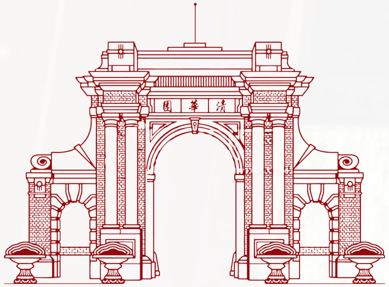
Vector-boson scattering in the ZZ(4ℓ)jj final state with full Run2 data

based on CMS-PAS-SMP-20-001
paper submitted to PLB: arxiv:2008.07013

Congqiao Li (李聪乔) (Peking University) *in collaboration with*

Roberto Covarelli* (Torino & INFN), **Ying An** (Peking), **Riccardo Bellan** (Torino), **Matteo Bonanomi** (LLR), **Claude Charlot** (LLR), **Elisa Fontanesi** (Bologna), **Duje Giljanovic** (LLR), **He He** (Wisconsin), **Damir Lelas** (FESB), **Qiang Li** (Peking), **Giacomo Ortona** (LLR), **Alexander Savin** (Wisconsin), **Toni Sculac** (FESB) ***@ CMS***

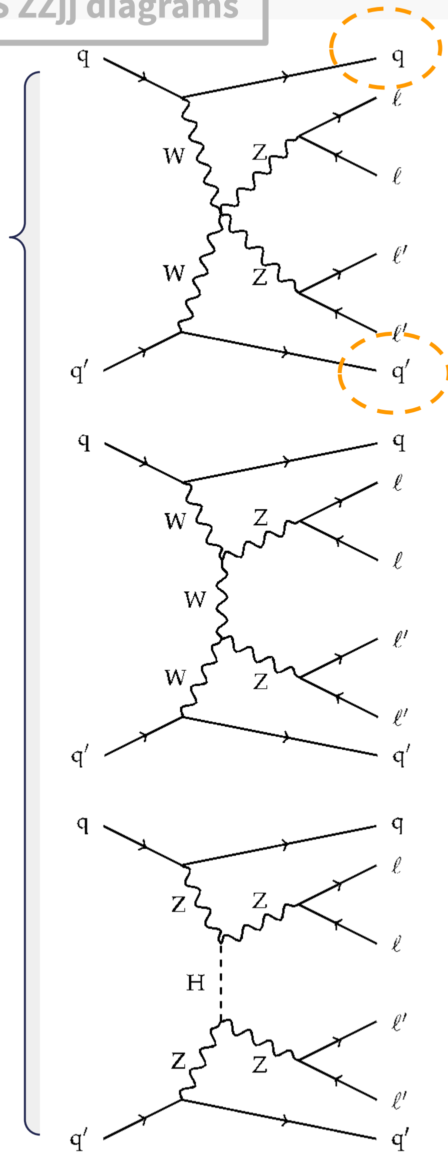


CLHCP 2020
Tsinghua University, Beijing

VBS topology and ZZjj channel

- ➔ **VBS @ LHC:** "the large boson-boson collider"
  - scattering of massive vector-boson is at the heart of **EW symmetry breaking** mechanism
  - measurement of σ crucial to the EW section of SM
  - however, rare in LHC collisions: one-in-trillion
- ➔ **VBS ZZjj topology**
  - characterized by **2 forward jets** and **ZZ produced in the central region** between them
- ➔ **Advantage**
  - clean, fully reconstructable final state for ZZ→ℓℓℓ'ℓ' channel
  - best sensitivity to certain aQGCs (f_{T8}, f_{T9})
- ➔ **Challenges**
  - large contamination from QCD-induced ZZ production
  - small σ for VBS signal (~0.3 fb) vs. main backgrounds (~10 fb)

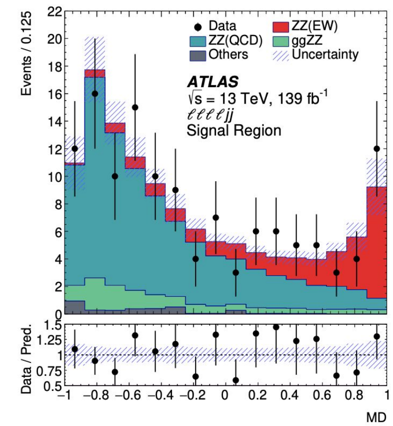
VBS ZZjj diagrams



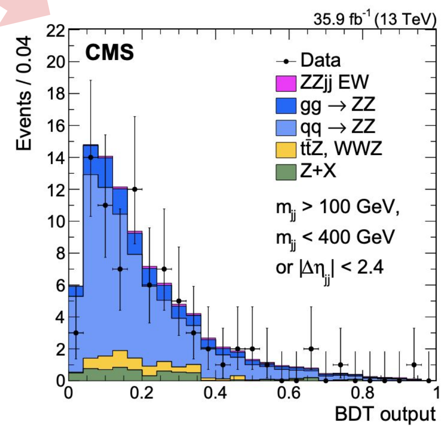
VBS ZZ measurement in Run2 at LHC

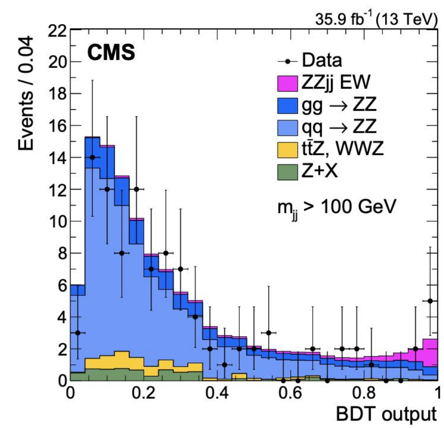
- ATLAS:
  - full Run2 (139 fb^{-1}); channel 4ℓjj + 2ℓ2vjj
    arxiv:2004.10612
  - observed (expected) significance = 5.5 (4.3)

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

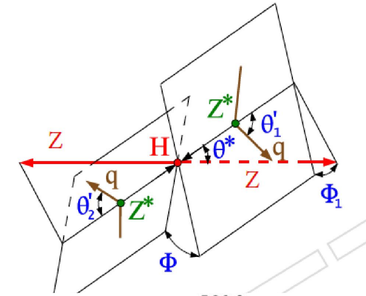


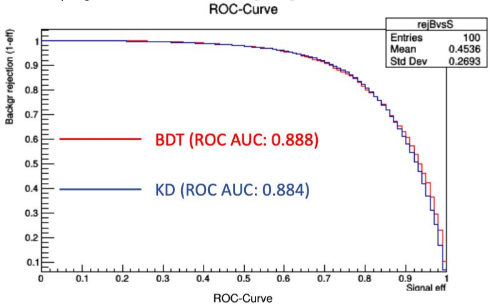
- CMS:
  - 2016 data (35.9 fb^{-1}); channel 4ℓjj
    arxiv:1708.02812
  - observed (expected) significance = 2.7 (1.6)
- CMS:
  - full Run2 (137 fb^{-1}); channel 4ℓjj
    arxiv:2008.07013 (this work)
  - observed (expected) significance = 4.0 (3.5)





Physics observables





- ➔ Cross-sections
  - ◆ Perform **EWK-discriminant-based analysis**
    - MELA: use MCFM ME probabilities:
      - event-by-event probabilities extracted from the respective leading-order matrix elements (MFCM)
      - Discriminant defined as **P(VBS)/[P(VBS)+ cP(QCD)]**
    - similar performance as validated with MVA techniques
  - ◆ Results
    - **Significance and signal strength of EWK signal** (fit 1D discriminant templates describing signal/background components) **in fiducial regions**, reproducing analysis cuts at particle level
      - **Baseline** (broadest dijet requirement, exploit all phase space)
      - **VBS-enriched** (dijet mass and Δη cuts used to define 2 signal-enriched regions)
    - **EWK+QCD cross-section in the same fiducial region** (cut-and-count)
- ➔ aQGCs in EFT approach
  - ◆ Fit high-$m_{4\ell}$ region
  - ◆ Results: set limits on relevant dim-8 operators

MC samples

signal (EWK)

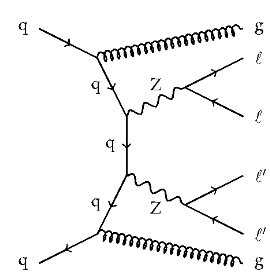
- Fully EWK (α_{EW}^6) 4ℓjj process
  - VBS (qq→ ZZq'q') and tribosons (qq→ ZZV→ ZZqq') not physically separable
  - Define a partonic invariant mass m_{jj} (LHE level); consider **m_{jj} > 100 GeV** as signal
- Both simulated with `MG5_aMC@NLO`
  - **VBS** (m_{jj} > 100 GeV) at LO QCD, Z→ ℓℓ with MadSpin
  - **Tribosons** (m_{jj} < 100 GeV) at NLO QCD (ZZW and ZZZ separately) with inclusive decays and full BW structure (off-shell included), then require m_{jj} < 100 GeV at LHE level
- Phantom generator for validation
  - good agreement in σ and shapes

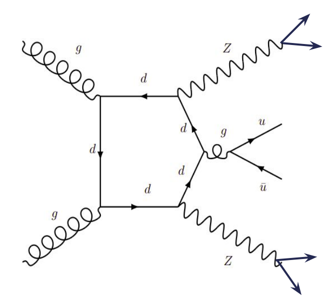
EWK-QCD interfence

- Interface simulated at LO with `MG5`
  - very small (< 5% of signal in sensitive high-MELA region)
  - take linear treatment
    $N = \mu_{EW} S + \sqrt{\mu_{EW}}\, I + B = \mu_{EW}(S + 0.5I) + (B + 0.5I)$
    ⇒ enable to perform a linear fit in μ_{EW}

background (QCD)

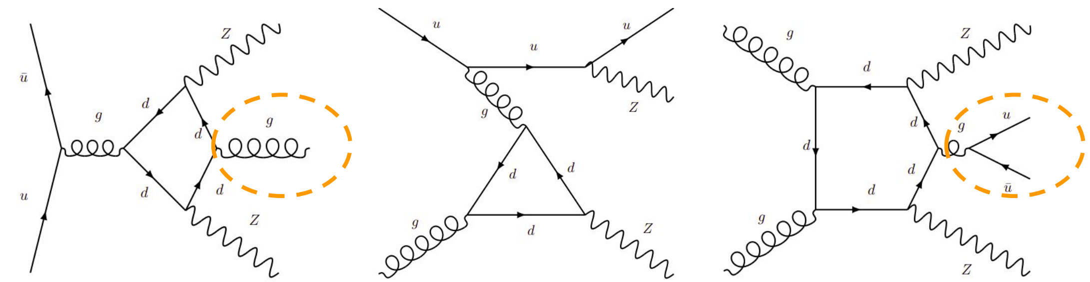
- QCD ($\alpha_{EW}^4 \alpha_s^2$) 4ℓjj process
- Separated into tree-level (qqbar initial state) and loop induced (gg initial state)
  - QCD qqbar→ ZZ + 0/1 jets at NLO with `MG5`
    - apply NNLO QCD k-factor
    - apply NLO EWK k-factor on top of it
  - QCD gg→ ZZ + 0/1/2 jets at LO with `MG5`
    - A new simulation expressly studied for this analysis
    - Pushing MadGraph capabilities to the limit
- Minor backgrounds
  - ttbar Z (LO `MG5`)
  - WWZ (NLO `MG5`)

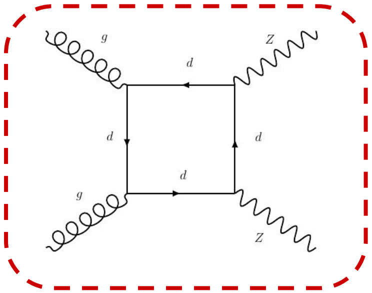




MC samples: the novel gg→ ZZ simulation

see details in arxiv:2006.12860
or CLHCP talk





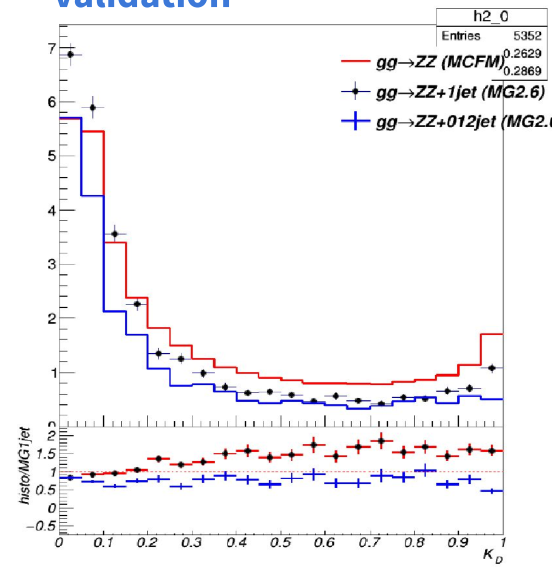
the gg→ZZ 0-parton diagram

- ➔ Technical challenges
  - **Advanced MadGraph usage**: loop-diagram filter
  - **CPU and computation time** (10 min/event) needs beyond what provided by CMS official production

➔ Why is it needed?

- Some diagrams of gg→ZZ+1,2-jet process which have different structures from the 0-jet
- jet can emit directly from the loop, which is not the ordinary ISR/FSR
  ⇒ expect a more accurate jet simulation using MadGraph, by producing the 0,1,2-jet processes and merging events with MLM matching

validation

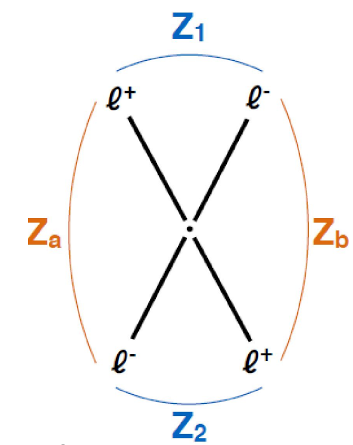


- ➔ **40% less contribution** in high-MELA region
- ➔ **Major improvement** over previous analyses
  - CMS 2016 paper used MCFM (+0 partons)
  - ATLAS uses SHERPA (+0,1 partons)

Object reconstruction

- ➔ Double + single lepton triggers (+ tri-electron)
- ➔ Lepton and jet reconstruction **inherit from H→ ZZ→ 4ℓ analysis**
  - ◆ Loose lepton reconstruction
    - Electrons: MVA-based with isolation included
    - Muon: Cut-based
    - Scale factors and tag-and-probe uncertainties from H→ ZZ→ 4ℓ analysis
  - ◆ Jets
    - anti-k_T, R = 0.4
    - **p_T > 30 GeV** , $|\eta| < 4.7$
    - migrate pile-up
    - jet energy correction appropriate for each year applied

ZZ candidate reconstruction



- ➔ Z candidate: opposite-sign same-flavor lepton pair satisfying **$60 < m_{\ell\ell(\gamma)} < 120$ GeV**
- ➔ $p_T(\ell_1) > 20$ GeV, $p_T(\ell_2) > 10$ GeV
- ➔ $\Delta R > 0.02$ between each of the four leptons
- ➔ $m_{\ell\ell} > 4$ GeV for OS pairs (regardless of flavor)
- ➔ reject 4μ and 4e candidates where the alternate pairing Z_a, Z_b satisfies
  - ◆ $|m(Z_a)-m(Z)|<|m(Z_1)-m(Z)|$ and $m(Z_b) < 12$ GeV
- ➔ **$m_{4\ell} > 180$ GeV**
- ➔ VBS-specific:
  - ◆ At least 2 jets passing requirements
  - ◆ $m_{jj} > 100$ GeV (isolates VBS from triboson ZZV → 4ℓjj)

Three regions for analysis:

Baseline: $m_{jj} > 100$ GeV

$m_{jj} > 400$ GeV && $\Delta\eta_{jj} > 2.4$?

Y → **VBS-enriched (loose)** → $\Delta\eta_{jj} > 5$ → **VBS-enriched (tight)**

N → **Control region**

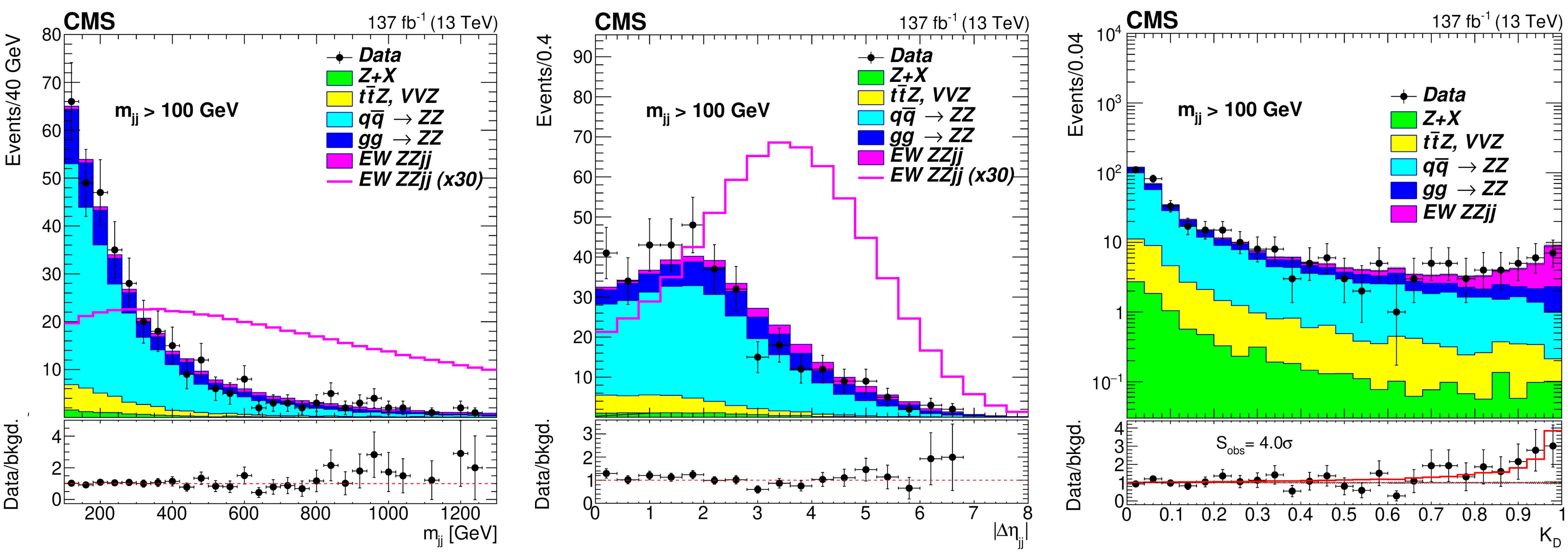
Background estimation & systematics

- Irreducible background:
  - From MC: qqbarZZ , ggZZ , ttbar Z, WWZ, WZZ, ZZZ
- Reducible background:
  - Z+X background: >1 non-prompt leptons in final state
    - main sources are non-isolated e and μ from decays of heavy-flavor mesons, mis-reconstructed jets and e from γ conversions (“fake leptons”)
  - fake rates estimated by measuring fake electron (muon) rates that pass final selection over those which pass a looser selection
  - same-sign (SS) method used from HZZ4ℓ analysis, use comparison of OS vs. SS to estimate systematics
- **Systematics**
  - taken in the form of shape / log-normal uncertainties (approximate numbers shown below)

Systematic source	qqZZ	ggZZ	VBS	Z+X	Shape	Years corr.?
QCD scales	10-12%	9-14%	6%	-	x	x
PDF + α_s	3.2%	5%	6.6%	-		x
Lepton trigger, reco, sel.	2.5-9%	2.5-9%	2.5-9%	-		x
L1 prefiring	0.6-1.0%	0.6%	1.8-3.0%	-	x	
Luminosity	2.3-2.5%	2.3-2.5%	2.3-2.5%	-		
JES	4.9-11.4%	3.6-10.2%	0.7-1.2%	-	x	
JER	2.2-6.3%	1.0-2.2%	0.2-0.4%	-		
MC samples	2.5-4.2%	3.2%	$\ll$ 1%	-	x	
Pileup	0.2-2.6%	0.4-2.7%	0.3-1.7%	-		
Reducible background	-	-	-	33-45%		

Baseline ZZjj region distributions

- ➔ Distribution of m_{jj}, $\Delta\eta_{jj}$ (variables to control the VBS phase-space) in the ZZjj region
  - ◆ large invariant mass and separation angle for the VBS signal
- ➔ MELA variable K_D used for the fit



Cross-section and signal strengths

- Signal extraction for pure EW signal and EW+QCD is implemented in the K_D discriminant distribution in the baseline ZZjj region, via a maximum-likelihood fit
  - measured (expected) signal strength

$$\mu_{\mathrm{EW}} = 1.22^{+0.47}_{-0.40}\,(1.00^{+0.44}_{-0.36}) \quad \mu_{\mathrm{EW+QCD}} = 0.99^{+0.13}_{-0.12}\,(1.00^{+0.13}_{-0.12})$$

  - background-only hypothesis is excluded with a significance of 4.0 (3.5 expected)
- Measured cross-sections in the three fiducial regions
  - SM σ obtained with particle-level cuts on generator-level MC samples used for the analysis and respective theory uncertainties
  - μ from binned ML fit using histogram templates with MINOS uncertainties
  - σ from analogous binned ML fits with frozen theory uncertainties

	Perturbative order	SM σ (fb)	Measured σ (fb)
		ZZjj inclusive	
EW	LO	0.275 ± 0.021	$0.33^{+0.11}_{-0.10}\,(\mathrm{stat})^{+0.04}_{-0.03}\,(\mathrm{syst})$
	NLO QCD	0.278 ± 0.017	
	NLO EW	$0.242^{+0.015}_{-0.013}$	
EW+QCD		5.35 ± 0.51	$5.29^{+0.31}_{-0.30}\,(\mathrm{stat}) \pm 0.47\,(\mathrm{syst})$
		VBS-enriched (loose)	
EW	LO	0.186 ± 0.015	$0.180^{+0.070}_{-0.060}\,(\mathrm{stat})^{+0.021}_{-0.012}\,(\mathrm{syst})$
	NLO QCD	0.197 ± 0.013	
EW+QCD		1.21 ± 0.09	$1.00^{+0.12}_{-0.11}\,(\mathrm{stat}) \pm 0.07\,(\mathrm{syst})$
		VBS-enriched (tight)	
EW	LO	0.104 ± 0.008	$0.09^{+0.04}_{-0.03}\,(\mathrm{stat}) \pm 0.02\,(\mathrm{syst})$
	NLO QCD	0.108 ± 0.007	
EW+QCD		0.221 ± 0.014	$0.20^{+0.05}_{-0.04}\,(\mathrm{stat}) \pm 0.02\,(\mathrm{syst})$

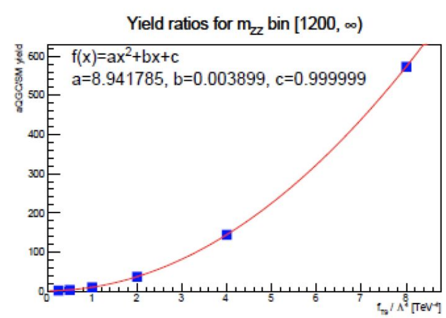
aQGC limits

- ➔ Dimension-8 operators in effective field theory
  - originate from covariant derivatives of the Higgs doublet and from charged and neutral field strength tensors associated to gauge bosons
  - ZZjj channel particularly sensitive to T0, T1, and T2 (charged-current operators), as well as the T8 and T9 (neutral-current operators)
- ➔ aQGC samples in the analysis
  - simulated by MG5 in the default coupling value f_{Ti}/Λ^4 = 2 TeV
  - other scenarios from reweighting
  - semi-analytic description of the expected m_{ZZ} distribution as a function of the aQGC couplings obtained by fitting quadratic functions

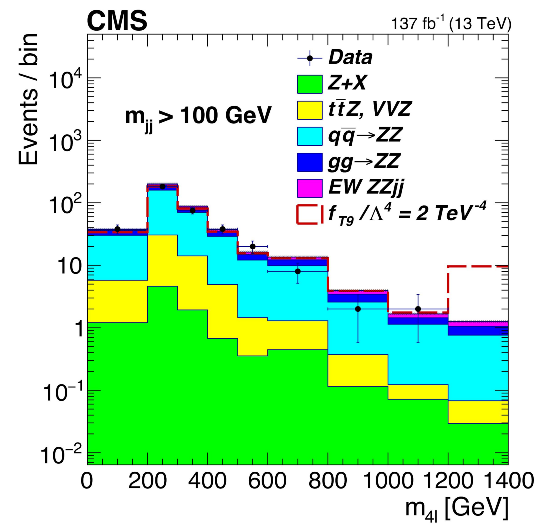
Coupling	Exp. lower	Exp. upper	Obs. lower	Obs. upper	Unitarity bound
f_{T0}/Λ^4	−0.37	0.35	−0.24 (−0.26)	0.22 (0.24)	2.4
f_{T1}/Λ^4	−0.49	0.49	−0.31 (−0.34)	0.31 (0.34)	2.6
f_{T2}/Λ^4	−0.98	0.95	−0.63 (−0.69)	0.59 (0.65)	2.5
f_{T8}/Λ^4	−0.68	0.68	−0.43 (−0.47)	0.43 (0.48)	1.8
f_{T9}/Λ^4	−1.5	1.5	−0.92 (−1.02)	0.92 (1.02)	1.8

expected lower and upper 95% C.L. limits (TeV^{-4}) on coupling of the quartic tensor operators T0-2 and neutral current operators T8 and T9

f_{T8}/Λ^4 and f_{T9}/Λ^4 achieve the world best result



m_{ZZ} distribution for SM and one aQGC scenario



Conclusion

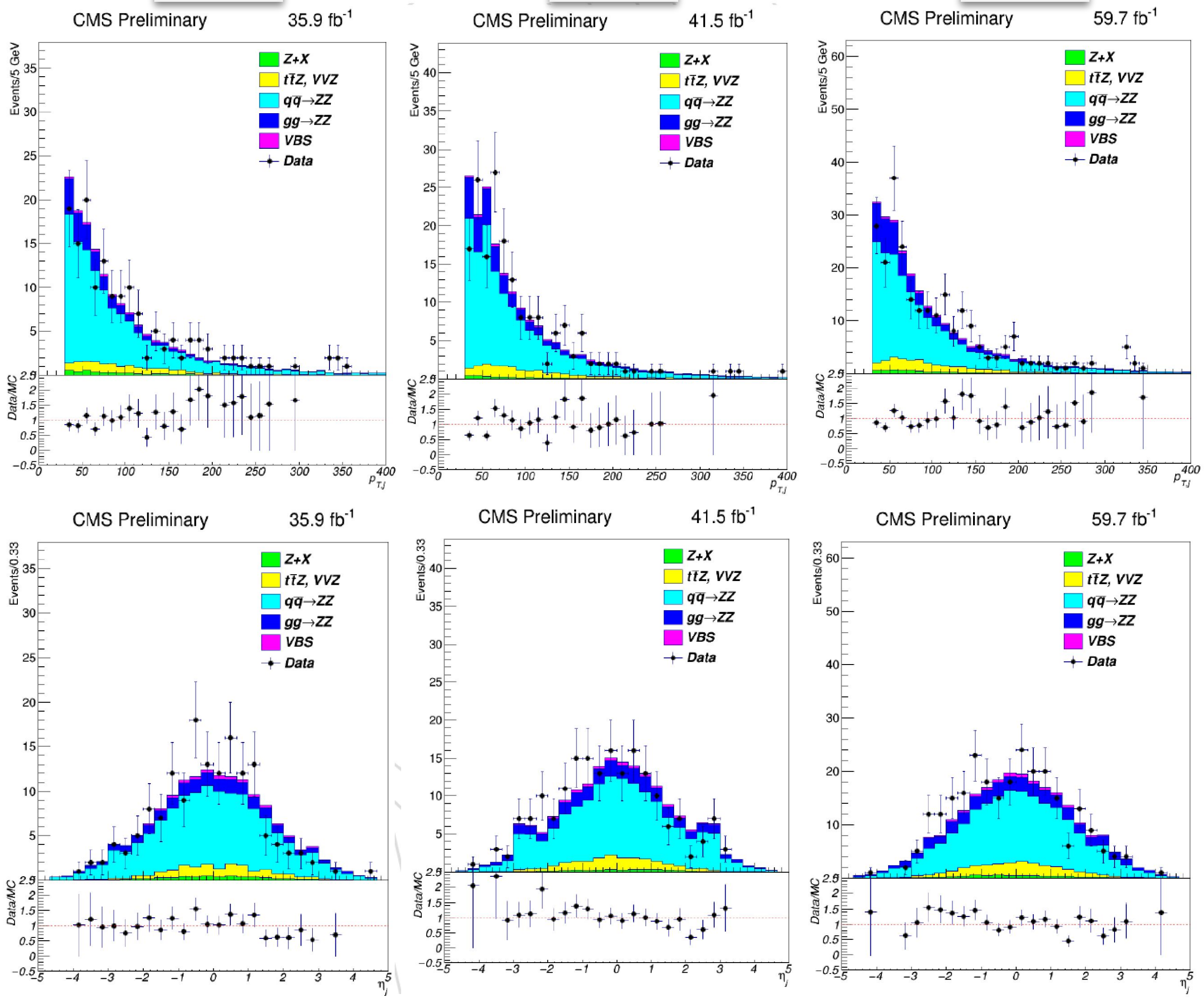
- ➔ VBS ZZ fully leptonic analysis of full Run2 data
  - ◆ data/MC in control region shows good agreement
  - ◆ search for EWK production performed using **MELA discriminant**
  - ◆ main changes w.r.t. 2016 CMS strategy
    - **advanced ggZZ simulation**
    - more accurate treatment of interference and triboson background
    - revision/improvement of some theory/experimental systematic effect
- ➔ Measurement results
  - ◆ **EWK signal significance** ⇒ **4 σ evidence**
  - ◆ **signal strength μ**, **EWK cross section** & **EWK+QCD cross section** for **three fiducial regions**
  - ◆ **limits on dim-8 Wilson coefficients** (aQGCs in EFT approach)

    ⇒ world best limits on f_{T8}/Λ^4 and f_{T9}/Λ^4

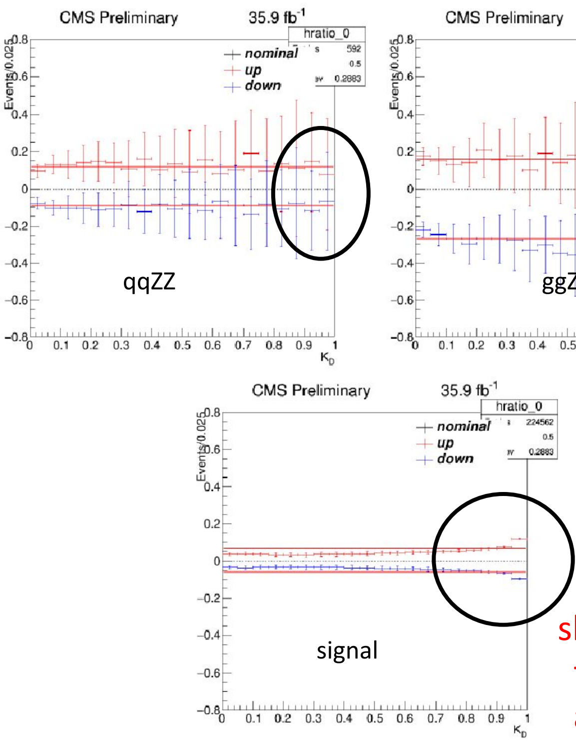
Backup

Data/MC agreement in CR

2016 | 2017 | 2018

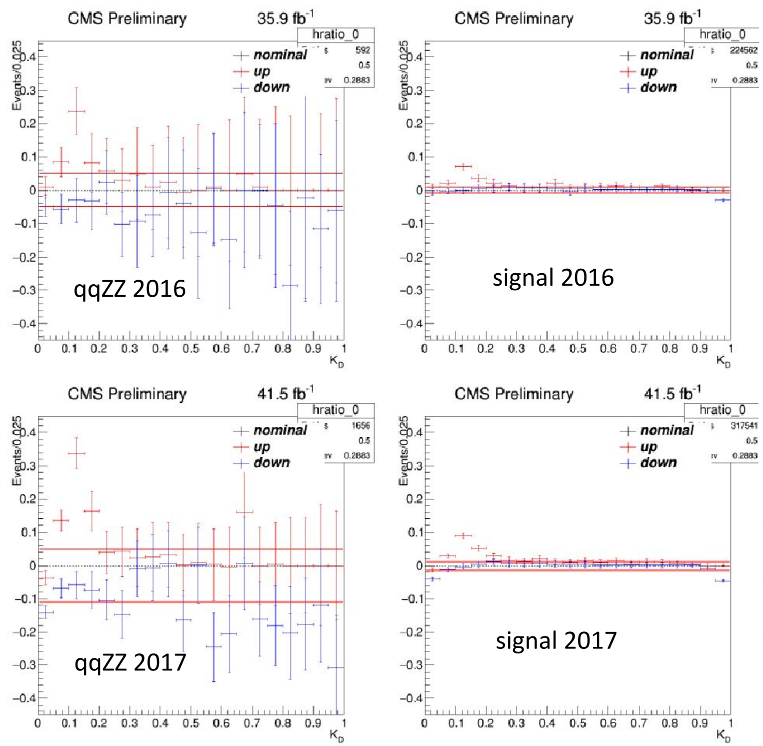


Systematic shape uncertainties

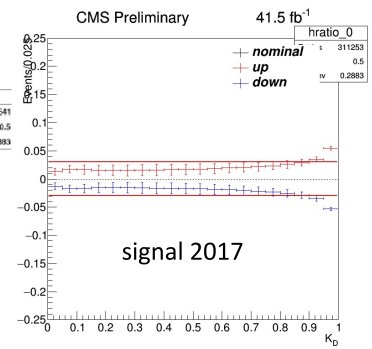


shape used for signal and qqZZ

QCD scale variation



JES variations



L1 prefiring weight variation