First observation of $VBS \rightarrow W^+W^-$

Flavia Cetorelli[*,1], Mattia Lizzo[*,2]
on behalf of the CMS collaboration

[1] Università & INFN, Milano-Bicocca (IT)
[2] Università & INFN, Firenze (IT)
*Primary authors

20/01/2022

Introduction

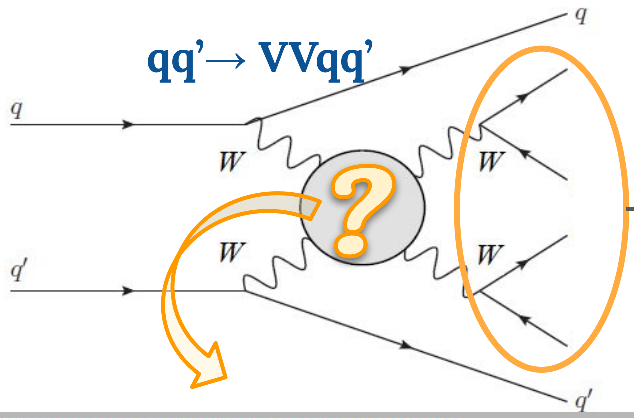
***"First observation of the electroweak production of a leptonically decaying W+W- pair in association with two jets in √s = 13 TeV pp collisions"*. CDS: 2791336**

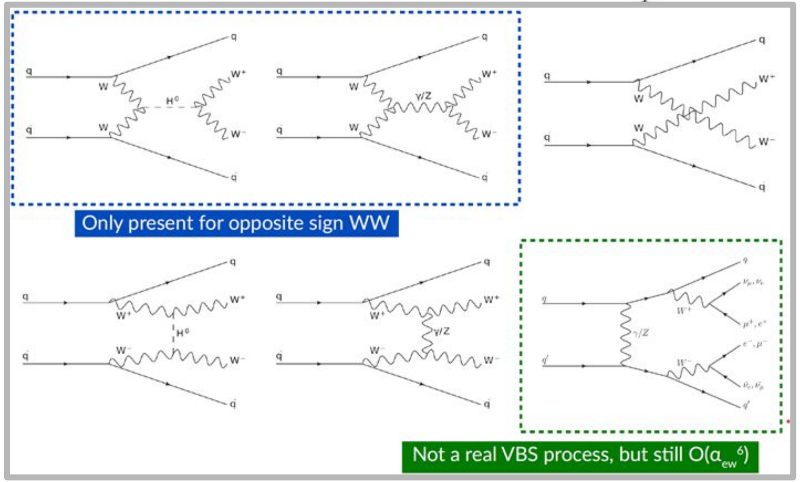
Why VBS → W+W-?

- ★ A **rare** process:
  - **not observed** yet;
  - **first cross section measurement** in this channel.
- ★ Provides complementary information to the **Higgs sector** and probes the **EWSSB** mechanism
- ★ The **first VBS observation** has been performed in the $\mathbf{W^{\pm}W^{\pm}}$ channel which has a good signal-to-background ratio → the $\mathbf{W^{+}W^{-}}$ channel is **as much interesting** from a theoretically point of view, but it is **more experimentally challenging**

Here we present **the first observation** of such mechanism, using the full Run 2 CMS data set that corresponds to an integrated luminosity of $\mathbf{138\ fb^{-1}}$.

VBS processes @LHC





The two massive bosons may decay hadronically or leptonically, leading to **3** possible **final states**:

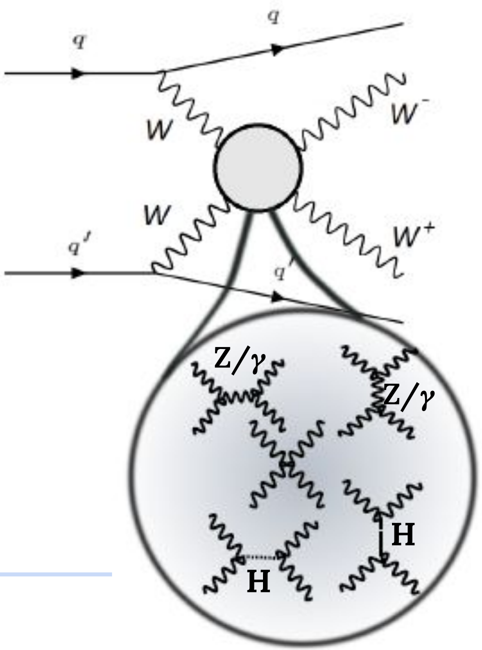
1. **leptonic VV→ lν lν** (this talk)
2. **semileptonic VV → lν q'q"**
3. **hadronic VV → qq' q"q'"**

+purity

+statistics

- ★ Production of a pair of W+W- bosons from a purely **electroweak process** @LO O($\alpha^6_{EWK}\ \alpha^0_S$)
- ★ Diagrams where an on-shell Higgs boson is exchanged (**VBF**) are considered as **backgrounds** and modeled with dedicated MC samples
- ★ The **interference** with the **QCD-induced WW** background O($\alpha^4_{EWK}\ \alpha^2_S$) is **negligible**

VBS → W^+W^-



★ A **rare** … :
- **~ fb cross section**

★ … **and challenging** process:
- **ttbar + tW** production;
- **DY** events (mostly in same flavour final state).

Typical leptonic signature

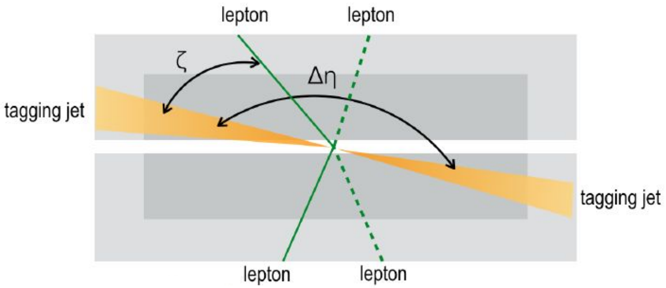
★ 2 high energy jet (tag jets) :
- great gap in **η and** high **m_{jj}**;
- **no QCD** activity between them.

★ 2 charged leptons and p_T^{miss}:
- **central** with respect to the tag jet.

m_{jj}, $\Delta\eta_{jj}$

$$z_\ell = \eta_\ell - \frac{1}{2}(\eta_{j_1} + \eta_{j_2})$$

$$z_{\ell\ell} = \frac{1}{2}|z_{\ell_1} + z_{\ell_2}|$$



Main backgrounds

ttbar - tW:

σ_{tt} ~ 1 nb

Main background of the analysis.

Strategy: b jets veto

QCD WW:

Strong interaction between the initial state quarks.

Same final state, but different kinematic.

Strategy: VBS selections.

Drell-Yan:

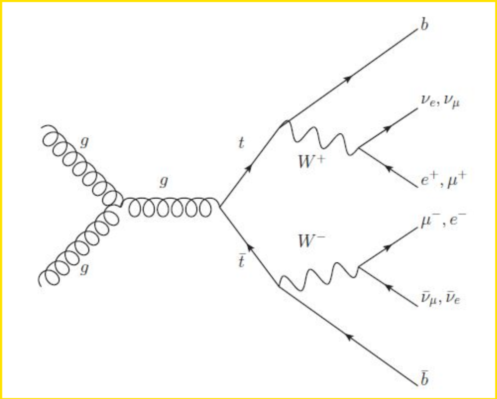
Mainly affecting the ee-μμ final state.

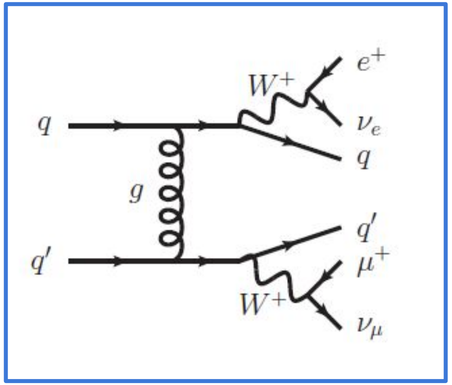
Strategy: selections on the invariant mass and pT of lepton pair, tighter selections on p_T^{miss} .

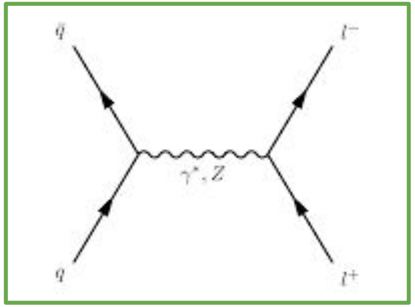
Nonprompt

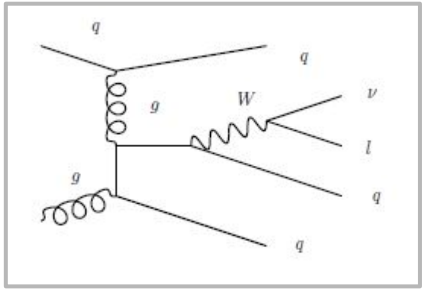
Mainly W+Jet :

jet misidentified as lepton **(fake lepton)**







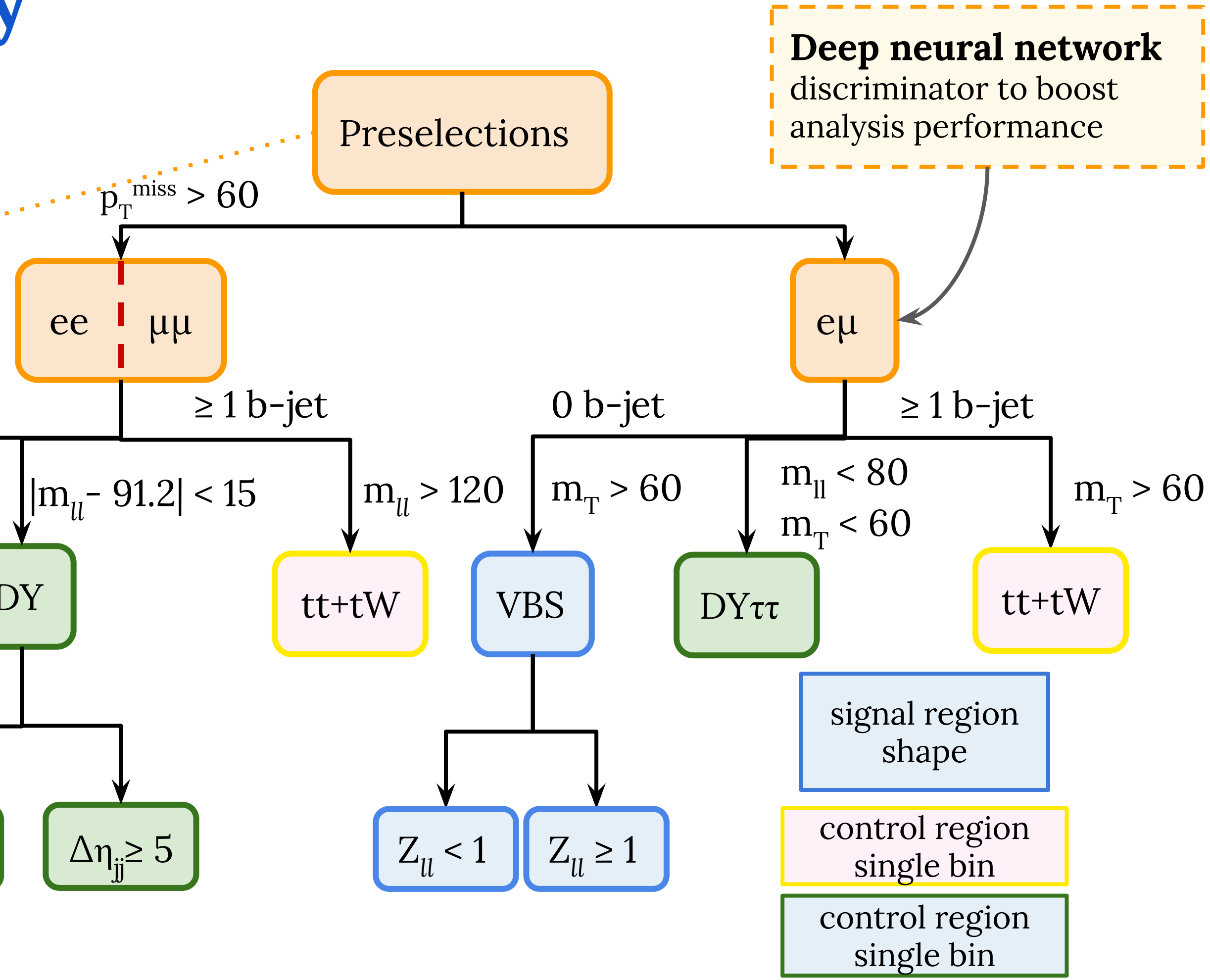


MC samples

Process	MC generator	QCD order + PS [PYTHIA 8]
VBS $W^+W^- \to 2\ell 2\nu$	MadGraph5_aMC@NLO v2.6.5 (2016) v2.4.2 (2017/2018)	2 jets @LO
$W^+W^- \to 2\ell 2\nu$ QCD-induced	Powheg v2 + MiNLO	inclusive NNLO precision
$W^+W^- \to 2\ell 2\nu$ gg-induced	MCFM v7.0	NLO (k-factor)
ttbar + tW $\to 2\ell 2\nu$	Powheg v2	NNLO (ttbar) / NLO (tW)
DY $\to \ell\ell$ + jets	MadGraph5_aMC@NLO	NLO FxFx
DY $\to \tau\tau$ + jets	MadGraph5_aMC@NLO	NLO FxFx
Z + 2 jets (EWK)	MadGraph5_aMC@NLO	LO
$H \to W^+W^- \to 2\ell 2\nu$	Powheg v2 + JHUGen v7.1.4	NLO
$V\gamma^*$ (V = W,Z)	MadGraph5_aMC@NLO	LO MLM ($W\gamma^*$) + NLO FxFx ($Z\gamma^*$)
$WZ \to 3\ell\nu$	Powheg v2	NLO ($m_{\ell\ell}$ > 100 MeV)
ZZ, WZ$\to 2\ell 2q$, VVV (V = W,Z)	MadGraph5_aMC@NLO	NLO

Analysis strategy

- ★ $p_T^{l1} > 25$ GeV && $p_T^{l2} > 13$ GeV
- ★ $p_T^{l3} < 10$ GeV
- ★ $m_{ll} > 50$ GeV && $p_T^{ll} > 30$ GeV
- ★ $p_T^{miss} > 20$ GeV
- ★ At least 2 jets with $p_T > 30$ GeV
- ★ $m_{jj} > 300$ GeV
- ★ $\Delta\eta_{jj} > 2.5$

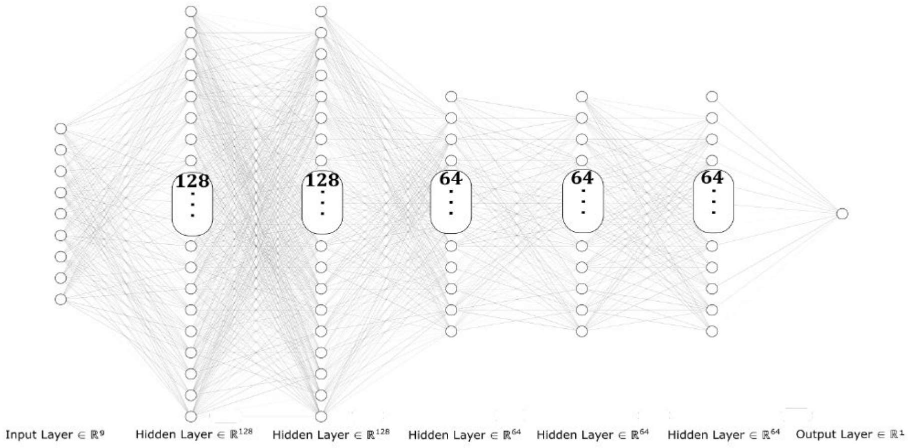


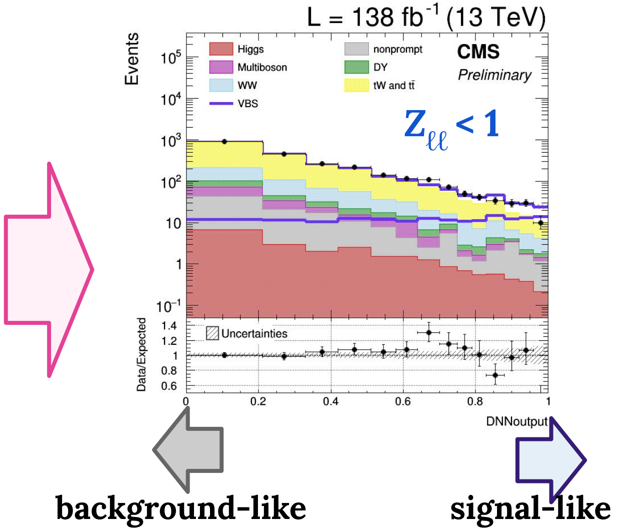
DNN

Variable	Description
m_{jj}	Invariant mass of the two VBS jets pair
$\Delta\eta_{jj}$	Pseudorapidity gap between the two VBS jets
p_{Tj_1}	p_T of the highest-p_T jet
p_{Tj_2}	p_T of the second highest p_T jet
$p_{T\ell\ell}$	p_T of the lepton pair
$\Delta\phi_{\ell\ell}$	Azimuthal angle between the two leptons
Z_{ℓ_1}	Zeppenfeld variable of the highest-p_T lepton
Z_{ℓ_2}	Zeppenfeld variable of the second highest p_T lepton
m_{TW_1}	Transverse mass of the $(p_{T\ell_1}, p_T^{miss})$ system

Deep neural network to disentangle **signal** from top and QCD-WW background:

- ★ different flavour final state (**eμ**);
- ★ **2 models** implemented:
  - ○ $Z_{ll} < 1$ phase space;
  - ○ $Z_{ll} \geq 1$ phase space.





DY→ℓℓ treatment

In **DY** CRs **two different contributions** are clearly visible and much sensitive to the $\Delta\eta_{jj}$ distribution itself:

- ★ **“hard”** DY events (**dark green**) populate the low $\Delta\eta_{jj}$ region, prefit data/MC discrepancy;
- ★ DY process + **at least 1 PU jet** (**light green**) peaks around $\Delta\eta_{jj} \sim 5$

⇒ Their normalisations are free to float in the fit and mainly driven by **dedicated CRs ($\Delta\eta_{jj} \gtrsim 5$)**

PU jet = reco jet not matched to any $p_T > 25$ GeV gen jet

Drell-Yan:

eμ:
- **DY→ ττ**

ee-μμ:
- **DY + no PU jets** — **CR: $\Delta\eta_{jj} < 5$**
- **DY + ≥ 1 PU jet** — **CR: $\Delta\eta_{jj} > 5$**

Three different contributions with different control regions

Signal extraction

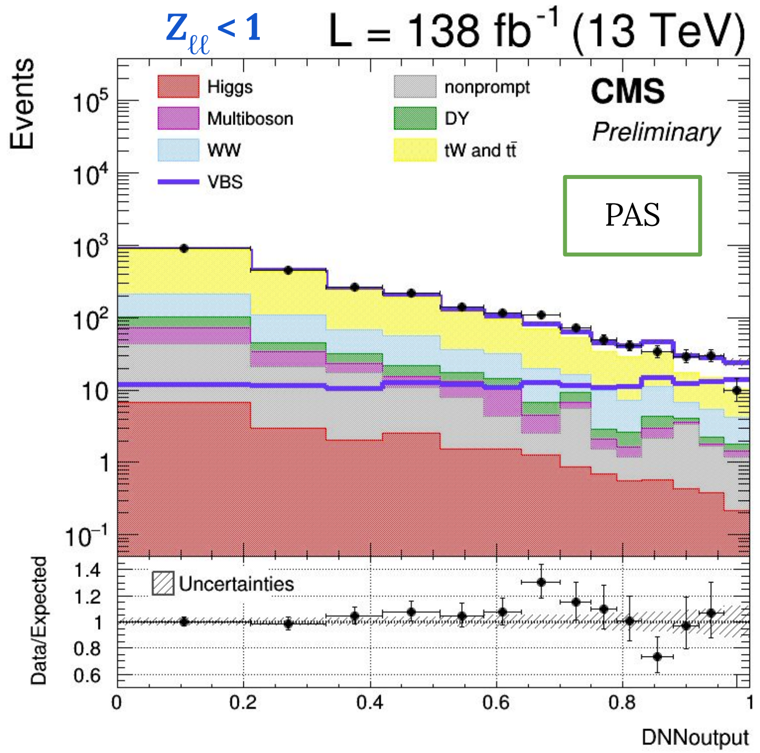
$$\mathcal{L}(\vec{n}|\mu,\theta) = p(\theta) \cdot \prod_{i=1}^{N} Poisson(n_i|\mu \cdot s_i(\theta) + b_i(\theta))$$

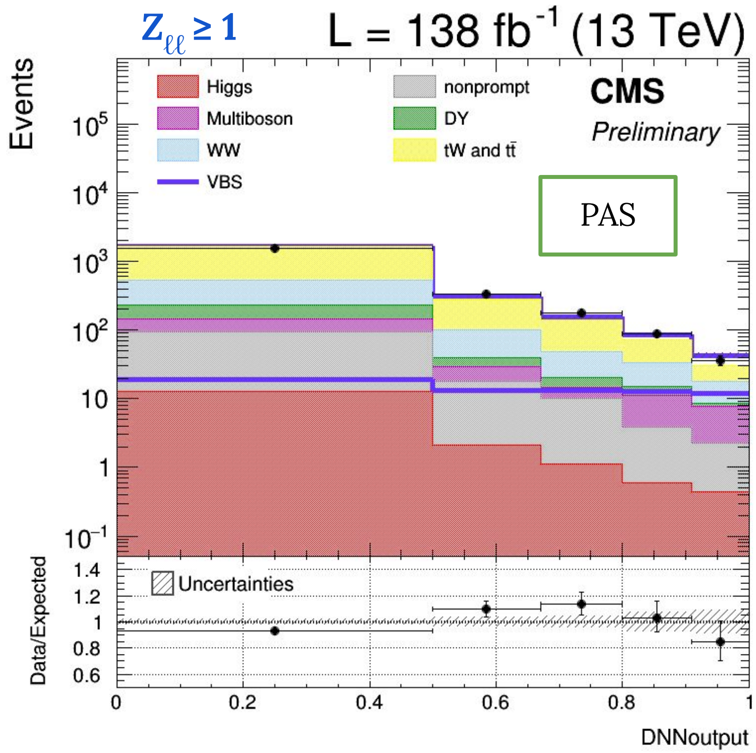
- ★ Combined **binned maximum likelihood fit** of the most discriminating variable distribution (m_{jj} or DNN output) with signal and background templates.
- ★ Performed simultaneously in all **signal region** categories ($Z_{\ell\ell} \lesssim 1$):
  - **SF:** divided into **four m_{jj} - $\Delta\eta_{jj}$** bins:
    - $2.5 < \Delta\eta_{jj} < 3.5$ and $300 < m_{jj} < 500$
    - $2.5 < \Delta\eta_{jj} < 3.5$ and $m_{jj} > 500$ GeV
    - $\Delta\eta_{jj} > 3.5$ and 300 GeV $< m_{jj} <$ 500 GeV

    } **Number of events**
    - $\Delta\eta_{jj} > 3.5$ and $m_{jj} > 500$ GeV → **purest region** → **m_{jj} distribution**.
  - **DF: DNN output**.
- ★ **DY** and **top CRs**: **single bin** categories → to constrain their **normalizations**

All categories reported in the scheme included **simultaneously** in the **fit**

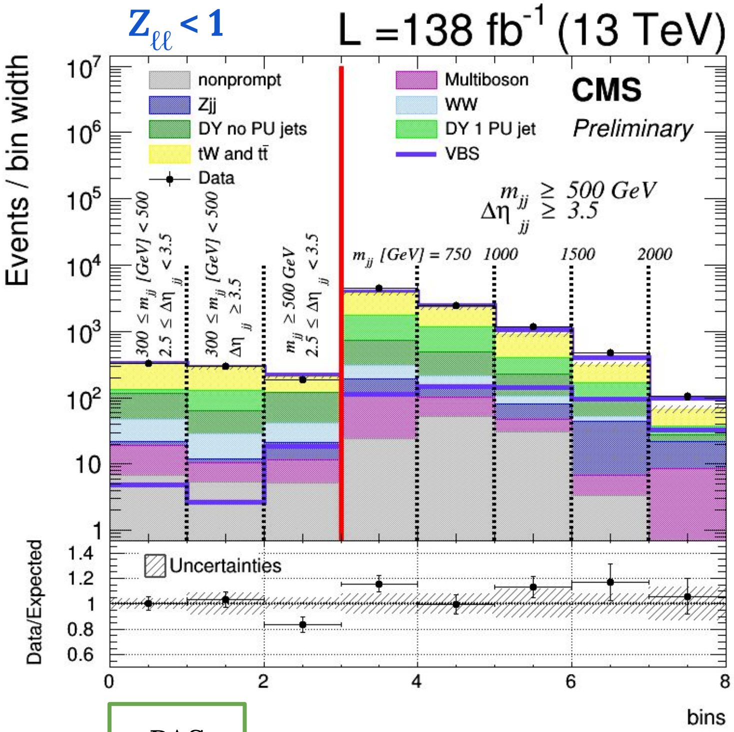
Signal regions *eμ*





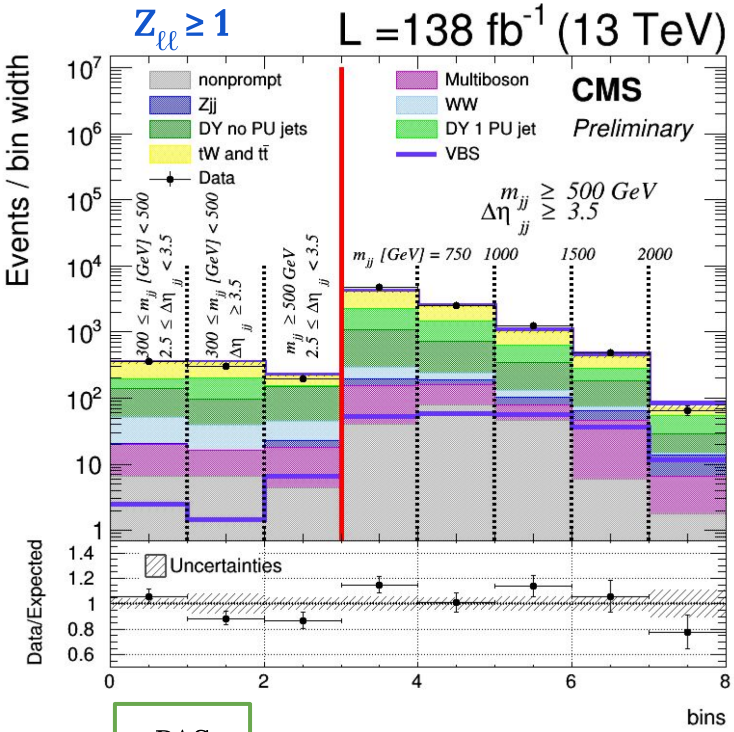
Signal regions *ee-μμ*

$Z_{\ell\ell} < 1$



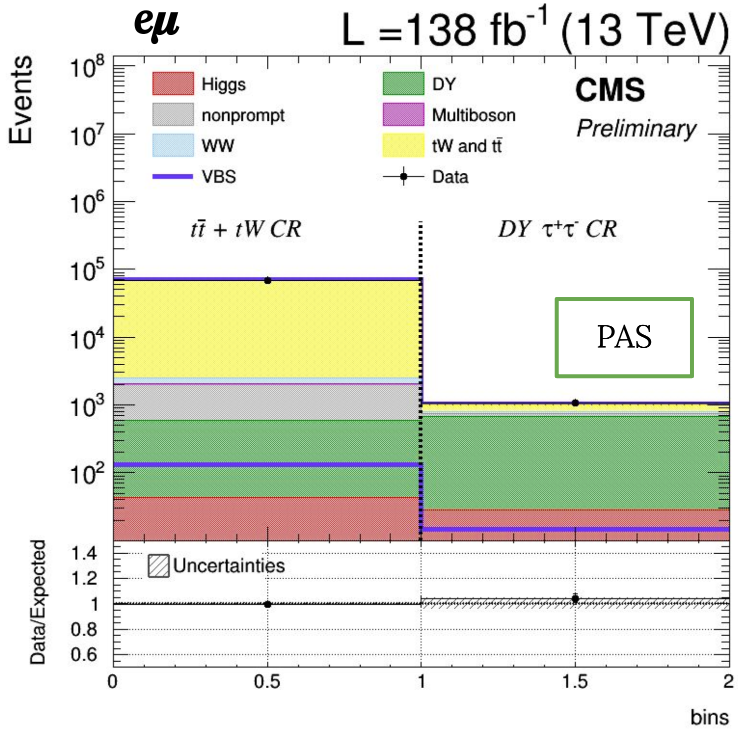
PAS

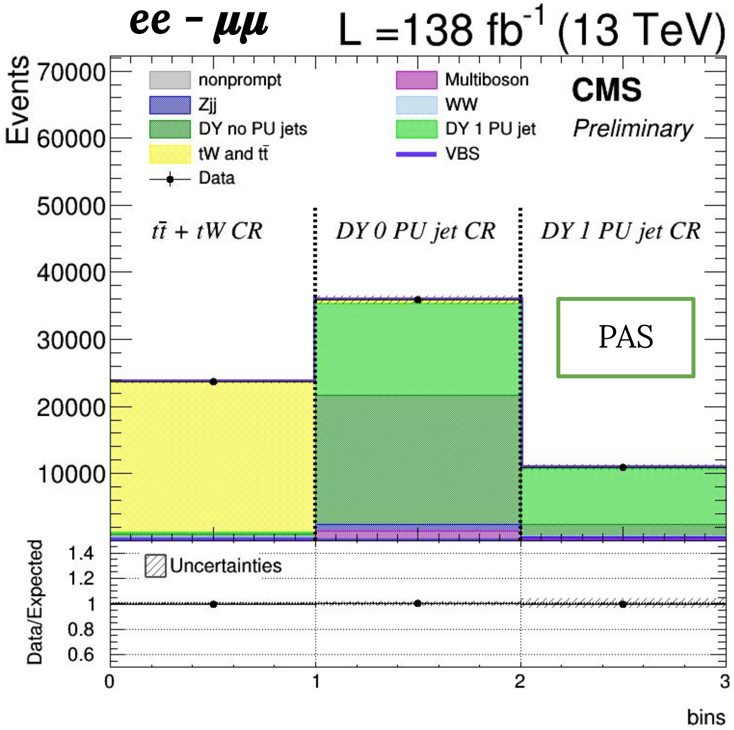
$Z_{\ell\ell} \geq 1$



PAS

Control regions





Systematic and statistical uncertainties

Systematic uncertainties

- ★ Represented by **individual nuisance** parameters with **log-normal** distributions;
- ★ could affect the:
  - **normalizations** of signal and backgrounds,
  - **shapes** of the predictions across the distributions of the observables.
- ★ Correlations taken into account.

Theoretical uncertainties

b tagging uncertainty

Dominated by **statistical** uncertainty

Uncertainty source	Impact
QCD-induced W^+W^- normalisation	5.3%
$t\bar{t}$ QCD scale	5.1%
QCD factorisation scale for VBS signal	5.0%
$t\bar{t}$ normalisation	4.9%
b tagging	3.5%
Prefiring corrections	3.3%
DY normalisation	2.9%
Jet energy scale + resolution	2.6%
p_T^{miss} energy scale	2.4%
QCD-induced W^+W^- QCD scale	2.1%
Luminosity	2.1%
Muon efficiency	2.0%
Pileup	1.8%
Electron efficiency	1.5%
Underlying event	1.3%
Parton shower	1.0%
Other	$< 1\%$
Total systematic uncertainty	13.1%
Total statistical uncertainty	14.9%
Total uncertainty	**19.8%**

Results

The **observed** (expected) **significance** with respect to the background-only hypothesis is **5.6 σ** (5.2 σ)

The cross section measurement of the W^+W^- EW production is performed in two fiducial volumes:

Inclusive phase space

Measured: **99 ± 20 fb**
LO prediction: 89 ± 5 fb

Parton-level observables:
$m_{qq} > 100$ GeV
$p_T > 10$ GeV

Good agreement with **SM** predictions!

Exclusive phase space

Measured: **10.2 ± 2.0 fb,**
LO prediction: 9.1 ± 0.6 fb.

Objects	Requirements
Leptons	e^+e^-, $\mu^+\mu^-$, $e^+\mu^-$, $e^-\mu^+$ $p_T^\ell = p_T^{bare\,\ell} + \sum_i p_T^{\gamma_i}$ if $\Delta R(\ell, \gamma_i) < 0.1$ $p_T^{\ell_1} > 25$ GeV, $p_T^{\ell_2} > 13$ GeV, $p_T^{\ell_3} < 10$ GeV $\lvert\eta\rvert < 2.5$ $p_T^{\ell\ell} > 30$ GeV, $m_{\ell\ell} > 50$ GeV
Jets	$p_T^j > 30$ GeV, $\lvert\eta\rvert < 4.7$ $\Delta R(j, \ell) > 0.4$ At least 2 jets, no b jets $m_{jj} > 300$ GeV, $\Delta\eta_{jj} > 2.5$
p_T^{miss}	$p_T^{miss} > 20$ GeV

More similar to the signal region.

Summary

- We reported the search of the **VBS** production of two **opposite sign W bosons** in the **fully leptonic** channel.

- **Deep neural network** built and trained to cope with the large top background and the irreducible WW QCD production (*e*μ final state):

- ***ee*-μμ categories also included to reach the 5 σ observation**, despite the additional contamination from DY events

- Results show the first observation of such mechanism:
  - **significance: 5.6 σ** (5.2 σ expected)
  - **exclusive cross section: 10.2 ± 2.0 fb** (9.1 ± 0.6 fb expected)
  - **inclusive cross section: 99 ± 20 fb** (89 ± 5 fb expected)

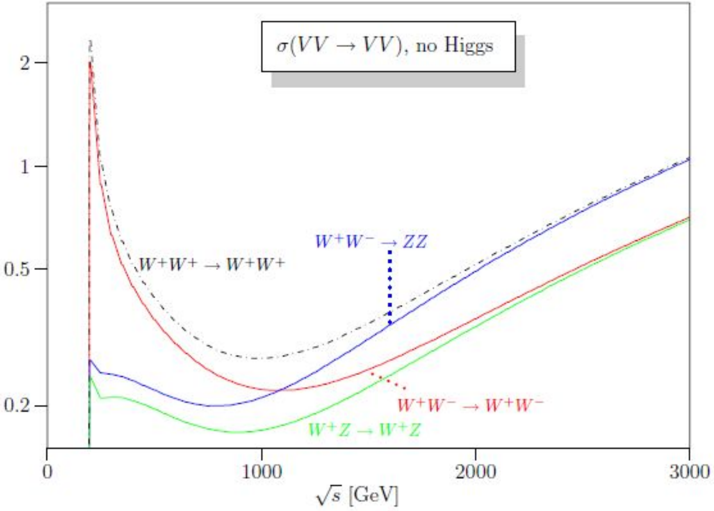
Good agreement with **SM** predictions!

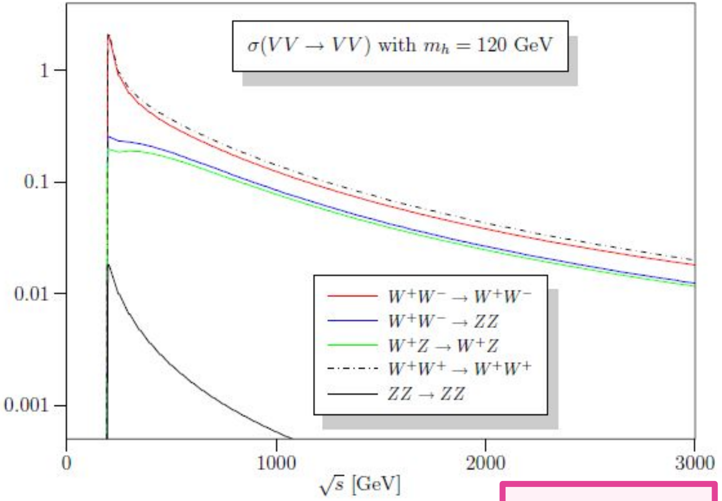
Analysis is expected to benefit from the larger **RUN III data set**

Backup

Unitarity of VBS cross sections

Cross sections (in nb) for the five different scattering processes of longitudinal weak gauge bosons.





Unitarity restored

Ana Alboteanu *et al* JHEP11(2008)010

Post-fit (pre-fit) yields table

Process	VBS $e\mu$	VBS $ee-\mu\mu$
WWewk	238.9 ± 21.9 (209.0 ± 5.4)	132.6 ± 6.9 (115.5 ± 2.2)
top	3081.9 ± 99.7 (2998.0 ± 189.3)	1152.3 ± 18.3 (1073.7 ± 33.7)
WW	736.3 ± 98.8 (1086.8 ± 89.0)	201.1 ± 22.6 (405.6 ± 22.0)
DY no PU jets	–	594.7 ± 19.9 (417.6 ± 25.9)
DY + 1 PU jet	–	436.1 ± 43.5 (370.4 ± 120.4)
DY$\tau\tau$	171.2 ± 7.4 (195.9 ± 6.2)	–
Non-prompt leptons	216.8 ± 24.6 (242.5 ± 31.7)	51.8 ± 6.1 (58.0 ± 7.8)
Multiboson	143.3 ± 9.8 (141.0 ± 15.9)	96.0 ± 6.0 (89.2 ± 7.8)
Higgs	46.6 ± 1.8 (43.2 ± 2.9)	–
Zjj	1.3 ± 0.2 (1.3 ± 0.3)	59.1 ± 4.3 (50.4 ± 6.5)

Selections

Preselections:

$p_T^{l1} > 25$ GeV && $p_T^{l2} > 13$ GeV
$p_T^{l3} < 10$ GeV
$m_{ll} > 50$ GeV && $p_T^{ll} > 30$ GeV
At least 2 jets with pT > 30 GeV
mjj > 300 GeV && $\Delta\eta jj > 2.5$

Categories selections:

VBS	$e\mu/\mu e$	$Z_{\ell\ell} < 1$	$m_T > 60$ GeV $m_{\ell\ell} > 50$ GeV no bjet with $p_T > 20$ GeV
		$Z_{\ell\ell} \geq 1$	
	ee	$Z_{\ell\ell} < 1$	$m_{\ell\ell} > 120$ GeV $p_T^{miss} > 60$ GeV no bjet with $p_T > 20$ GeV
		$Z_{\ell\ell} \geq 1$	
	$\mu\mu$	$Z_{\ell\ell} < 1$	
		$Z_{\ell\ell} \geq 1$	
$t\bar{t}$ and tW	$e\mu/\mu e$		$m_{\ell\ell} > 50$ GeV no b-jet with $p_T > 20$ GeV
	ee		$m_{\ell\ell} > 120$ GeV $p_T^{miss} > 60$ GeV at least one b-jet with $p_T > 20$ GeV
	$\mu\mu$		
DY	$e\mu/\mu e$		$m^{T} < 60$ GeV 50 GeV$< m_{\ell\ell} < 80$ GeV no b-jet with $p_T > 20$ GeV
	ee	$\Delta\eta_{jj} < 5$	$\|m_{\ell\ell} - m_Z\| < 15$ GeV $p_T^{miss} > 60$ GeV no b-jet with $p_T > 20$ GeV
		$\Delta\eta_{jj} \geq 5$	
	$\mu\mu$	$\Delta\eta_{jj} < 5$	
		$\Delta\eta_{jj} \geq 5$	

Comparison with standard samples

- ★ We drew a comparison between the "**standard**" and the "**dipole recoil**" signal samples across all data sets @reco-level

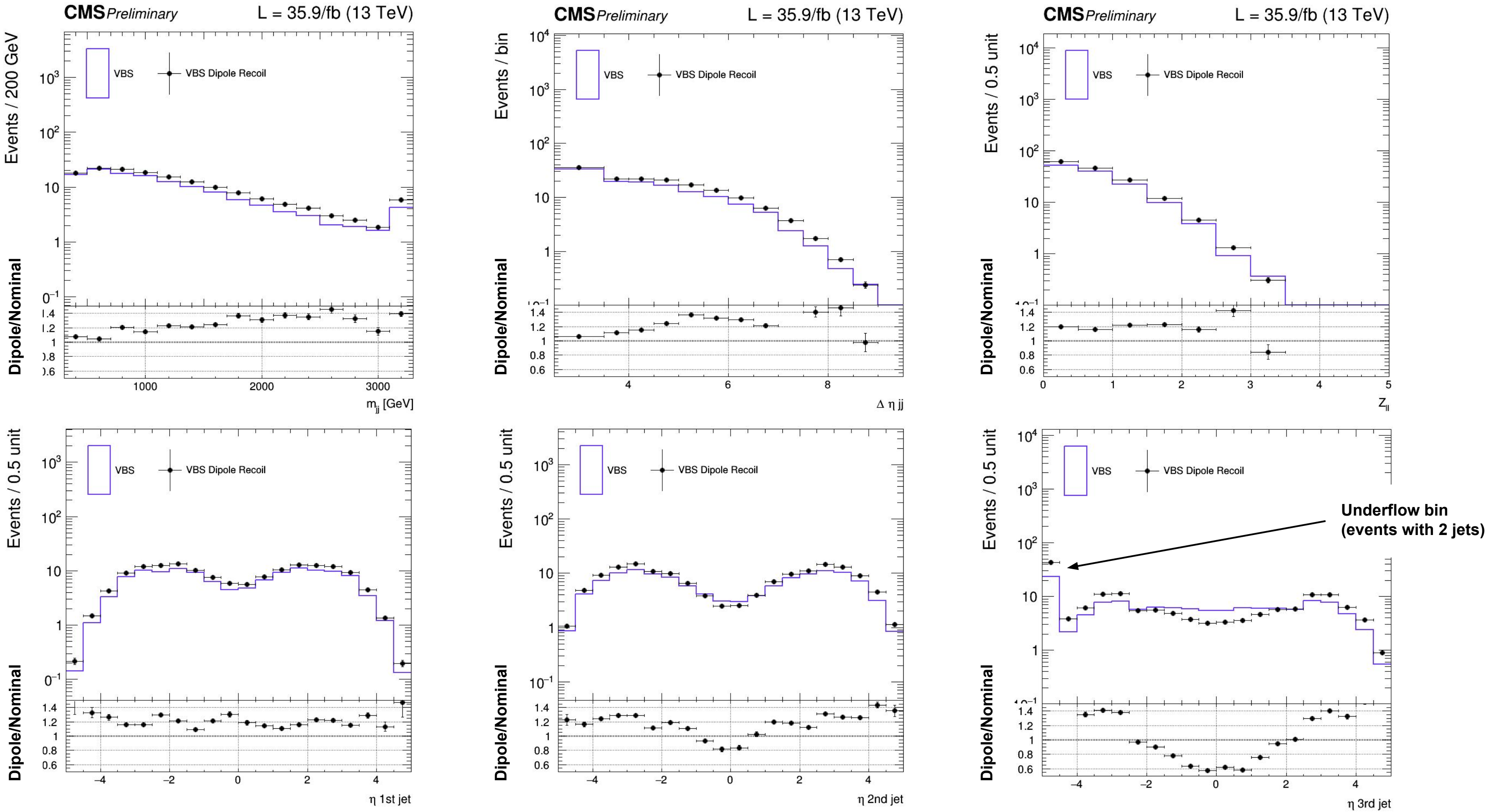
- ★ **Inclusive selection**:
  - $m_{\ell\ell} > 50$ GeV
  - $p_T^{miss} > 20$ GeV
  - $Z_{\ell\ell}$ categories merged
  - eμ-ee-μμ categories merged

- ★ **2016** samples are those showing more differences → in order to heal the **unphysical `bwcutoff = 15000`** setting **in the standard sample**, we used to cut @LHE level on W masses (no similar cut is applied to the dipole recoil sample instead)

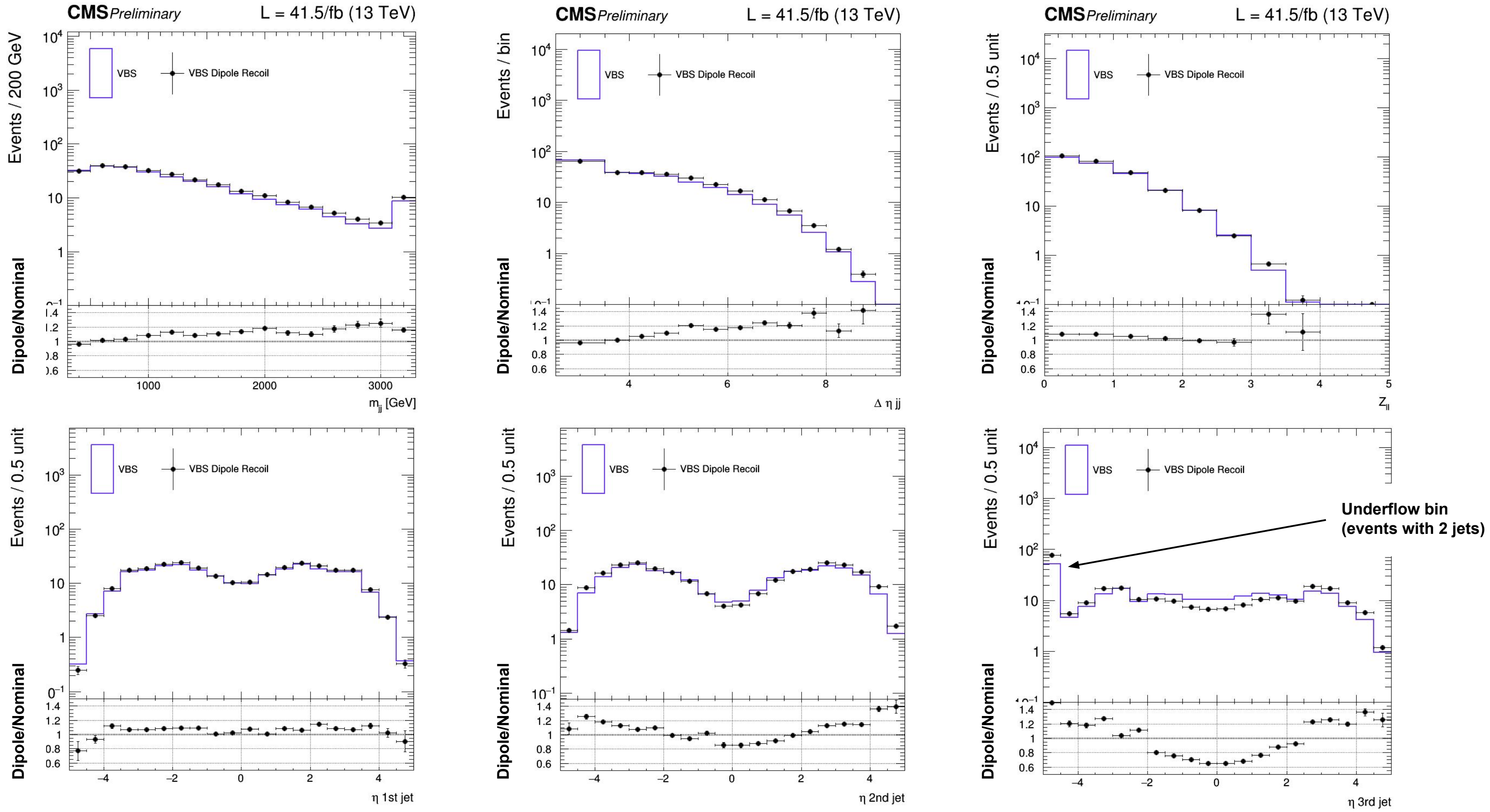
- ★ **2017/2018** samples are rather **similar** as far as **2jets-related variables** are concerned

- ★ **3jets variables** are significantly **affected** by the dipole recoil setting, but they **do not have a large impact on the analysis** (see for instance $Z_{\ell\ell}$ distributions)

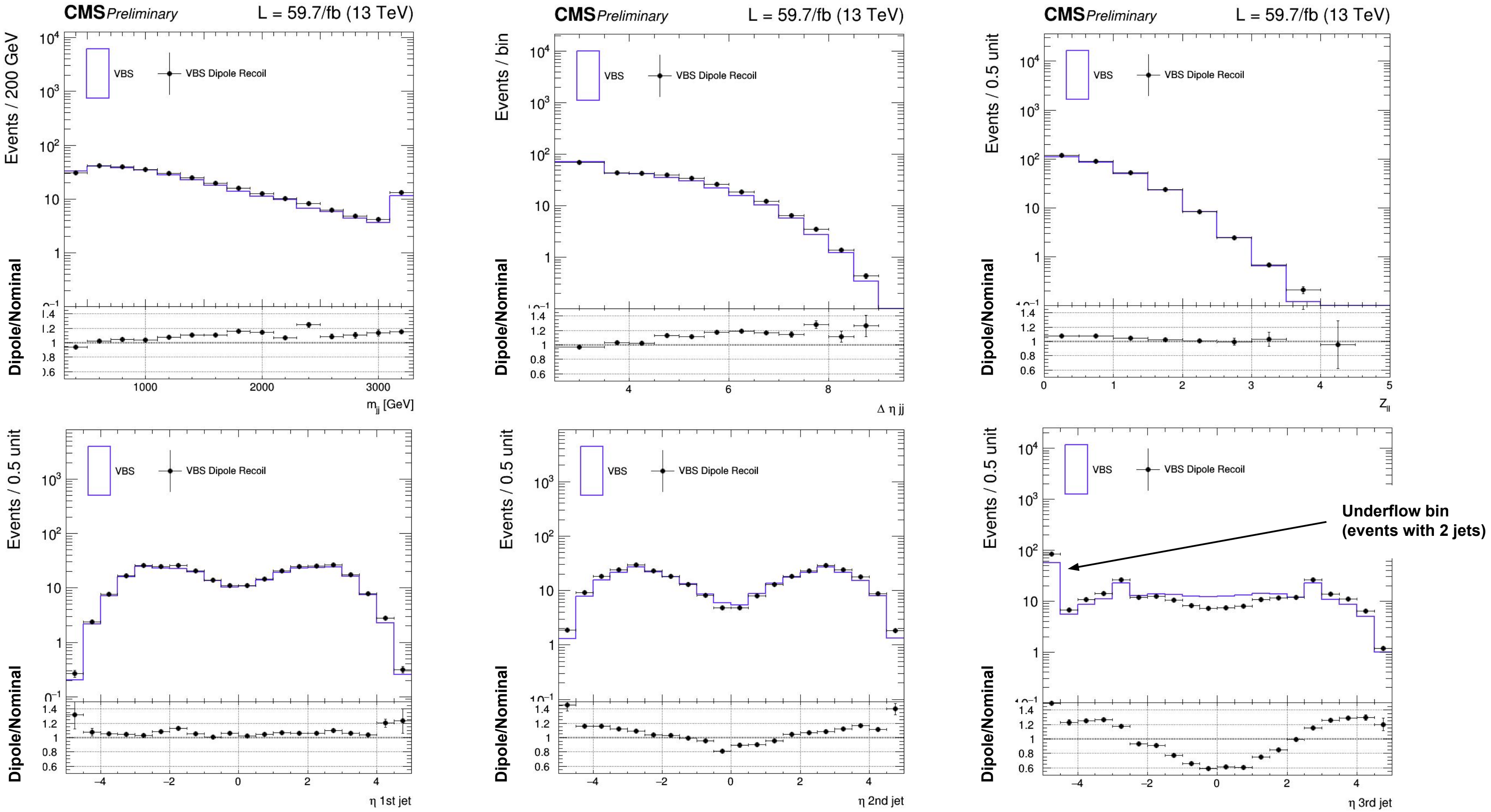
2016


CMS Preliminary
L = 35.9/fb (13 TeV)
VBS
VBS Dipole Recoil
Events / 200 GeV
Dipole/Nominal
m_jj [GeV]
Events / bin
Δ η jj
Events / 0.5 unit
Z_ll
η 1st jet
η 2nd jet
η 3rd jet
Underflow bin
(events with 2 jets)

2017


CMS Preliminary
L = 41.5/fb (13 TeV)
VBS
VBS Dipole Recoil
Events / 200 GeV
Events / bin
Events / 0.5 unit
Dipole/Nominal
m_{jj} [GeV]
Δ η jj
Z_{ll}
η 1st jet
η 2nd jet
η 3rd jet
Underflow bin
(events with 2 jets)

2018


CMS Preliminary
L = 59.7/fb (13 TeV)
VBS
VBS Dipole Recoil
Events / 200 GeV
Dipole/Nominal
m_{jj} [GeV]
Events / bin
$\Delta\eta$ jj
Events / 0.5 unit
Z_{ll}
η 1st jet
η 2nd jet
η 3rd jet
Underflow bin
(events with 2 jets)