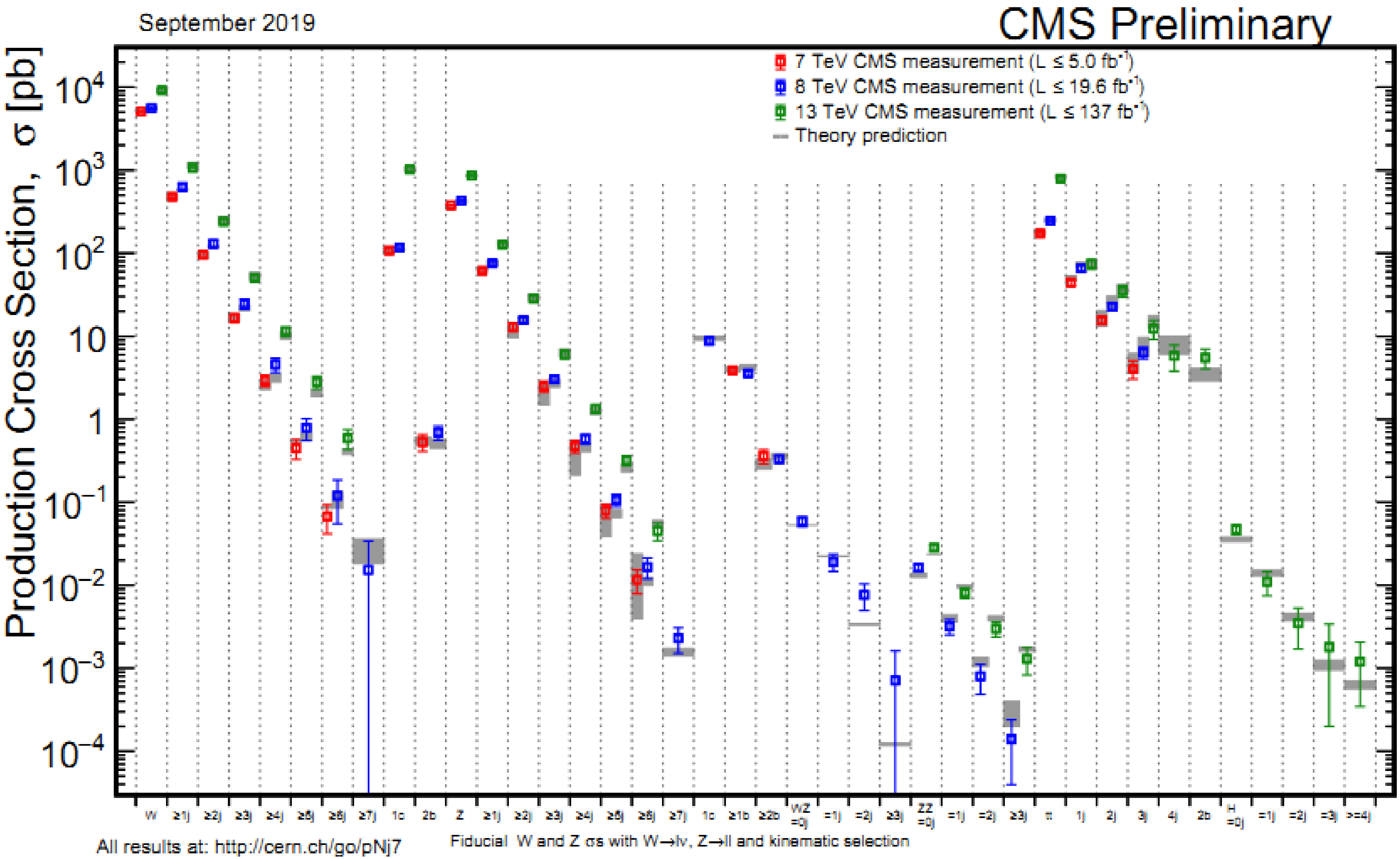
- Although the ZZjj process is very rare, but it's **fully reconstructed final state** in $llll jj$ channel **provides maximal information** to probe SM and interpret BSM.

- The observation of EW ZZjj production is a new milestone reached in the study of EW VVjj production. The precision measurements can help us understand the nature of EWSB.

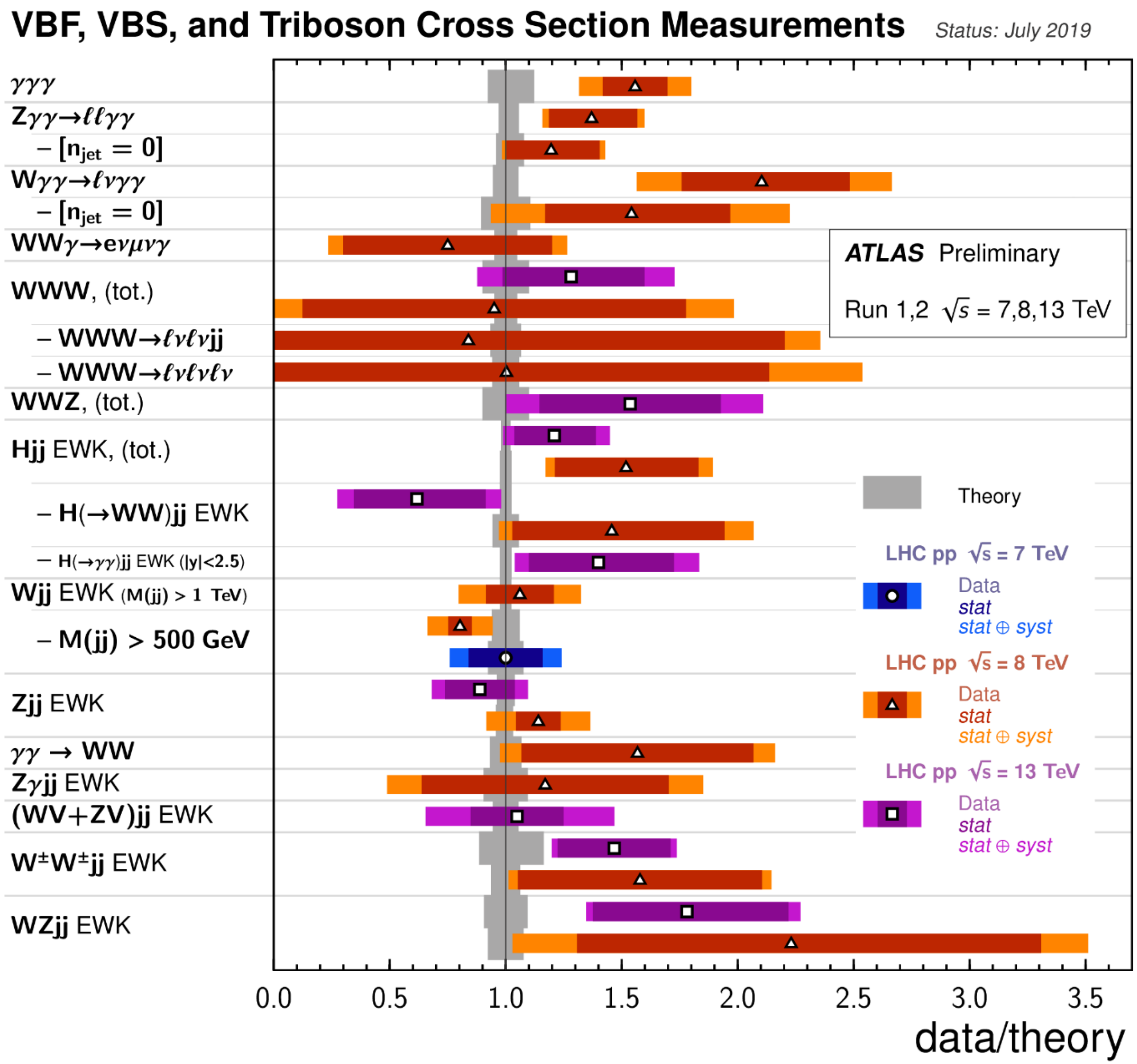
- A nice poster will be presented by Jing Li tonight!

backup

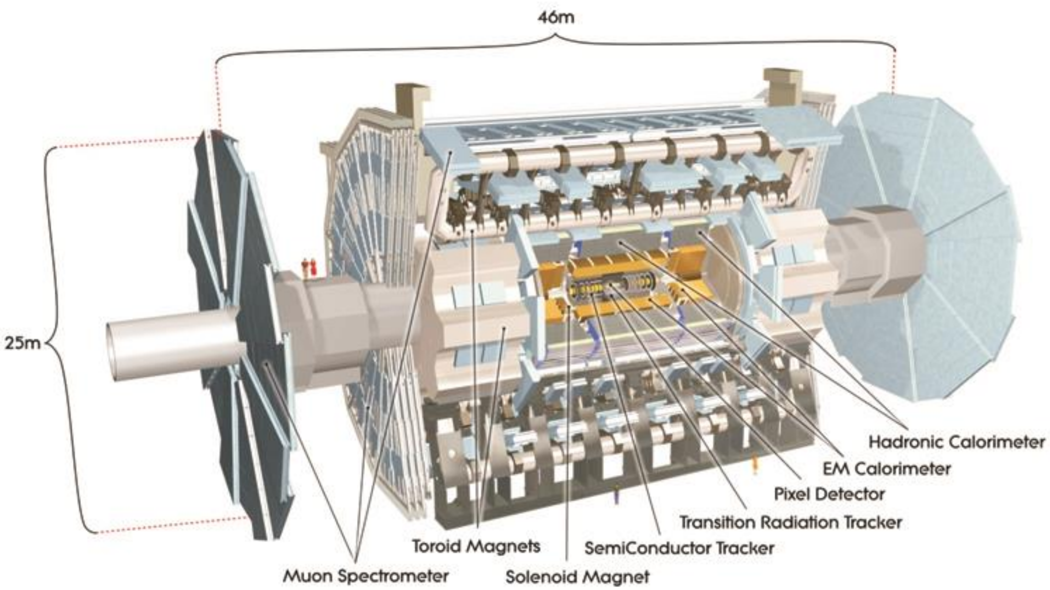
Cross-section measurements in CMS



Cross-section measurements in ATLAS



The ATLAS detector



Calorimeters (CALO)
-Pb/LAr accordion structure
-e/γ trigger identification and measurement : $\sigma/E \sim 10\%/\sqrt{E}$
-HAD: trigger and measurement of jets and E_T^{miss}
-Forward calorimeters(FCAL): covers up to |η|< 4.9

Inner detector (ID)
-|η|< 2.5
-Si pixels, Si strips, TRT
-Precise tracking and vertexing(in 2014, add Insertable B-layer)
-e/π separation

Muon Spectrometer (MS)
-Triggering|η|< 2.4
-Precision Tracking |η|< 2.7
-Magnetic filed produced by toroids
-Muon momentum resolution < 10% up to 1 TeV

Background estimation

- **$llll jj$ background: Fake lepton backgroun**

  -- Fake-factor measured in jet-enriched samples

  - Z+jets CR: same-flavor opposite-charge lepton pair under Z mass and two additional jets.
  - $t\bar{t}$ CR: two jets (at least one bjet) and two high p_T isolated leptons forming an $e\mu$ pair with $E_T^{miss} > 50\text{GeV}, m_T^W < 60GeV$.
  - $f = N_{good}/N_{poor}$ (flavor and pT/eta dependent).
  - Poor electrons are defined by reverting isolation or eleID cuts.
  - Poor muons are defined by reverting isolation or d0 significance cut (but still pass |d0significance|<10).

  -- $lllljj$ fake CR : SR with 1 or 2 leptons passing poor lepton definition.

  -- Fake contribution in signal region:

  - $N_{fake} = \left(N_{gggp} - N_{gggp}^{ZZ}\right) \times f - \left(N_{ggpp} - N_{ggpp}^{ZZ}\right) \times f \times f$
  - ZZ contribution is subtracted.
  - The second term is due to double counting of N_{gggp} and N_{ggpp}.

  -- Systematics: varying “poor” lepton definition, MC contamination, use one bin fake factor instead of p_T/η dependent ones, fake factor difference from data and MC

3l CR & eμ CR

The $e\mu$CR selections are listed as:

- two different-flavour opposite-charge leptons
- veto events with any additional lepton with Loose ID and $P_T > 7$GeV
- $80 < M_{\ell\ell} < 100$GeV
- $P_T^{\ell_1} > 30$GeV, $P_T^{\ell_2} > 20$GeV,$|\eta_\ell| < 2.5$
- $n_{jets} \geq 2$, $P_T^{j_1} > 60$GeV, $P_T^{j_2} > 40$GeV, $|\eta_j| < 4.5$
- $M_{jj} > 400$GeV, $\Delta Y_{JJ} > 2$, $Y_{j1} \times Y_{j2} < 0$
- B-jet veto
- MET Significance > 12

The definition of 3lCR is:

- $80 < M_{\ell\ell} < 100$GeV
- $P_T^{\ell_1} > 30$GeV, $P_T^{\ell_2} > 20$GeV,$|\eta_\ell| < 2.5$, medium
- $p_{\mathrm{T}}^{\ell 3rd} > 20$GeV, $|(\eta^{\ell 3rd})| < 2.5$, medium
- Transverse mass $m_T^W > 40$GeV
- B-jet veto: 85% working point
- $n_{jets} \geq 2$
- $P_T^{J_1} > 60$GeV, $P_T^{J_2} > 40$GeV
- MET Significance > 3

Z+jets

$$m_{\mathrm{T}}^{W} = \sqrt{2 P_T^{\ell 3} E_{\mathrm{T}}^{\mathrm{miss}} \left[1 - \cos\left(\Delta\phi\left(P_T^{\ell 3}, E_{\mathrm{T}}^{\mathrm{miss}}\right)\right)\right]}$$

To keep more statistics, no m_{jj} or $|\Delta Y_{jj}|$ cuts and loose MET significance.

Uncertainties

- **Theoretical uncertainties:**

-- PDF, QCD scale, α_s, parton showering (PS).

- PDF: the envelope of the NNPDF internal errors and the differences between the nominal and alternative PDFs.

- QCD scales: 7-point scale variations of the renormalization (μ_r) and the factorization scale (μ_f) ({0.5,0.5}, {1,0.5}, {0.5,1}, {1,1}, {2,1}, {1,2}, {2,2}). The largest deviation is chosen as the uncertainty.

- PS: comparing the nominal Pythia8 parton showering with the alternative Herwig7 algorithm.

- α_s: varying the α_s value within $\pm$0.001.

-- Interference effect between the EW and QCD processes is 7%(2%) in $lllljj$($ll\nu\nu jj$) channel. Treat as an extra uncertainty in the EW signal predictions.

-- Generator modelling uncertainty: estimated by comparing Sherpa with MadGraph5 _aMC@NLO 2.6.1 predictions at particle level.

- **Experimental uncertainties:**

-- luminosity: 1.7%.

-- The momentum scale and resolution of leptons and jets, lepton reconstruction and selection efficiencies, trigger selection efficiency, the calculation of the E_T^{miss} soft-term, the pile-up correction, and the b-jet identification efficiency: 5-10%.

-- Jet pile-up uncertainty.