Observation of electroweak production of two jets and a Z-boson pair on the LHC ATLAS

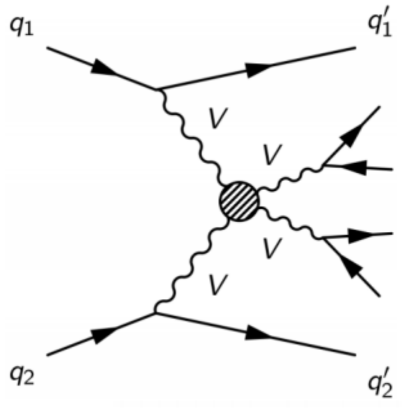
MBI 2022

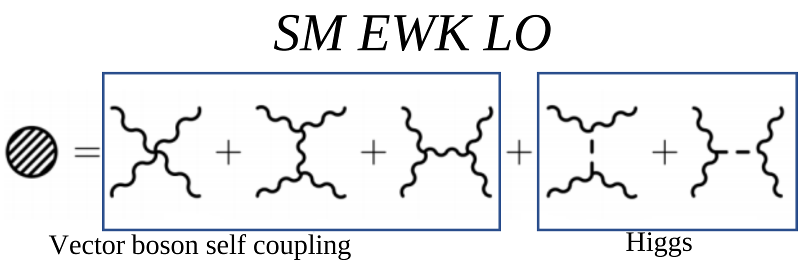
Aug. 22ed, 2022

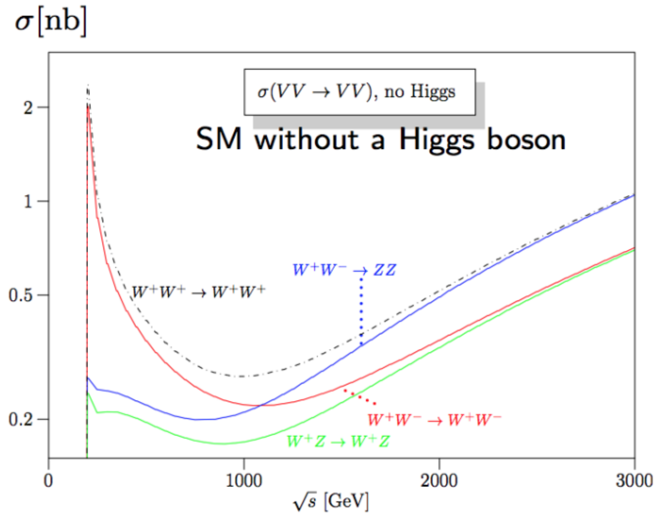
Mingyi Liu[1]
mingyi.liu@cern.ch

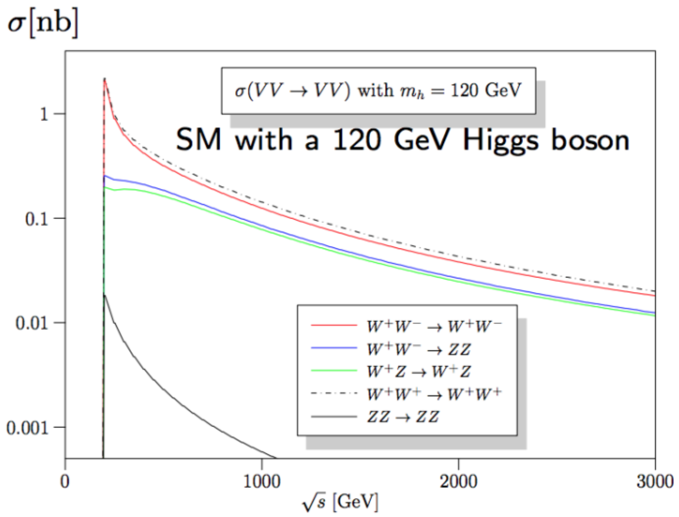
1. Univ. of Sci. & Tech. of China

Vector Boson Scattering









Resonances and unitarity in weak boson scattering at the LHC - IOPscience

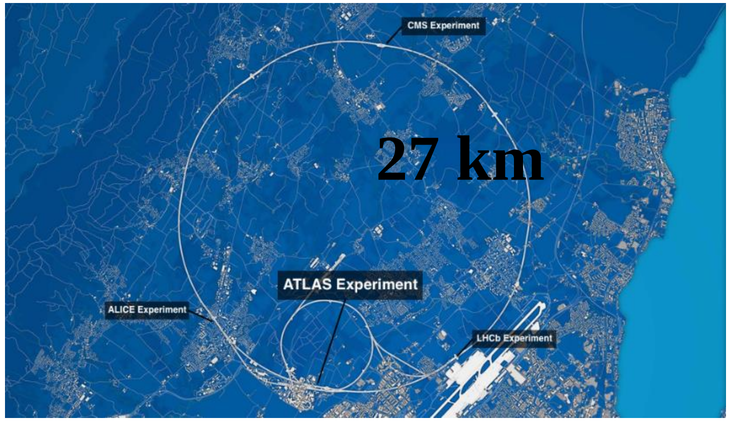
- Higgs boson acts as "moderator" to unitarize high-energy longitudinal vector boson scattering
- Unitarity: If only Z/W are exchanged, the amplitude of vector boson scattering violates unitarity
- Higgs boson restore unitarity of total amplitude
- **Vector boson scattering is key process to experimentally probe the SM nature of electroweak symmetry breaking.**

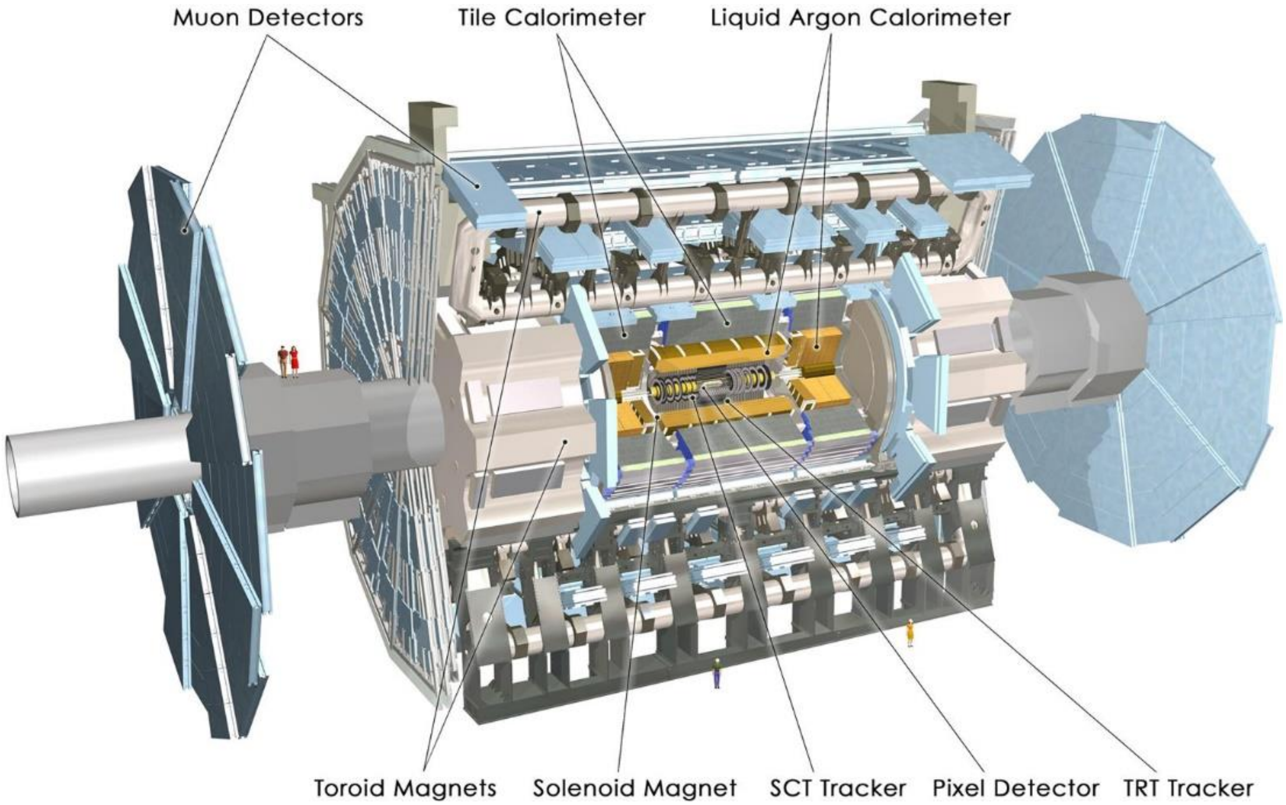
Overview of VBS measurements on ATLAS

✓ **Status of Aug. 2022**

Channel	Obs. (Exp.) significance	Experimental Challenges
ssWW	**6.5(4.9) σ** 36.1 fb^{-1}	"Golden channel": first observation of VBS in this channel, very good EW/QCD ratio
WZ	**5.3(3.2) σ** 36.1 fb^{-1}	Similar cross section as ssWW, but larger QCD backgrounds
ZZ	**5.7(4.8) σ** First 139 fb^{-1} analysis	Very clean 4l channel, low background but small cross section
Zγ	**~10 σ** 139 fb^{-1}	Higher statistics due to photon

The ATLAS detector





CERN, Geneva, Switzerland
ATLAS, CMS, LHCb, ALICE
100 meters underground
ATLAS size: 46m×25m

Inner detector:
- $|\eta| < 2.5$
- Momentum, electrical charge
- Pixel detector, tracker (semi-conductor), TRT

Calorimeters:
- EM Cal.($|\eta| < 2.5$):

e&γ, lead absorber submerged by LAr
- Hadron Cal.($|\eta| < 4.9$):

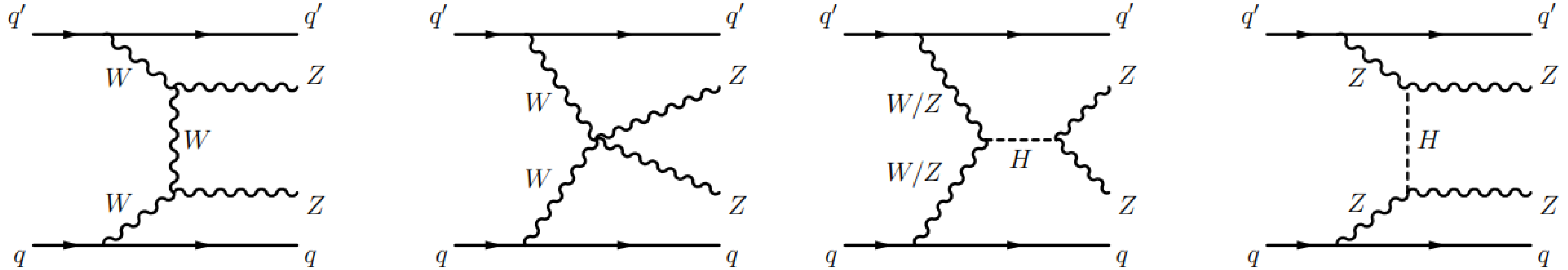
LAr with copper/tungsten absorber (forward)
Scintillator tile with steel absorber (central)

Muon spectrometer:
- Tigger($|\eta| < 2.4$): TGC, RPC
- Tracking($|\eta| < 2.7$): MDT, CSC

Magnet system:
- Solenoid Magnet: 2T
- Toroid Magnets: 4T

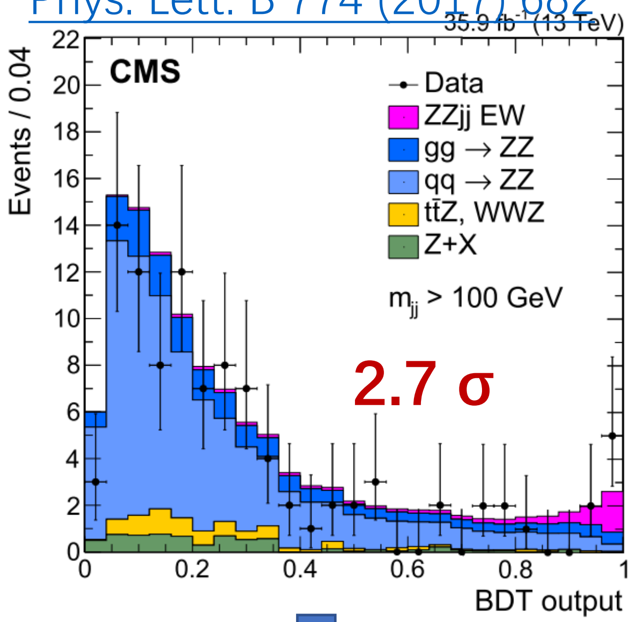
Introduction

- The EW ZZjj production was searched for by the CMS collaboration using 35.9 fb^{-1} of 13 TeV pp collision data (2015~2016), but no evidence was found.
- With more data collected between 2017~2018, luminosity increased to 139 fb^{-1}, more precise searches are presented by the LHC.
- Search for EW ZZjj production with 4l and llνν final states
  - 4l: very clean experimental signature, high signal-to-background ratio
  - llνν: larger branching ratio, but more complex background

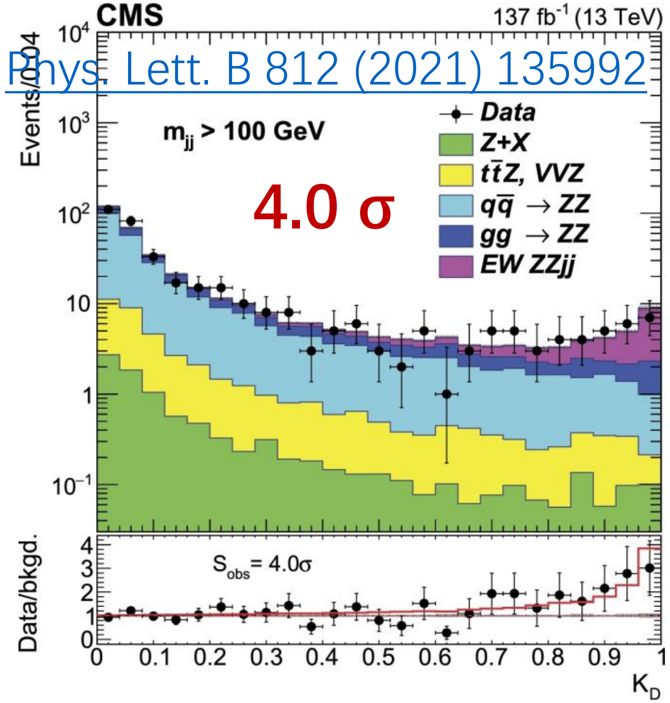


- This ATLAS analysis presents the first VBS ZZjj observation (5.7σ) with 139 fb^{-1} full run-2 datasets

Phys. Lett. B 774 (2017) 682



Phys. Lett. B 812 (2021) 135992



Analysis strategy

- Working on ATLAS full run-2 data with 139 fb^{-1}
- Experimental signature: 2j (Back-to-back, forwarded) + 4l/2l with E_T^{miss} from neutrinos

➢ Step 1: Inclusive cross-section measurements in the fiducial volume

➢ Step 2: MVA (BDTG)-based analysis is then used to extract the EW VBS ZZ signal from background for:

  ✓ Observation of electroweak production of ZZjj

  ✓ Observed signal strength of the EW processes

MC simulation

EWK ZZjj (NLO)	**Powheg-Box v2**
QCD ZZjj	Sherpa222
QCD ggZZjj	4l:Sherpa222; llvv:gg2VV
WZ	Sherpa222
ttbar, single top	Powheg
ttV	MG5_aMC@NLO
Z+jets	Sherpa221
VVV	Sherpa222
WW, WZ Semi-leptonic	Powheg

The analysis is performed independently in $ZZ \to 4l$ and $ZZ \to ll\nu\nu$ final states. Then combined together

Interference between EW and QCD processes is treated as systematic on the EW VBS ZZ production measurement

Event selection

4l

✓ Pairing:

M_Z as reference for Mll

llvv

✓ 3rd lepton veto to reduce ZZ→4l, WZ→3l+v;

✓ MET requirement to veto Z+jets

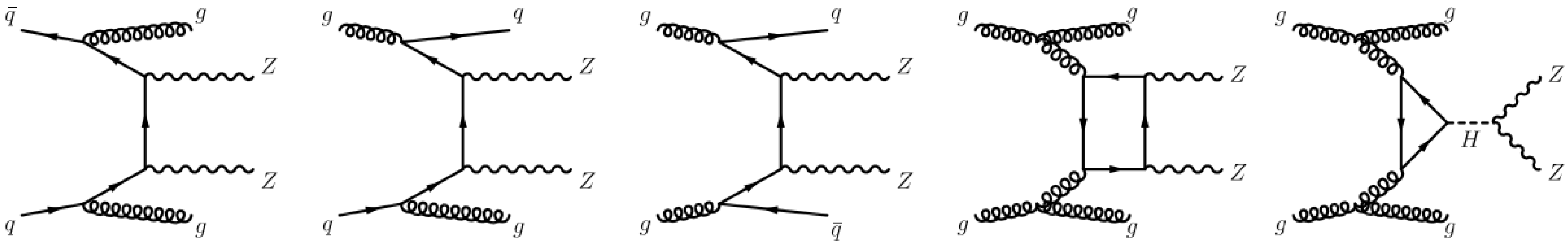
	$\ell\ell\ell\ell jj$	$\ell\ell\nu\nu jj$
Electrons	$p_T > 7$ GeV, $\|\eta\| < 2.47$ $\|d_0/\sigma_{d_0}\| < 5$ and $\|z_0 \times \sin\theta\| < 0.5$ mm	
Muons	$p_T > 7$ GeV, $\|\eta\| < 2.7$ $\|d_0/\sigma_{d_0}\| < 3$ and $\|z_0 \times \sin\theta\| < 0.5$ mm	$p_T > 7$ GeV, $\|\eta\| < 2.5$
Jets	$p_T > 30$ (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 $\Delta R(\ell, \ell') > 0.2$ $66 < m_{\ell^+\ell^-} < 116$ 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 No b-tagged jets E_T^{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 $\Delta y(jj) > 2$

Generally tighter selections for llvv channel to get rid of the complex BKG

- Benchmark feature for the EW VBS process: A pair of back-to-back jet
- Optimizations of jet selection implemented to enhance EW VBS
- 4l channel: QCD CR for BKG strength constrain: Mjj<300 GeV or dYjj<2

Background: 4l channel

- Irreducible BKGs: qqZZ and ggZZ (Including ggF)



  - QCD vertex induced processes
  - BKG yield constrained by QCD CR (Normalization factor)

- Reducible BKGs (Fake)
  - Fake leptons from Z + jets / tt , poor modeling: Data-driven **fake factor** method
  - Fake enriched regions (Z+jets CR and ttbar CR) defined to obtain the fake factors (F.F.)
  - F.F.s from Z+jets CR and ttbar CRs are combined based on the individual ratios (Z+jets over ttbar)

Fake factor method (Data-driven)

- Fake factor measured in fake-enriched regions: Z+jets CR and ttbar CR
  - Poor electrons: fail the isolation or fail the loose electron identification (But still pass the VeryLooseLH)
  - Poor muons: fail the isolation or fail the d0 significance cut (but still pass |d0significance<10|)

- 4l fake CR (F.F. application region)

  SR/CR with 1 or 2 poor leptons

- Fake contribution
  - F = N(Good lepton)/N(Poor lepton)
  - $N_{fake} = (N_{gggp} - N^{ZZ}_{gggp}) \times F - (N_{ggpp} - N^{ZZ}_{ggpp}) \times F^2$
  - With prompt ZZ contribution subtracted
  - Events with more than 2 poor leptons are neglected due to tiny contribution

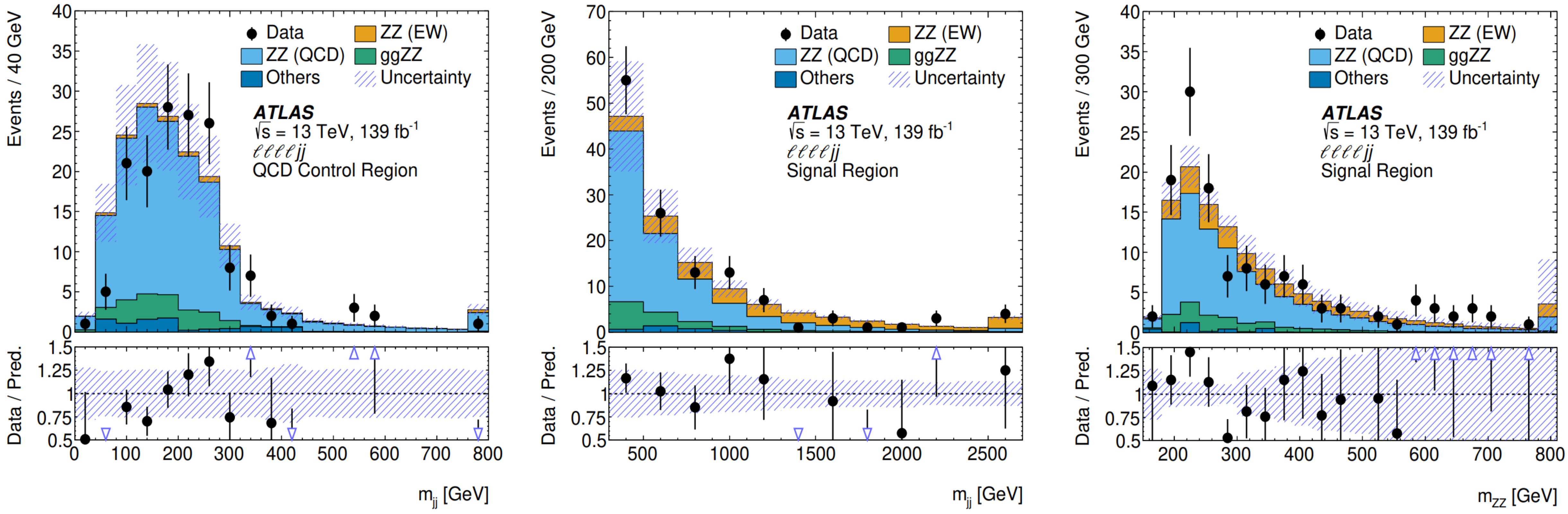
Subtraction for double counting

Fake factors:
- Separate by flavor
- 2D Functions of Pt / η
- F.F.s from Z+jets and ttbar CRs are combined based on the individual ratios (Z+jets over ttbar)

Fake lepton BKG yield is small, yielding less than 4 events (together with other minor processes like triboson and ttV).

Yields and distributions in 4l channel

139 fb^{-1}	EW ZZjj	QCD ZZjj	ggZZjj	Others	Total	Data
4l SR	22.4±2.5	77±25	13.1±4.4	3.2±2.1	115±26	127



CR for BKG strength constrain

Together with the SGN, the MC prediction agrees with data within 1σ deviation

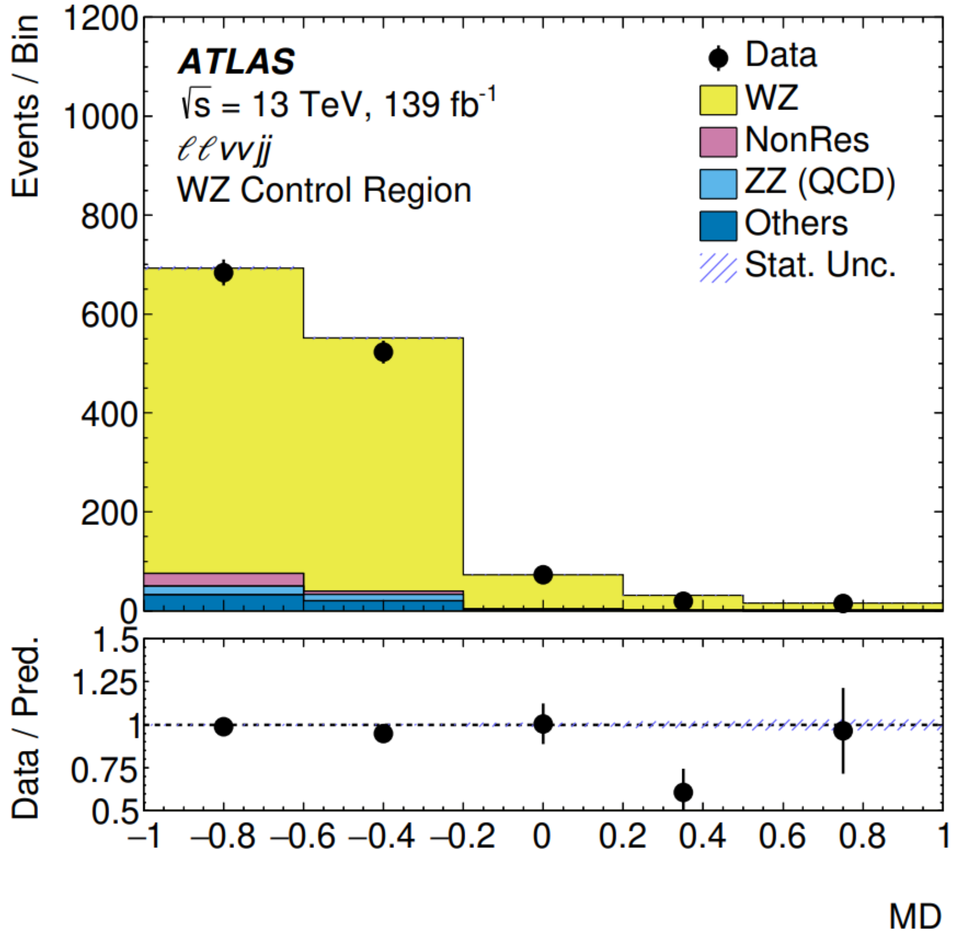
Background: llvv channel

- Irreducible BKGs from ZZ:
  - Estimated from MC simulation

- WZjj BKG:
  - Data-driven, yield from dedicated 3l-CR (Discriminated by lepton flavor), via transfer factor:

$$N_{WZ} = N_{Data-non\,WZ\,MC}^{3l\,CR} \times \frac{N_{MC}^{2l\,SR}}{N_{MC}^{3l\,CR}}$$

  - The EW WZjj MC is scaled by $1.77^{+0.51}_{-0.45}$ (The μ_{EW} measured from ATLAS VBS WZjj paper)



Background: llvv channel

- Non-resonant-ll (top, WW, Wt, Zττ) BKG:
  - Data-driven, estimated with events in dedicated eμ-CR (Non-res-ll enriched region)
  - e/μ reconstruction and selection efficiency difference reflected by the epsilon factor ε^2 = Nee/Nμμ
  - The number of "non-res-ll" BKG in the kinetic region Q(p_T, η) is obtained from eμ-CR:

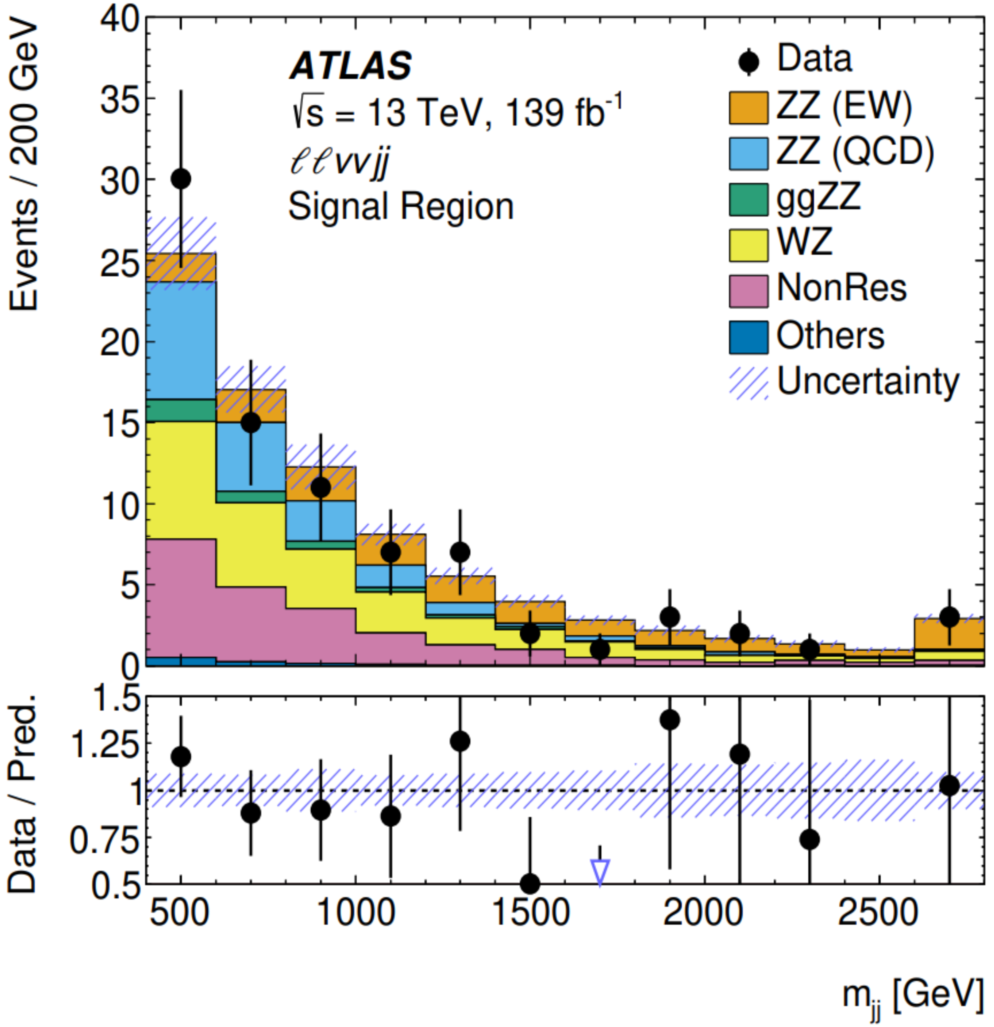
By Z→ee and Z→μμ events from data

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

- Z+jets BKG:
  - Z+jets background is largely reduced by the MET-significance requirement
  - Shape from MC, yield is estimated from data-driven in low MET-significance region

Yields and distributions in llvv channel

139 fb^{-1}	*llvv* SR
EW ZZjj	13.6±0.7
QCD ZZjj	17.2±3.5
QCD ggZZjj	3.5±1.1
Non-resonant-ll	21.4±4.8
WZ	24.6±1.1
Others	1.2±0.9
Total	81.5±6.4
Data	82



Together with the SGN, the MC prediction agrees with data within 1σ deviation

Systematics

- Theoretical uncertainty
  - PDF: Following the PDF4LHC procedure
  - α_S: Varying the α_S value by ± 0.001
  - QCD scale: Varying the renormalization and factorization scales (Envelope), conservatively treated as uncorrelated between the SR and the QCD CR in the 4l channel.
  - Parton shower: Comparing the nominal Pythia 8 parton showering with the alternative Herwig 7

~10% and 30% for the EW and QCD ZZjj process

- Experimental uncertainty
  - Luminosity uncertainty
  - The energy measurements of leptons and jets
  - The lepton reconstruction and selection efficiencies
  - Trigger selection efficiency
  - Pile-up correction
  - etc.

~10% and 5% in the 4l and llvv channel

- MC statistical uncertainty

Systematics: Interference

Interference between EW and QCD processes is treated as systematic on the EW VBS ZZ production measurement

Two approaches to estimate the interference:

1. Subtract the EWK and QCD process from the inclusive one

2. Generate the interference process directly via Madgraph (Baseline)

$$|M|^2 = |M_{EWK} + M_{QCD}|^2 = |M_{EWK}|^2 + |M_{QCD}|^2 + 2 \times Re(M^*_{EWK} \cdot M_{QCD})$$

Varying from 10% to 2% in different MD regions

The ratio for Interference over EWK signal is treated as the systematic.

Systematics: 4l background

Modeling of the QCD ZZjj

- QCD ZZjj modelling is very important for the VBS analysis.
- Thus we compared different generators → Additional shape syst.: Compare Sherpa with Madgraph
- Also tried different correlation approaches, and eventually chose the most convective one: Uncorrelated between the low and high MD regions.

Ranges from 30% to 20% in different MD regions

Pile-up modeling (Only considered for the QCD ZZjj BKG)

- Compare shape difference using low/high mu QCD MC events to cover the possible bias due to pile-up effect

In different MD regions are below 10%
Except in the last bin in the QCD ZZjj CR where it reaches 50% due to the statistical uncertainty of the simulated events.

Fake uncertainty

- Fake factor calculation
- 4l fake region where the fake factors are applied.

Systematics: llvv background

- **WZjj background:**
  - Systematics from: data statistic in 3l CR (dominant one, 5%);
  - Uncertainty from the quoted VBS WZjj paper on the 1.77 factor
- **Non-resonant-ll (WW, top, Zττ) Background:**
  - Systematics include the data statistical uncertainty (dominant one, 20%);
  - The shape difference between MC and data-driven in SR
- **Z+jets background:**
  - Systematics include the MC and data-driven difference (dominant),
  - Different fitting functions, and different fitting range

Inclusive Cross-section measurement

- Cross sections are measured for the **inclusive processes**, in individual 4l and llvv channels in **fiducial volume**.

- **Fiducial regions** are defined closely following the detector level selections, except
  - 4l channel: Z window loose to [60, 120] GeV (is [66, 116] GeV for detector-level).
  - llvv channel: truth MET > 130 GeV instead of MET-significance > 12

- **Measures cross-section:** $\boldsymbol{\sigma^{F.V.} = \frac{N_{data} - N_{bkg}}{C \times Lumi}}$

- **C-factor:** $\boldsymbol{C = \frac{N_{Reco}}{N_{F.V.-truth}}}$

	Measured fiducial σ [fb]	Predicted fiducial σ [fb]
$\ell\ell\ell\ell jj$	$1.27 \pm 0.12(\text{stat}) \pm 0.02(\text{theo}) \pm 0.07(\text{exp}) \pm 0.01(\text{bkg}) \pm 0.02(\text{lumi})$	$1.16 \pm 0.04(\text{stat}) \pm 0.20(\text{theo})$
$\ell\ell\nu\nu jj$	$1.13 \pm 0.28(\text{stat}) \pm 0.04(\text{theo}) \pm 0.06(\text{exp}) \pm 0.15(\text{bkg}) \pm 0.02(\text{lumi})$	$1.07 \pm 0.01(\text{stat}) \pm 0.12(\text{theo})$

Statistical Fit for EW Processes

- Use **Gradient Boost Decision Tree (BDTG)**
  - Jet-related ones are important for 4l
  - Both jet and dilepton related ones are important for llvv
- Fit 3 regions simultaneously: 4l SR, 4l QCD CR, llvv SR
- Parameter of interest (POI): $\boldsymbol{\mu_{EW}}$
- Normalization factor constrained in 4l QCD CR: $\boldsymbol{\mu_{QCD}}$
- The full set of systematics are included
  - Experimental systematics: Fully correlated in the 3 regions for each NP
  - Theoretical systematics: Different correlation approached have been tested and the most conservative one is chosen.

Observation!

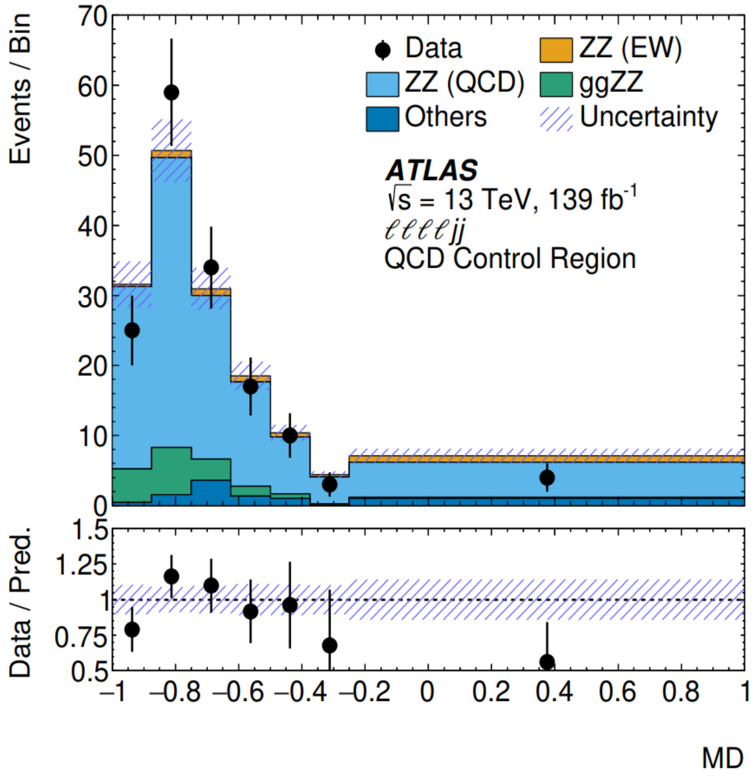
	μ_{EW}	$\mu_{\mathrm{QCD}}^{\ell\ell\ell\ell jj}$	Significance Obs. (Exp.)
$\ell\ell\ell\ell jj$	1.4 ± 0.4	0.98 ± 0.22	$5.5\ (4.4)\ \sigma$
$\ell\ell\nu\nu jj$	0.8 ± 0.6	–	$1.3\ (2.0)\ \sigma$
Combined	1.21 ± 0.31	0.99 ± 0.22	$5.7\ (4.8)\ \sigma$

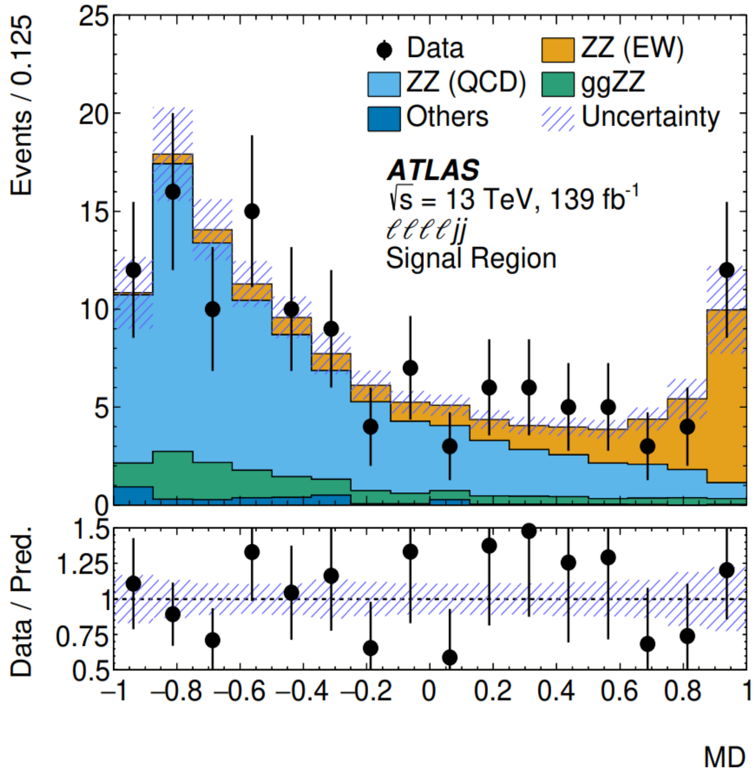
BDT variables

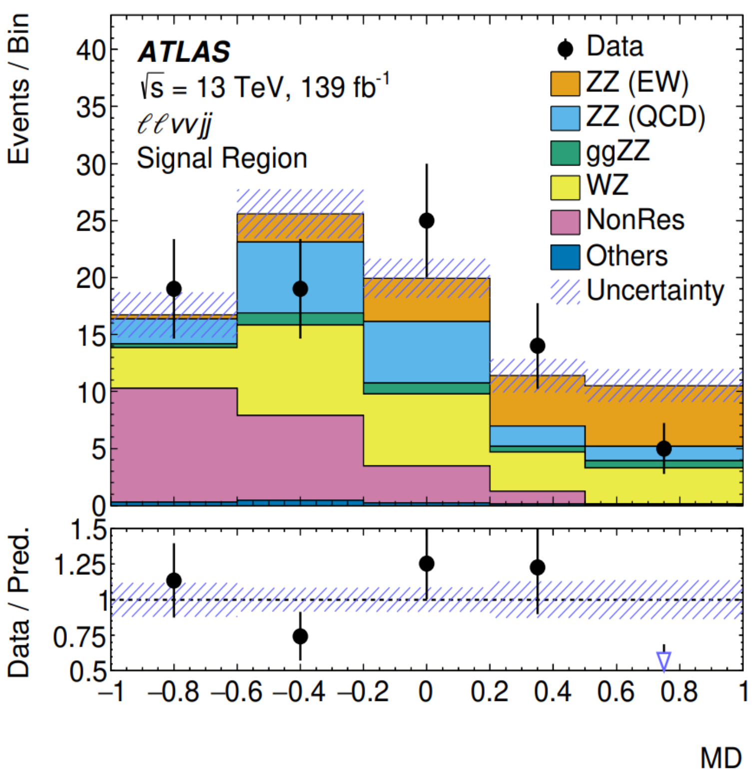
4l	llvv
Mjj	$\Delta\eta(ll)$
leading jet Pt	Mll
2^{nd} jet Pt	$\Delta\Phi(ll)$
$p_T(ZZjj)/H_T(ZZjj)$	Mjj
$Y(j1)\times Y(j2)$	MET-sig
$\lvert\Delta Y(jj)\rvert$	$\lvert\Delta Y(jj)\rvert$
Y^*_{Z2}	$Y(j1)\times Y(j2)$
Y^*_{Z1}	H_T
p_T^{ZZ}	$\Delta R(ll)$
m_{ZZ}	2^{nd} jet Pt
p_T^{Z1}	MET
p_T^{l3}	p_T^{l2}
	leading p_T^{l}

Post-fit distributions

- **BDT distributions are after statistical fit.**
- **The data distributions are well reproduced by the predicted contributions.**







Conclusion

- The observation of VBS (VV→V'V') process in ZZ channel, with full Run 2 data of ATLAS is presented.
  - Observed (expected) significance of **5.7 (4.8)σ**.
  - Inclusive cross section in 4l and llvv channels has been measured individually. Compatible with SM prediction and still dominant by data statistical uncertainty.

- This result completes the observation of weak boson scattering for massive bosons.

- **Accepted by Nature Physics!**