(Some) Details on ATLAS VBS ZZ Results

Yusheng Wu
University of Science & Technology of China

For ATLAS VBS ZZ team

LHC diboson meeting, Nov 24, 2020

Introduction

- Long interests in vector boson scattering even since 1970s, due to its obvious connection with Higgs mechanism and BSM

- Among all VBS processes, **VBS ZZ** is the kind of **rarest process**, with O(0.1 fb) effective fiducial cross-sections.

- Before LHC start, many didn't expect we could find this so fast => owing to **excellent accelerator/detector/analysis work**, and **great theory advances**

- **ATLAS and CMS have both measured** VBS ZZ (in its hosting channels, so called electroweak ZZjj) with **good significance and found SM compatibility**, providing another triumph and milestone in testing the SM to its extremity.

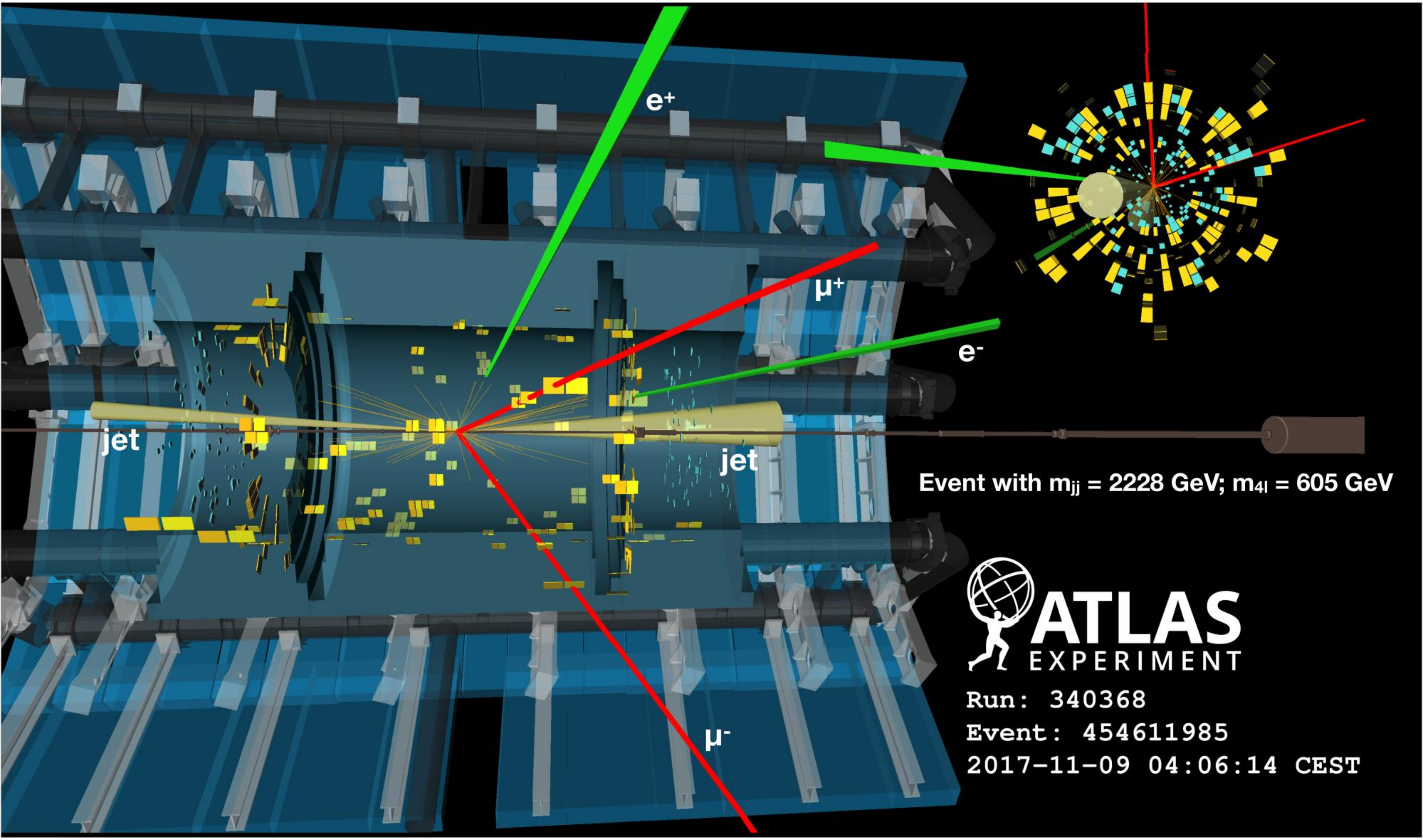
- Although we cannot use this limited number of events to tell much into the ultimate goal, this observation nevertheless **signified that the door to the future is opening …**

The plan to explain a bit the ATLAS results

- All materials shown are based on public ATLAS results
  - arXiv:2004.10612
  - or https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PAPERS/STDM-2017-19/, including extra plots/tables

- From time to time, personal interpretations and remarks will be provided, focusing on why certain actions were taken in the data analysis

An event display



Relevant diagrams

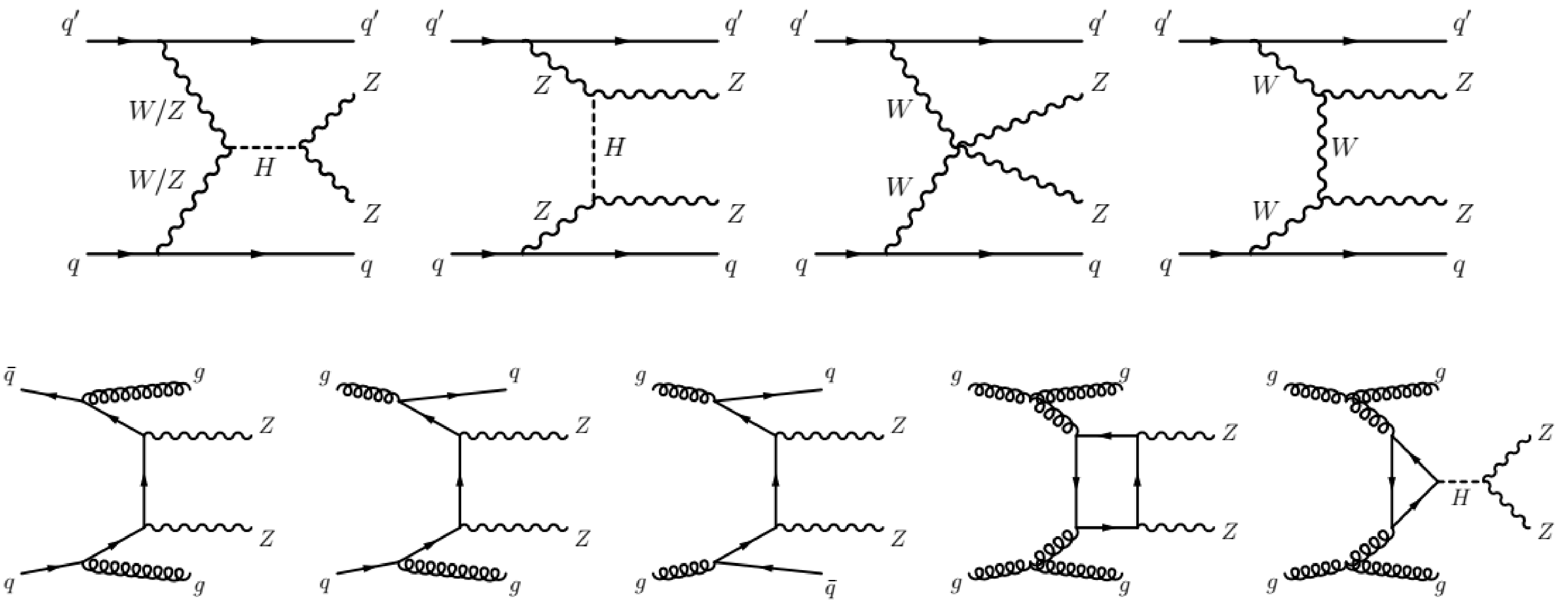
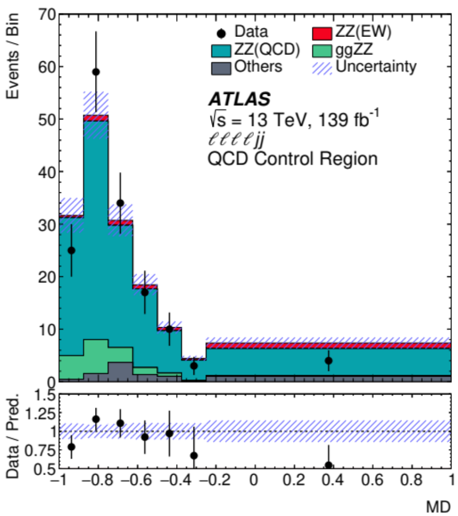


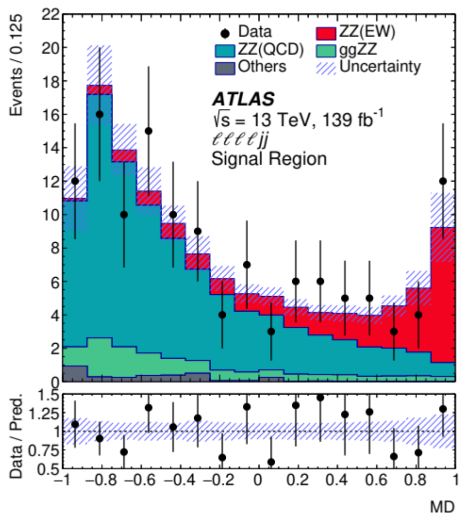
Figure 1: Typical diagrams for the production of $ZZjj$, including the relevant EW VBS diagrams (first row) and QCD diagrams (second row).

One needs to work on good S/N ratio channels (llllјј and llννjj) and take advantage of the VBS characteristics (more forward two jets, less color flow between jets)

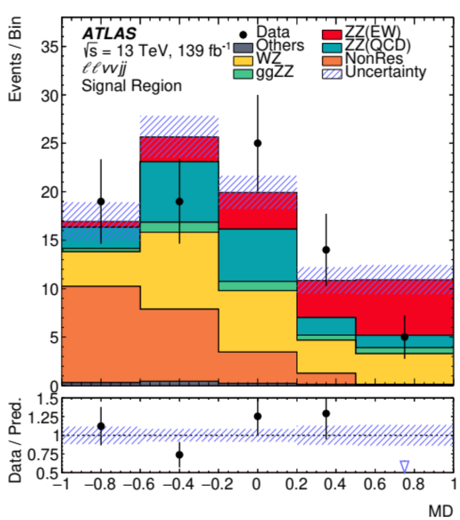
ATLAS Results



(a)



(b)



Used ATLAS full Run-II data, 139 fb^{-1}, based on llllјj and llννjj final states

Obtained 5.5 σ observation significance of electroweak ZZjj

Fitted 3 regions simultaneously based on BDT discriminants: QCD control region, signal region for lllljj, and SR for llννjj

Use of two final states

- We knew $ll\nu\nu jj$ would have worse sensitivity, but still decided to use it as a cross-check (especially if anything extraordinary might be observed in the llllјј channel)

lllljj (main channel):

- clean with four charged leptons
- mainly fight against QCD ZZjj (use of start-of-art modelling, and cross-check with data)
- fully reconstructable final state
- benefit from precise lepton calibration
- reply on understanding of jets (calibration and systematic uncertainties)

$ll\nu\nu jj$:

- suffer from large backgrounds (QCD ZZjj, WZ, top, WW, Z+jets …), constraints with data
- although larger BR, must harshly restrict the phase space
- still has advantages in large mass/pT ranges for BSM sensitivity
- rely on understanding of MET
- rely on understanding of jets

Selection of detector-level events

- Largely shaped by the detector acceptance and signal characteristics over backgrounds

<table>
<tr><th></th><th>$\ell\ell\ell\ell jj$</th><th>$\ell\ell\nu\nu jj$</th></tr>
<tr><td>Electrons</td><td colspan="2">$p_{\rm T} > 7$ GeV, $|\eta| < 2.47$
$|d_0/\sigma_{d_0}| < 5$ and $|z_0 \times \sin\theta| < 0.5$ mm</td></tr>
<tr><td rowspan="2">Muons</td><td>$p_{\rm T} > 7$ GeV, $|\eta| < 2.7$</td><td>$p_{\rm T} > 7$ GeV, $|\eta| < 2.5$</td></tr>
<tr><td colspan="2">$|d_0/\sigma_{d_0}| < 3$ and $|z_0 \times \sin\theta| < 0.5$ mm</td></tr>
<tr><td>Jets</td><td>$p_{\rm T} > 30$ (40) GeV for $|\eta| < 2.4$ ($2.4 < |\eta| < 4.5$)</td><td>$p_{\rm T} > 60$ (40) GeV for the leading (sub-leading) jet</td></tr>
<tr><td>ZZ selection</td><td>$p_{\rm T} > 20, 20, 10$ GeV for the leading, sub-leading and third leptons
Two OSSF lepton pairs with smallest $|m_{\ell^+\ell^-} - m_Z| + |m_{\ell'^+\ell'^-} - m_Z|$
$m_{\ell^+\ell^-} > 10$ GeV for lepton pairs
$\Delta R(\ell, \ell') > 0.2$
$66 < m_{\ell^+\ell^-} < 116$ GeV</td><td>$p_{\rm T} > 30$ (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
$E_{\rm T}^{\rm miss}$-significance > 12</td></tr>
<tr><td rowspan="2">Dijet selection</td><td colspan="2">Two most energetic jets with $y_{j_1} \times y_{j_2} < 0$</td></tr>
<tr><td>$m_{jj} > 300$ GeV and $\Delta y(jj) > 2$</td><td>$m_{jj} > 400$ GeV and $\Delta y(jj) > 2$</td></tr>
</table>

More precise description in the paper

- Difference between the two channels motivated/determined by experiment aspects
- Pileup contribution (often concerned) largely suppressed and thought to have negligible impact
- Selection chosen to be effective but not too strict, leaving room for later machine learning improvement

Observed yields and backgrounds

Table 1: Observed data and expected event yields in 139 fb^{-1} of data in the $\ell\ell\ell\ell jj$ and $\ell\ell\nu\nu jj$ signal regions. All the minor backgrounds are summed together as 'Others'. Uncertainties in the predictions include both the statistical and systematic components.

Process	$\ell\ell\ell\ell jj$	$\ell\ell\nu\nu jj$
EW $ZZjj$	20.6 ± 2.5	12.3 ± 0.7
QCD $ZZjj$	77 ± 25	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 ± 26	78.4 ± 6.2
Data	127	82

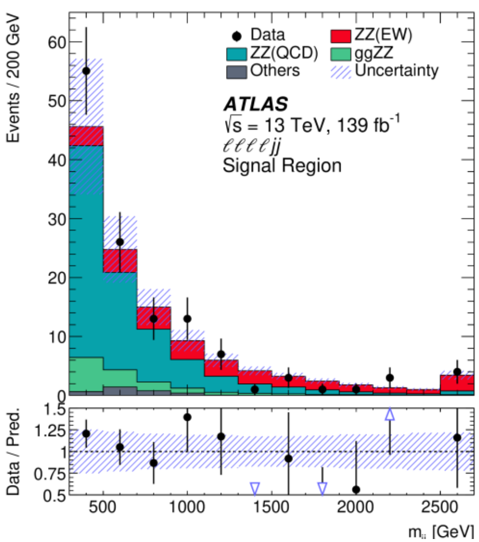
Mostly top and WW → Non-resonant-$\ell\ell$

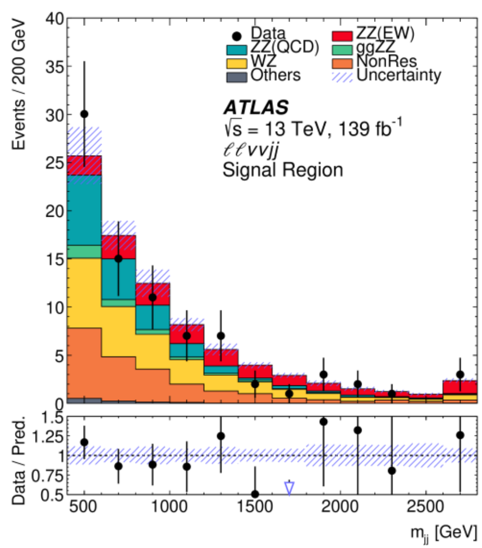
Backgrounds due to non-prompt leptons, fake MET, triboson, ttV → Others

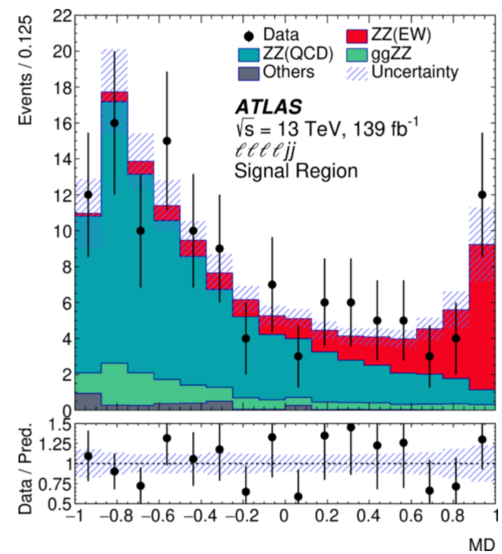
- A reasonable agreement between data and predictions was found
- Further discrimination between signal and backgrounds necessary to reach at observation threshold

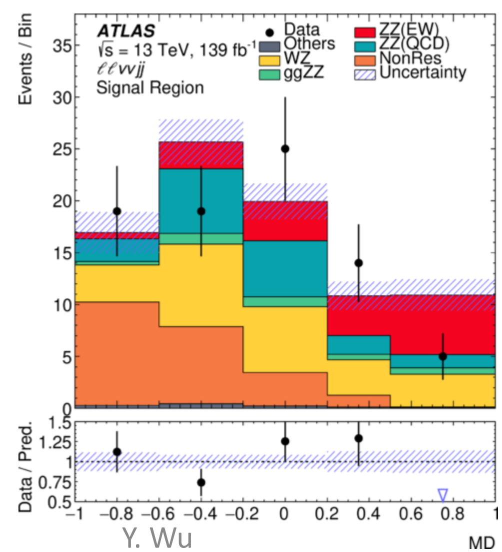
Multivariant Discriminants

Boost of sensitivity from m(jj) to multivariable:









Based on Gradient Boosted Decision Trees

Input variables are typical jet and lepton kinematics

Hyperparameters scanned for optimal performance

Binning was optimized too

Low MD → background-like

High MD → signal-like

MC simulation

❑ EW ZZjj

- ❖ MadGraph5_aMC@NLO 2.6.1, LO in QCD + PS
- ❖ NLO calculation not included in time for initial publication, to be incorporated for revisions and Run-II legacy results, won't affect data observation

❑ QCD ZZjj

- ❖ Sherpa 2.2.2, NLO in QCD for up to 1 parton, LO in QCD for 2-3 partons
- ❖ box diagram gluon-gluon QCD ZZjj, with Sherpa/gg2VV, LO in QCD + PS; cross-sections corrected to NLO in QCD

❑ EW-QCD ZZjj interference

- ❖ modelled with MadGraph5_aMC@NLO 2.6.1
- ❖ found to be small w.r.t. EW ZZjj; treated as uncertainty in modelling of EW ZZjj

❑ Other processes

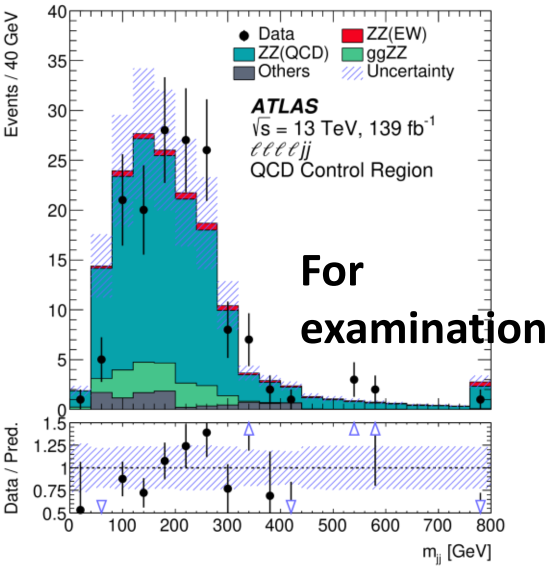
- ➢ Sherpa 222: WWjj, WZjj, triboson; Sherpa 221: Z+jets (including EW process)
- ➢ Powheg-Box V2: ttbar; Powheg-Box V1: Wt; MG5_aMC@NLO: ttV and ttVV

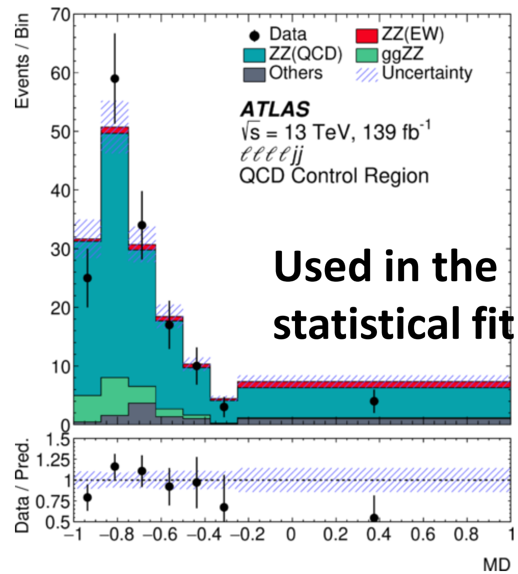
QCD ZZjj

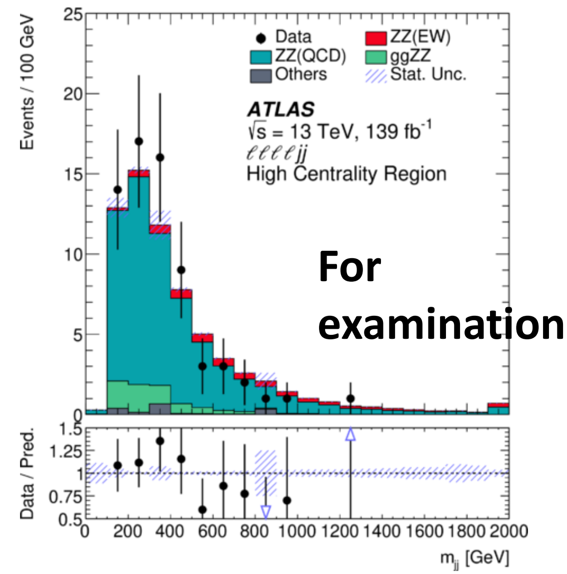
- based on MC simulation (full experimental uncertainties and typical modelling uncertainties – QCD scale, PDF, PS), with examination provided by data CR and VR (lllljj only), and with constraints from sensitive events in the final statistical fit

QCD ZZjj CR: fail either mjj or deta(jj) cut

QCD ZZjj VR: in SR but large centrality







⇒ Data and simulation was found to be compatible, even up to high mjj

⇒ Full uncertainty on QCD ZZjj amounts to 30% level or larger depending on regions

⇒ QCD CR provides a normalization factor of QCD ZZjj for lllljj channel; while llννjj channel relies on pure MC for QCD ZZjj modelling

⇒ Correlation scheme for modelling uncertainties been treated conservatively

Other Backgrounds

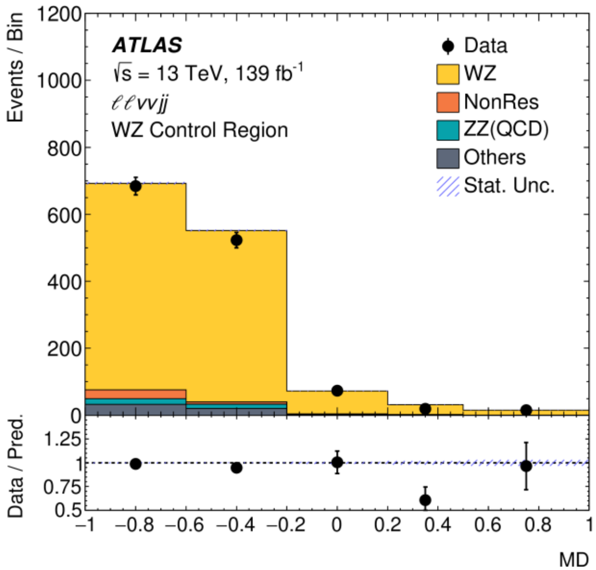
❑ lllljj channel

- ❖ **Fake background** yield estimated via fake-factor method and shape via simulation, > 50% uncertainty
- ❖ **Triboson, ttV(V)** via simulation

❑ llννjj channel

- ❖ **WZjj**: based on simulation

and corrected with 3-lepton CR



Non-resonant-ll contribution: WW, ttbar, Wt, estimated using e+μ+jj CR for the yield and simulation for shape

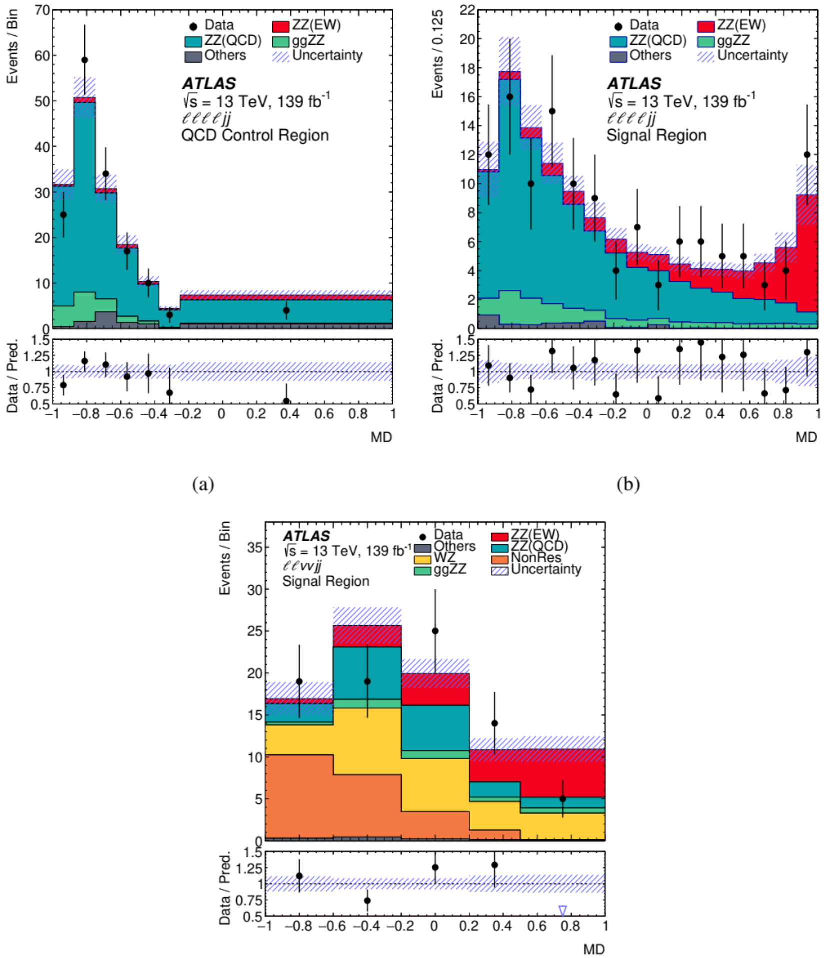
Others:

Z+jets largely suppressed, estimated via low-to-high MET extrapolation (>100% unc.)

ttV/VVV estimated via simulation

Statistical Fitting

❑ A simultaneous fit to three regions



POI:
σ(EW ZZjj) normalized to SM prediction

Free Normalization factor for QCD ZZjj only applied in llll jj: σ(QCD ZZjj) normalized to SM prediction

All experimental and modelling systematic uncertainties treated as nuisance parameters in the fit, affecting signal and background distributions if applicable

Uncertainties

Experimental systematic uncertainties

- luminosity uncertainty, 1.7% constant
- uncertainties due to lepton, jet, and MET reconstruction, and pileup condition modelling => inclusive uncertainty typically 5% for simulated processes
- uncertainties due to the same sources correlated among processes and channels
- background uncertainties and MC statistical uncertainties
- Additional uncertainty in QCD ZZjj pileup modelling as difference in QCD ZZjj shape between low and high pile-up conditions (~10% difference in shape)

Modelling uncertainties

- For EW and QCD ZZjj processes, included uncertainties due to PDF, QCD scale, parton-showering, and alpha(S) variation: up to 30% impact on the yield and shape of QCD ZZjj
- Conservative approach was taken to un-correlate modelling uncertainties between regions and channels (see the paper for further description): choosing a more aggressive approach could lead to a further 10-20% improvement to the significance
- Additional uncertainty assigned to cover difference between Sherpa and MG5 in the QCD ZZjj shape modelling
- EW-QCD interference considered as modelling uncertainty on the EW process

Results

Table 2: Observed μ_{EW} and $\mu_{\text{QCD}}^{\ell\ell\ell\ell jj}$, as well as the observed and expected significance of EW $ZZjj$ processes from the individual $\ell\ell\ell\ell jj$ and $\ell\ell\nu\nu jj$ channels, and the combined fits. The full set of statistical and systematic uncertainties is included.

	μ_{EW}	$\mu_{\text{QCD}}^{\ell\ell\ell\ell jj}$	Significance Obs. (Exp.)
$\ell\ell\ell\ell jj$	1.5 ± 0.4	0.95 ± 0.22	5.5 (3.9) σ
$\ell\ell\nu\nu jj$	0.7 ± 0.7	–	1.2 (1.8) σ
Combined	1.35 ± 0.34	0.96 ± 0.22	5.5 (4.3) σ

- The observed significance is larger than the expected owing to the slight excess of data in high MD
- The llννjj channel is of a cross-checking purpose as expected
- Results shown a compatibility with the SM
- The final significance is statistically limited; the systematic uncertainties on the signal strength are about 10% and 30% for lllljj and llννjj

Fiducial phase space

	$\ell\ell\ell\ell jj$	$\ell\ell\nu\nu jj$
Electrons	$p_T > 7$ GeV, $\|\eta\| < 2.47$	$p_T > 7$ GeV, $\|\eta\| < 2.5$
Muons	$p_T > 7$ GeV, $\|\eta\| < 2.7$	$p_T > 7$ GeV, $\|\eta\| < 2.5$
Jets	$p_T > 30$ (40) GeV for $\|\eta\| < 2.4$ ($2.4 < \|\eta\| < 4.5$)	$p_T > 60$ (40) GeV for the leading (sub-leading) jet
ZZ selection	$p_T > 20, 20, 10$ GeV for the leading, sub-leading and third leptons Two OSSF lepton pairs with smallest $\|m_{\ell^+\ell^-} - m_Z\| + \|m_{\ell'^+\ell'^-} - m_Z\|$ $m_{\ell^+\ell^-} > 10$ GeV for lepton pairs $\Delta R(\ell, \ell') > 0.2$ $60 < m_{\ell^+\ell^-} < 120$ GeV	$p_T > 30$ (20) GeV for the leading (sub-leading) lepton One OSSF lepton pair and no third leptons $80 < m_{\ell^+\ell^-} < 100$ GeV $E_T^{miss} > 130$ GeV
Dijet selection	Two most energetic jets with $y_{j_1} \times y_{j_2} < 0$ $m_{jj} > 300$ GeV and $\Delta y(jj) > 2$	Two most energetic jets with $y_{j_1} \times y_{j_2} < 0$ $m_{jj} > 400$ GeV and $\Delta y(jj) > 2$

Similar to detector-level selection, with simplifications
(e.g. MET vs MET-significance)

For a measurement of EW ZZjj cross-section
- Multiple the μ(EW ZZjj) derived in the fit with the SM prediction

For a measurement of inclusive ZZjj
- use the simple counting experiment, treat EW+QCD ZZjj as the signal
- fiducial to detector correction factors: 70% for lllljj, and 22% for llννjj
- worse factor for llννjj due to the migration and resolution effect of using MET at truth level

Fiducial cross-sections

Measured EW ZZjj cross-section combining two channels
0.82 +/- 0.21 fb, in comparison to prediction of 0.61 +/- 0.03 fb

Note: the fiducial region is broader than the most sensitive region (i.e. high MD) used to extract the EW ZZjj signal

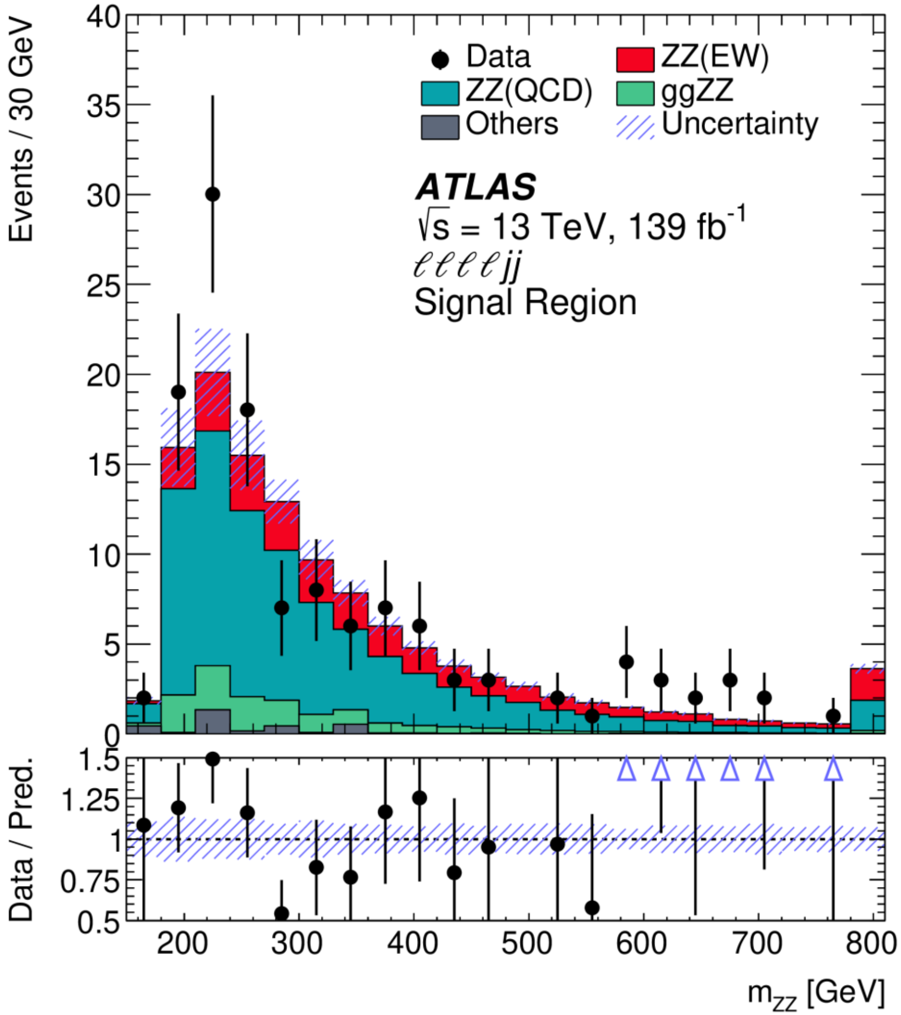
Measured EW+QCD ZZjj cross-section

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

A reasonable precision has been achieved for the two fiducial measurements

Looking forward

My favorite plot in the paper



With the full Run-II data, we not only have observed the rare process,

=> but also been able to have a peek at the high s-hat region, which was the original target that attracted us to seek for anomalies

Summary

- A brief discussion on the analysis details has been provided for ATLAS VBS ZZ measurement
- This is a typical measurement, seeking for a rare process and relying on many factors to succeed
- Luckily, excellent theory advances and experimental scrutinies made it possible
- Extracting further insights using this channel would be difficult without a big leap in the data statistics
- Nevertheless, next phase of studies shall and have already been started: differential measurements, direct constraints on resonance models in VBS channel, polarization studies, effective field theory studies, etc.

Thank you for your attention!