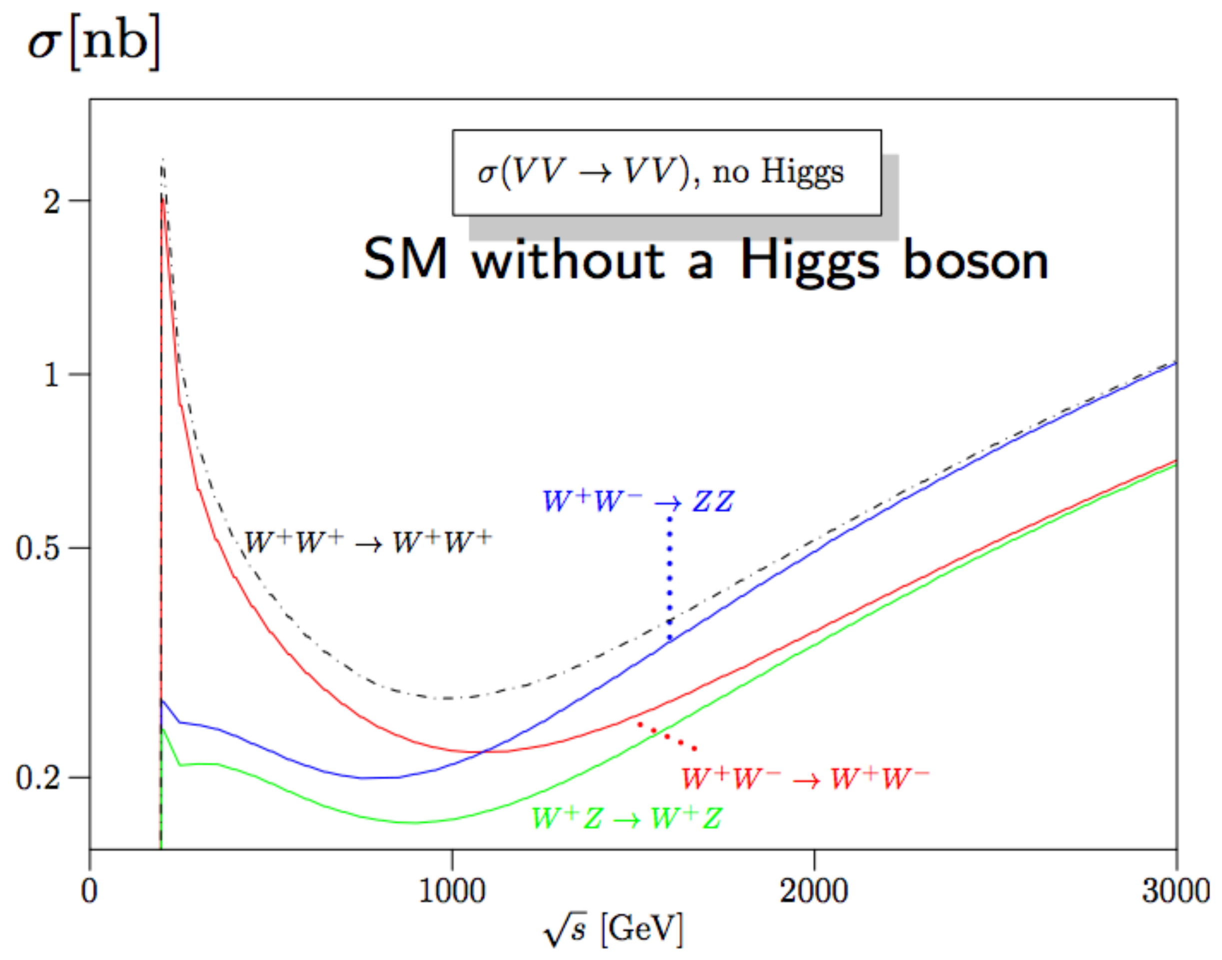
Measurement of Inclusive 4l (llνν) + 2-jet Cross Section and Observation of EW Component in 13 TeV Proton-Proton Collisions with the ATLAS Detector

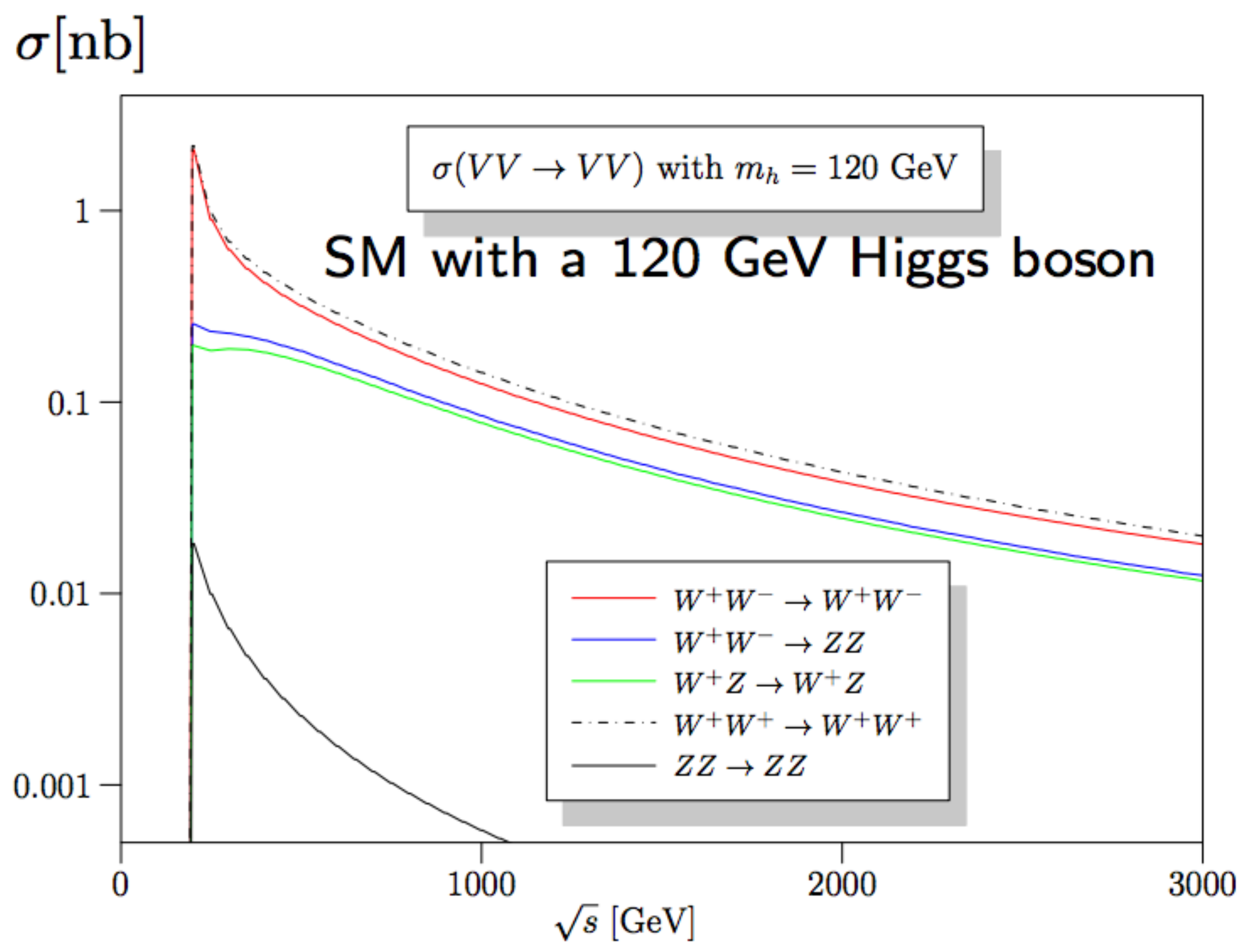
ATLAS Conference note: [ATLAS-CONF-2019-033]()

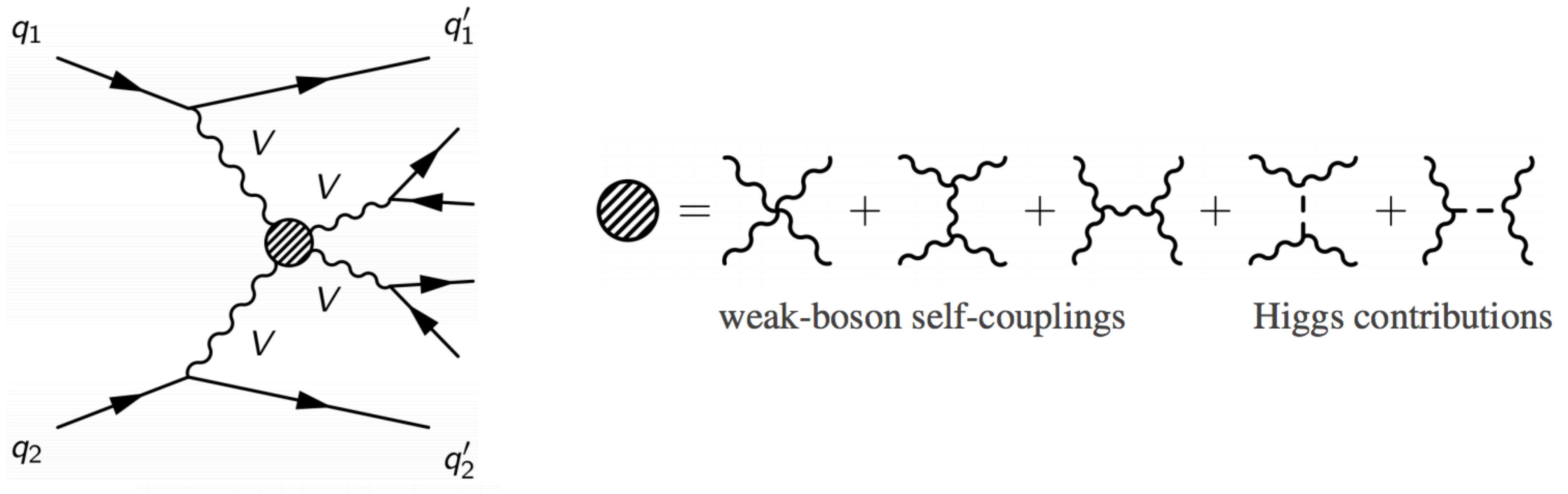
Heling Zhu
On behalf of the ATLAS Collaboration
University of Sci. and Tech. of China / Brookhaven National Lab
Sep. 2nd, 2019 @ Corfu

Vector Boson scattering

- 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.







SM-EWK diagrams: $O(\alpha^6_{EW})$ @ LO

- VBS-production

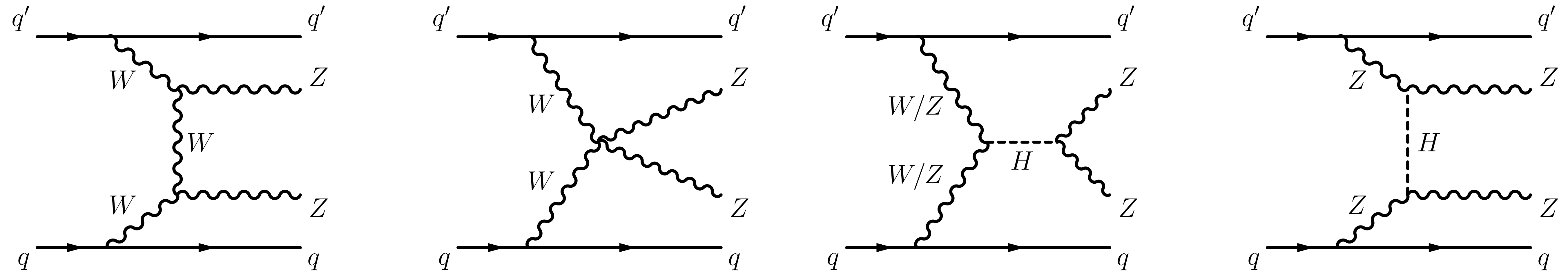
Overview of VBS measurements in LHC @ 13TeV

- Status of August 2019

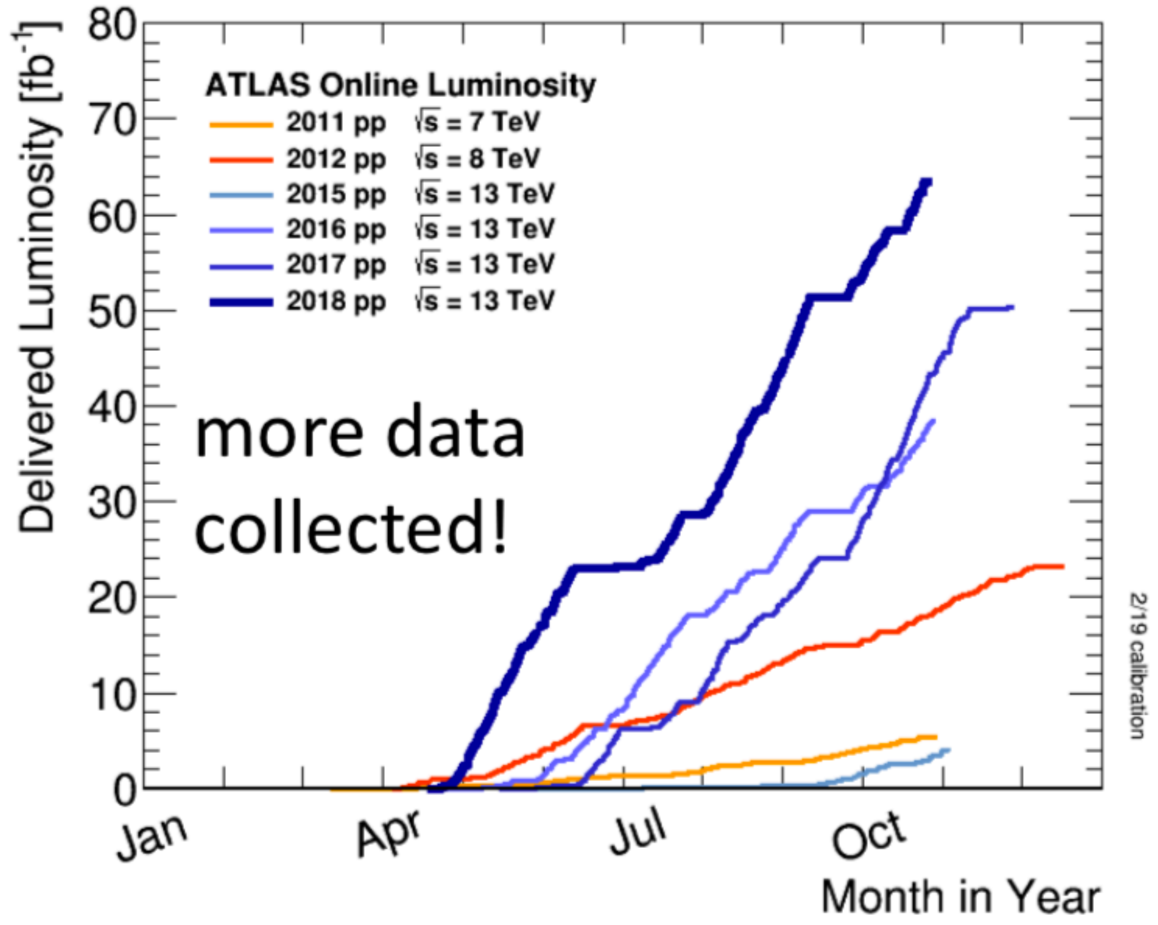
Channel	Observed (expected) significance		Experimental Challenges
	ATLAS	CMS	
ssWW	6.5 (4.9)σ	5.5 (5.7)σ	"Golden channel": first observation of VBS in this channel, very good EW/QCD ratio
WZ	5.3 (3.2)σ	1.9 (2.7)σ	Similar cross section as ssWW, but larger QCD backgrounds
ZZ	5.5 (4.3)σ First 139 fb^{-1} analysis	2.7 (1.6)σ llllljj channel only	Very clean 4l channel, low background but small cross section
Zγ	4.1 (3.8)σ	3.9 (5.2)σ +8TeV: 4.7 (5.5)σ	Higher statistics due to photon, but no sensitivity to BSM EWSB

Introduction

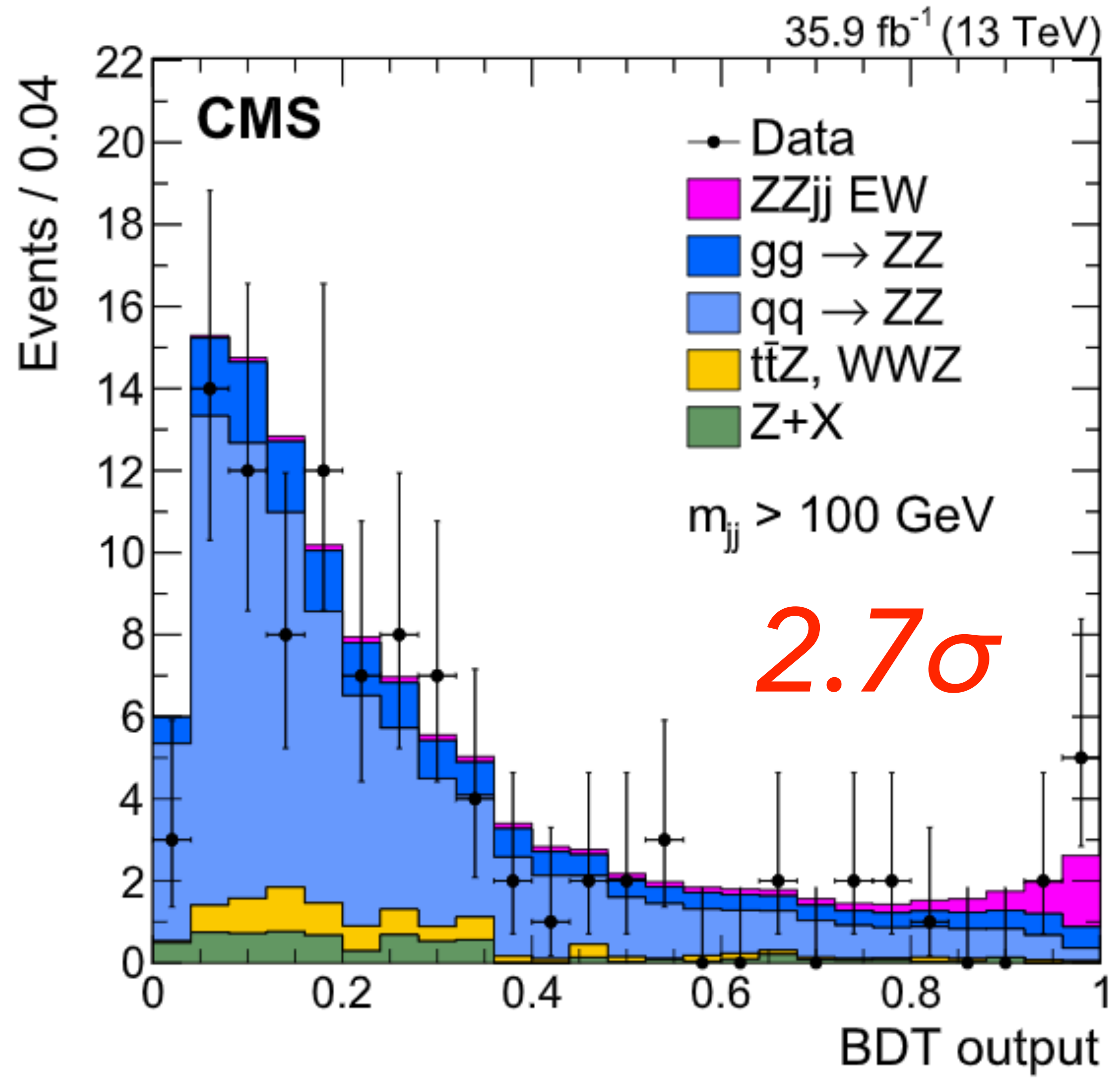
- This analysis measures the inclusive cross-section of ZZjj processes and searches for **EW ZZjj production** with 4l and llνν final states
    - 4l: very clean experimental signature, high signal-to-background ratio
    - llνν: larger branching ratio



- The EW ZZjj production was searched for by the CMS collaboration using 36.1 fb^{-1} of 13 TeV pp collision data, but no evidence was found
- This analysis presents the first VBS ZZjj observation with 139 fb^{-1} full run-2 datasets



Time for rare processes → benchmark SM measurements!



[Phys. Lett. B 774 (2017) 682]

Analysis Strategy

- Working on ATLAS full run-2 data with 139 fb^{-1}
- Experimental signature: 4 leptons, or 2 leptons with E_T^{miss} + two forward jets
- First step: inclusive cross-section measurements with cut-based analysis
- MVA (BDTG)-based analysis is then used to extract the EW VBS ZZ signal from background for
  - Observation of electroweak production of ZZ+2j
  - Observed signal strength of the EW processes
- The analysis is performed independently in ZZ→4l and ZZ→llvv final states, then combine together.
- Interference between EW and QCD processes is treated as systematic on the EW VBS ZZ production measurement

MC simulation

EWK ZZjj	*MadGraph5*
QCD ZZjj	*Sherpa222*
QCD ggZZjj	*4l:Sherpa222; llvv:gg2VV*
WZ→ llvv	*Sherpa222*
ttbar, singleTop	*Powheg*
ttbarV	*MG5_aMC@NLO*
Zjets	*Sherpa221*
Triboson	*Sherpa222*
WW, WZ semileptonic	*Powheg*

Event selection

	$\ell\ell\ell\ell jj$	$\ell\ell\nu\nu jj$
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 $\Delta y(jj) > 2$

Generally tighter selections in llvv channel due to more backgrounds

- The optimizations of Jet selection have been studied to enhance the EW VBS contribution in forward region.

4l

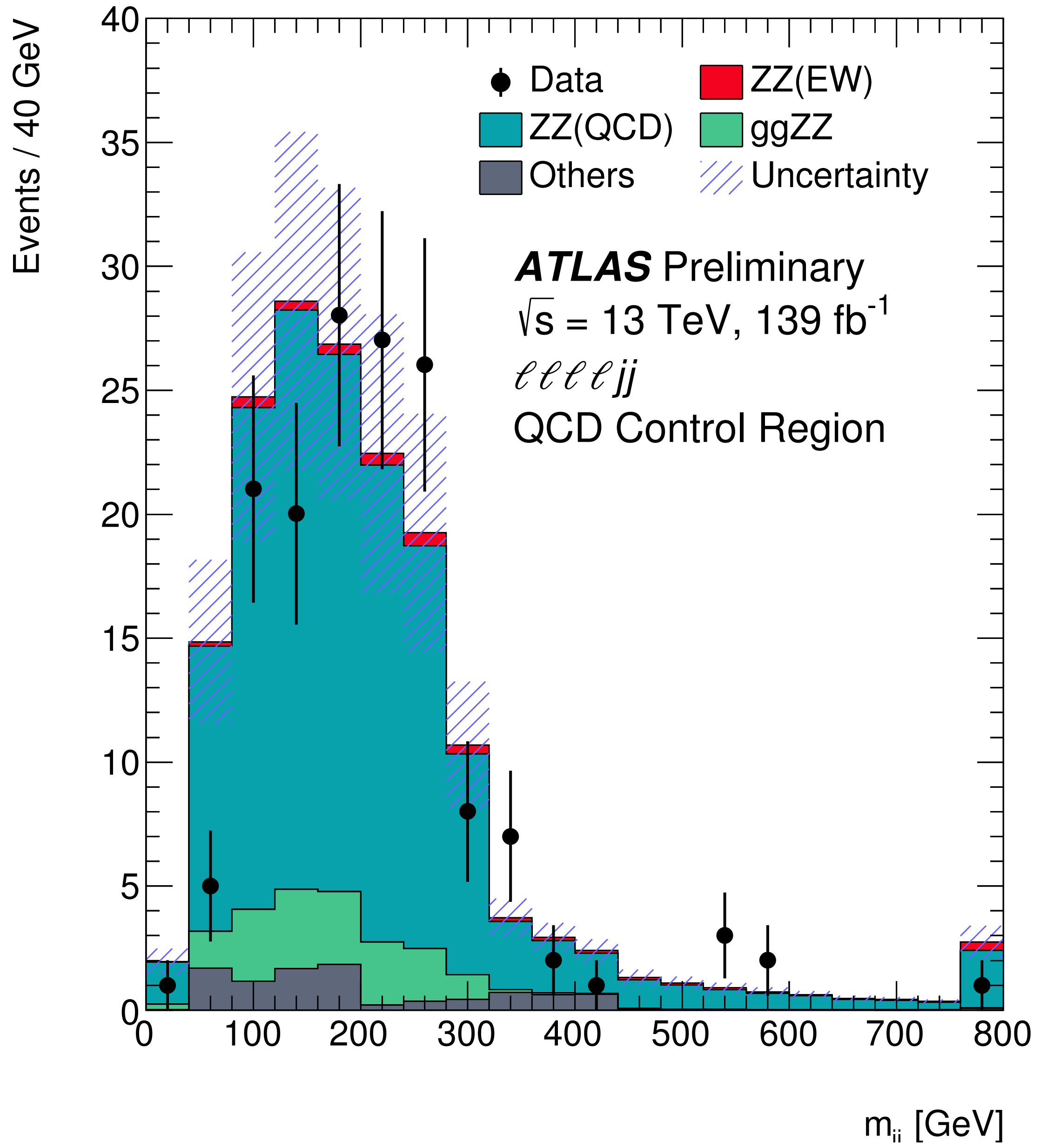
- QCD CR to constrain QCD strength:
  - Mjj < 300GeV or dYjj < 2

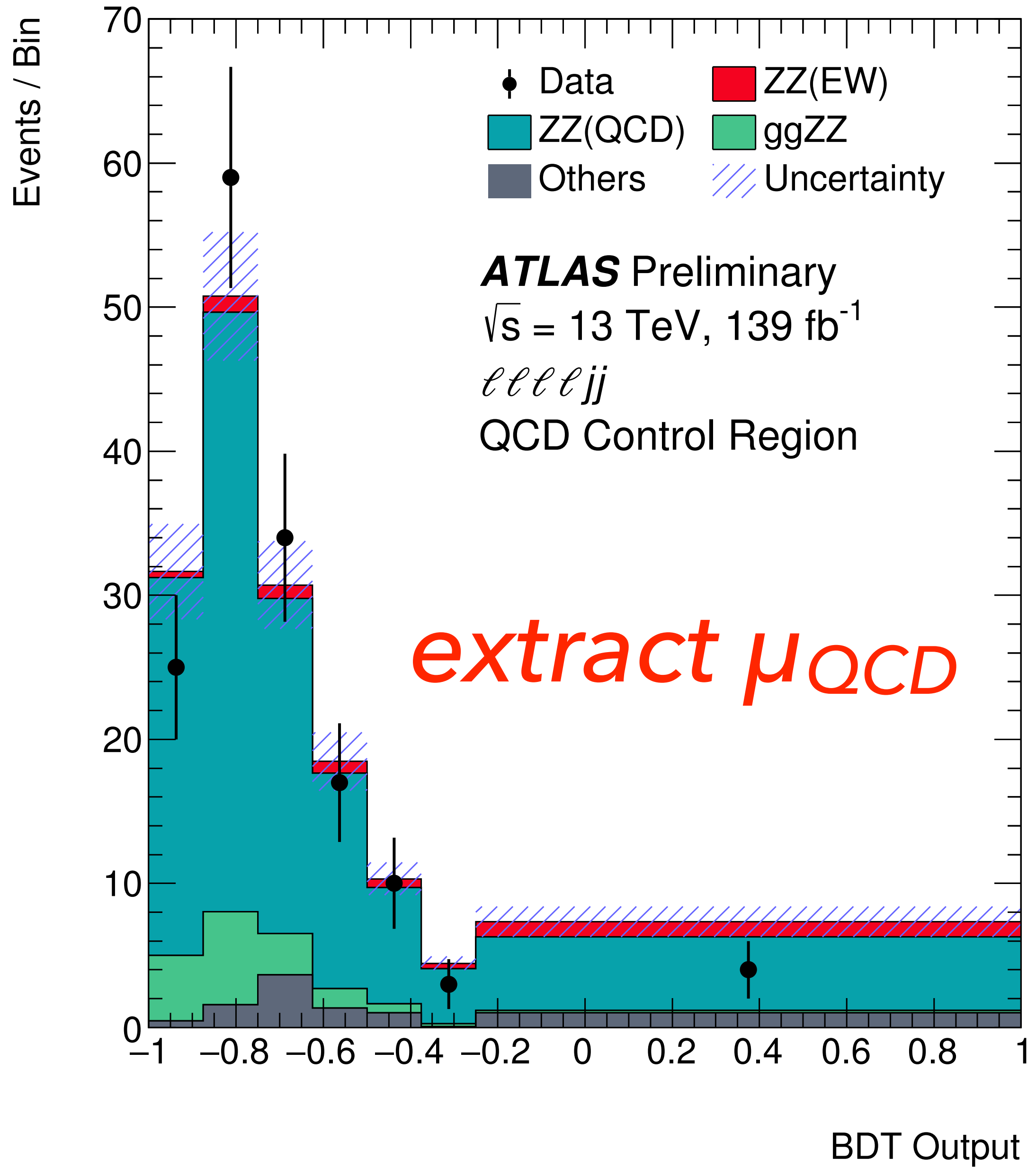
llvv

- 3rd lepton veto to reduce ZZ→4l, WZ→3l+v
- MET-significance has high suppression for Zjets

ZZ→4l: Background

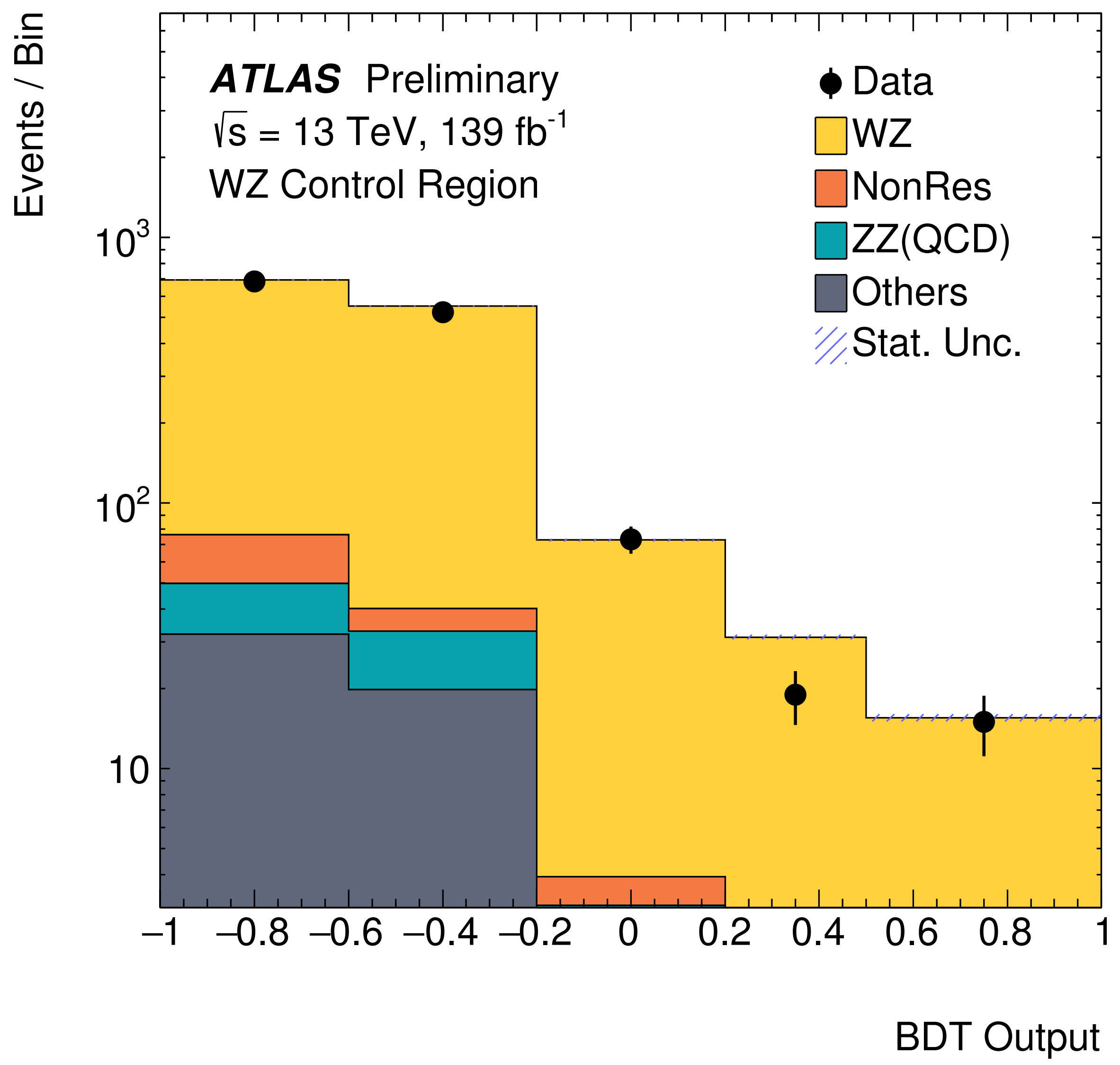
- **ZZ irreducible, qq→ZZ and gg→ZZ**
  - Yield constraint by dedicated QCD CR
  - Additional systematics included for the shape modeling
    - Difference between MG and sherpa
- **Fake background: Zjets, $t\bar{t}$**
  - Fake factor method used, derived from dedicated CRs.
  - Systematics:
    - Variations on bad lepton definitions
    - Data MC difference: due to the very limited statistics even in fake control region, a very conservative approach is used here
    - Statistical uncertainty on fake factor

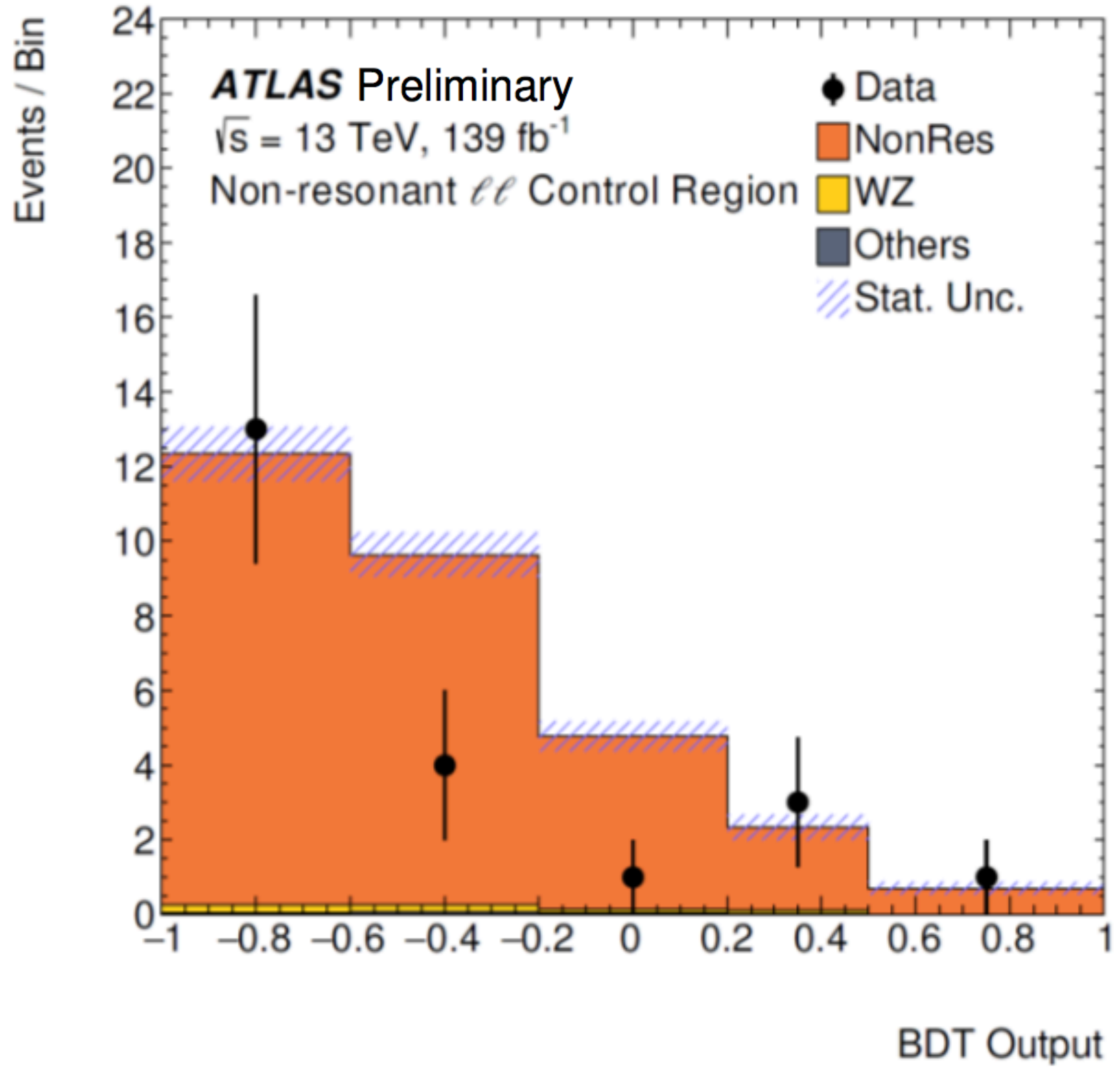




ZZ→llvv: Background

- **ZZ irreducible:**
  - Estimated from simulations, systematics are same as 4l
- **WZjj background:**
  - Yield from dedicated 3l-CR, the EW WZjj is scaled by $1.77^{+0.51}_{-0.45}$ (the μ_{EW} measured from latest ATLAS VBS WZjj paper)
- **Non-resonant-ll (top, WW, Wt, Zττ) Background**
  - Estimated with events in dedicated eμ-CR
  - e/μ reconstruction and selection efficiency difference is taken care of by the epsilon factor $\varepsilon^2 = N_{ee}/N_{\mu\mu}$
- **Zjets background:**
  - *Zjets* background is largely reduced by the *MET*-significance requirement
  - Shape from MC, yield is estimated from data-driven in low MET-significance region





Inclusive Cross-section Measurements

- 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

pre-fit event yields

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

- Measured cross-Section: $\sigma^{F.V} = \frac{N_{data} - N_{bkg}}{C \times Lumi}$, while the C-factor is $C = \frac{N_{Reco}}{N_{F.V-truth}}$

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

Statistical Fit for EW Processes

- Use Gradient Boost Decision Tree (BDTG) for both 4l and llvv channels
  - Trained with leptons and jets kinematics after SR selection
- Fit 3 regions simultaneously: 4l SR, 4l QCD CR, llvv SR
- μ_{EW} as parameter of interest;
  μ_{QCD} (for 4l) is constrained by QCD CR.
- The full set of systematic uncertainties are included
  - Experimental systematic uncertainties are fully correlated.
  - Theoretical systematics are uncorrelated due to different fiducial definitions.

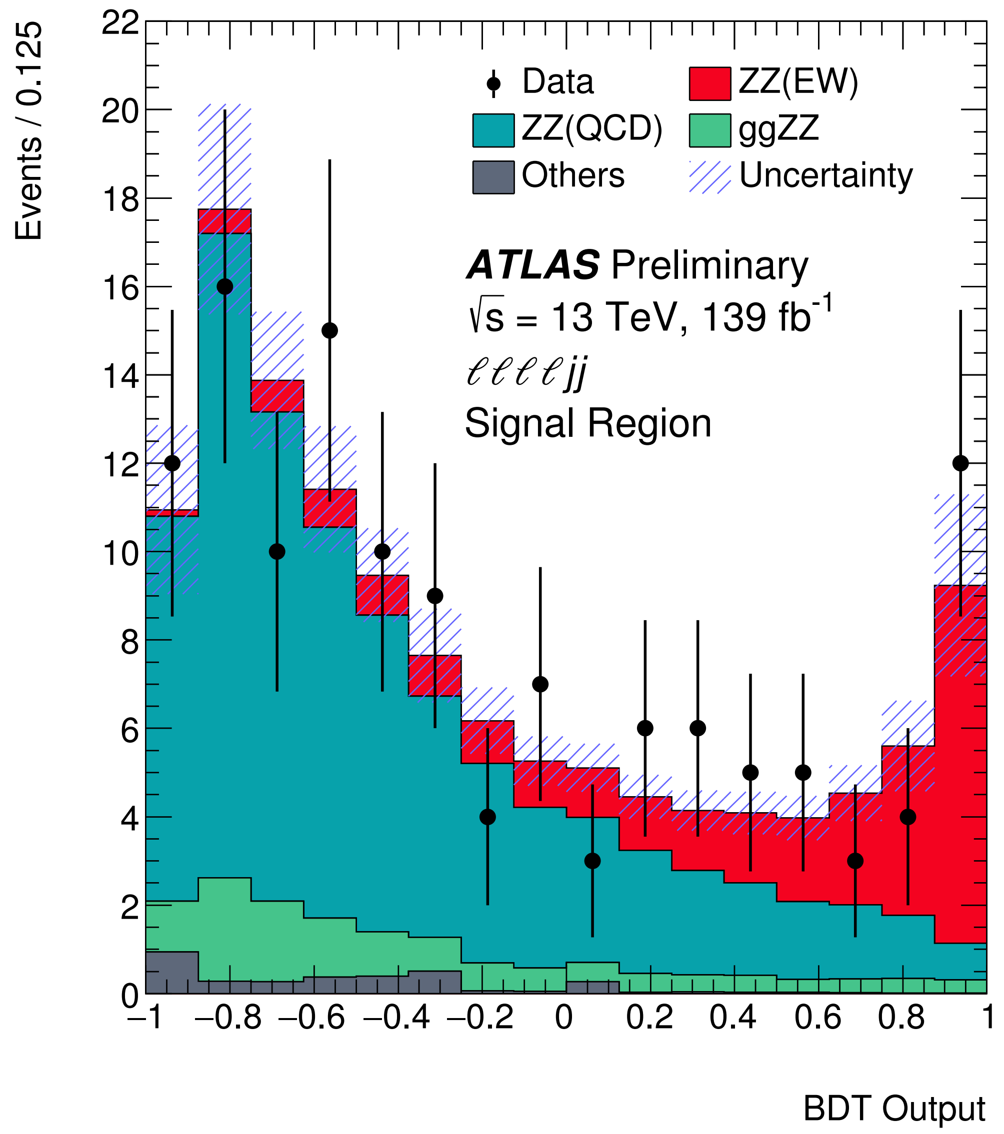
Observation!

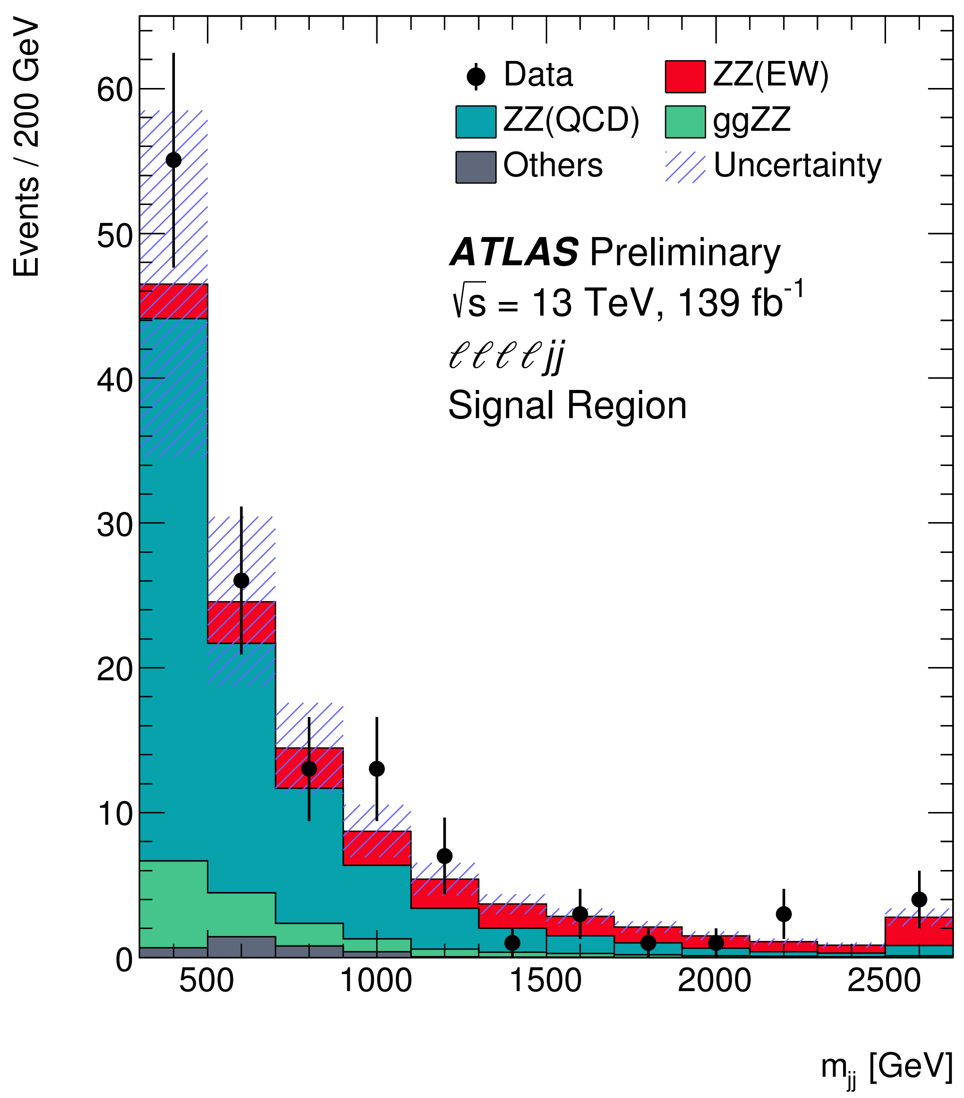
	μ_{EW}	$\mu_{\mathrm{QCD}}^{\ell\ell\ell\ell jj}$	Significance Obs. (Exp.)
$\ell\ell\ell\ell jj$	1.54 ± 0.42	0.95 ± 0.22	$5.48\ (3.90)\ \sigma$
$\ell\ell\nu\nu jj$	0.73 ± 0.65	-	$1.15\ (1.80)\ \sigma$
Combined	1.35 ± 0.34	0.96 ± 0.22	$5.52\ (4.30)\ \sigma$

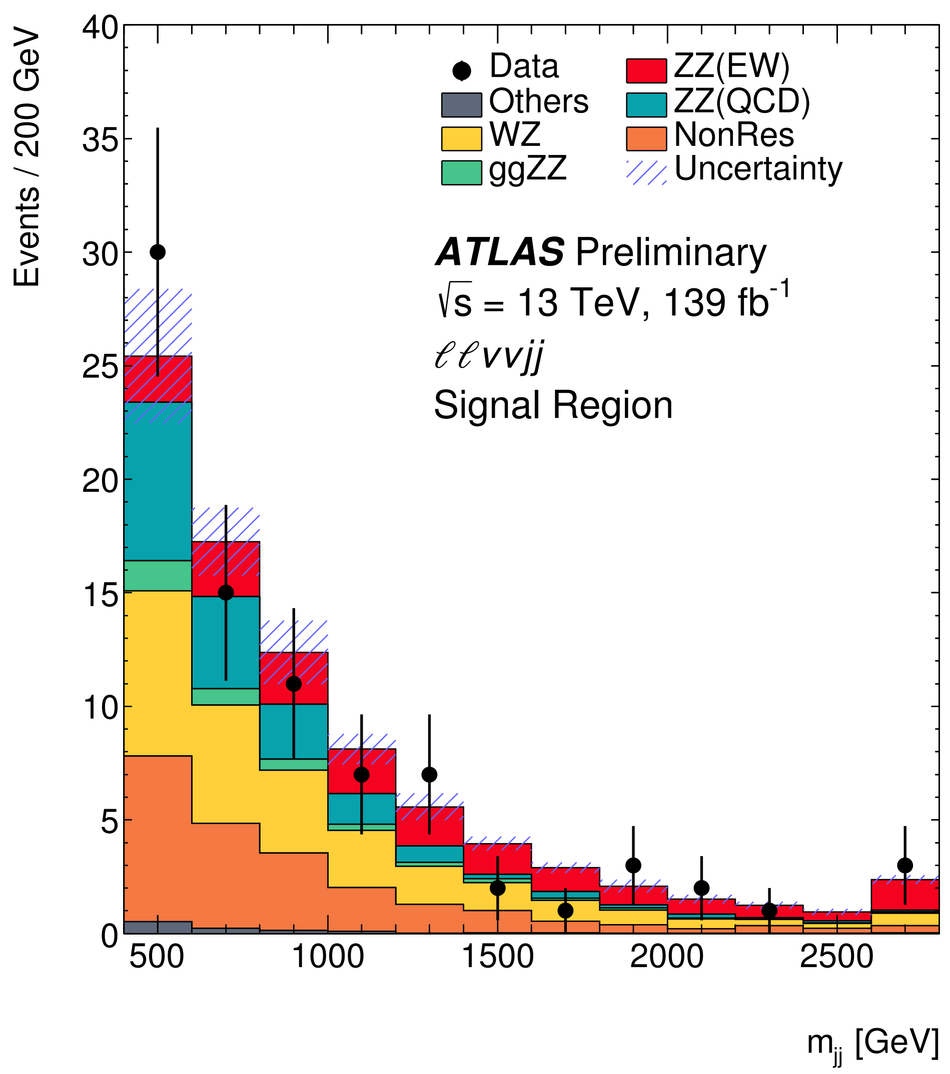
Rank	4l	llvv
1	Mjj	$\Delta\eta$(ll)
2	leading p_Tj	m_{ll}
3	2nd p_Tj	$\Delta\Phi$(ll)
4	p_T(ZZjj)/H_T(ZZjj)	Mjj
5	Y(j1)×Y(j2)	MET-sig
6	|ΔY(jj)|	ΔY(jj)
7	Y^*_{Z2}	Y(j1)×Y(j2)
8	Y^*_{Z1}	HT
9	p_T^{ZZ}	ΔR(ll)
10	m_{ZZ}	2nd p_Tj
11	p_T^{Z1}	MET
12	p_T^{l3}	2nd p_T^l
13	-	leading p_T^l

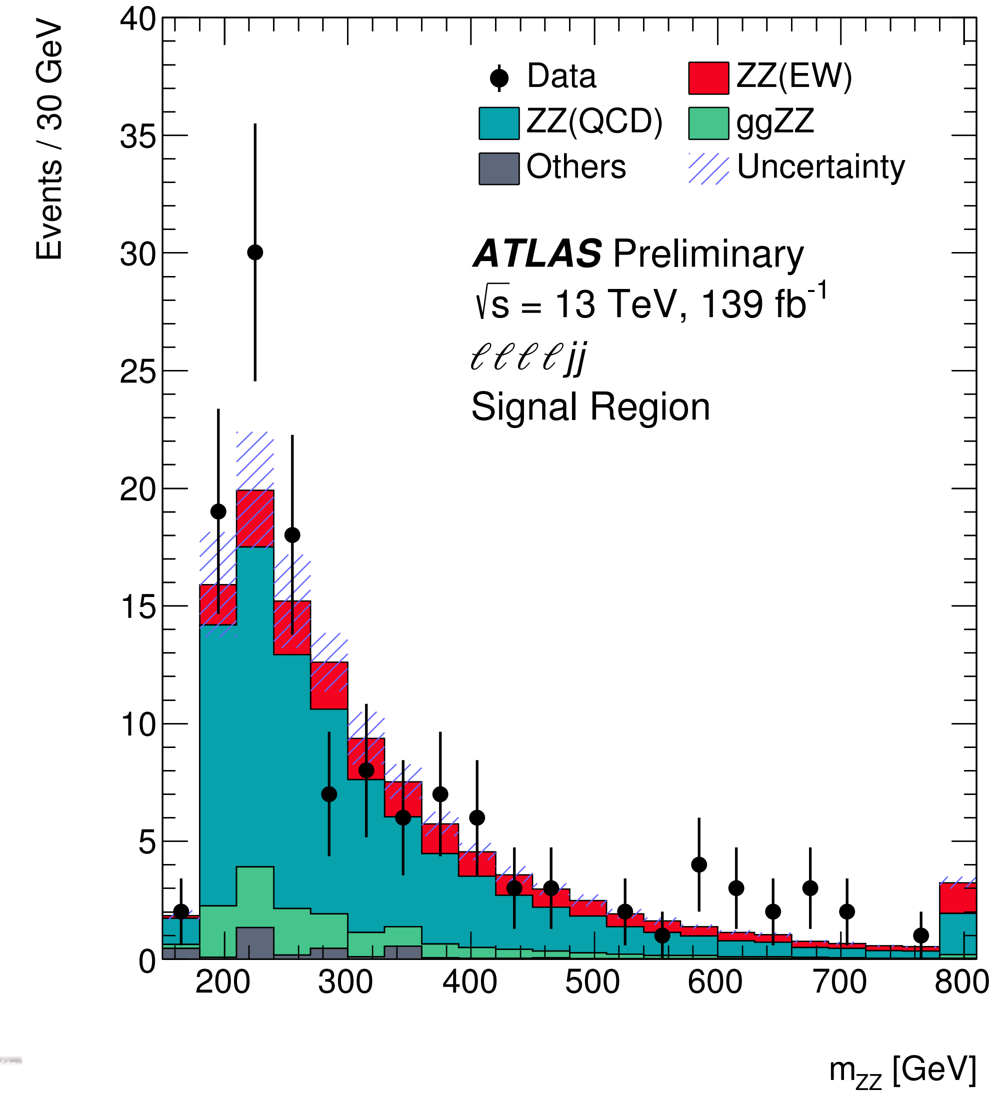
Post-fit distributions

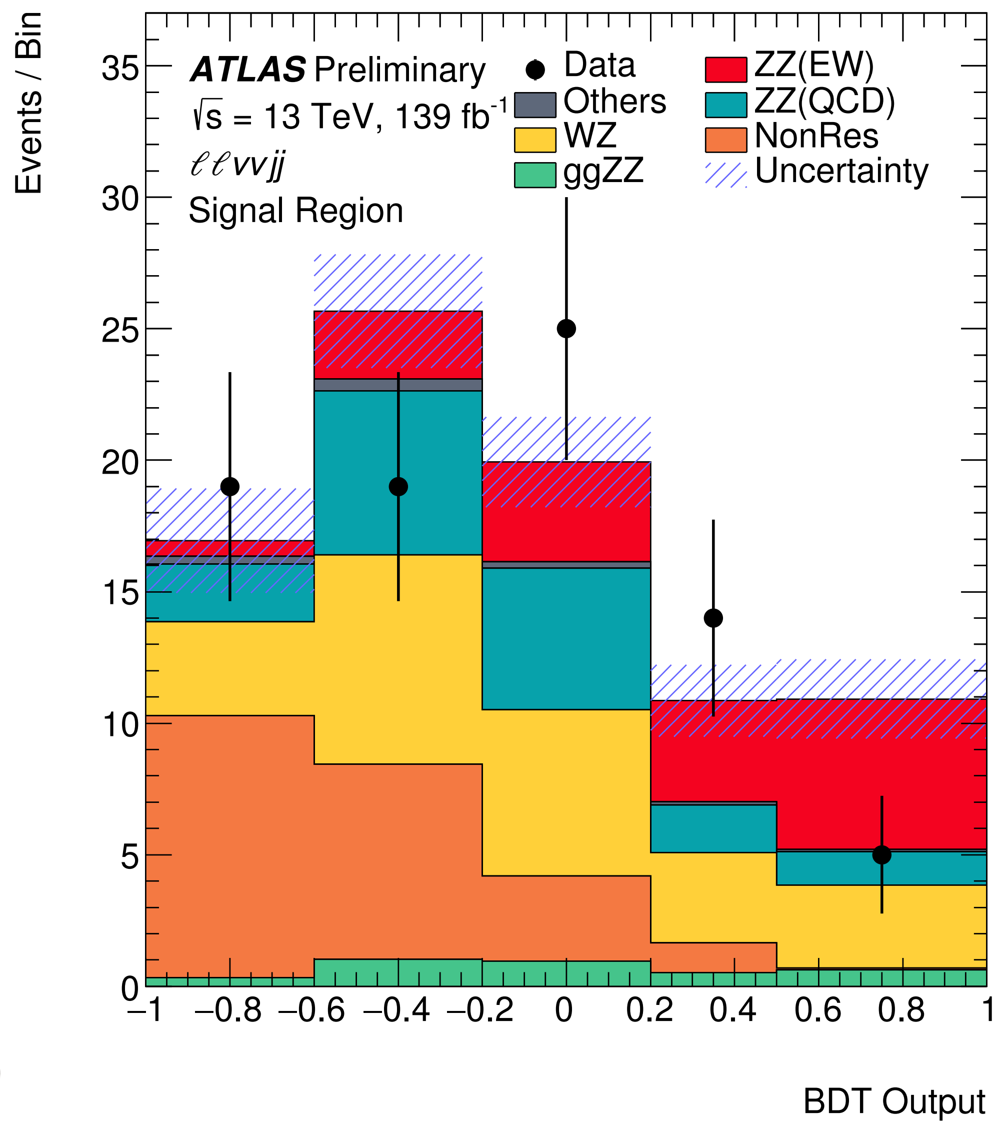
- Mjj/MZZ: predictions have been scaled with μ_{EW}=1.35 and μ_{QCD}=0.96
- BDT distributions are after statistical fit.
- The data distributions are well reproduced by the predicted contributions.











Conclusion

Physics Briefing

New milestone reached in the study of electroweak symmetry breaking

The new ATLAS result provides a firm observation of the electroweak production of two jets in association with ZZ, with a statistical significance of 5.5 standard deviations.

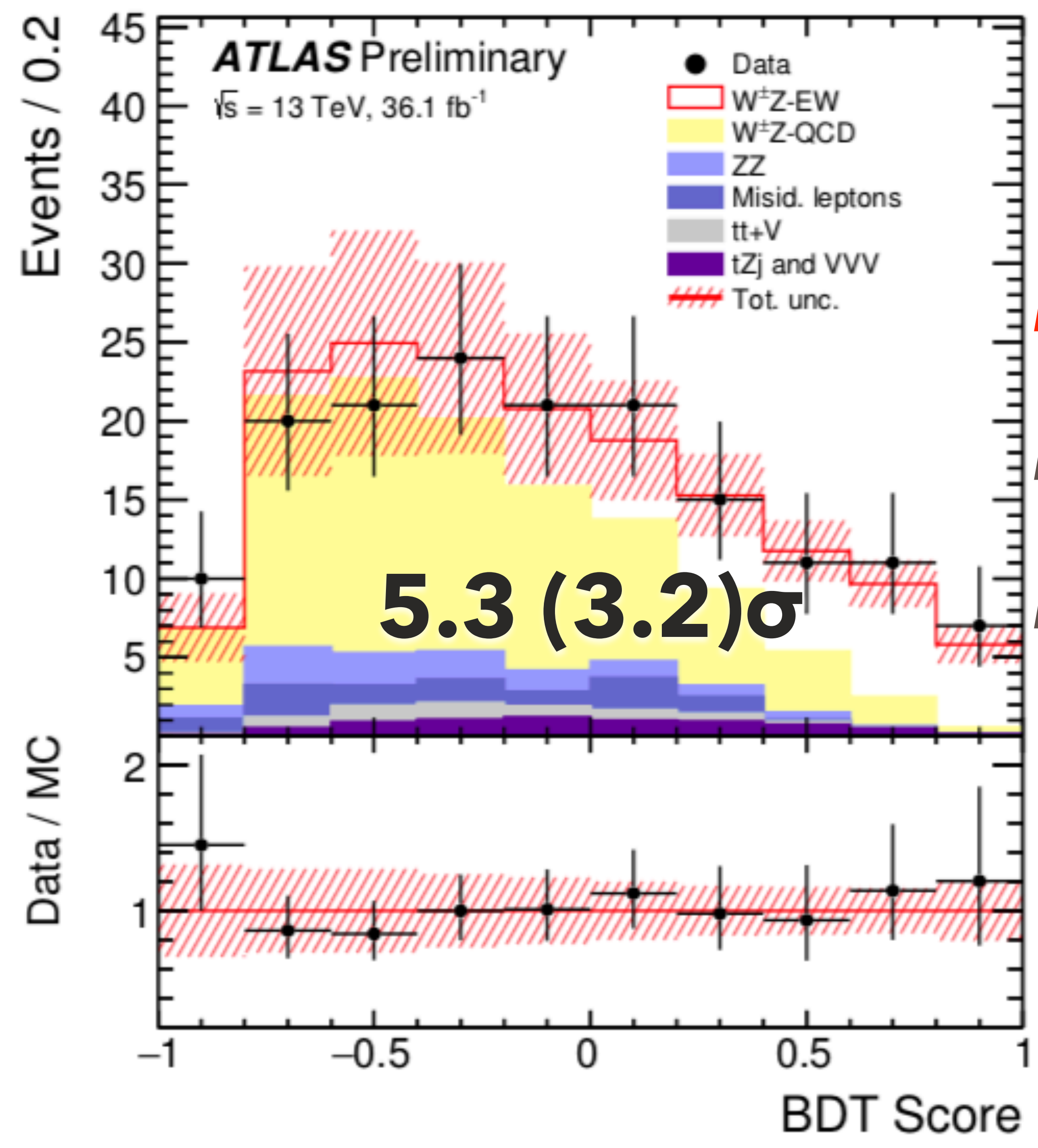
- The observation of VBS (VV->V'V') process in ZZ channel, with full Run 2 data of ATLAS is presented. ***With more data collected in LHC, it's time for rare processes!***
  - Observed (expected) significance of 5.5 (4.3)σ.
  - Inclusive cross section in 4l and llvv channels has been measured individually. Compatible with SM prediction and still dominant by data statistical uncertainty.
- This result completes the observation of weak boson scattering for massive bosons, and sparks new ways to test EWSB.

Backup

WZ scattering

- Three leptons final states
- QCD production of WZjj is the dominant background.
- Analysis strategies:
  - **CMS, conservative approach:** use features of EW vs. QCD processes that are well-understood and robust against limitations of theoretical predictions
  - **ATLAS, aggressive approach**: train BDT for EW vs QCD discrimination on 15 variables (selected from 33 studied); background normalization factor were constrained by CRs.

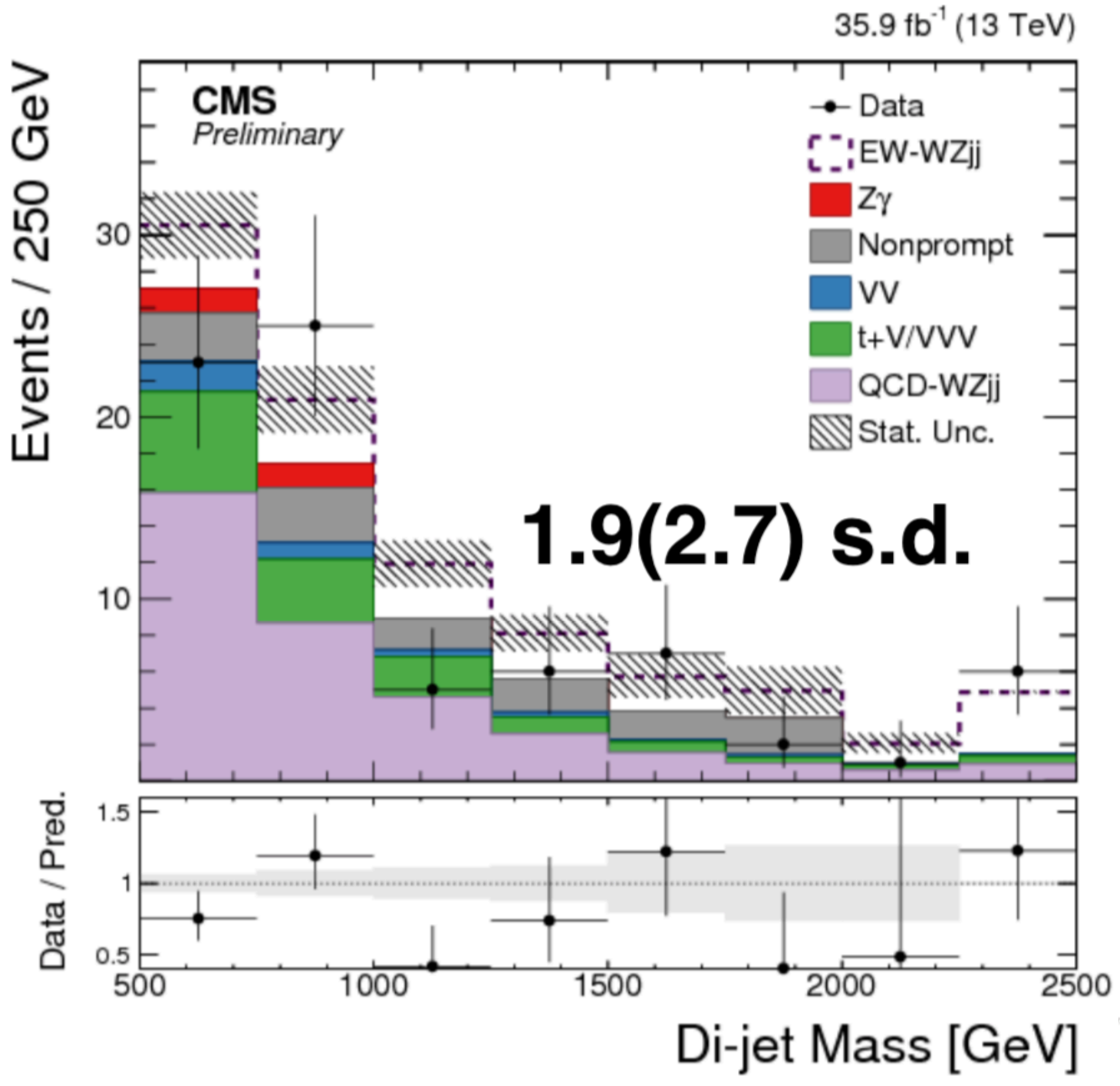
post-fit discriminants



$\mu_{QCD\text{-}WZ} = 0.60 \pm 0.25$

$\mu_{ttV} = 1.18 \pm 0.19$

$\mu_{ZZ} = 1.34 \pm 0.29$



ZZ→4l: QCD ZZjj bkg, Additional systematics

- Modeling of the QCD ZZjj
  - The shape differences between different generators can not be covered by the standard QCD up/down systematics
  - Additional shape systematic derived by comparing Sherpa with MG samples at truth level
- Pile-up modeling
  - Compare shape difference using low/high mu QCD MC events to cover the possible bias due to pile-up effect

ZZ→4l: background, Fake

- Fake-factor measured in jet-enriched samples: Z+jets CR and ttbar CR
  - 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)
- 4l fake CR : SR with 1 or 2 leptons passing poor lepton definition
- Fake contribution in signal region:
  - F = N("Good Lepton")/N("Poor Lepton")
  - $N_{fake} = (N_{gggp} - N^{ZZ}_{gggp}) \times F - (N_{ggpp} - N^{ZZ}_{ggpp}) \times F^2$
  - ZZ contribution is subtracted
  - The second term is due to double counting of N_{gggp} and N_{ggpp}

- *Fake factor vs **flavor and p_T/η dependent***
- *Fake lepton background nominal estimation: **2.306±1.016***

ZZ→llvv: Background Systematics

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

Interference

- Additional uncertainty due to the interference between EW and QCD processes
  - Checked with private MG sample. 7(2)% in 4l (llvv) channel, w.r.t the EW component. Difference mostly due to different m_{JJ} cut between two channels
  - In the 4l channel this is developed vs. m_{JJ} and then translated to BDT at truth level
  - In llvv channel a global factor is used due to difficulty in rebuilding the BDT at truth level (mostly due to no MET- significance variable in truth level)

Decouple the systematics in 4l channel

- For the QCD modeling, 4 different configurations have been tested (2x2 of, decouple 4l SR and 4l CR or no, decouple 4l low BDT and high BDT or no).
- The most conservative approach is chosen (i.e, decouple both)

Table1	Expected	Observed
μ_{EWK}	0.995 ± 0.347	1.575 ± 0.416
μ_{QCD}	0.932 ± 0.269	0.894 ± 0.254
Significance	4.119	5.949

Table2	Expected	Observed
μ_{EWK}	0.940 ± 0.351	1.509 ± 0.420
μ_{QCD}	0.976 ± 0.233	0.934 ± 0.22
Significance	3.866	5.490

Table3	Expected	Observed
μ_{EWK}	1.020 ± 0.346	1.611 ± 0.413
μ_{QCD}	0.940 ± 0.272	0.904 ± 0.257
Significance	4.144	5.993

Table4	Expected	Observed
μ_{EWK}	0.961 ± 0.351	1.540 ± 0.418
μ_{QCD}	0.980 ± 0.233	0.946 ± 0.224
Significance	3.897	5.480

- *Table1: no decouple*
- *Table2: decouple QCD scale in 4l SR and 4l CR*
- *Table3: decouple QCD shape in low BDT region and high BDT region*
- *Table4: table 2+3*