ATLAS CMS results in semi-leptonic VBS

G. Boldrini - CNRS/IN2P3 - LLR, École polytechnique

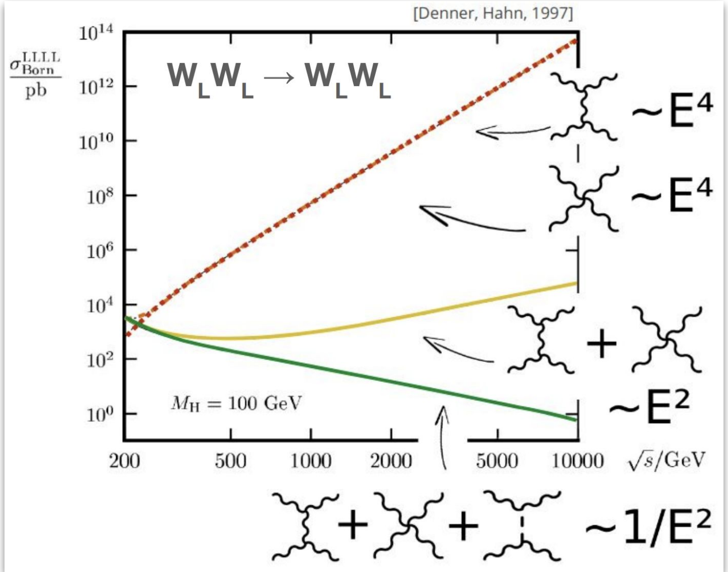
giacomo.boldrini@cern.ch

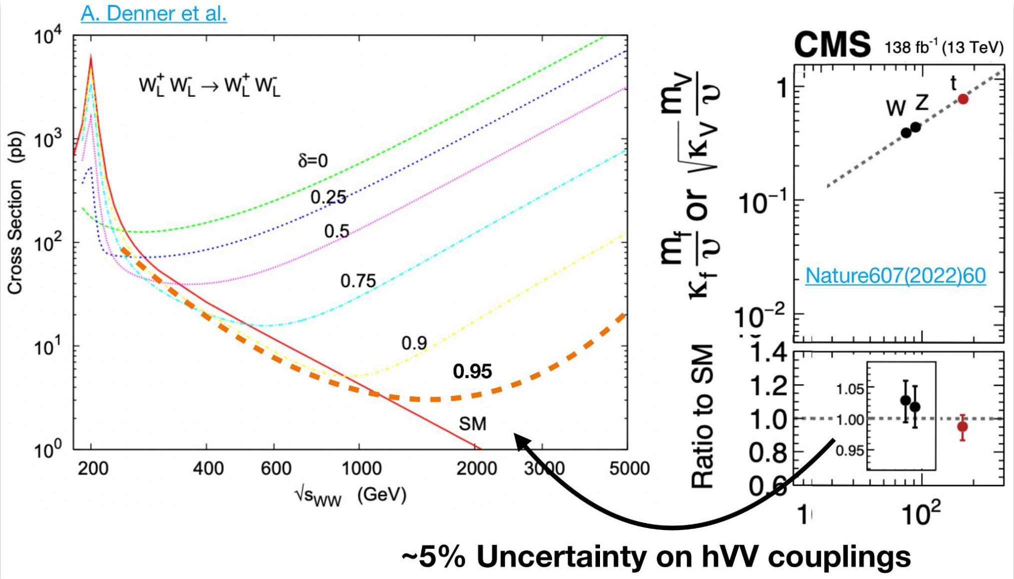
Introduction

VBS is a fundamental probe to understand the electroweak symmetry breaking mechanism

→ The presence of the **Higgs field regularizes the VBS cross-section** by canceling exactly the E^2 behaviour of bosonic-only processes.

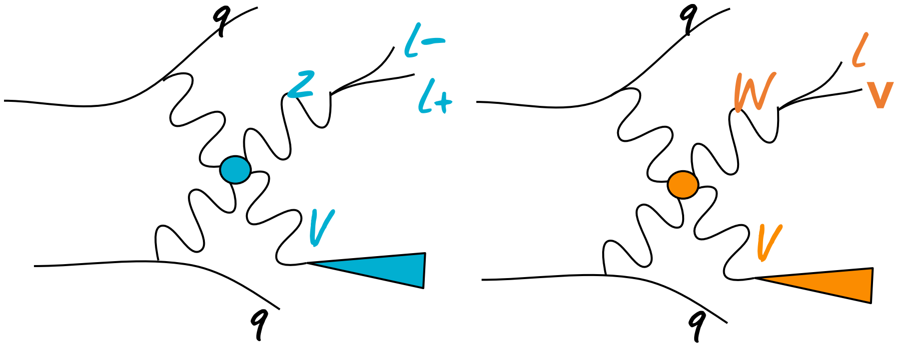
→ **A delicate equilibrium:** if Higgs boson not SM one (δ), energy-growth of $V_L V_L \rightarrow V_L V_L$ cross section hint of new physics


[Denner, Hahn, 1997]


A. Denner et al.

~5% Uncertainty on hVV couplings

Outline



Ongoing activities from ATLAS and CMS on semileptonic VBS signatures

→ VBS $W^{\pm}V$ → lνjj
→ VBS ZV → lljj

Instead of focusing on the specific analysis strategies, the **current limitations** as well as **possible future developments** will be discussed

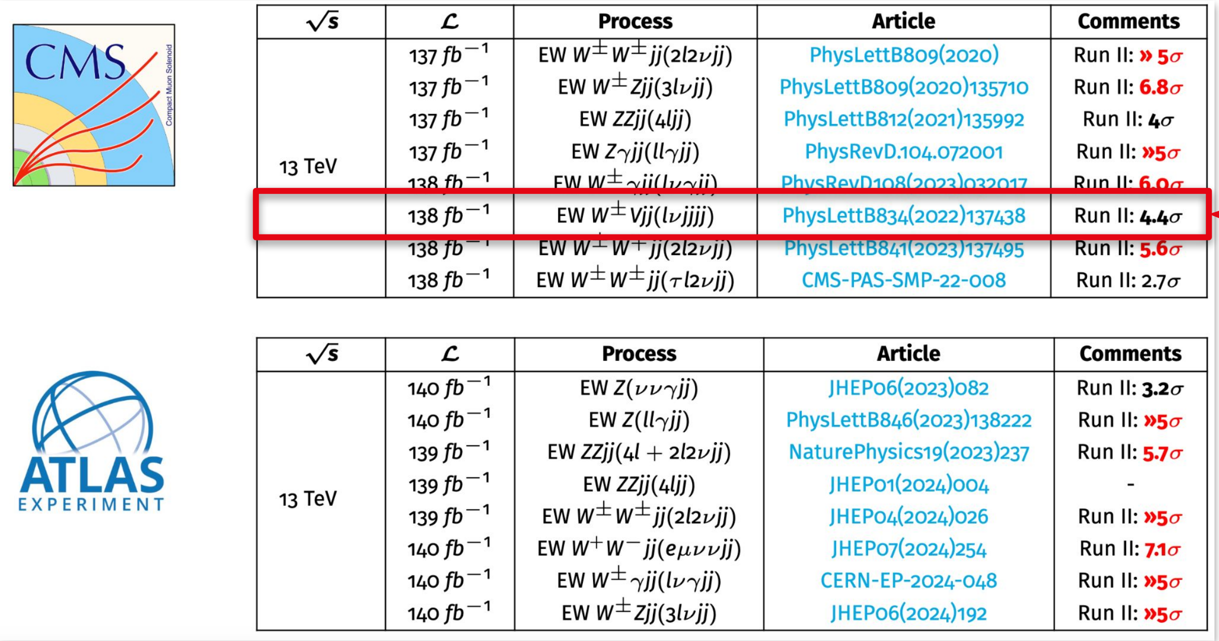
→ Modelling of the VBS signal
→ Modelling of the dominant background sources
→ Semileptonic VBS in a global view

Vector Boson Scattering at the LHC

Vector Boson Scattering (VBS) pivotal measurements in the ATLAS/ CMS EW landscape.

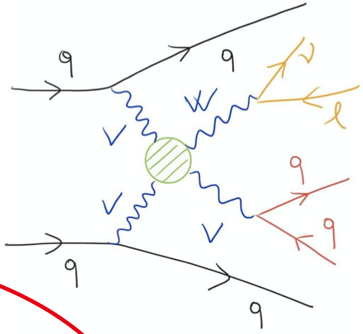
→ First interesting results with 35.9 fb^{-1} from the LHC Run-II (2016)

→ Increasing interest in VBS analyses with complex final states at the LHC thanks to the full Run-II luminosity (138 fb^{-1})



$\sqrt{s}$	$\mathcal{L}$	Process	Article	Comments
13 TeV	$137\ fb^{-1}$	EW $W^{\pm}W^{\pm}jj(2l2\nu jj)$	PhysLettB809(2020)	Run II: **» 5σ**
	$137\ fb^{-1}$	EW $W^{\pm}Zjj(3l\nu jj)$	PhysLettB809(2020)135710	Run II: **6.8σ**
	$137\ fb^{-1}$	EW $ZZjj(4ljj)$	PhysLettB812(2021)135992	Run II: **4σ**
	$137\ fb^{-1}$	EW $Z\gamma jj(ll\gamma jj)$	PhysRevD.104.072001	Run II: **»5σ**
	$138\ fb^{-1}$	EW $W^{\pm}\gamma jj(l\nu\gamma jj)$	PhysRevD108(2023)032017	Run II: **6.0σ**
	$138\ fb^{-1}$	EW $W^{\pm}Vjj(l\nu jjjj)$	PhysLettB834(2022)137438	Run II: **4.4σ**
	$138\ fb^{-1}$	EW $W^{\pm}W^{\mp}jj(2l2\nu jj)$	PhysLettB841(2023)137495	Run II: **5.6σ**
	$138\ fb^{-1}$	EW $W^{\pm}W^{\pm}jj(\tau l2\nu jj)$	CMS-PAS-SMP-22-008	Run II: 2.7σ

$\sqrt{s}$	$\mathcal{L}$	Process	Article	Comments
13 TeV	$140\ fb^{-1}$	EW $Z(\nu\nu\gamma jj)$	JHEP06(2023)082	Run II: **3.2σ**
	$140\ fb^{-1}$	EW $Z(ll\gamma jj)$	PhysLettB846(2023)138222	Run II: **»5σ**
	$139\ fb^{-1}$	EW $ZZjj(4l+2l2\nu jj)$	NaturePhysics19(2023)237	Run II: **5.7σ**
	$139\ fb^{-1}$	EW $ZZjj(4ljj)$	JHEP01(2024)004	-
	$139\ fb^{-1}$	EW $W^{\pm}W^{\pm}jj(2l2\nu jj)$	JHEP04(2024)026	Run II: **»5σ**
	$140\ fb^{-1}$	EW $W^{+}W^{-}jj(e\mu\nu\nu jj)$	JHEP07(2024)254	Run II: **7.1σ**
	$140\ fb^{-1}$	EW $W^{\pm}\gamma jj(l\nu\gamma jj)$	CERN-EP-2024-048	Run II: **»5σ**
	$140\ fb^{-1}$	EW $W^{\pm}Zjj(3l\nu jj)$	JHEP06(2024)192	Run II: **»5σ**

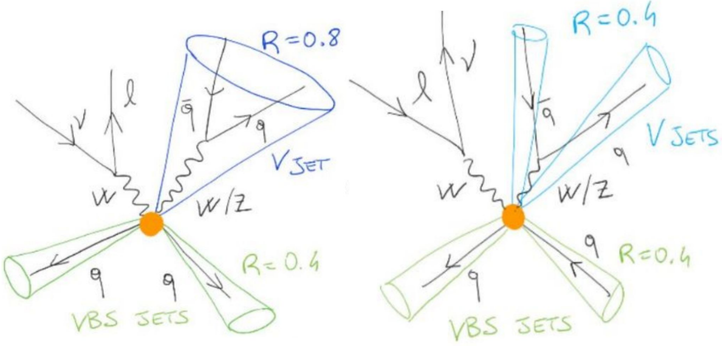


First LHC evidence of a semileptonic VBS process.

CMS Semileptonic VBS W±V → lvjj PhysLettB834(2022)137438

First LHC evidence of a semileptonic VBS process. Final state with 4 jets, one charged lepton + MET. Search for WV VBS where the W± → l±vl and V(W±/Z) → qq

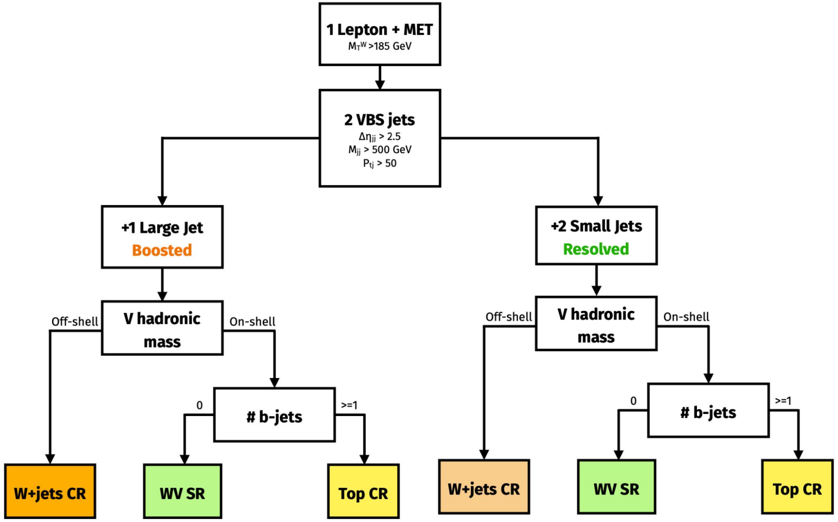
- **Resolved regime**: Four R = 0.4 jets resolved in ΔR
- **Boosted regime**: Two R = 0.4 and one R = 0.8 jets for boosted decays of the V-boson



Backgrounds

- **Dominant W+jets production** → data driven based corrections needed to simulations
- **QCD induced diboson and triboson production**
- **Drell Yan + jets**
- **Semileptonic tt and single top**
- **Non-prompt** mainly from QCD-multijet, data driven estimate

---



1 Lepton + MET
$M_T^W > 185$ GeV

2 VBS jets
$\Delta\eta_{jj} > 2.5$
$M_{jj} > 500$ GeV
$P_{tj} > 50$

+1 Large Jet
Boosted

+2 Small Jets
Resolved

V hadronic mass
Off-shell
On-shell

 # b-jets
0
>=1

W+jets CR
WV SR
Top CR
W+jets CR
WV SR
Top CR

Simultaneous fit of 12 regions

$$\mu_{EW} = 0.85 \pm 0.12(\text{stat})^{+0.19}_{-0.17}(\text{syst}) = 0.85^{+0.23}_{-0.21}$$

$$\mu_{EW+QCD} = 0.97 \pm 0.06(\text{stat})^{+0.19}_{-0.21}(\text{syst}) = 0.97^{+0.20}_{-0.22}$$

CMS Semi-leptonic VBS W±V → lνjj PhysLettB834(2022)137438

Three different measurements from the same data

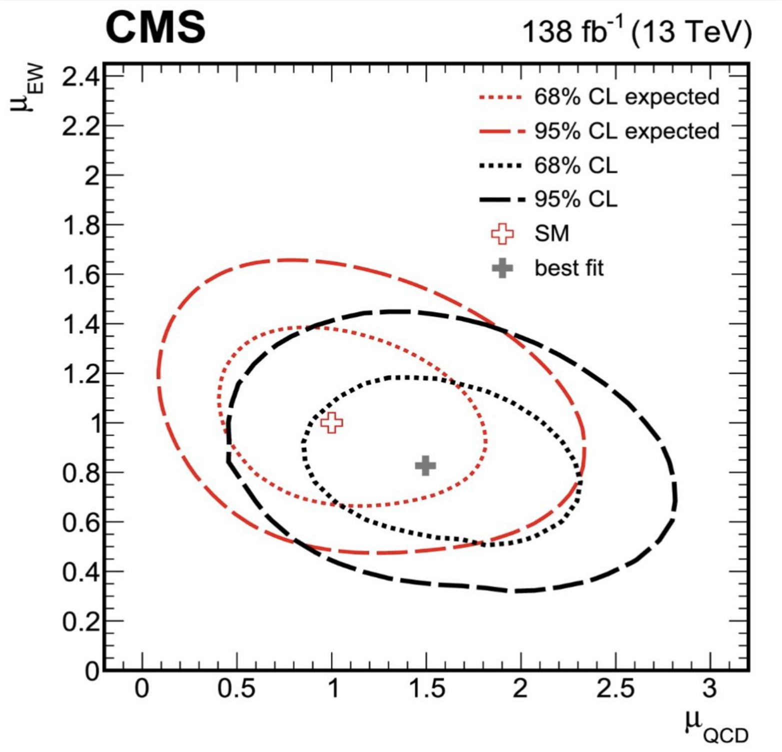
→ SM electroweak signal strength (**4.4σ significance**)

$$\mu_{EW} = 0.85 \pm 0.12(\text{stat})^{+0.19}_{-0.17}(\text{syst}) = 0.85^{+0.23}_{-0.21}$$

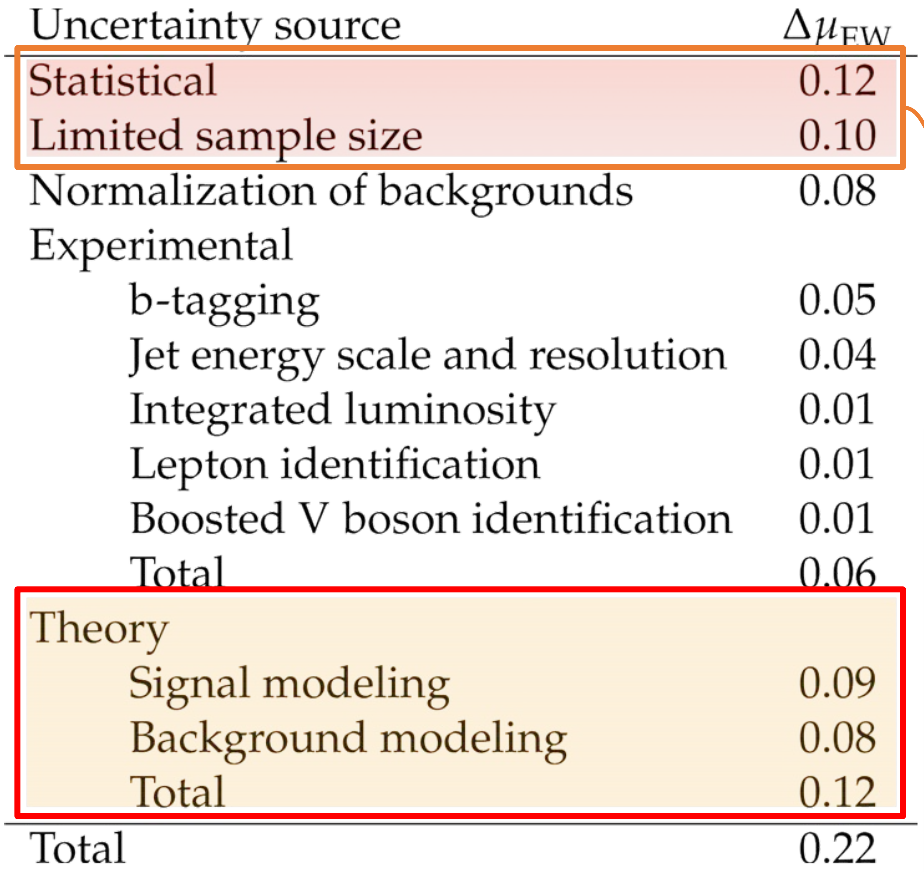
→ Considering **EW and QCD WV production as signal**

$$\mu_{EW+QCD} = 0.97 \pm 0.06(\text{stat})^{+0.19}_{-0.21}(\text{syst}) = 0.97^{+0.20}_{-0.22}$$

→ **Simultaneous 2D fit** of the EW and QCD WV signal strengths



VBS W±V → lvjj - Limitations statistics

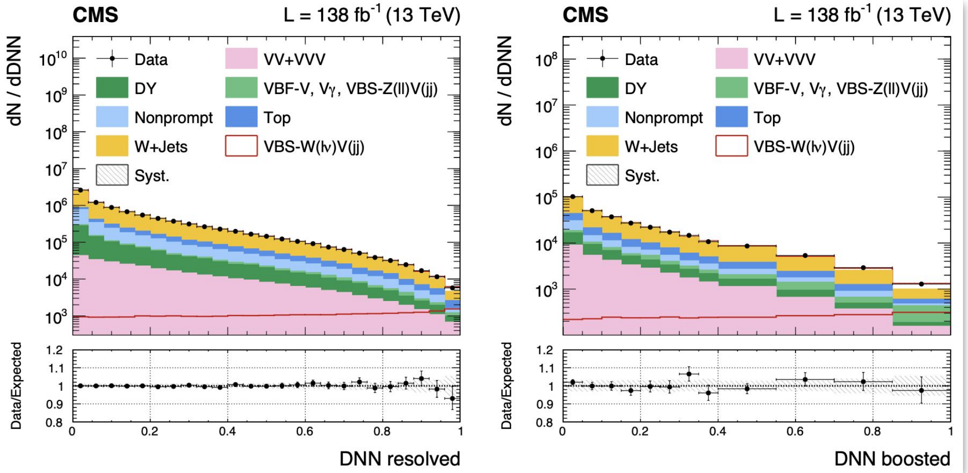


Uncertainty source	$\Delta\mu_{EW}$
Statistical	0.12
Limited sample size	0.10
Normalization of backgrounds	0.08
Experimental	
b-tagging	0.05
Jet energy scale and resolution	0.04
Integrated luminosity	0.01
Lepton identification	0.01
Boosted V boson identification	0.01
Total	0.06
Theory	
Signal modeling	0.09
Background modeling	0.08
Total	0.12
Total	0.22

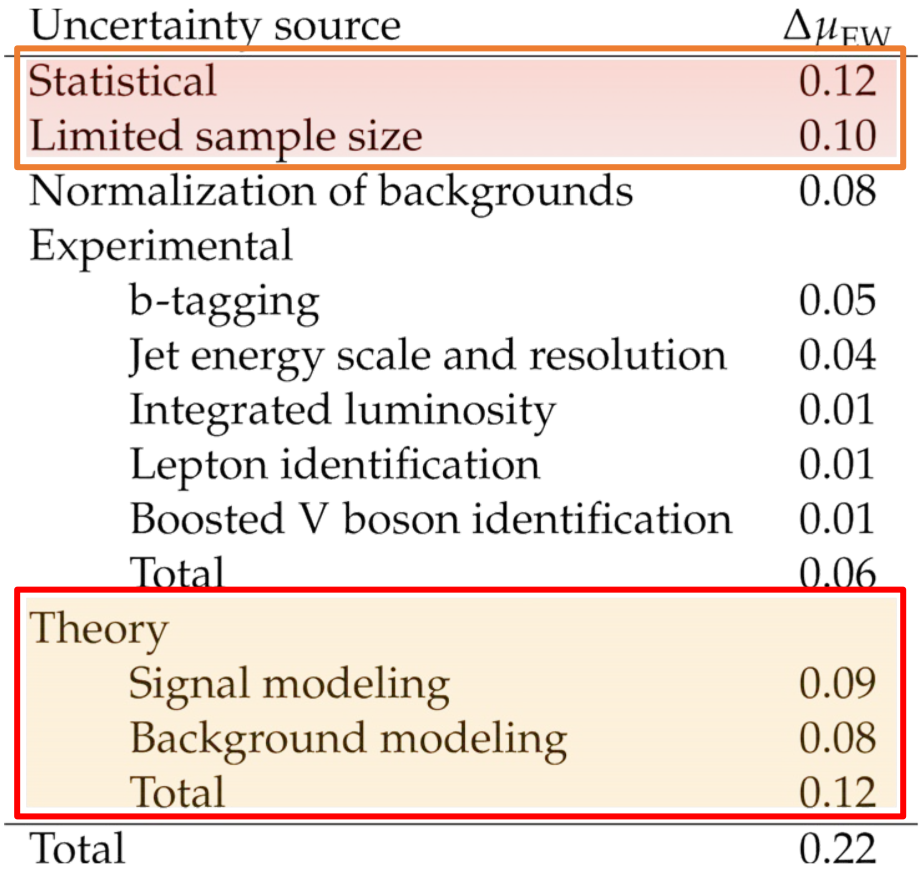
VBS EW WV sensitivity enhanced with DNN models

Limited data and Monte Carlo in tails of DNN spectra lead to important uncertainties

→ These uncertainties **will scale with** the increasing integrated **luminosity** (e.g. LHC Run-III or HL-LHC)

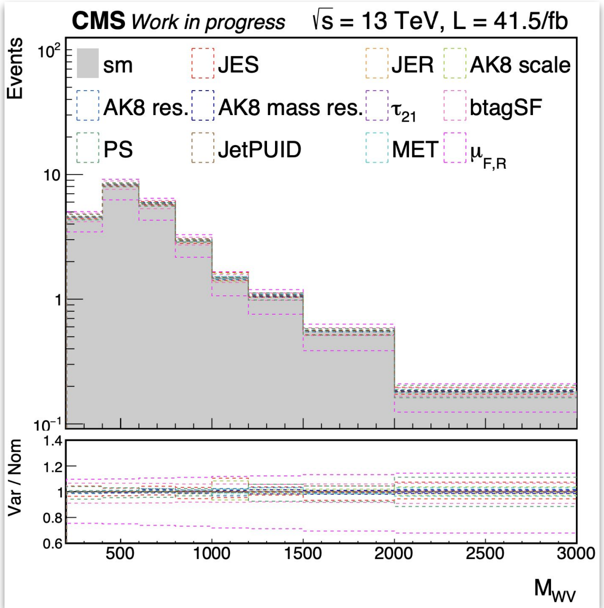


VBS W±V → lvjj - Limitations signal uncertainties



Uncertainty source	$\Delta\mu_{EW}$
Statistical	0.12
Limited sample size	0.10
Normalization of backgrounds	0.08
Experimental	
b-tagging	0.05
Jet energy scale and resolution	0.04
Integrated luminosity	0.01
Lepton identification	0.01
Boosted V boson identification	0.01
Total	0.06
Theory	
Signal modeling	0.09
Background modeling	0.08
Total	0.12
Total	0.22

*endorsed result from CMS-TS-2024-002



Theory uncertainties dominant both on VBS EW signal and on backgrounds

→ VBS WV signal simulated at **LO in QCD** (MadGraph v2.6.5)**: large dependence on $\mu_{R,F}$**

→ **No improvement with luminosity**

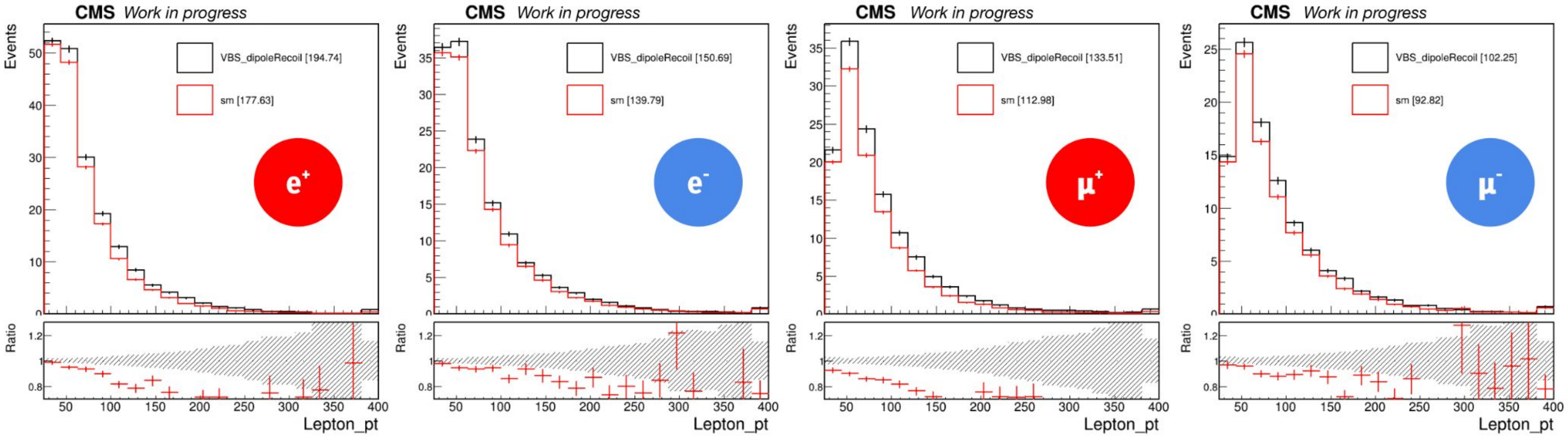
VBS W±V → lvjj - Limitations signal modelling

LO signal modelling sensitive to process definition: 2→4 + NWA used for public result or **i**nclusive **2→6**. **Large shape and normalization disagreement observed in SRs between the two generations**: **2→6 cross-section ~15% smaller** compatible with the observed signal strength of 0.85 using the 2→4 sample

🟥 2→6 VBS EW sample

⬛ 2→4 VBS EW sample

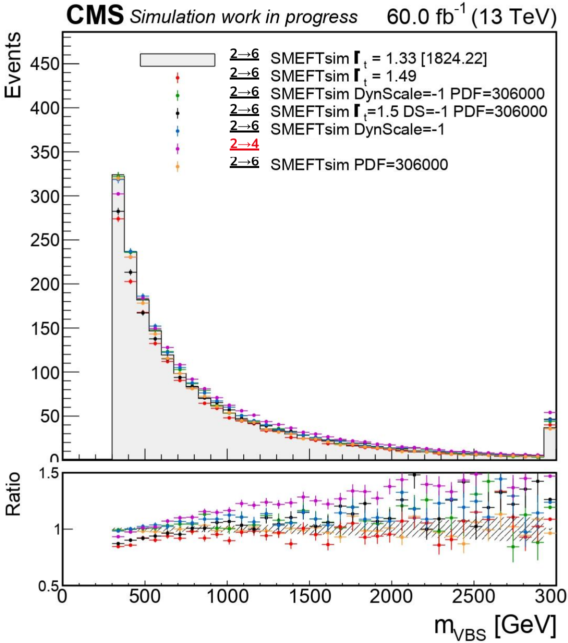
*endorsed result from CMS-TS-2024-002

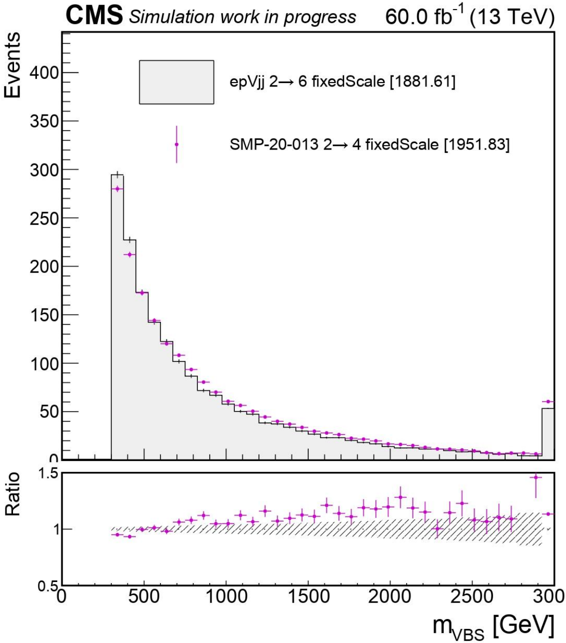


VBS $W^{\pm}V \rightarrow l\nu jj$ - Limitations signal modelling

Origin of the disagreement: dynamical computation of hard scattering scale differs between 2→6 and 2→4, covered by $\mu_{F,R}$

Fixing the hard scattering scale to a constant value (m_Z ~ 91 GeV) removes the energy-growth behaviour but the ~15% normalization disagreement remains suggesting that **a 2→6 modelling LO@QCD of the signal might be in better agreement with observations.**





VBS W±V → lvjj - Limitations background modelling

VBS W±V → lvjj present an harsh multijet background.

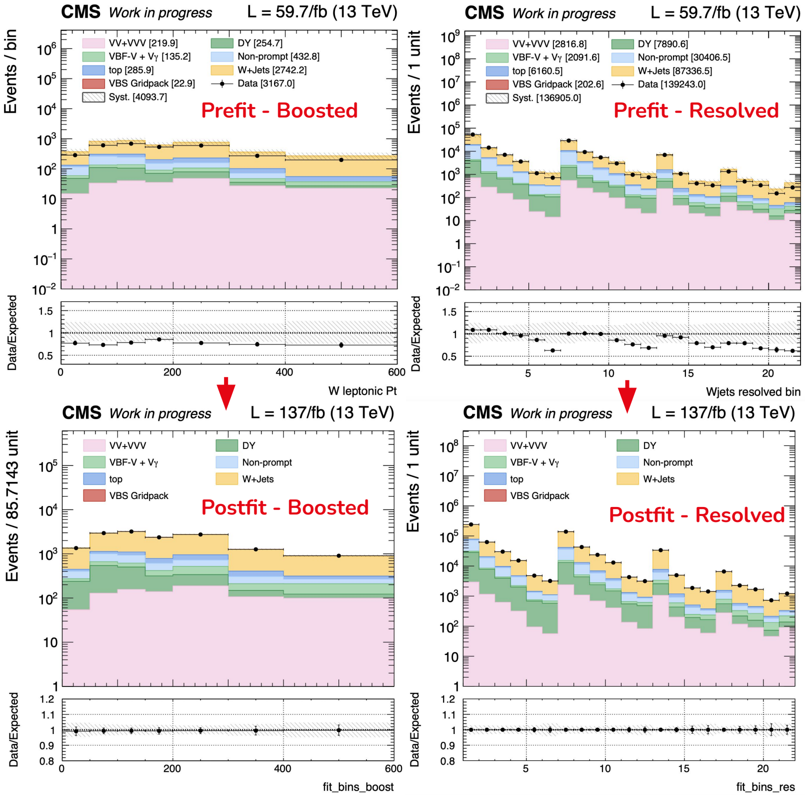
Main source from W+jets production

→ **W+Jets MC:** Madgraph+Pythia8 samples, HT binned up to 4 partons LO@QCD. **NLO W+Jets** (up to 2 partons at NLO@QCD) brings no improvement (except in $\Delta\eta^{jj}$) + larger stat. uncertainty due to negative weights

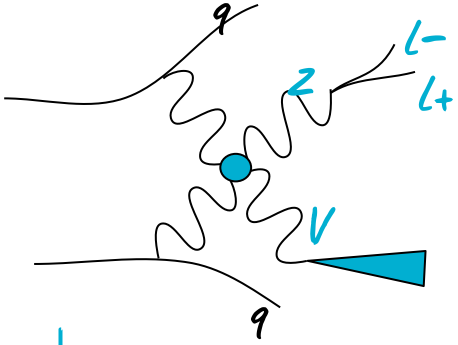
→ **Data-driven correction to the W+Jets LO sample** in both boosted and resolved categories: measure W+jets normalization in bins of p_T^W and subleading VBS jet p_T from a dedicated control region included in the global fit

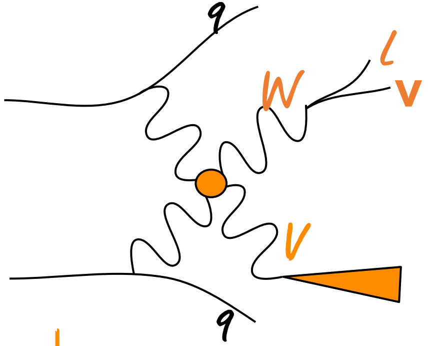
Trading off systematic uncertainties for a better background description → Need for precise background predictions

Each bin has a freely floating normalization parameter



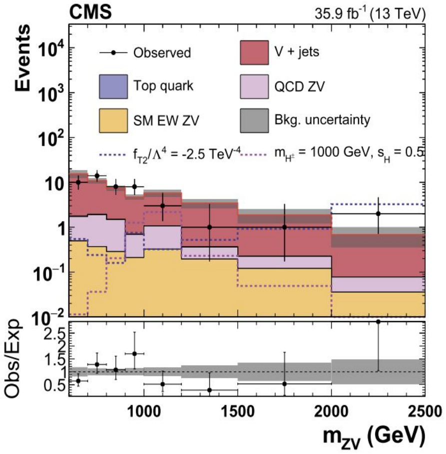
CMS Semileptonic VBS ZV → 2l2j

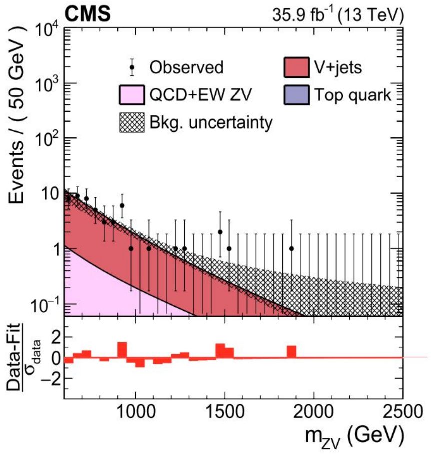




Run-II ongoing

Run-II published





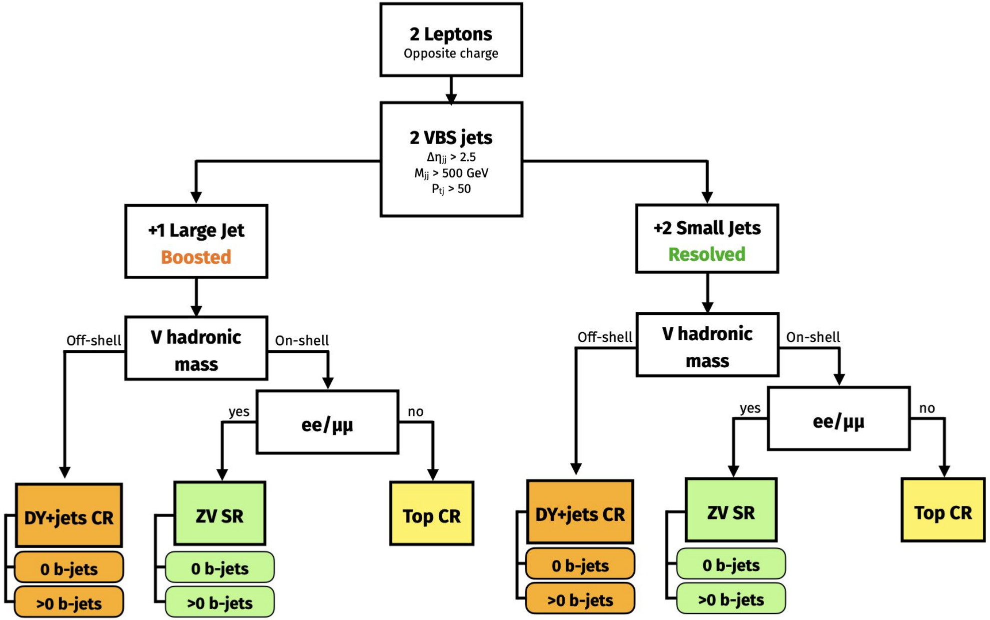
PhysicsLettersB798(2019)134985

While the VBS WV analysis is published, **CMS is also studying the VBS ZV counterpart with full Run-II luminosity**. A preliminary result with 35.9 fb^{-1} is public with focus on aQGCs.

VBS-ZV and VBS-WV present a very similar topology and **face the same problems regarding background and signal modelling**

The VBS-ZV is disadvantaged by the Z→2l branching ratio, leading to a lower S/B ratio

CMS Semileptonic VBS ZV → 2l2j



2 Leptons
Opposite charge

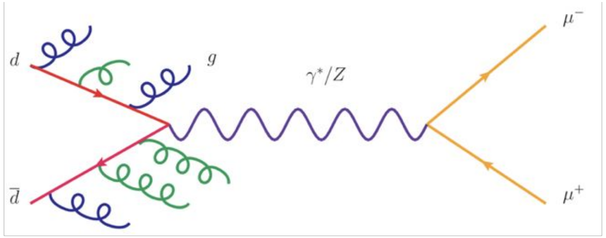
2 VBS jets
$\Delta\eta_{jj} > 2.5$
$M_{jj} > 500$ GeV
$P_{tj} > 50$

+1 Large Jet
Boosted

V hadronic mass
Off-shell
On-shell
ee/μμ
yes
no
DY+jets CR
0 b-jets
>0 b-jets
ZV SR
0 b-jets
>0 b-jets
Top CR

+2 Small Jets
Resolved

V hadronic mass
Off-shell
On-shell
ee/μμ
yes
no
DY+jets CR
0 b-jets
>0 b-jets
ZV SR
0 b-jets
>0 b-jets
Top CR



$\mathcal{O}(\alpha_{EW}^2\alpha_S^4)$ DY+jets

Similar analysis with respect to VBS WV Final state with 4 jets, 2 charged leptons

- **Resolved regime**: Four R = 0.4 jets resolved in ΔR
- **Boosted regime**: Two R = 0.4 and one R = 0.8 jets for boosted decays of the V-boson

Backgrounds

- **Dominant DY+jets production** → data driven based corrections needed to simulations
- **QCD induced diboson and triboson production**
- **Semileptonic tt and single top** → normalization floating in top enriched control regions

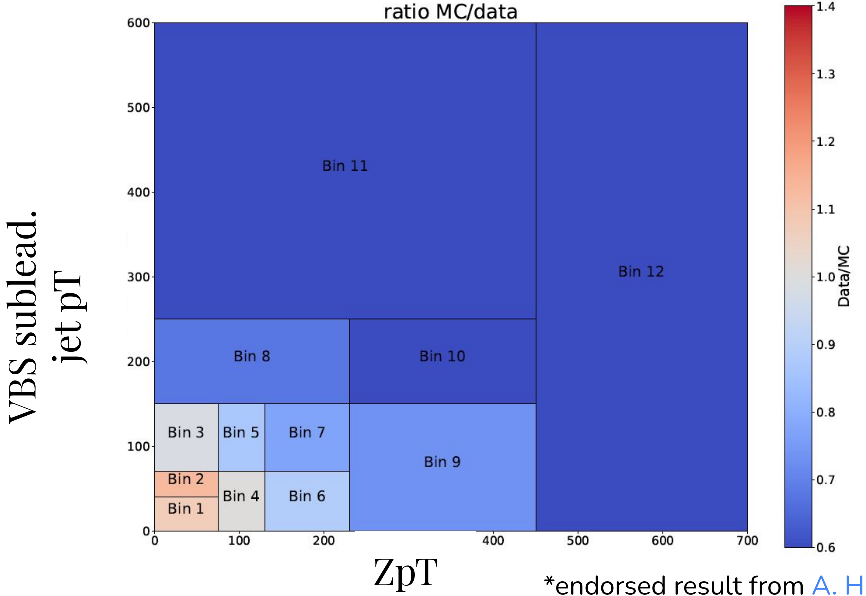
*endorsed result from A. Hakimi

VERY PRELIMINARY		BOOSTED	RESOLVED	Combined
RUN 2	b-veto	0.71	1.12	1.28
	b-tag	0.84	0.84	1.12
	combined	1.02	1.35	**1.8**

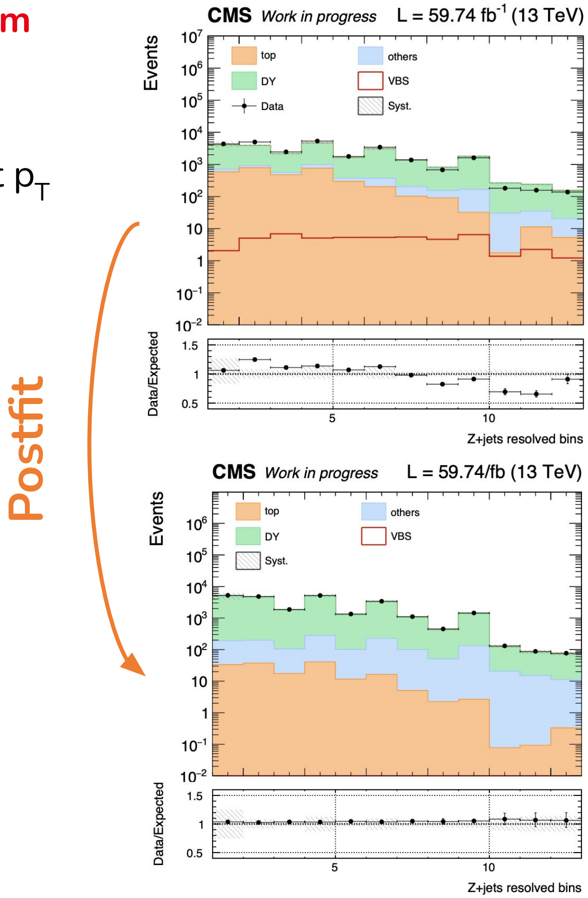
VBS ZV → 2l2j - Limitations background modelling

VBS ZV → 2ljj present an harsh multijet background. **Main source from Z+jets production**

→ **Z+Jets MC:** Madgraph+Pythia8 samples, HT binned up to 4 partons LO@QCD. **Data-driven correction** in bins of p_T^Z and subleading VBS jet p_T from a dedicated control region included in the global fit



*endorsed result from A. Hakimi

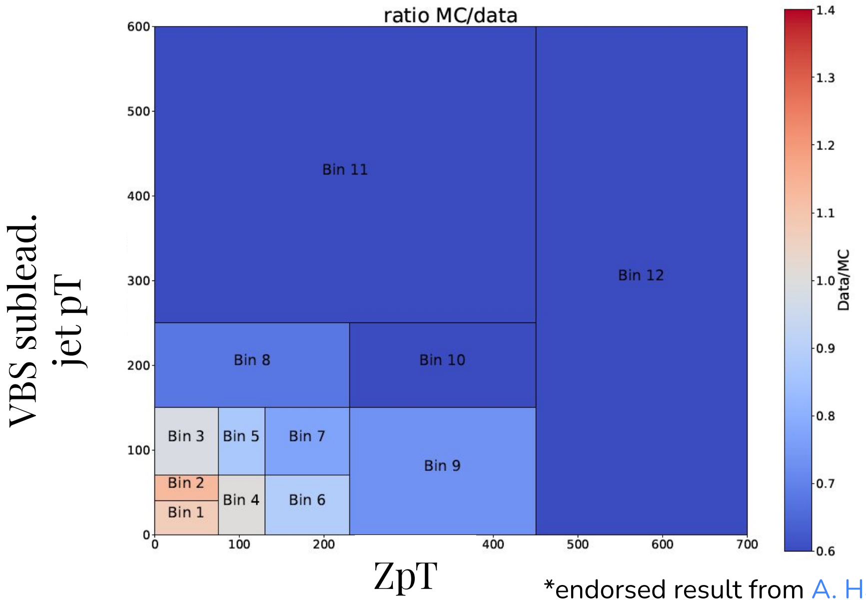


Resolved

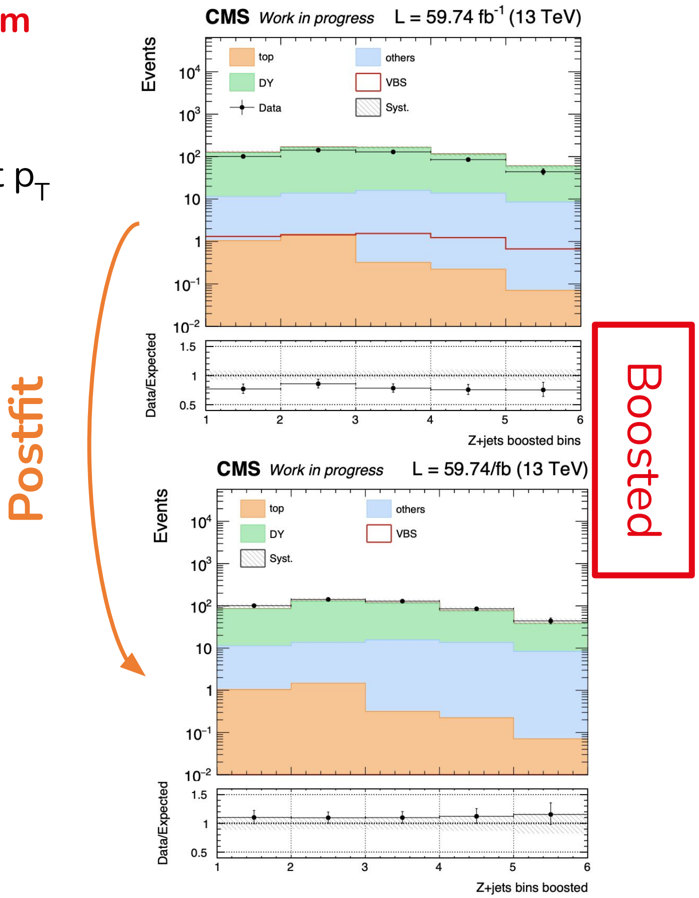
VBS ZV → 2l2j - Limitations background modelling

VBS ZV → 2ljj present an harsh multijet background. Main source from Z+jets production

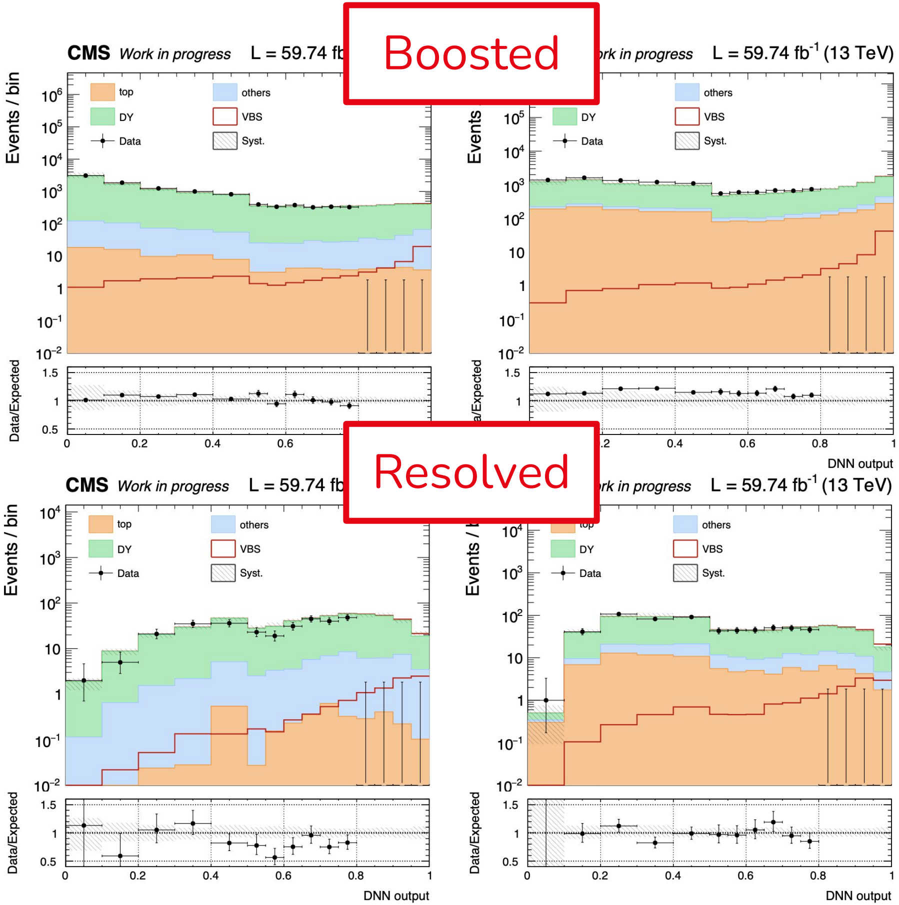
→ **Z+Jets MC:** Madgraph+Pythia8 samples, HT binned up to 4 partons LO@QCD. **Data-driven correction** in bins of p_T^Z and subleading VBS jet p_T from a dedicated control region included in the global fit



*endorsed result from A. Hakimi



VBS ZV → 2l2j - Limitations background modelling


Boosted

Resolved

As for VBS-WV, **multiple DNN models in boosted / resolved regions are used to extract the VBS-ZV signal**

→ Measurement dominated by **statistical uncertainty**, **DY-correction**, **background** and **signal modelling** (QCDscales DY, QCD-VV, VBF-V).

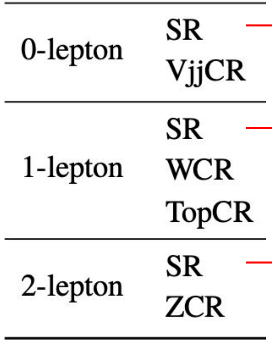
→ Room for improvement with LHC Run III and more accurate theory predictions

Source	**Uncertainty**
Statistical	+0.417 -0.410
Z+jets correction	+0.327 -0.312
Theoretical	+0.272 -0.243
Simulation statistics	+0.174 -0.164
Experimental	+0.199 -0.154
Total	**+0.652 - 0.612**

*endorsed result from A. Hakimi

ATLAS Semileptonic VBS VV PhysRevD.100.032007

ATLAS has only one semileptonic public result with limited Run II dataset (35.5 fb^{-1}). Full Run II analysis is ongoing. **Simultaneous measurement of WW/WZ/ZZ VBS where one vector boson decays hadronically**

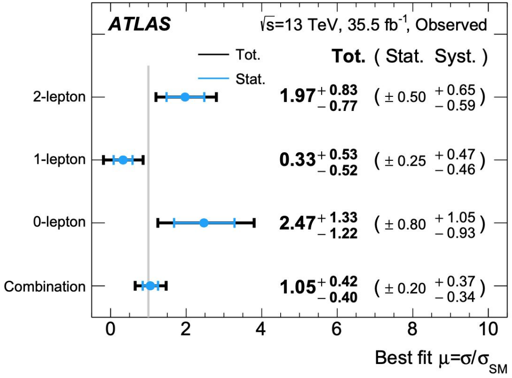


0-lepton	SR VjjCR
1-lepton	SR WCR TopCR
2-lepton	SR ZCR

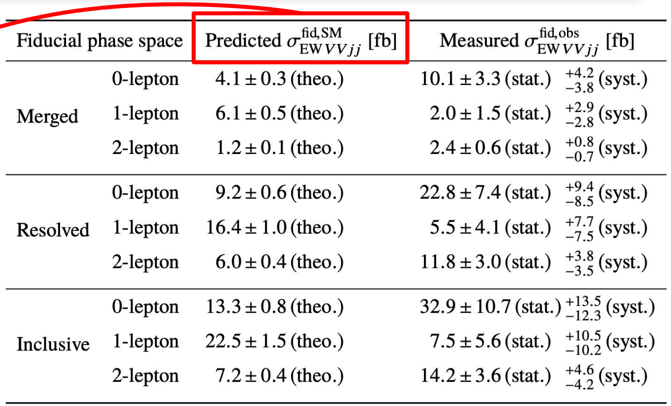
ZZ → νν2j

W±V → lνjj

ZV → 2l2j

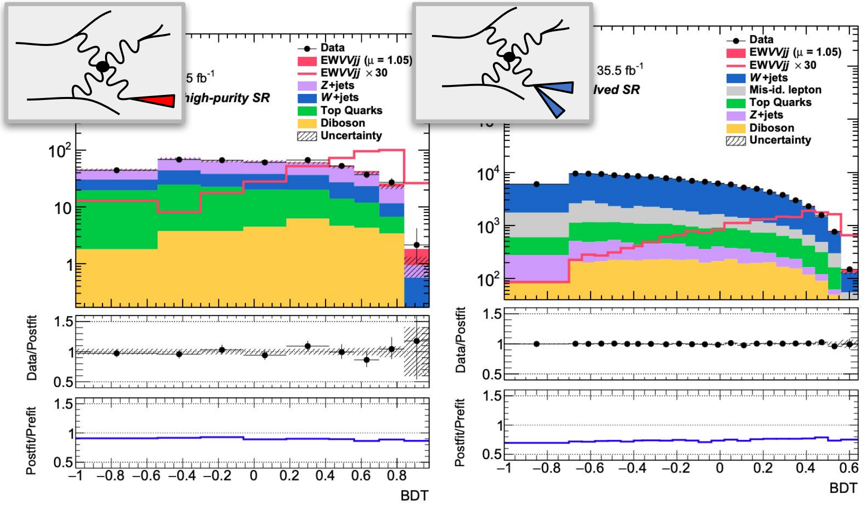


As CMS analysis, **boosted and resolved regimes** analyzed, **BDT used to extract the signal** in each regime

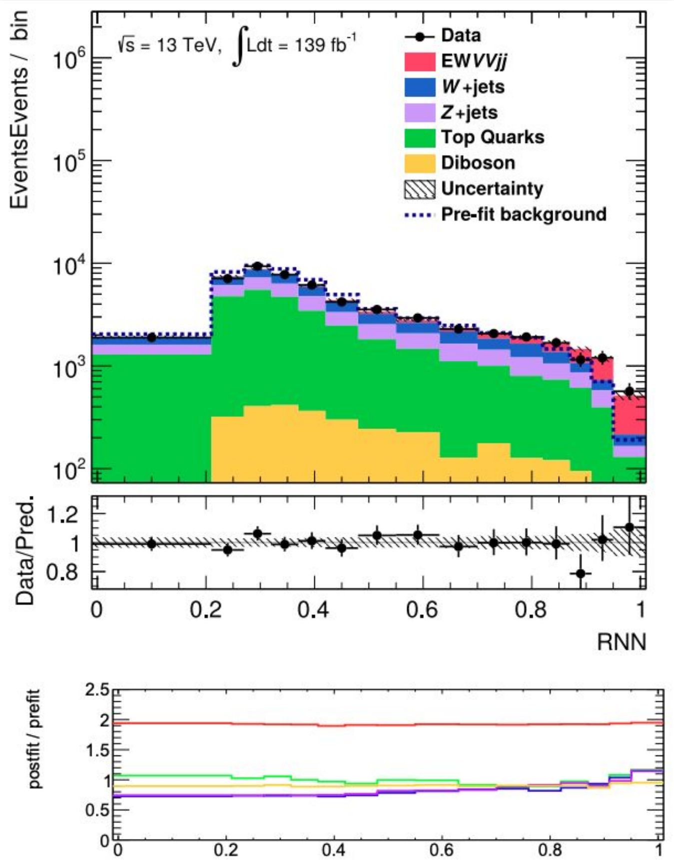


Fiducial phase space		Predicted $\sigma^{\text{fid,SM}}_{\text{EW}\,VVjj}$ [fb]	Measured $\sigma^{\text{fid,obs}}_{\text{EW}\,VVjj}$ [fb]
Merged	0-lepton	4.1 ± 0.3 (theo.)	10.1 ± 3.3 (stat.) $^{+4.2}_{-3.8}$ (syst.)
	1-lepton	6.1 ± 0.5 (theo.)	2.0 ± 1.5 (stat.) $^{+2.9}_{-2.8}$ (syst.)
	2-lepton	1.2 ± 0.1 (theo.)	2.4 ± 0.6 (stat.) $^{+0.8}_{-0.7}$ (syst.)
Resolved	0-lepton	9.2 ± 0.6 (theo.)	22.8 ± 7.4 (stat.) $^{+9.4}_{-8.5}$ (syst.)
	1-lepton	16.4 ± 1.0 (theo.)	5.5 ± 4.1 (stat.) $^{+7.7}_{-7.5}$ (syst.)
	2-lepton	6.0 ± 0.4 (theo.)	11.8 ± 3.0 (stat.) $^{+3.8}_{-3.5}$ (syst.)
Inclusive	0-lepton	13.3 ± 0.8 (theo.)	32.9 ± 10.7 (stat.) $^{+13.5}_{-12.3}$ (syst.)
	1-lepton	22.5 ± 1.5 (theo.)	7.5 ± 5.6 (stat.) $^{+10.5}_{-10.2}$ (syst.)
	2-lepton	7.2 ± 0.4 (theo.)	14.2 ± 3.6 (stat.) $^{+4.6}_{-4.2}$ (syst.)

MadGraph LO@QCD

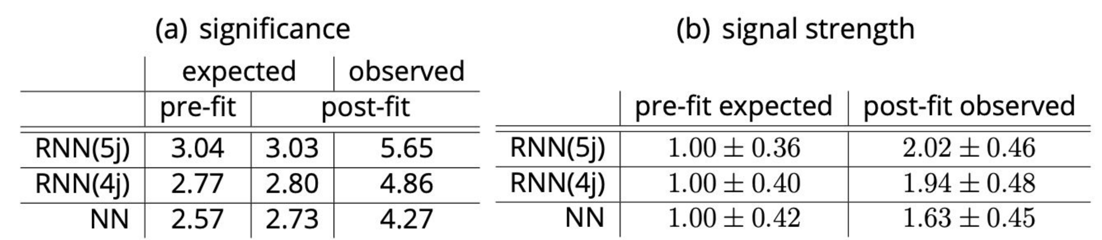


ATLAS Semileptonic VBS VV Run II



Promising preliminary results toward the analysis of the full Run II dataset from ATLAS [CERN-THESIS-2022-165]

→ **Novel RNN methods** exploited to extract the signal showing **more discriminative** power with respect to traditional **NN**.
Attention to susceptibility to modelling uncertainties in the training

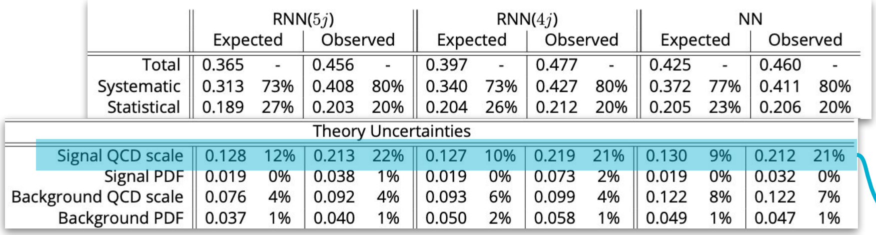


(a) significance

	expected		observed
	pre-fit	post-fit	
RNN(5j)	3.04	3.03	5.65
RNN(4j)	2.77	2.80	4.86
NN	2.57	2.73	4.27

(b) signal strength

	pre-fit expected	post-fit observed
RNN(5j)	1.00 ± 0.36	2.02 ± 0.46
RNN(4j)	1.00 ± 0.40	1.94 ± 0.48
NN	1.00 ± 0.42	1.63 ± 0.45



	RNN($5j$)				RNN($4j$)				NN			
	Expected		Observed		Expected		Observed		Expected		Observed	
Total	0.365	-	0.456	-	0.397	-	0.477	-	0.425	-	0.460	-
Systematic	0.313	73%	0.408	80%	0.340	73%	0.427	80%	0.372	77%	0.411	80%
Statistical	0.189	27%	0.203	20%	0.204	26%	0.212	20%	0.205	23%	0.206	20%
Theory Uncertainties												
Signal QCD scale	0.128	12%	0.213	22%	0.127	10%	0.219	21%	0.130	9%	0.212	21%
Signal PDF	0.019	0%	0.038	1%	0.019	0%	0.073	2%	0.019	0%	0.032	0%
Background QCD scale	0.076	4%	0.092	4%	0.093	6%	0.099	4%	0.122	8%	0.122	7%
Background PDF	0.037	1%	0.040	1%	0.050	2%	0.058	1%	0.049	1%	0.047	1%

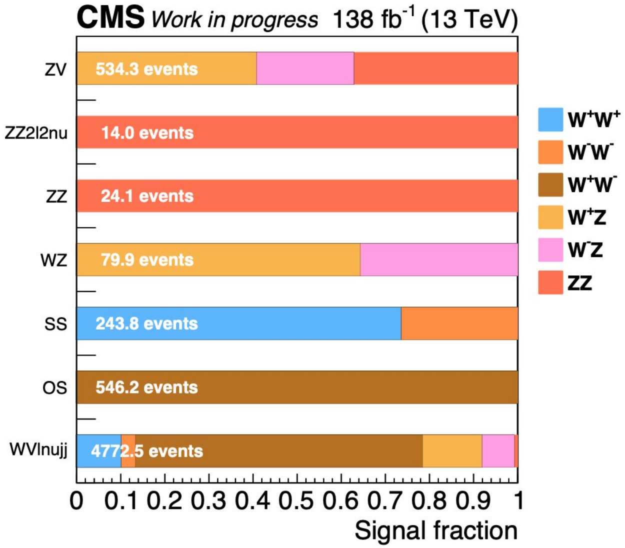
QCDscales dominant syst. uncertainty

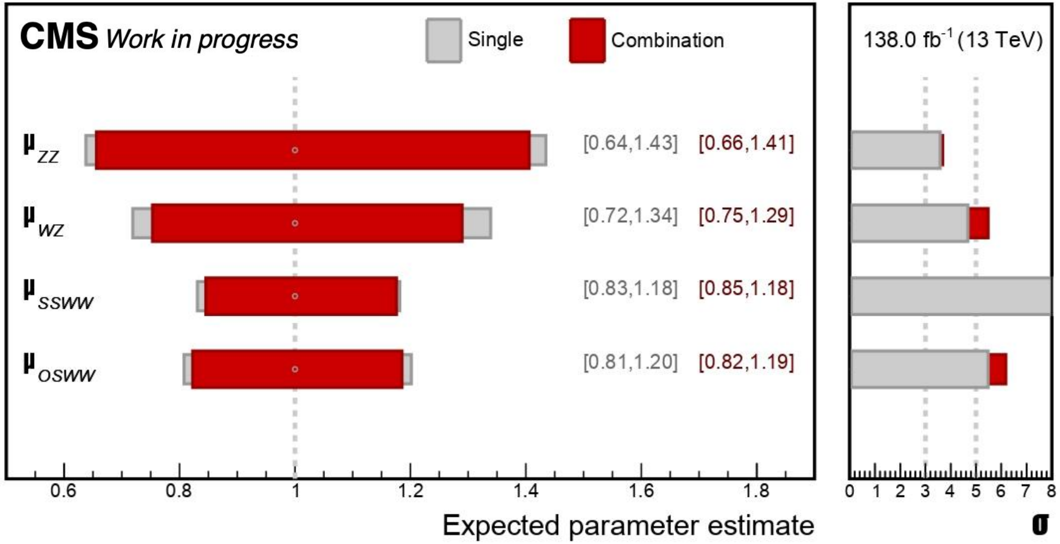
Semileptonic processes in the global context

The **measurement of polarized vector boson scattering is a long term goal** of both ATLAS and CMS collaborations → promising result for HL-LHC. **If we want to obtain insights into the EW sector before we need a statistical combination of different VBS channels** (Run-II + Run-III). Semileptonic ZV and WV channels act as a link between fully-leptonic (and more pure) channels

*endorsed result from CMS-TS-2024-002

The combined measurement of VBS signal strength is a first step toward a combined measurement of polarization fractions → Semileptonic VBS can play an important role

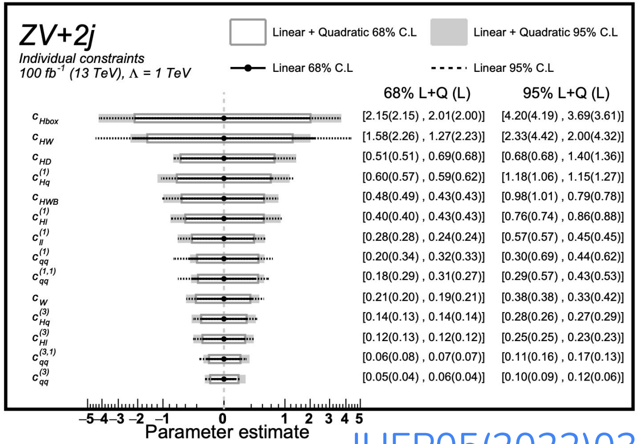




Semileptonic processes in the global context

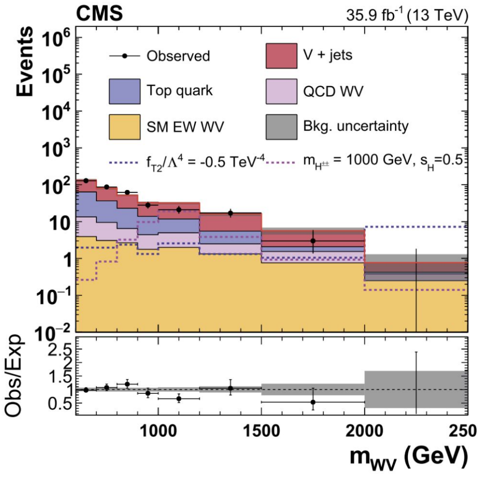
Semileptonic VBS channels lead the sensitivity to **dimension-8 operators / aQGC** + increasing interest in EFT at dimension 6→ LO predictions might overshoot in high-mass tails, NLO corrections needed for a reliable inference

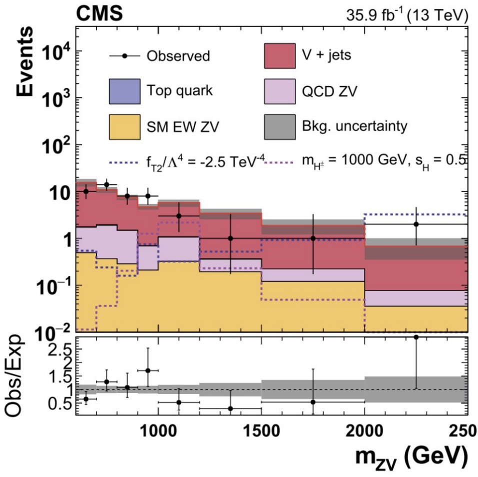
→ **Run II results ongoing from CMS**



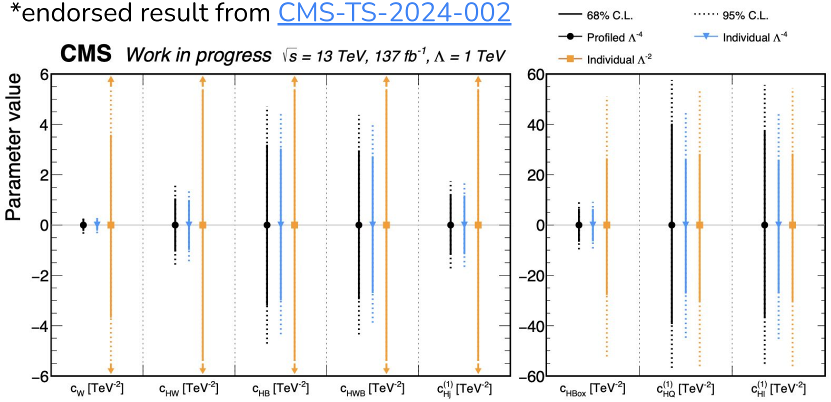
	68% L+Q (L)	95% L+Q (L)
c_{Hbox}	[2.15(2.15) , 2.01(2.00)]	[4.20(4.19) , 3.69(3.61)]
c_{HW}	[1.58(2.26) , 1.27(2.23)]	[2.33(4.42) , 2.00(4.32)]
c_{HD}	[0.51(0.51) , 0.69(0.68)]	[0.68(0.68) , 1.40(1.36)]
$c_{Hq}^{(1)}$	[0.60(0.57) , 0.59(0.62)]	[1.18(1.06) , 1.15(1.27)]
c_{HWB}	[0.48(0.49) , 0.43(0.43)]	[0.98(1.01) , 0.79(0.78)]
$c_{Hl}^{(1)}$	[0.40(0.40) , 0.43(0.43)]	[0.76(0.74) , 0.86(0.88)]
$c_{ll}^{(1)}$	[0.28(0.28) , 0.24(0.24)]	[0.57(0.57) , 0.45(0.45)]
$c_{qq}^{(1)}$	[0.20(0.34) , 0.32(0.33)]	[0.30(0.69) , 0.44(0.62)]
$c_{qq}^{(1,1)}$	[0.18(0.29) , 0.31(0.27)]	[0.29(0.57) , 0.43(0.53)]
c_{W}	[0.21(0.20) , 0.19(0.21)]	[0.38(0.38) , 0.33(0.42)]
$c_{Hq}^{(3)}$	[0.14(0.13) , 0.14(0.14)]	[0.28(0.26) , 0.27(0.29)]
$c_{Hl}^{(3)}$	[0.12(0.13) , 0.12(0.12)]	[0.25(0.25) , 0.23(0.23)]
$c_{qq}^{(3,1)}$	[0.06(0.08) , 0.07(0.07)]	[0.11(0.16) , 0.17(0.13)]
$c_{qq}^{(3)}$	[0.05(0.04) , 0.06(0.04)]	[0.10(0.09) , 0.12(0.06)]

JHEP05(2022)039





*endorsed result from CMS-TS-2024-002



Conclusions and future perspectives

The **interest in Semileptonic VBS signatures is growing in both ATLAS and CMS** collaborations **thanks to the Run-II integrated luminosity**: the first evidence for VBS WV is a promising first step toward a complete understanding of the VBS process.

However the analyses are complex
→ Tiny signal overwhelmed by multijet backgrounds
→ Use of advanced Machine Learning techniques to extract the signals
→ Difficult modelling of the EW signal and large uncertainties limit the sensitivity
→ Multijet background hard to model with MC, data driven techniques limit the sensitivity
→ Limited statistical power can be cured by analyzing Run III data

Semileptonic VBS signatures play an important role in the global picture
→ In global combinations, they act as a link between pure fully leptonic channels
→ Vector boson polarization taggers interesting tool for polarized scattering
→ Boosted regimes sensitive to BSM physics at dimension-8 or even 6