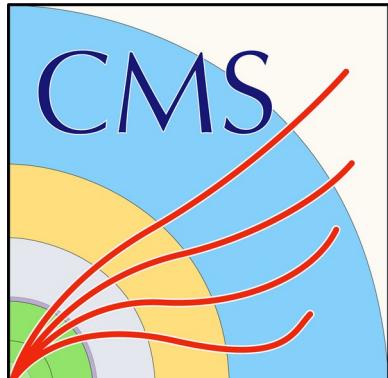


# Vector Boson Scattering and forward Pileup Jets in ATLAS

**Louis Portales**  
LLR, CNRS

Monday March 29, 2021



# **Vector Boson Scattering and forward Pileup Jets in ATLAS**

**Louis Portales**  
**LLR, CNRS**

**A few months back:**

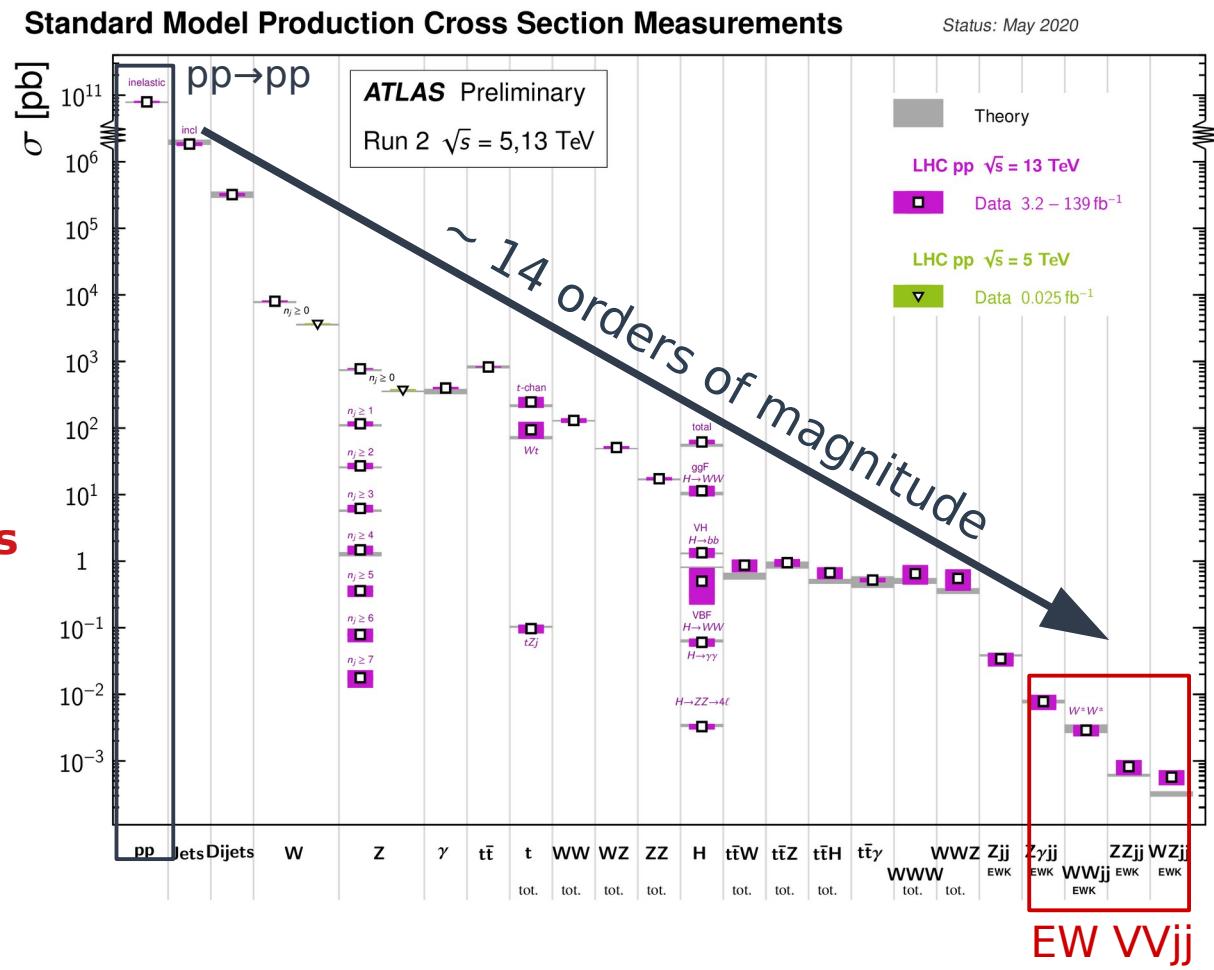
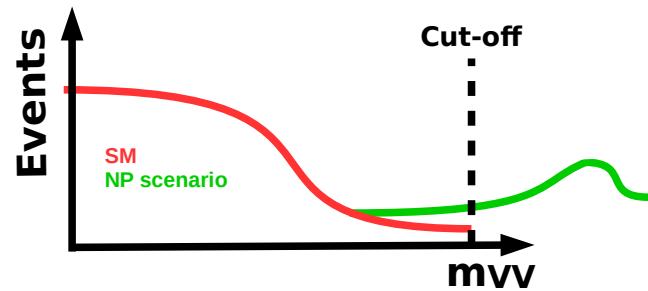


# The Standard Model - Experimental status

> **SM extensively tested**

## > Current physics goals

- Higgs boson properties
  - Search for new physics
    - **Accessible through rare SM processes**
    - **Seen as resonance or excess in high energy tails**



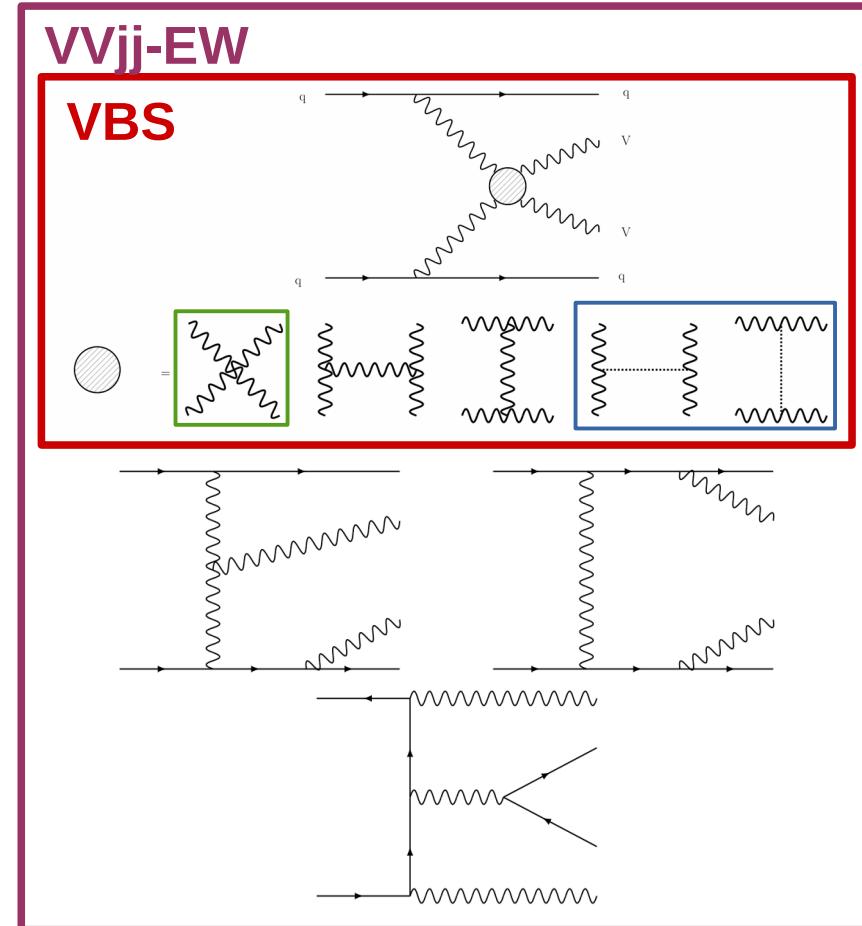
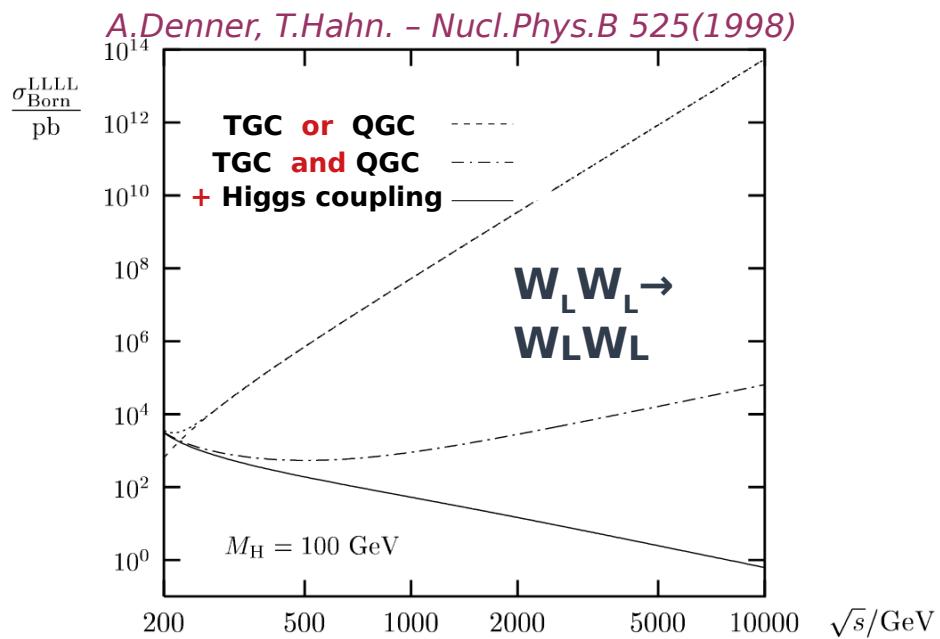
# Vector Boson Scattering

## > VVjj-EW production

- $O(\alpha^6)$  processes ( $V=W,Z$ )
- Includes **Vector Boson Scattering**

## > Important process to constrain (B)SM

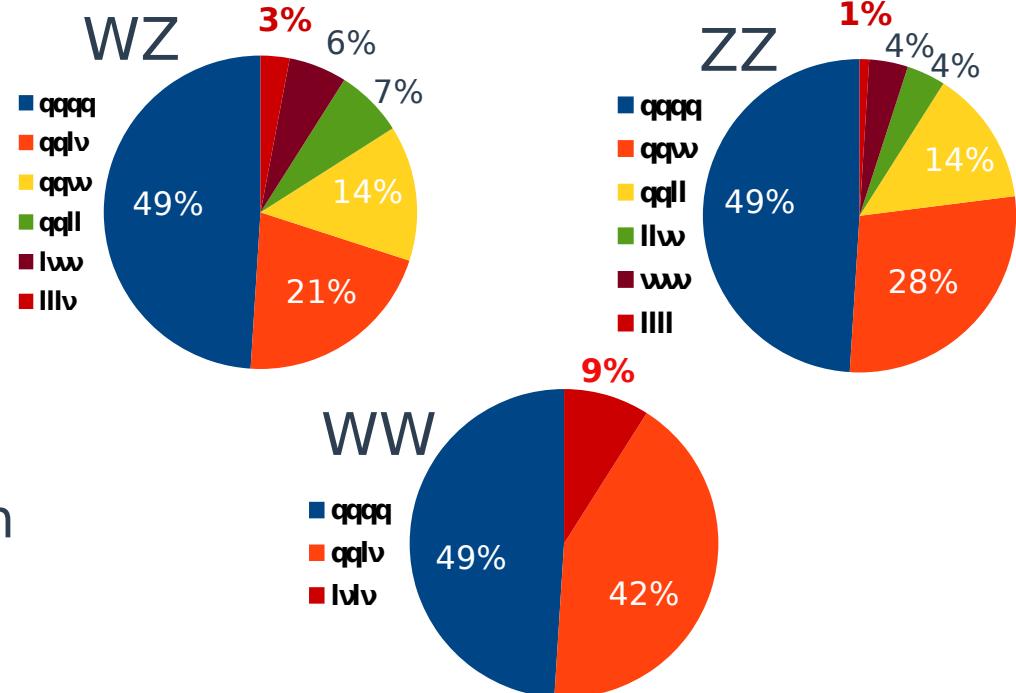
- **(Anomalous) quartic gauge coupling**
- **Higgs sector:**  $VV \rightarrow VLVL$  unitarity



# Experimental considerations

## > Which boson decay modes?

- **(Semi)hadronic** decays
  - Large backgrounds
  - Low reconstruction efficiency
- **Fully leptonic** decays:
  - High efficiency
  - Best candidates for observation



## > Which boson pair?

First observation (CMS,  $36 \text{ fb}^{-1}$ ): PLB 120 (2018)

- 2 neutrinos
- Large misid. leptons impact
- + Low QCD ( $\mathcal{O}(\alpha^6)$ ) background
- + Largest cross-section

Observation (ATLAS,  $36 \text{ fb}^{-1}$ ): PLB 793 (2019)

- Very large QCD background
- + “Medium” cross-section
- + Low impact from misid. leptons

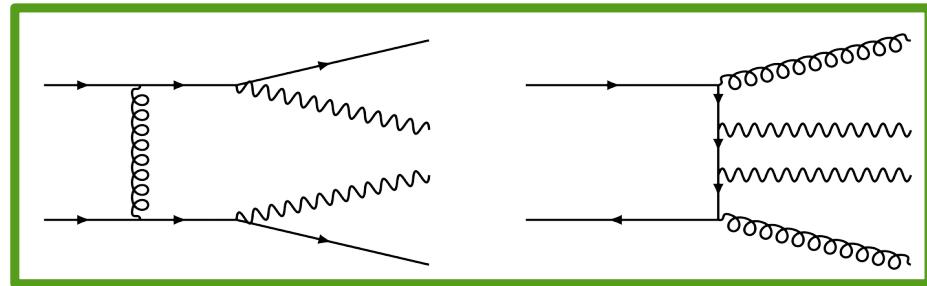
Observation (ATLAS,  $139 \text{ fb}^{-1}$ ) arXiv:2004.10612

- Tiny cross-section
- + Fully reconstructed final state

Process	$W^\pm W^\pm jj \rightarrow \ell'^\pm \ell^\pm \nu_\ell \bar{\nu}_\ell$	$WZjj \rightarrow \ell^+ \ell^- \ell' \nu_\ell \bar{\nu}_{\ell'}$	$ZZjj \rightarrow \ell^+ \ell^- \ell' \ell''$
$\sigma(pp \rightarrow X) [\text{fb}] (\text{EW})$	3.97	2.34	0.098
$\sigma(pp \rightarrow X) [\text{fb}] (\text{QCD})$	0.35	4.38	0.1
$EW/QCD$	$\sim 10$	$\sim 0.5$	$\sim 1$

# QCD background and VBS signature

> **QCD background mitigation is challenging**

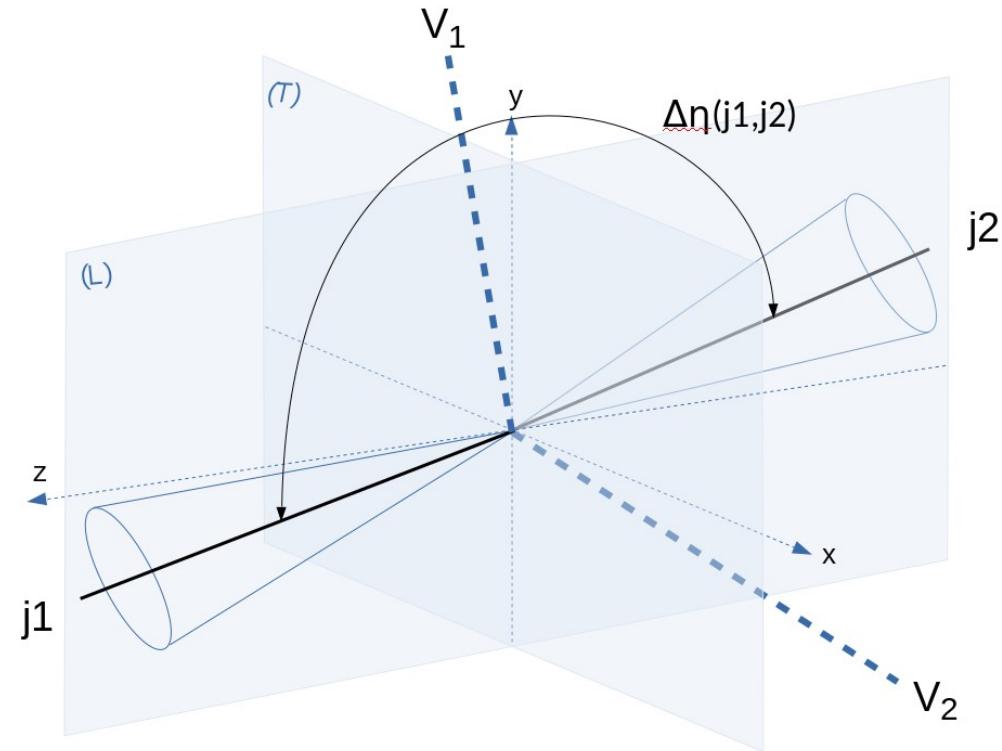


> **WZjj-EW has specific signature**

> **Vector bosons centrally emitted**

> **Forward quark-jets**

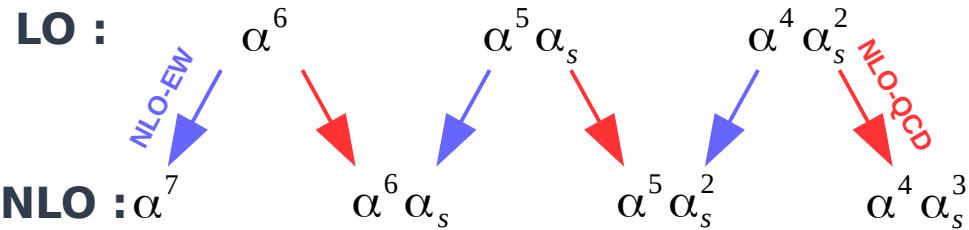
- Large angular separation
- High invariant mass
- **Used to mitigate WZjj-QCD**



# Theoretical considerations - Modelling

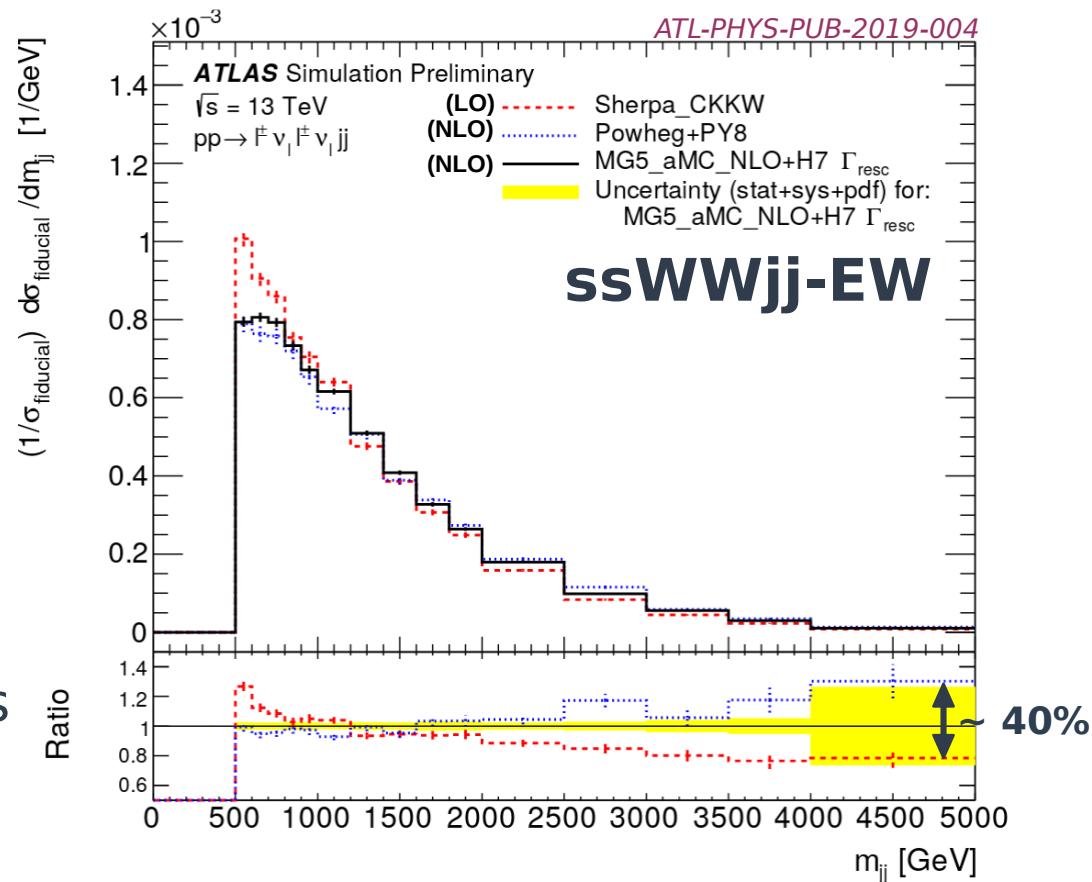
- > Rely on MC simulation to study EW/QCD separation
- > LO generators still used in most cases

## WZjj-EW Interference WZjj-QCD



- > Large difference between generators

- Mostly QCD-related
  - NLO-QCD & approximations disagree
  - Concerns jet kinematics & multiplicity



# Theoretical considerations - Higher order corrections

## > Large NLO-EW corrections

**WZjj-EW Interference WZjj-QCD**

LO :  $\alpha^6$

NLO :  $\alpha^7$

$\alpha^6 \xrightarrow{\text{NLO-EW}}$

$\alpha^6 \alpha_s \xrightarrow{\text{NLO-EW}}$

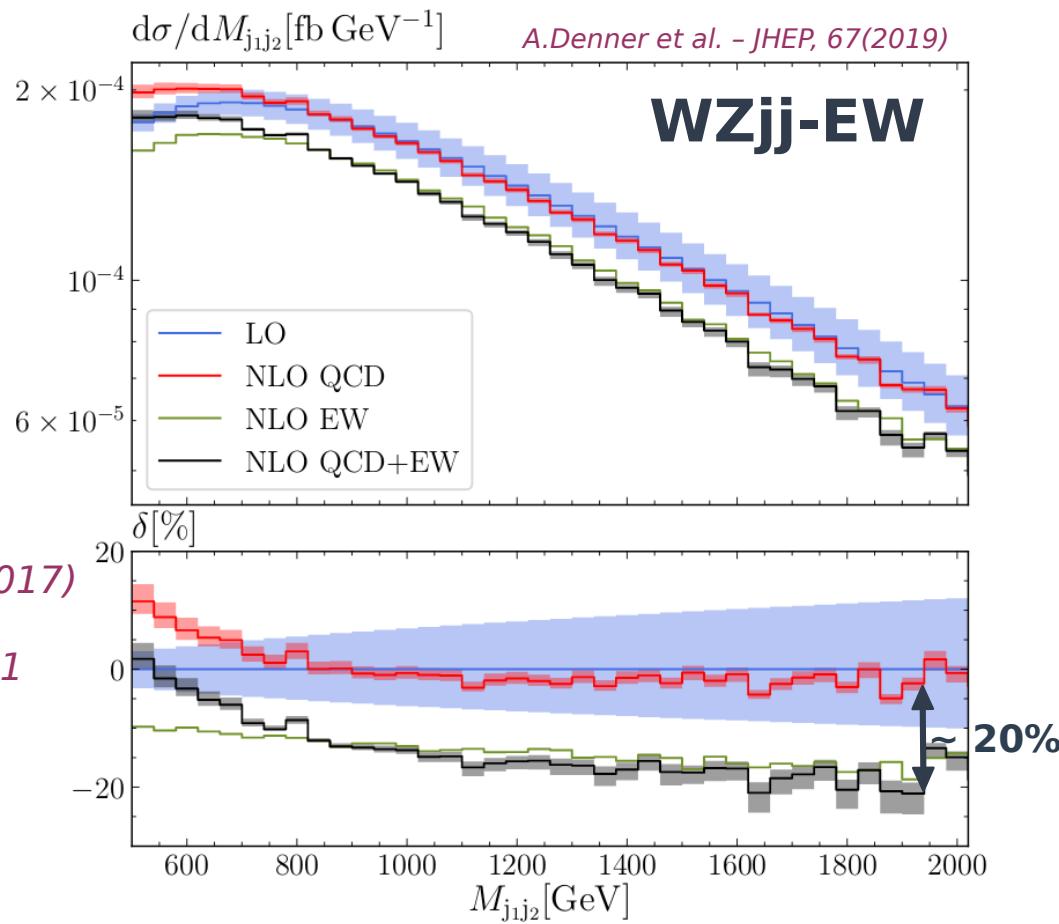
$\alpha^5 \alpha_s \xrightarrow{\text{NLO-EW}}$

$\alpha^5 \alpha_s^2 \xrightarrow{\text{NLO-QCD}}$

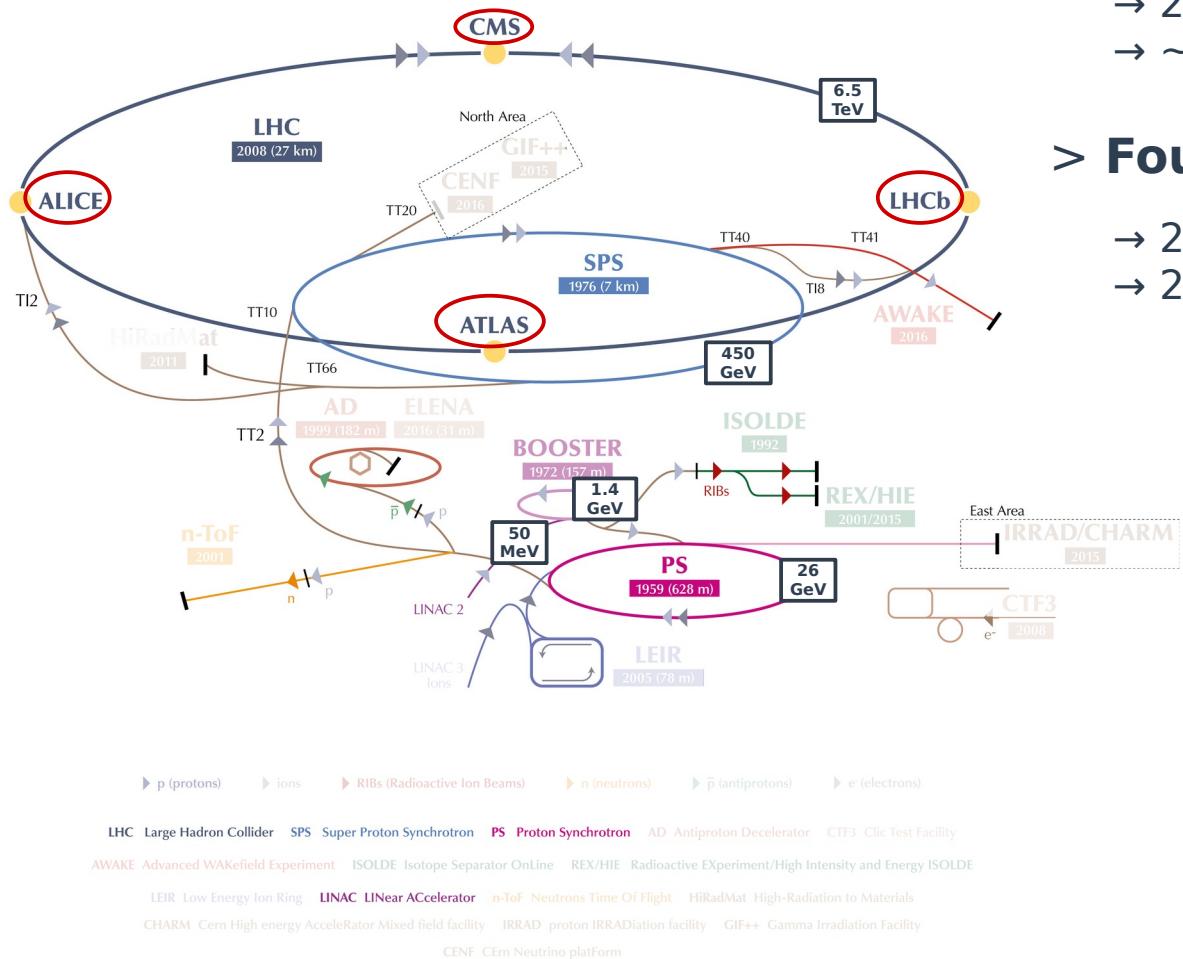
$\alpha^4 \alpha_s^2 \xrightarrow{\text{NLO-QCD}}$

$\alpha^4 \alpha_s^3 \xrightarrow{\text{NLO-QCD}}$

- Only recently computed
  - WW (2017): *B.Biedermann et al. - JHEP, 124(2017)*
  - WZ (2019): *A.Denner et al. - JHEP, 67(2019)*
  - ZZ (2020): *A.Denner et al. - arXiv:2099.00411*



# The Large Hadron Collider

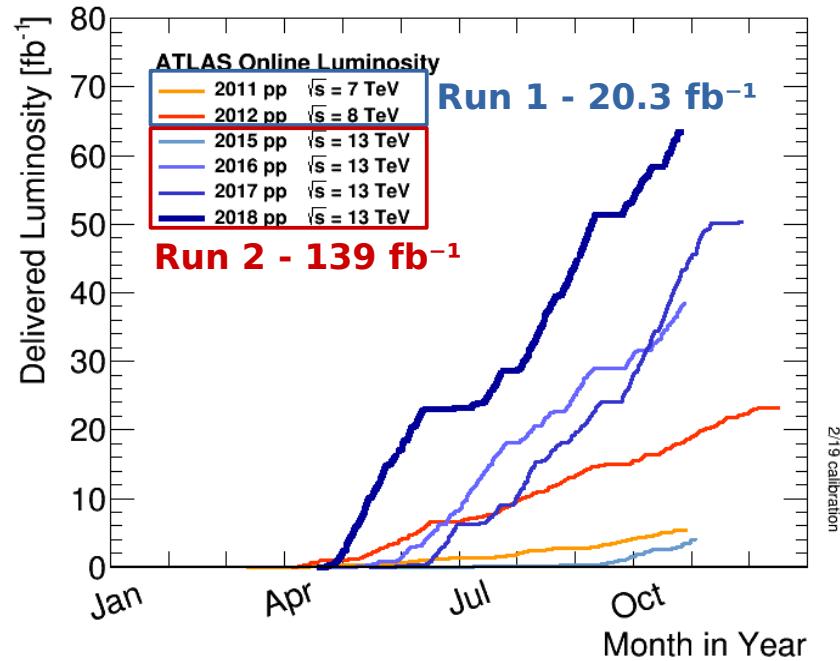


## > Proton collisions, up-to $\sqrt{s}=13$ TeV

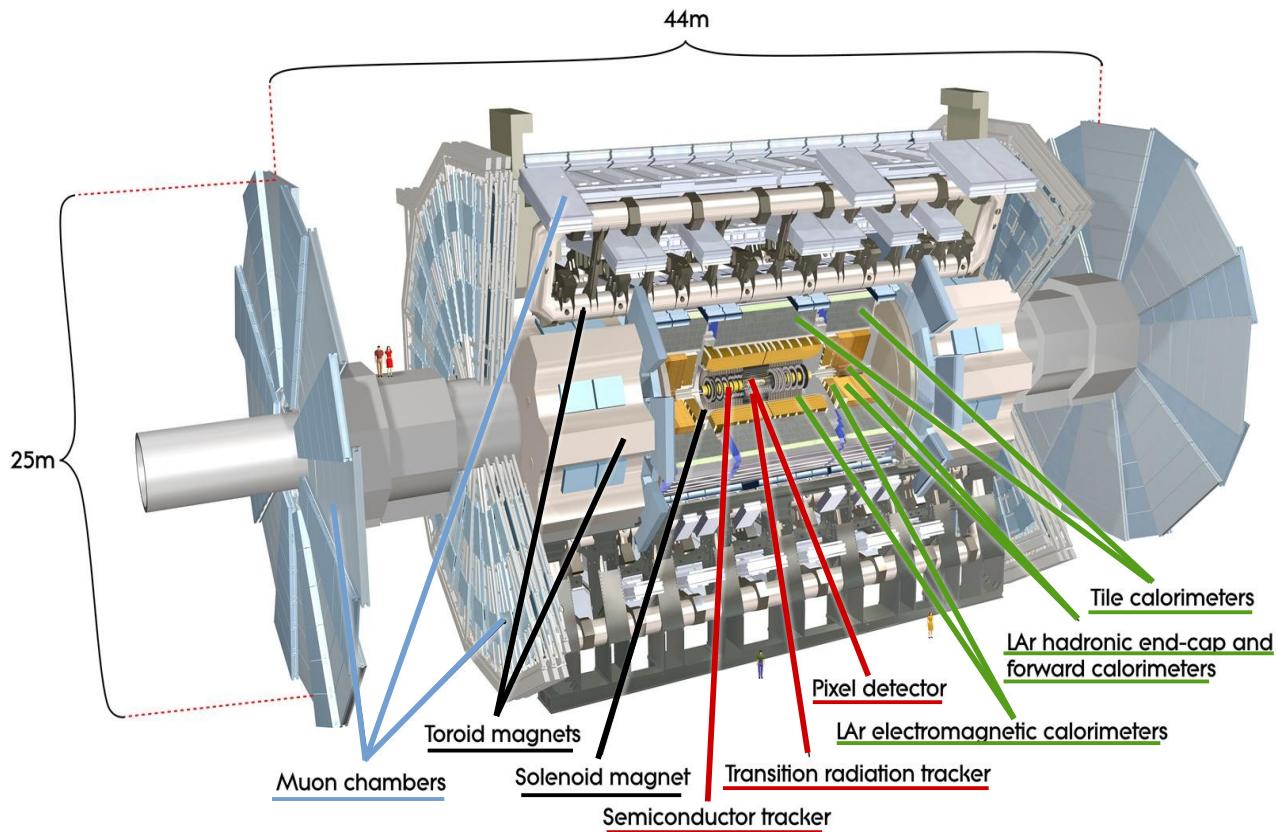
- High collision frequency  $\sim 40$  MHz
  - 25 ns between each bunch crossing
  - $\sim 10^{11}$  protons per bunch

## > Four large-scale experiments

- 2 “specialised”: ALICE, LHCb
  - 2 “general purpose”: ATLAS, CMS

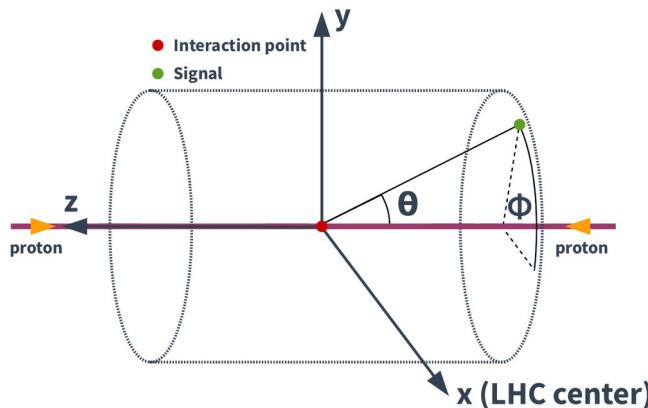


# The ATLAS detector



→ Coordinates expressed as  $(\eta, \phi)$

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right)$$



## > Electromagnets

- Solenoid:  $B = 2 \text{ T}$
- Toroids:  $B = 4 \text{ T}$

## > Inner Detector

- Charged particles tracks

$$\frac{\sigma_{p_T}}{p_T} = 5 \times 10^{-5} p_T / \text{GeV} \oplus 1 \%$$

## > Calorimeters

- EMCal:  $e/\gamma$  energy

$$\frac{\sigma_E}{E} = \frac{10 \%}{\sqrt{E [\text{GeV}]}} \oplus 0.7 \%$$

- HCal: hadrons energy

$$\frac{\sigma_E}{E} = \frac{53 \%}{\sqrt{E [\text{GeV}]}} \oplus 3 \%$$

## > Muon Spectrometer

- Muon tracks

$$\frac{\sigma_{p_T}}{p_T} \simeq 1 \% - 3 \%$$

# Particle detection and reconstruction

## > Electrons

- ID track + EMCal clusters

## > Muons

- Track in all subdetectors

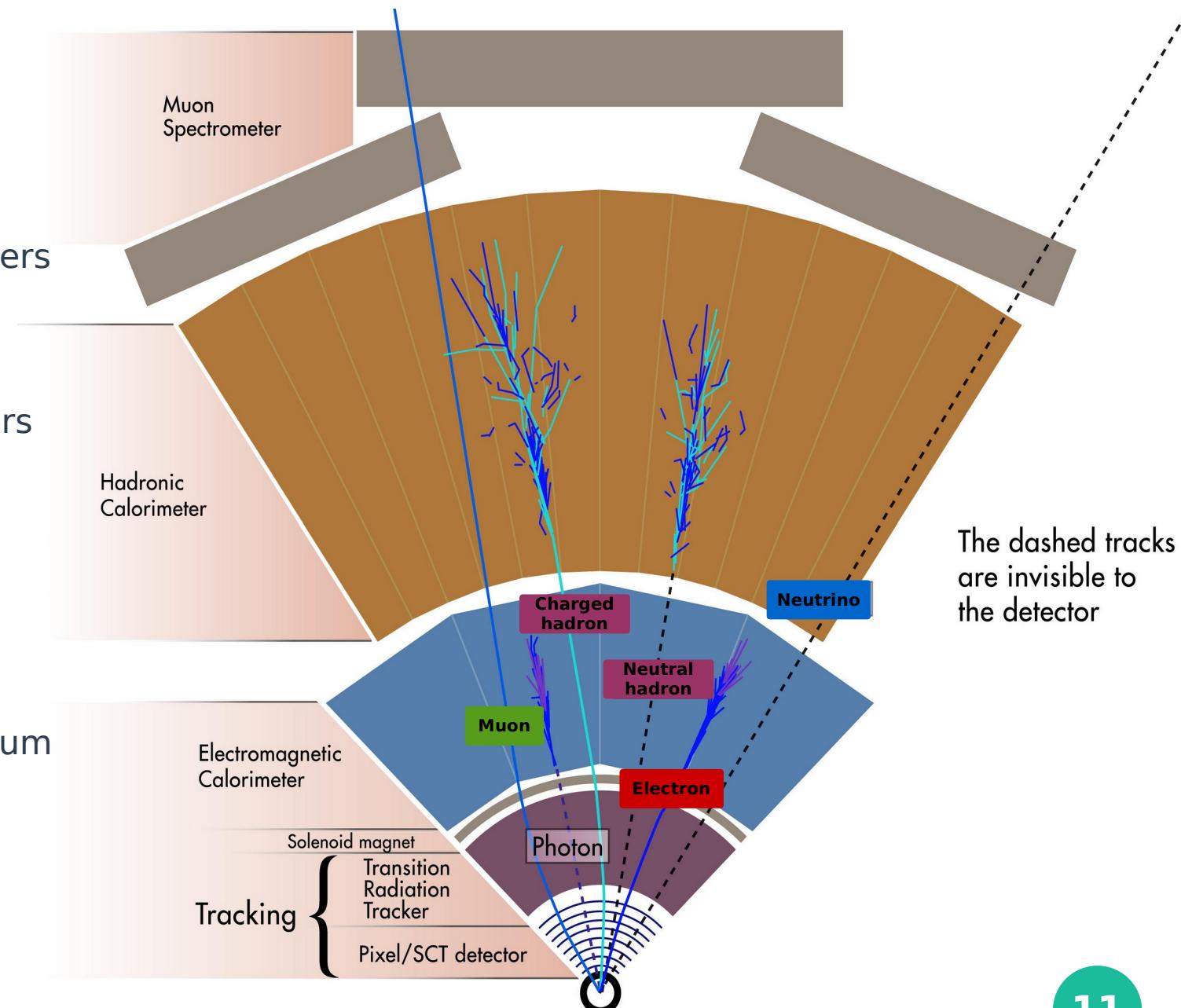
## > Hadrons (Jets)

- EMCal/HCal clusters  
(+ ID tracks)

## > Neutrinos

- From missing momentum

$$E_T^{\text{miss}} = \left| - \sum_i \vec{p}_T^i \right|$$

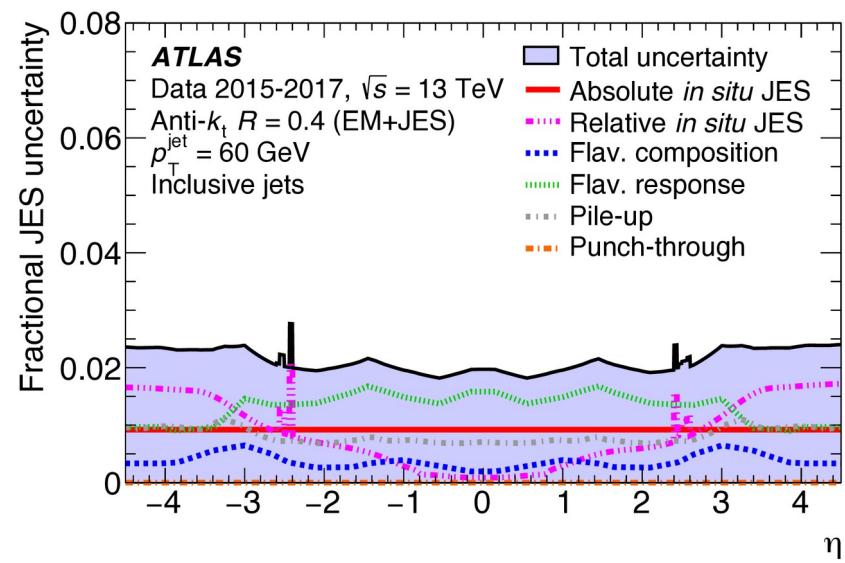
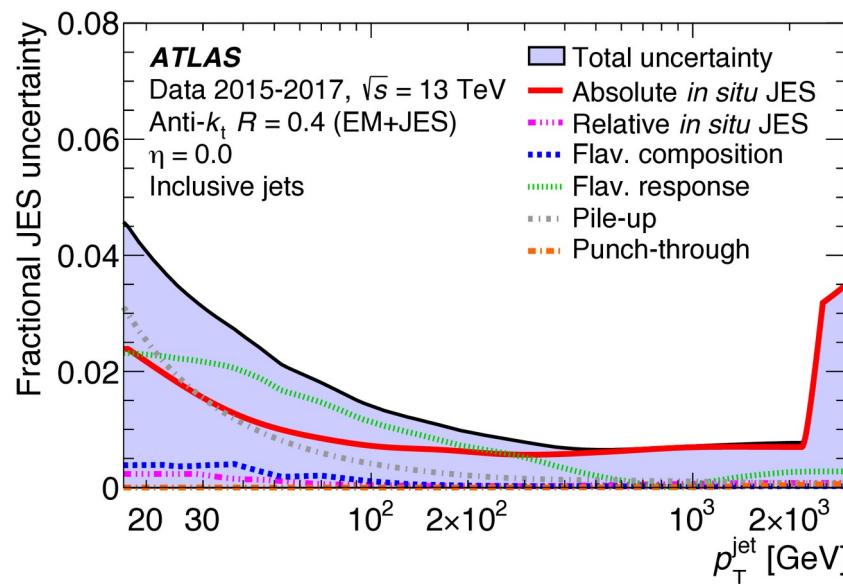
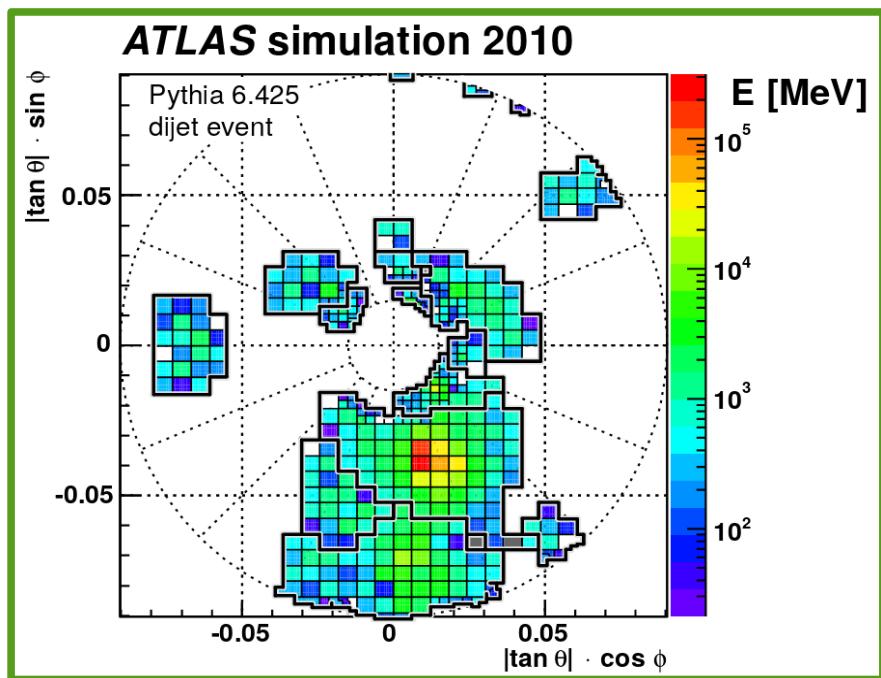


# Jet reconstruction & performance

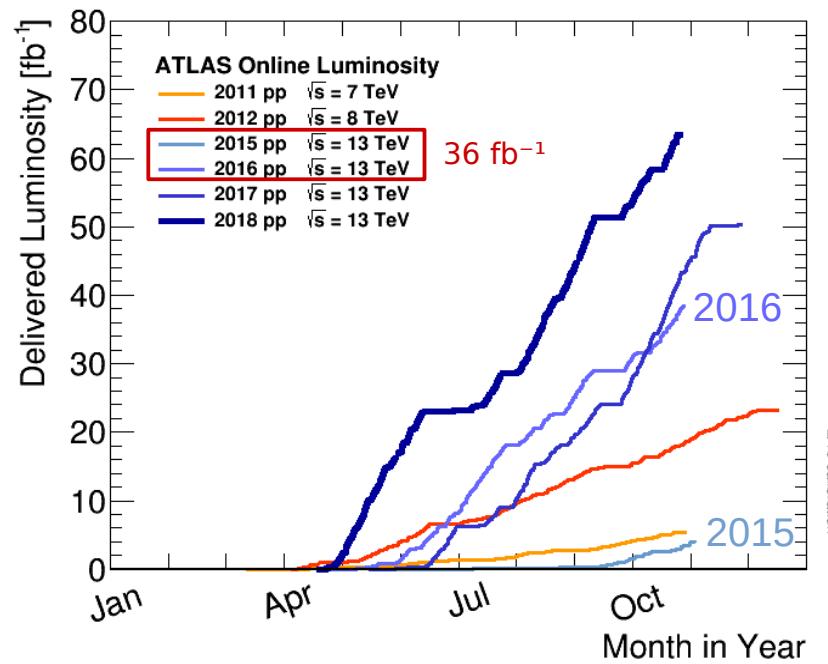
**EMTopo** jets: “historical” approach

- 1 Build **Topoclusters** from neighbouring calorimeter cells ( $|E_{\text{cell}}| > 2\sigma_{\text{noise}}$ )
- 2 Identify most energetic cluster, combine neighbouring clusters using **Anti- $k_T$**  ( $R=0.4$ ):

$$\min(k_{T,i}^{-2}, k_{T,j}^{-2}) \frac{\Delta R_{ij}}{R^2} < k_{T,i}^{-2}$$



## Observation of the WZjj-EW production



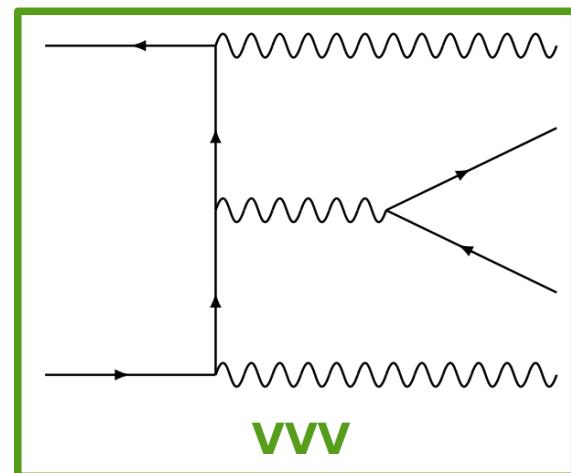
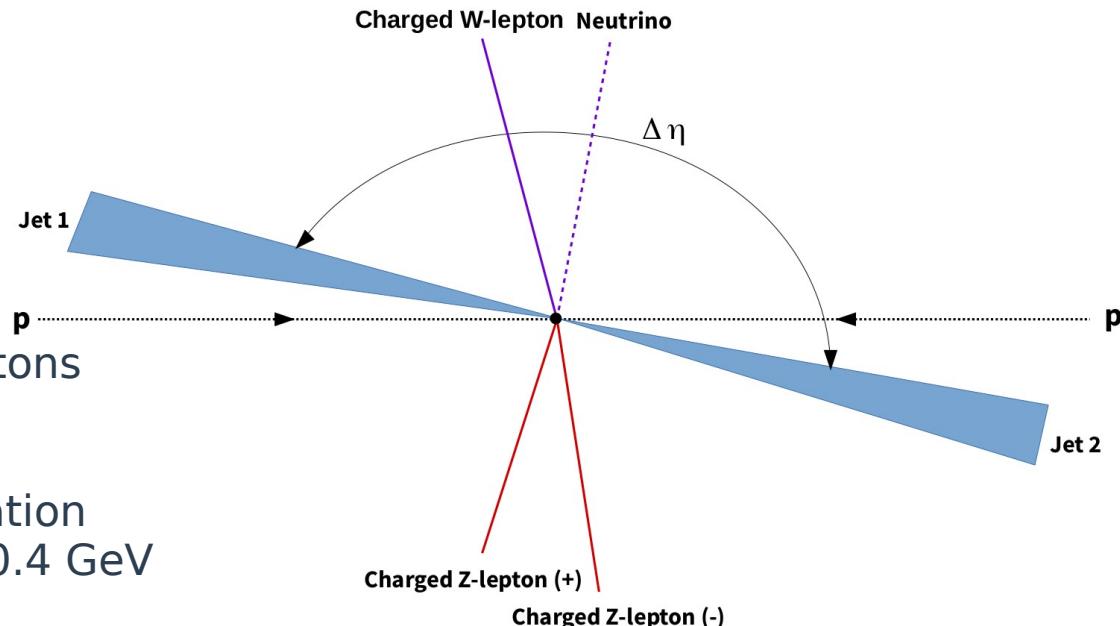
# VBS selection

## > Leptons & Bosons

- Single-lepton triggers
- Exactly 3 isolated leptons
- Z boson:
  - 2 same-flavor, opposite-charge leptons
  - $|m_{ll} - m_Z| < 10 \text{ GeV}$
- W boson:
  - 3<sup>rd</sup> lepton:  $p_T > 20 \text{ GeV}$  + tight isolation
  - matched to  $E_T^{\text{miss}}(\nu)$ , fixing  $m_W = 80.4 \text{ GeV}$

## > Jets

- At least two (EMTopo) jets
- $p_T > 40 \text{ GeV}$ 
  - **Jet uncertainties impact reduction**
- In opposite hemispheres ( $\eta_{j1}\eta_{j2} < 0$ )
  - **"Tagging jets"**
- $m_{jj} > 150 \text{ GeV}$ 
  - **Triboson background suppression**
  - **"VBS selection"**



# VBS selection

## > **WZjj-EW signal**

- Subdominant

## > **WZjj-QCD background**

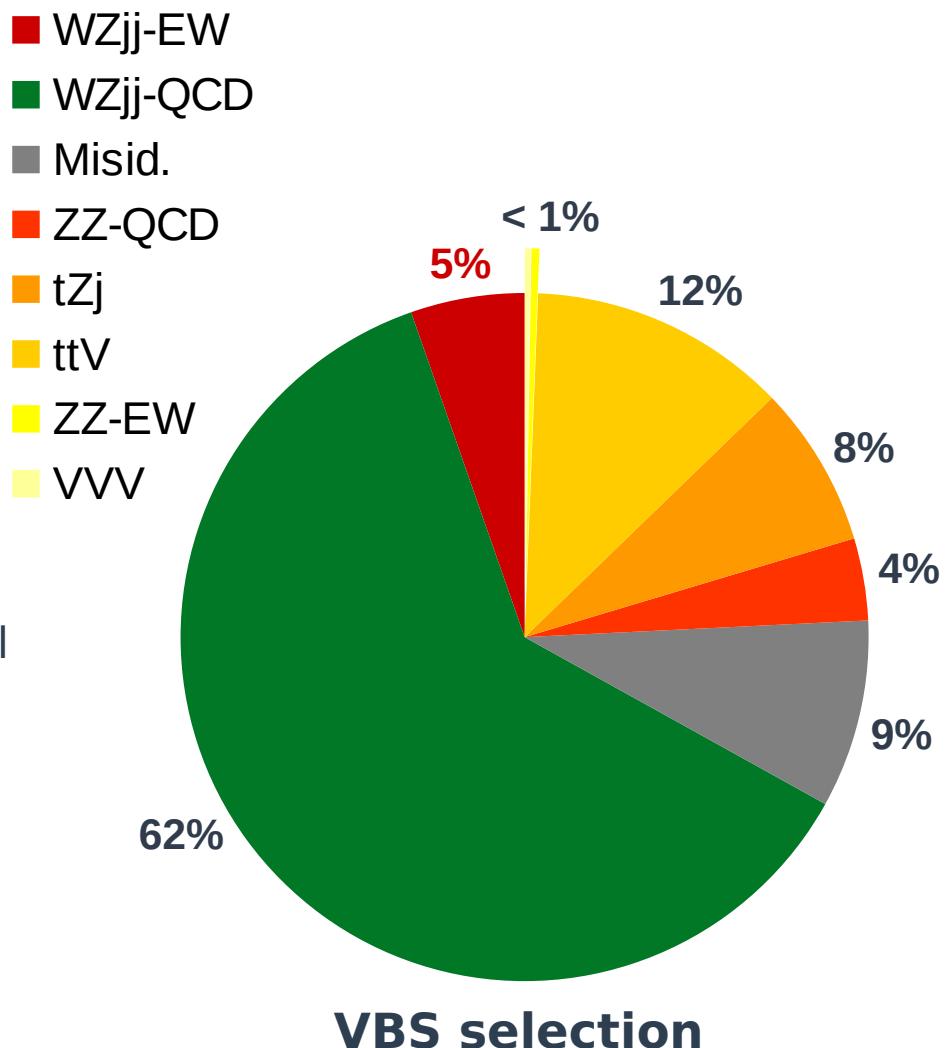
- Largely dominant

## > **“Irreducible” background**

- Same final-state particle content as signal
- **Or** Additional particles not reconstructed  
→ **tZj, ttV, ZZ**

## > **“Reducible” background**

- Misidentified leptons  
→ **tt, Z+jets, Zγ**
- Data-driven estimate



# Analysis challenges

- > **Signal extraction strategy designed to deal with low signal purity**
- > **Necessity to control backgrounds**
  - High- $m_{jj}$  Signal Region (SR) + additional regions for background control
- > **WZjj-EW and WZjj-QCD separation**
  - Multivariate discriminant in SR
- > **Low statistics in signal-pure regimes**
  - Combined likelihood fit (SR + CR)

# VBS selection splitting

> VBS selection split in three regions

- **Signal Region**

→  $N_{b\text{-jet}} = 0, m_{jj} > 500 \text{ GeV}$

- **QCD-CR**

→  $N_{b\text{-jet}} = 0, m_{jj} < 500 \text{ GeV}$

- **b-CR**

→  $N_{b\text{-jet}} > 0$

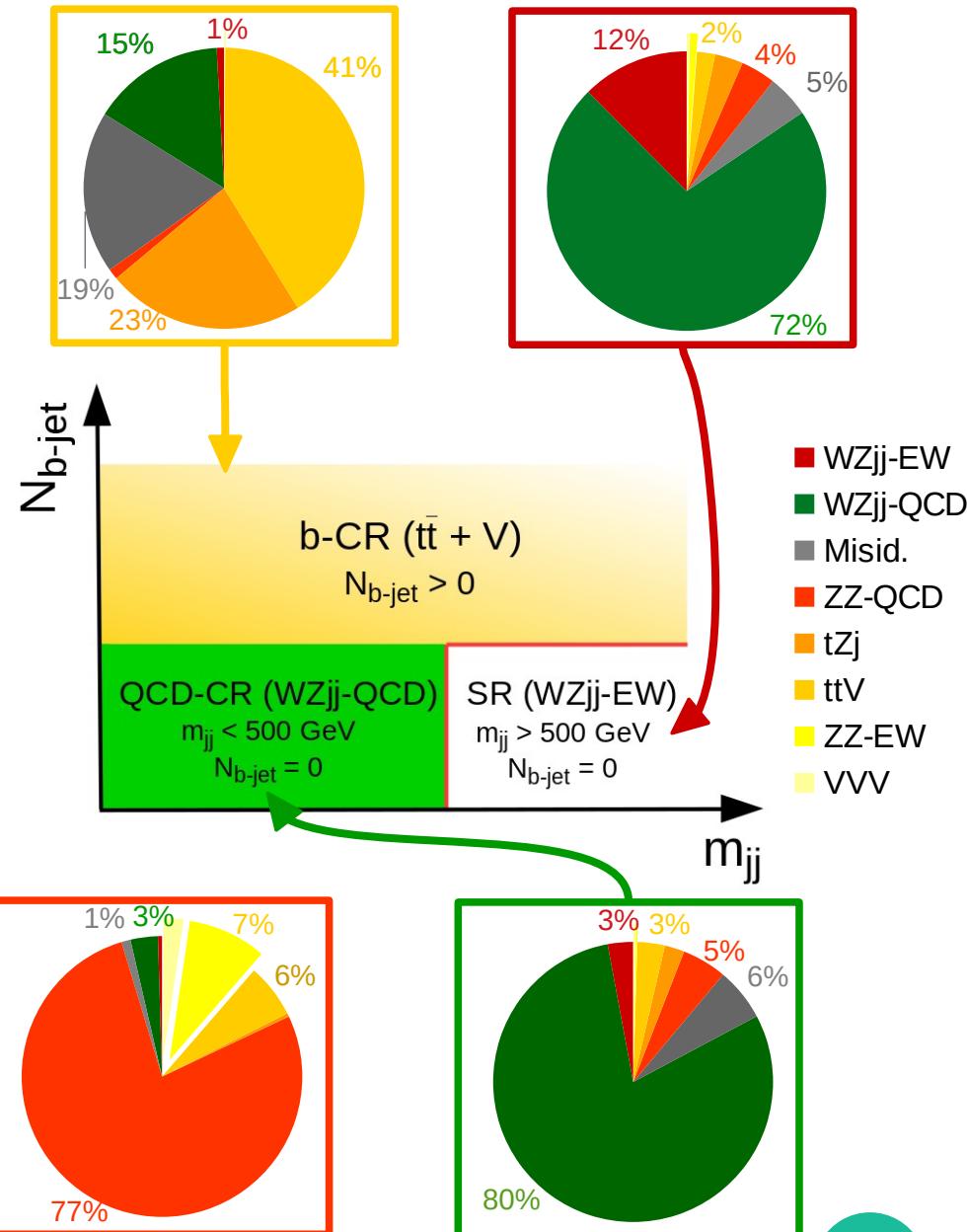
> Additional region for ZZ background

- **ZZ-CR**

→ Inverted 4<sup>th</sup> lepton veto

> Still low SR purity

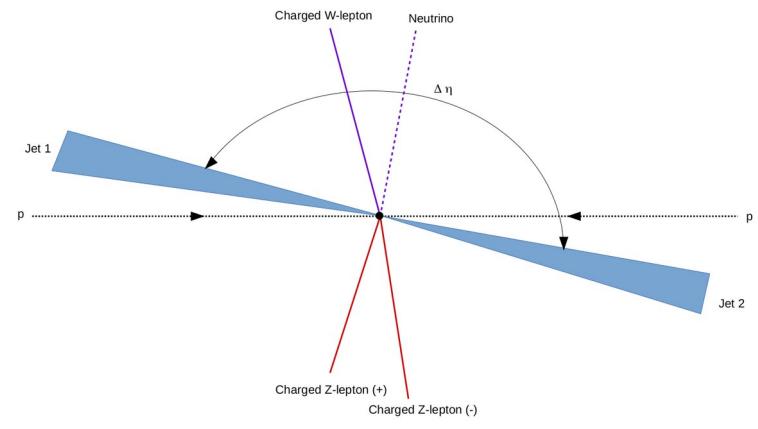
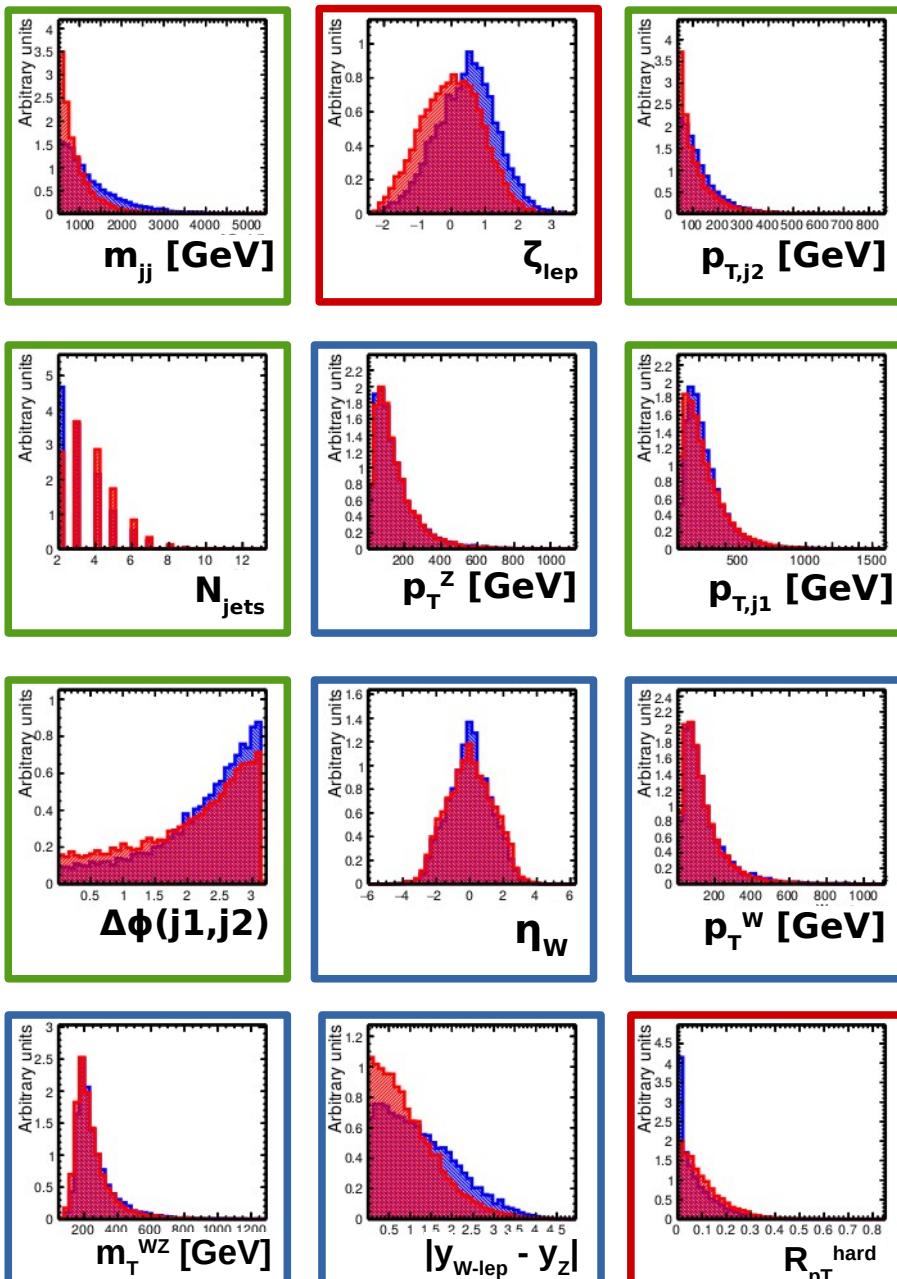
→ Multivariate discriminant



# Discriminative variables

$\sqrt{s} = 13 \text{ TeV}, 36.1 \text{ fb}^{-1}$   
WZjj-EW Signal Region

WZjj-EW  
Irred. background



## > Jet-related variables

$m_{jj}$ ,  $p_T^{j1,j2}$ ,  $N_{jets}$ ,  $\Delta\eta(j1,j2)$ ,  $\Delta\phi(j1,j2)$ ,  $n_{j1}$

## > Bosons kinematics

$p_T^Z$ ,  $p_T^W$ ,  $\eta_Z$ ,  $\eta_W$ ,  $m_T^{WZ}$ ,  $|y_{W-lep} - y_Z|$

## > "Global" variables

$\zeta_{lep}$ ,  $\Delta R(j1,Z)$ ,  $R_{pT}^{hard}$

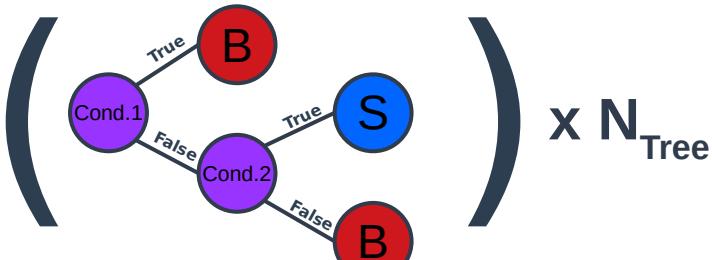
→ Low discrimination power with single variables

# Multivariate discriminant - Definition & Performance

## > Multivariate discriminant

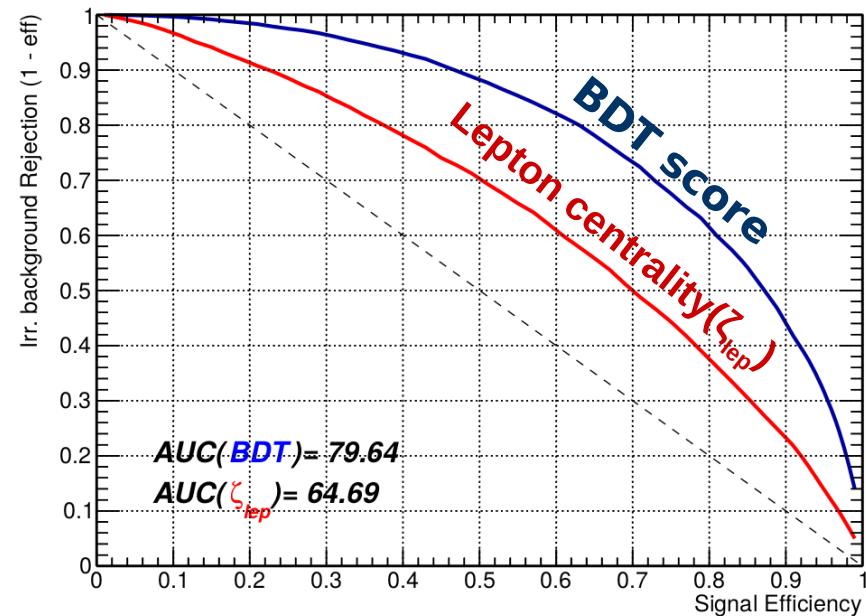
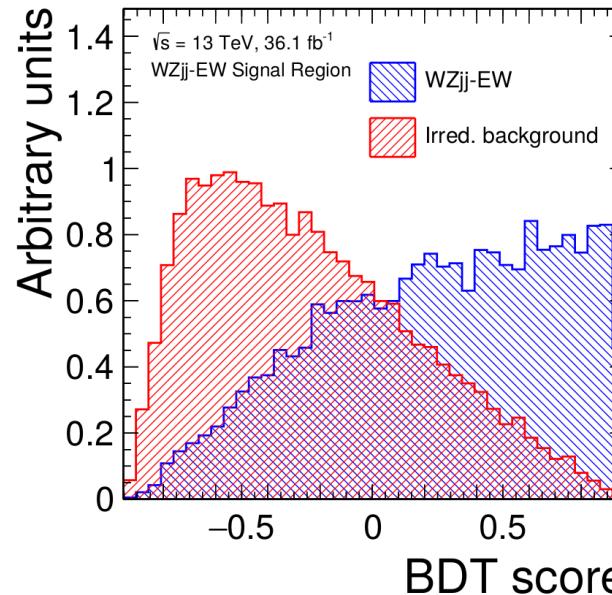
- Optimised combination of variables
- Accounts for both **single-variable discrimination** and **correlations**

## > Using Boosted Decision Trees



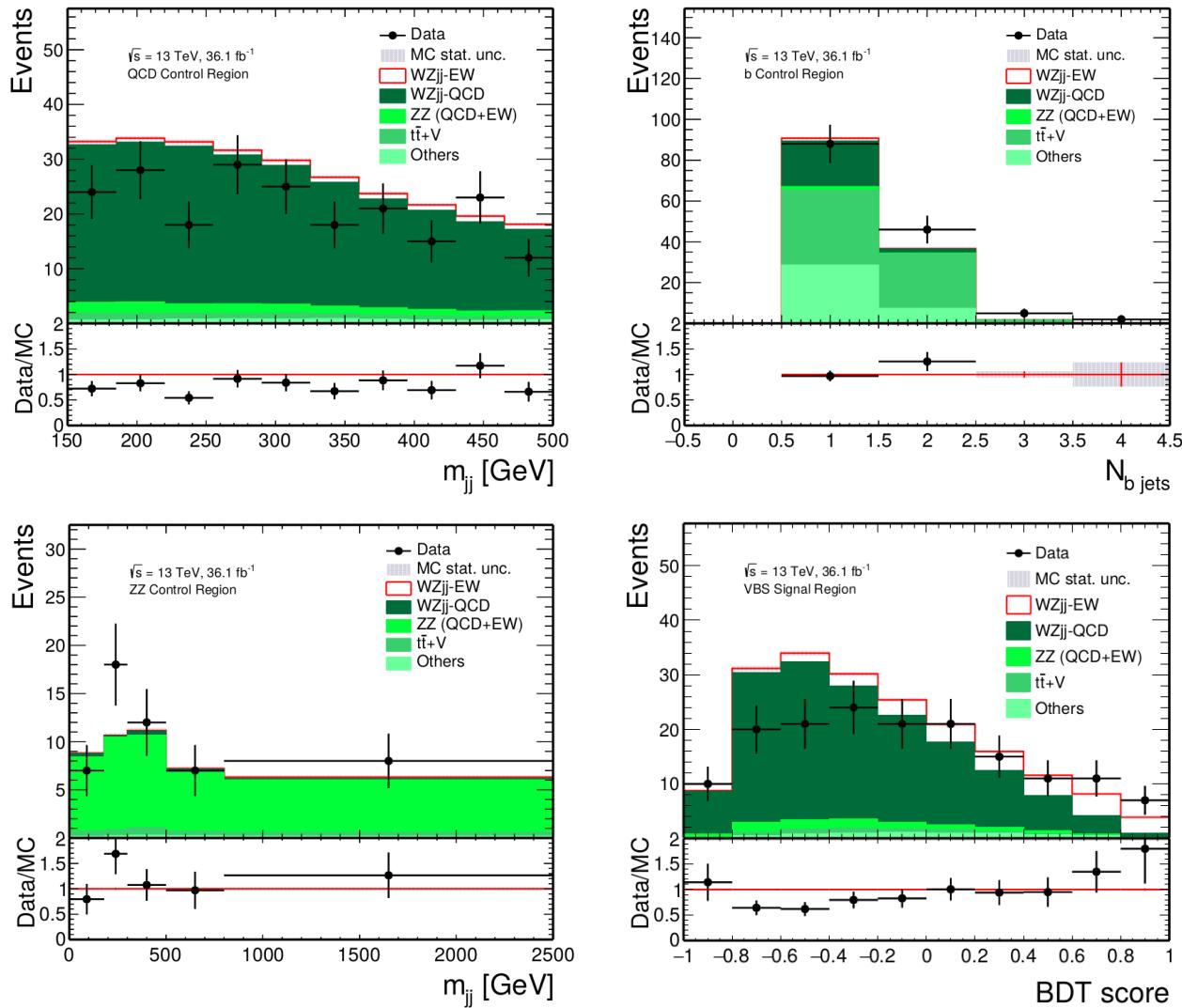
## > 23% improvement seen in MC

- Compared to lepton centrality

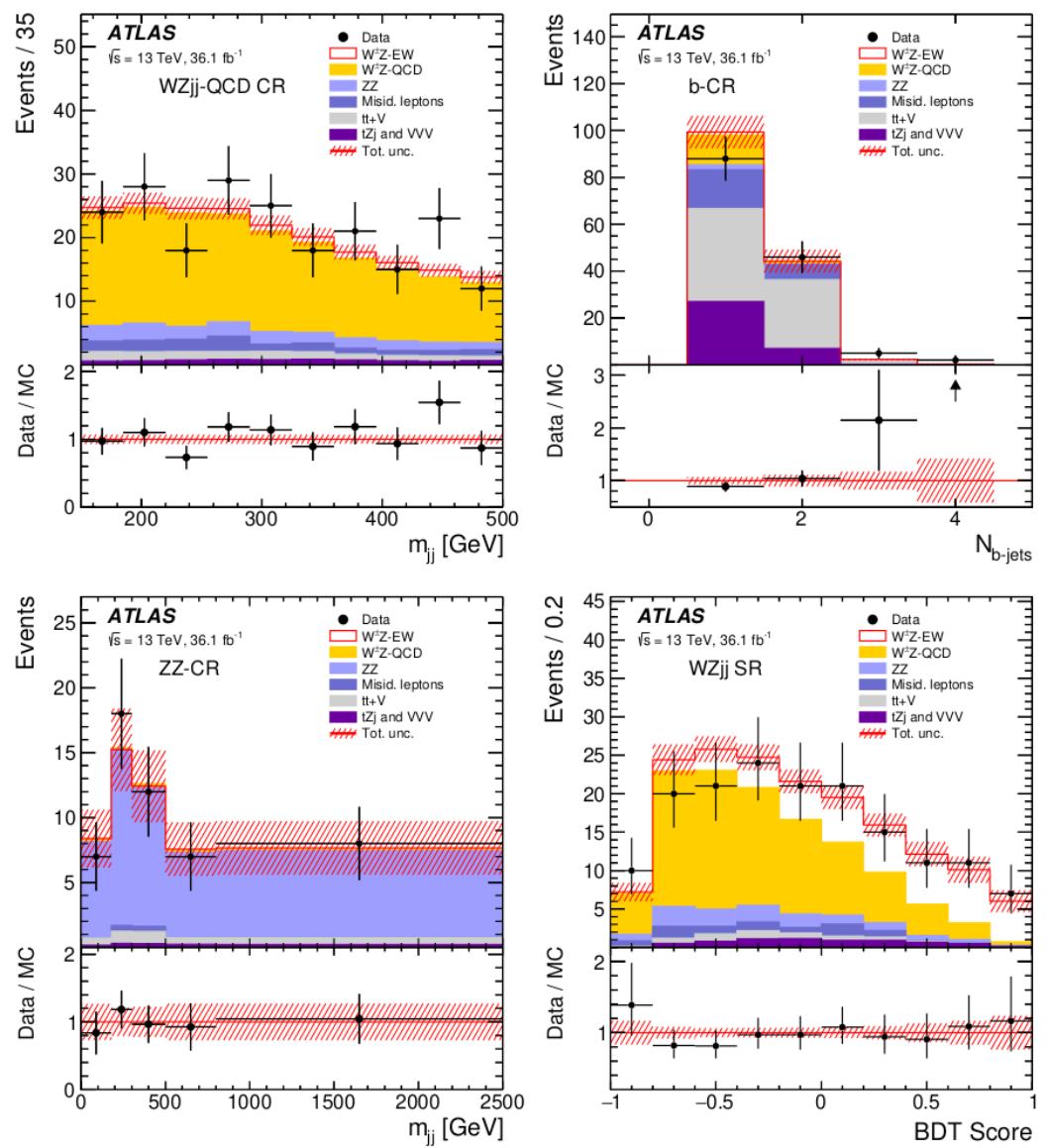


# Likelihood fit setup

- > Combined likelihood fit
- > Using the four regions
  - **m<sub>jj</sub>** in ZZ-CR and QCD-CR
  - **Nb-jets** in b-CR
  - **BDT score** in SR
- > 4 free parameters
  - Signal cross-section (**SR**)
  - WZjj-QCD, t̄t+V and ZZ-QCD corrections (**SR + CRs**)



# Fit results - Parameters constraints



## > Background correction factors

Correction factor	Value
$\mu_{WZjj\text{-QCD}}$	$0.56 \pm 0.16$
$\mu_{t\bar{t}+V}$	$1.07 \pm 0.28$
$\mu_{ZZ\text{-QCD}}$	$1.34 \pm 0.44$

## > Uncertainties

- Statistically dominated
- Non-negligible systematics impact:
  - Theory and Modelling
  - Jet reconstruction and calibration

Source	Uncertainty [%]
WZjj-EW theory modelling	4.8
WZjj-QCD theory modelling	5.2
WZjj-EW and WZjj-QCD interference	1.9
Jets	6.6
Pile-up	2.2
Electrons	1.4
Muons	0.4
b-tagging	0.1
MC statistics	1.9
Misid. lepton background	0.9
Other backgrounds	0.8
Luminosity	2.1
Total Systematics	10.7

# Fit results - Cross section measurement

> Background-only hypothesis rejected at **5.3  $\sigma$**

→ **First observation of WZjj-EW production**  
*Physics Letters B 793 (2019), 469-492*

> **WZjj-EW cross-section derived from fit result**

→ Extracted in fiducial phase-space close to SR definition

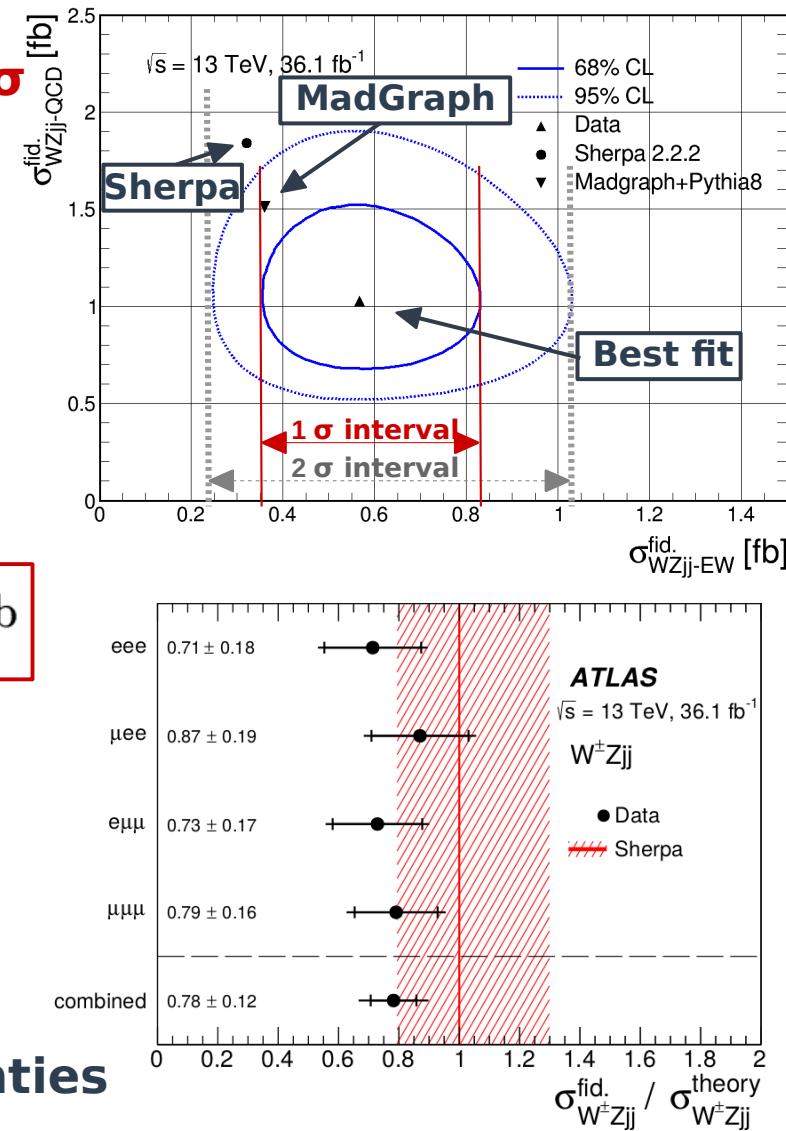
$$\sigma_{WZjj-EW}^{fid.} = 0.57^{+0.14}_{-0.13} (\text{stat.})^{+0.07}_{-0.06} (\text{syst.}) \text{ fb} = 0.57^{+0.16}_{-0.14} \text{ fb}$$

$$\sigma_{\text{Sherpa}}^{\text{fid., EW th.}} = 0.321 \pm 0.002 (\text{stat.})^{+0.005}_{-0.005} (\text{PDF})^{+0.027}_{-0.023} (\text{scale}) \text{ fb}$$

$$\sigma_{\text{MadGraph}}^{\text{fid., EW th.}} = 0.366 \pm 0.004 (\text{stat.}) \text{ fb}$$

→ **Measurement within 2  $\sigma$  of MC predictions**

> **Inclusive WZjj measurement within uncertainties**



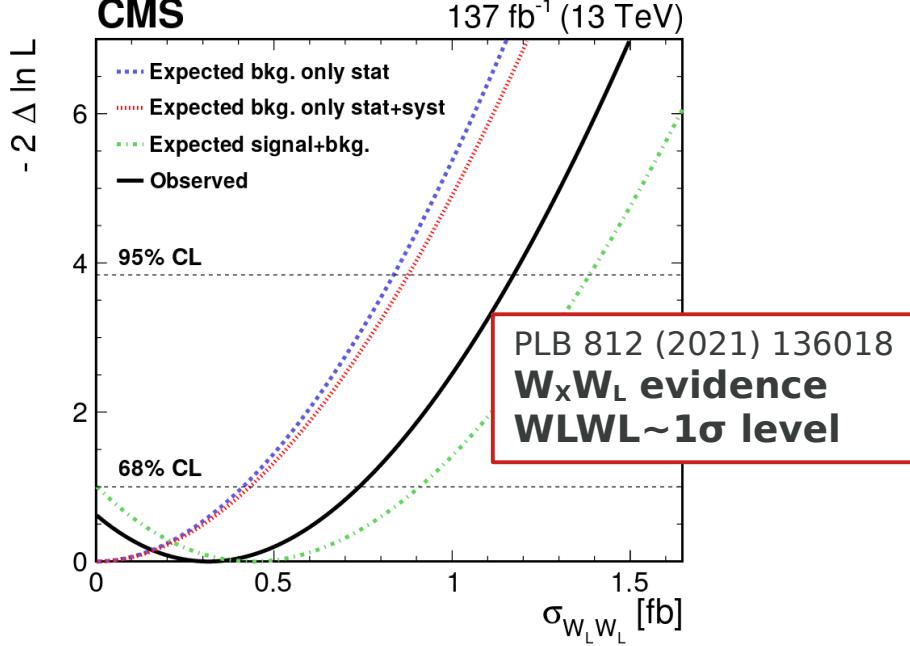
# VBS prospects

## > WZjj-EW observed with partial Run 2 dataset

- “Only” 36  $\text{fb}^{-1}$  of data
- **140  $\text{fb}^{-1}$  now available**: ongoing analysis effort

## > Looking further: Run 3, HL-LHC & beyond

- Potential for in-depth studies:  $V_L V_L$  scattering, QGC (+SM-EFT)



$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{f_i^{(6)}}{\Lambda^2} \mathcal{L}_i^{(6)} + \sum_i \frac{f_i^{(8)}}{\Lambda^4} \mathcal{L}_i^{(8)}$$

### Run 2 CMS (WZ+WW)

	Observed ( $\text{TeV}^{-4}$ )	Expected ( $\text{TeV}^{-4}$ )
$f_{T0}/\Lambda^4$	[-1.1, 1.6]	[-1.6, 2.0]
$f_{T1}/\Lambda^4$	[-0.69, 0.97]	[-0.94, 1.3]
$f_{T2}/\Lambda^4$	[-1.6, 3.1]	[-2.3, 3.8]
$f_{M0}/\Lambda^4$	[-11, 12]	[-15, 15]
$f_{M1}/\Lambda^4$	[-15, 14]	[-18, 20]
$f_{M6}/\Lambda^4$	[-22, 25]	[-31, 30]
$f_{M7}/\Lambda^4$	[-16, 18]	[-22, 21]
$f_{S0}/\Lambda^4$	[-34, 35]	[-31, 31]
$f_{S1}/\Lambda^4$	[-86, 99]	[-91, 97]

PLB 809 (2020), 135710

### HL-LHC

	14 TeV	
	$WZjj$	$W^\pm W^\pm jj$
$f_{S_0}/\Lambda^4$	[-8,8]	[-6,6]
$f_{S_1}/\Lambda^4$	[-18,18]	[-16,16]
$f_{T_0}/\Lambda^4$	[-0.76,0.76]	[-0.6,0.6]
$f_{T_1}/\Lambda^4$	[-0.50,0.50]	[-0.4,0.4]
$f_{M_0}/\Lambda^4$	[-3.8,3.8]	[-4.0,4.0]
$f_{M_1}/\Lambda^4$	[-5.0,5.0]	[-12,12]

Arxiv:1902.04070

# Content

## (Forward) pileup jet tagging

# Luminosity and Pileup

## > Physics over ~14 orders of magnitude

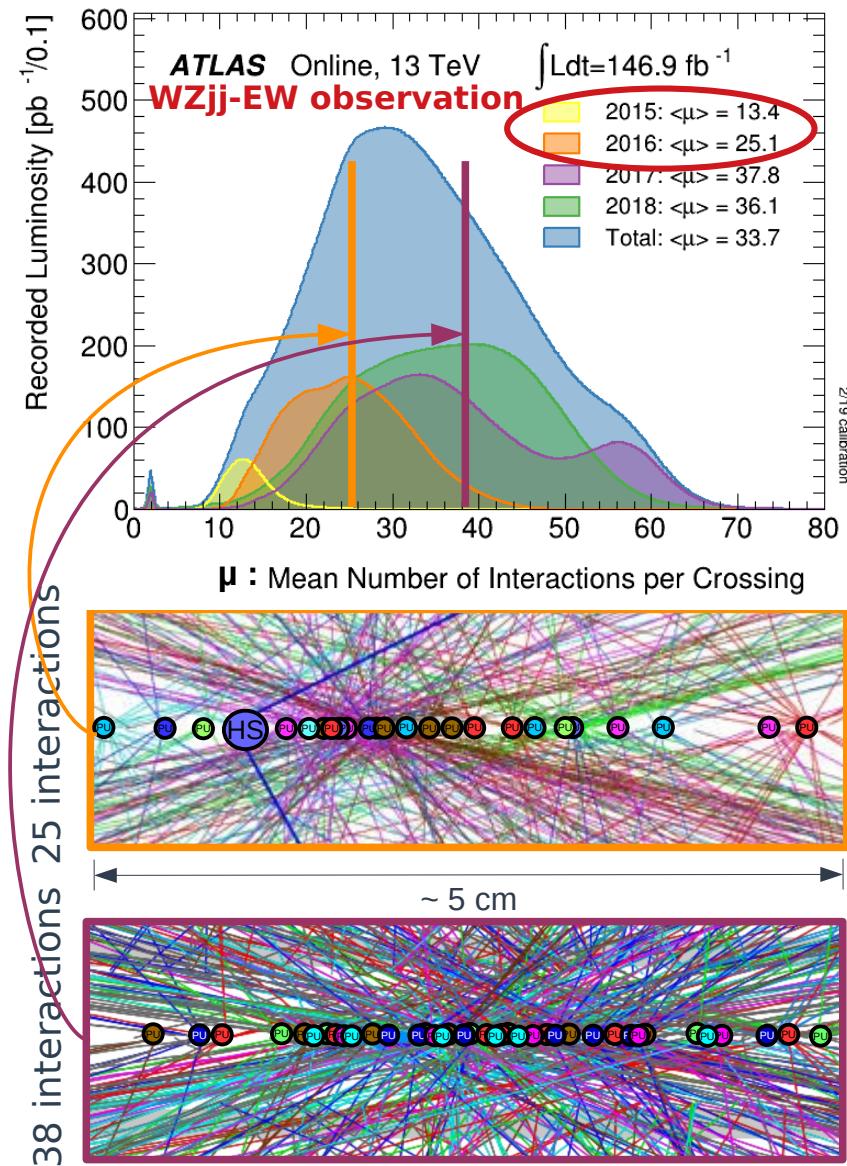
- Possible thanks to high luminosity:

$$L = \frac{N_b^2 n_b f_{rev} \gamma}{4 \pi \epsilon_n \beta^*} F \simeq 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$$

$N_b \simeq 10^{11}$  → Number of protons in a “bunch”

## > Downside: Pileup

- Up-to ~70 simultaneous collision in an event
- Can impact reconstruction performance
- Yields additional **jets from pileup**



# Pileup jets

## > HS jet (signal jets)

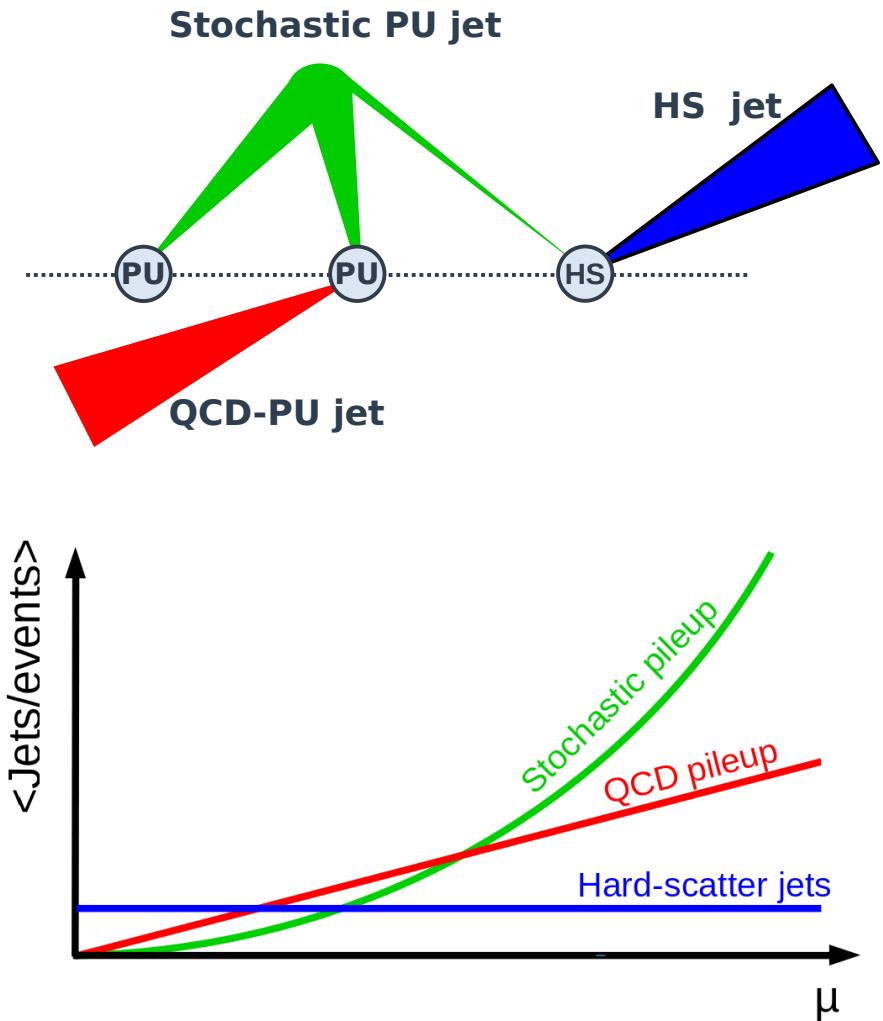
- Originates from the main interaction vertex  
→ HS vertex
- Nature depends on the targeted interaction  
→ **Independent of  $\mu$**

## > QCD-PU

- Well defined hadronic jet
- Originates from a PU vertex  
→ **Linear increase with  $\mu$**

## > Stochastic PU

- Jet composed of uncorrelated components
- Typically not associable to a single vertex  
→ **Faster than linear**



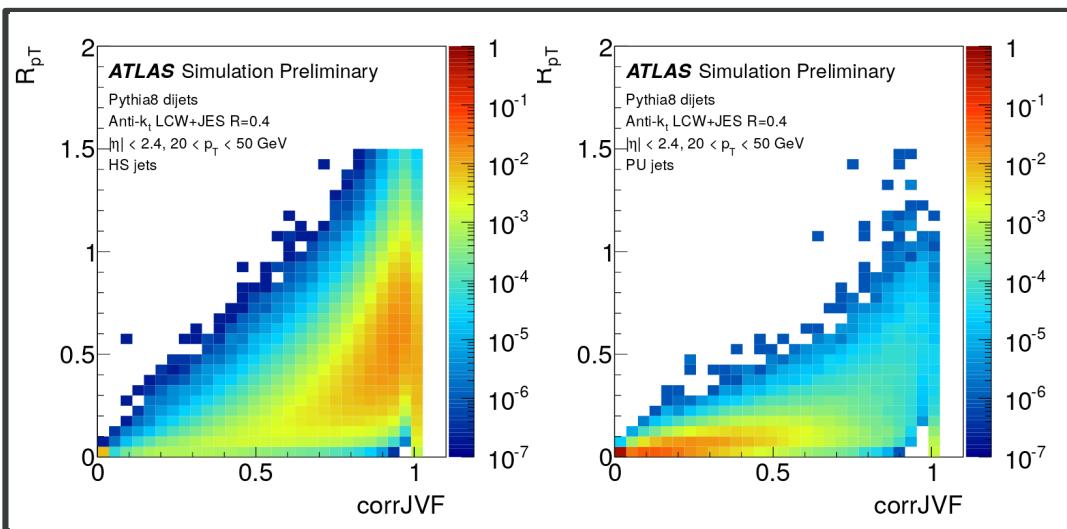
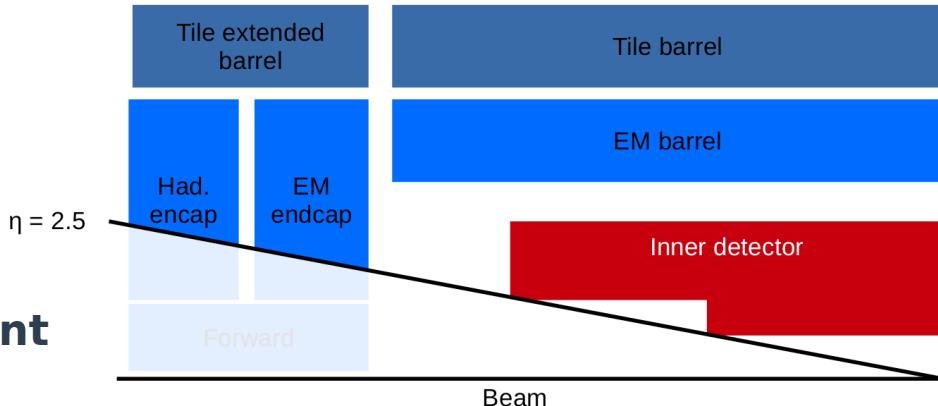
# The Jet-Vertex-Taggers - JVT

## > Pileup-jet suppression strategy

- Match jets to the interaction vertices
- Remove jets originating from PU vertices

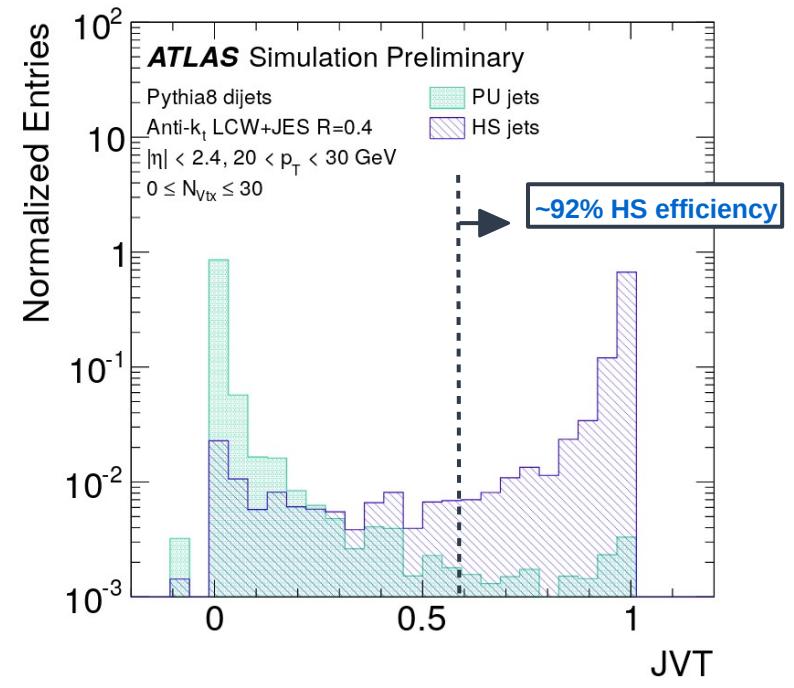
## > Within tracker acceptance ( $|\eta| < 2.5$ )

→ JVT: Highly efficient **track-based discriminant**



$$R_{pT} = \frac{\sum_k p_T^{track_k}(PV_0)}{p_T^{jet}}$$

$$corrJVF = \frac{\sum_k p_T^{track_k}(PV_0)}{\sum_l p_T^{track_l}(PV_0) + \frac{p_T^{PU}}{k \cdot n_{tracks}}}$$



# The Jet-Vertex-Taggers - fJVT

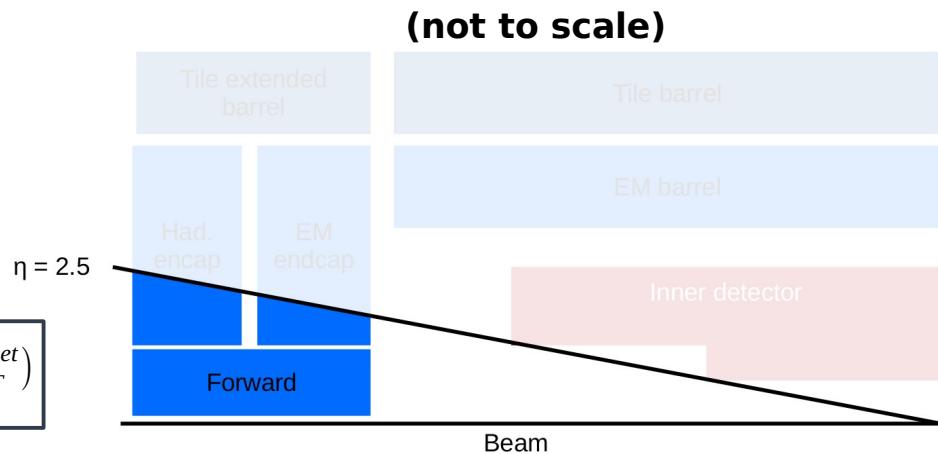
- > Outside tracker acceptance ( $|\eta| > 2.5$ )
  - **fJVT**: Match forward jets to vertices using **full-event information**

- 1 Match central jets to PU vertices (JVT,  $R_{pT}$ )

- 2 For each PU vertex:  $-p_{T,i}^{\text{miss}} = \frac{1}{2} \left( k \sum_{\text{tracks} \in PV_i} p_T^{\text{track}} + \sum_{\text{jets} \in PV_i} p_T^{\text{jet}} \right)$

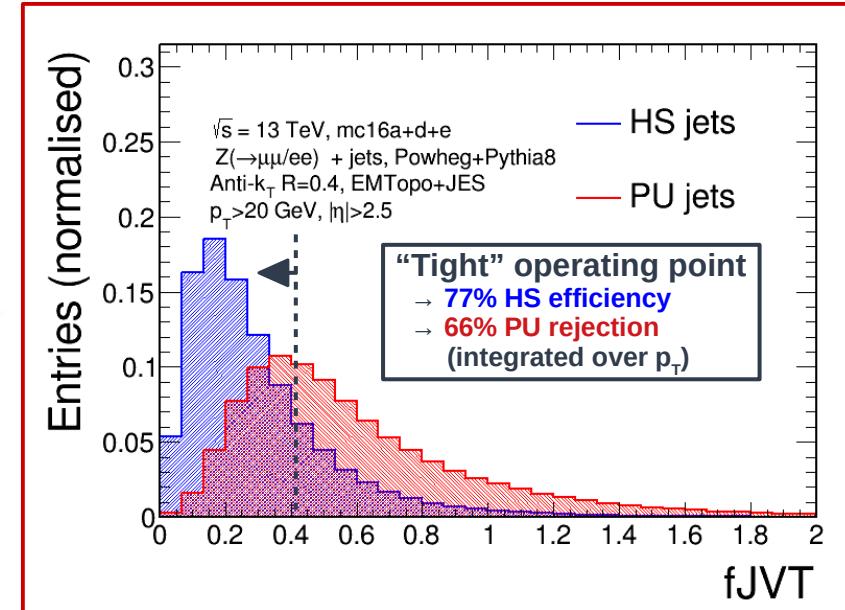
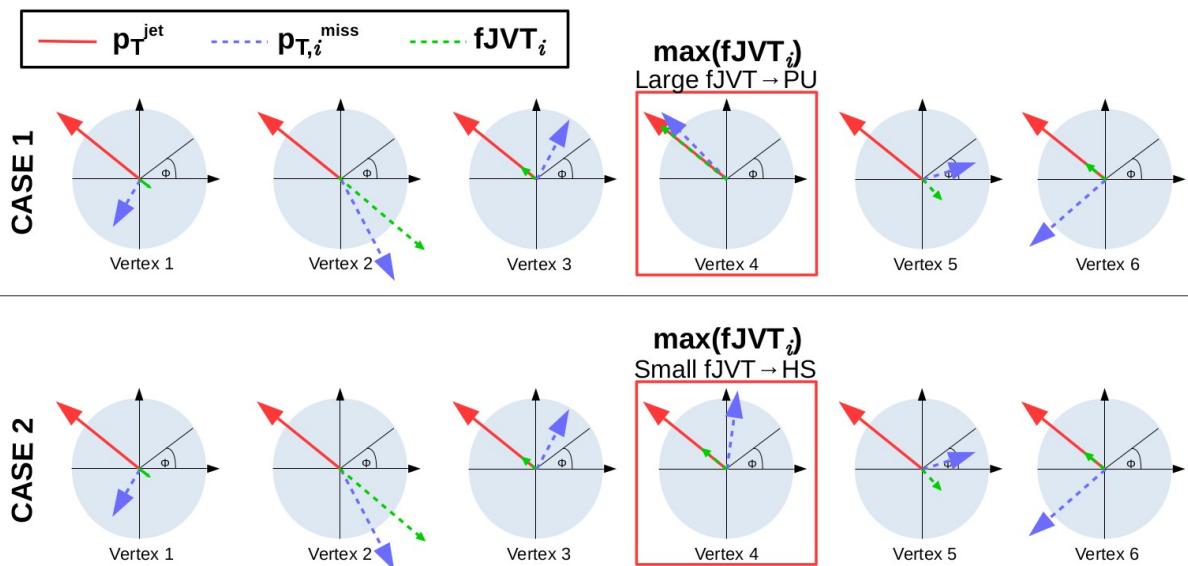
- 3 For each forward jet & PU vertex:

$$fJVT_i = \frac{-p_{T,i}^{\text{miss}} \cdot p_T^{\text{jet}}}{p_T^{\text{jet}} \cdot p_T^{\text{jet}}}$$



- 4

$$fJVT = \max(fJVT_i)$$

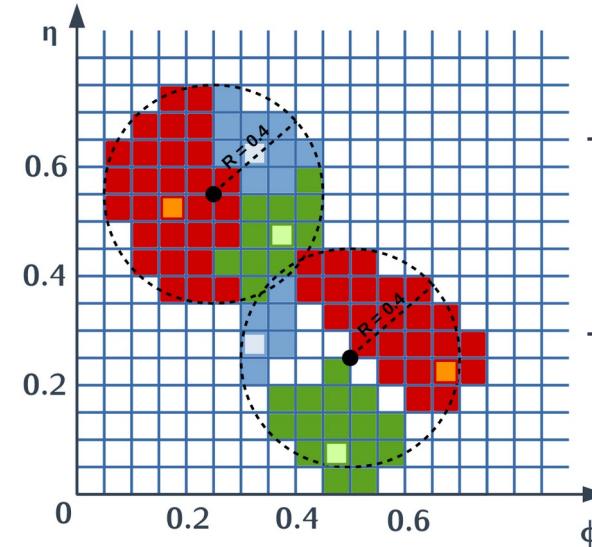


# Improving the fJVT

> **fJVT not adapted for stochastic PU**

- Other jet properties can help

> **Jet shape and structure**



> **Jet timing**

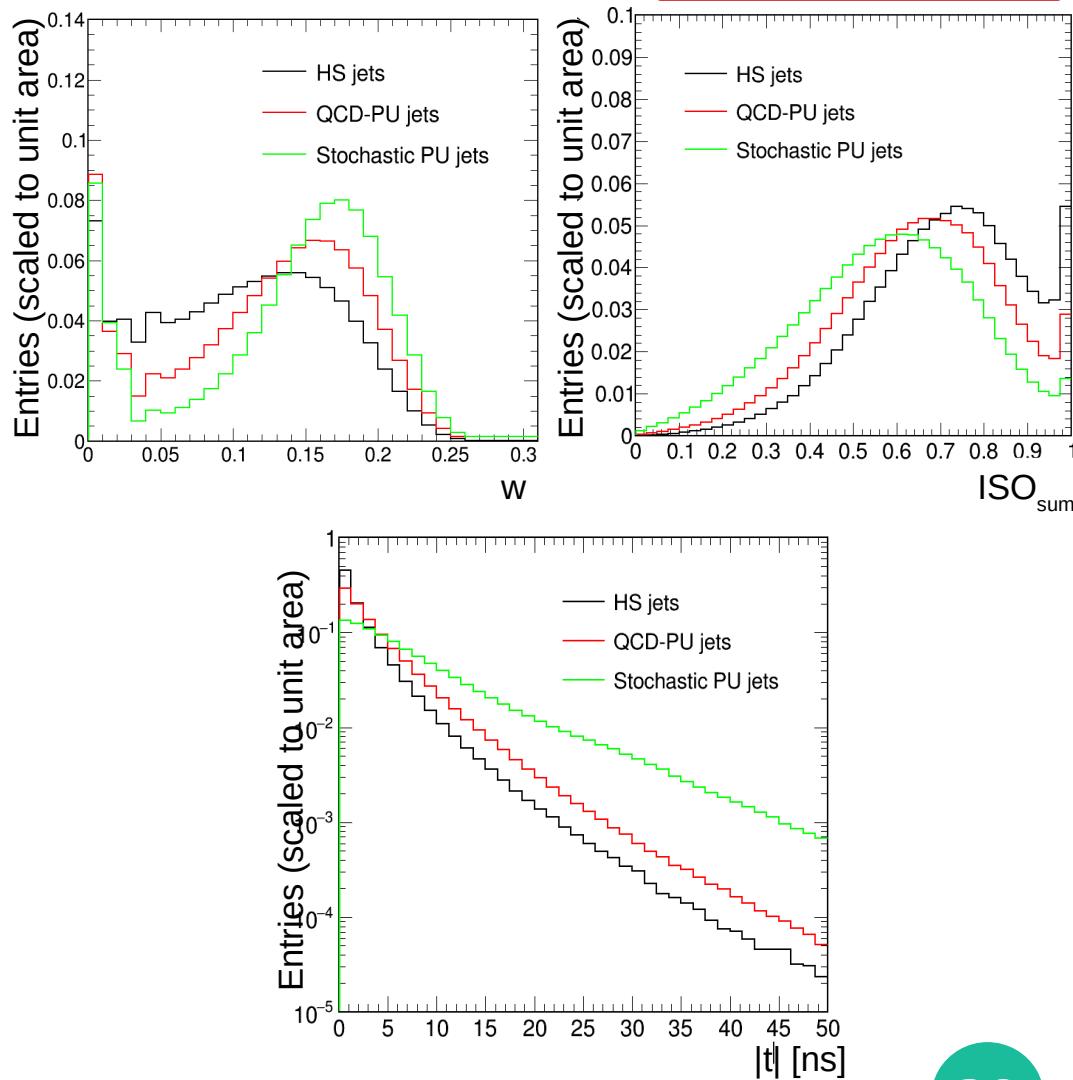
- Already used with fJVT,  $|t|<10$  ns cut

$$t_{\text{cluster}} = \frac{\sum t_{\text{cell}} E_{\text{cell}}^2}{\sum E_{\text{cell}}^2}$$

$$\rightarrow t = \frac{\sum t_{\text{cluster}} E_{\text{cluster}}^2}{\sum E_{\text{cluster}}^2}$$

$$w = \frac{\sum \Delta R(jet, cluster) p_T^{\text{cluster}}}{\sum p_T^{\text{cluster}}}$$

$$ISO_{\text{sum}} = \frac{\sum \left( \frac{N_{\text{cells}}^{\text{clustered}}}{N_{\text{cells}}^{\text{outer}}} \right) E_{\text{cluster}}^2}{\sum E_{\text{cluster}}^2}$$



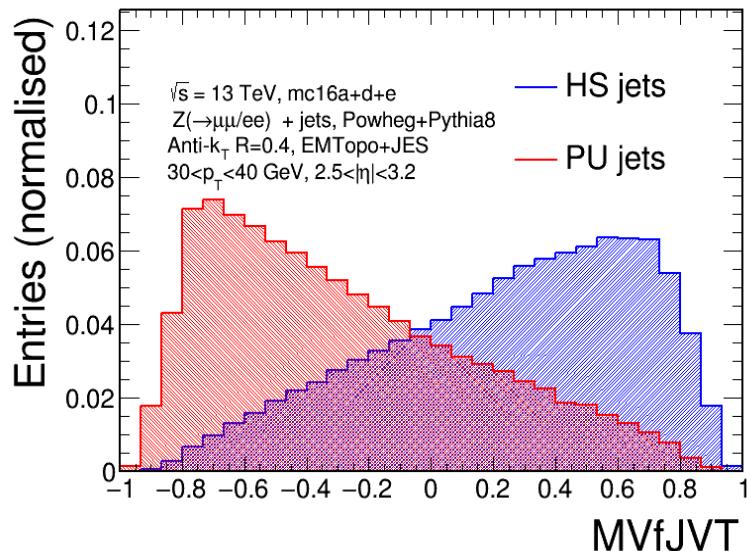
# MultiVariate fJVT

## > Multivariate combination of 8 variables

- Using **Boosted Decision Trees**
- Selected for optimal PU discrimination
- Trained for HS/PU separation  
→ No QCD/stochastic PU distinction

## > Trained using Z+jets MC

- High statistics
- Flattened  $\mu$  distribution,  $pT, |\eta|, \mu$  categories  
→ Reduced dependence
- Significant performance gain w.r.t. fJVT



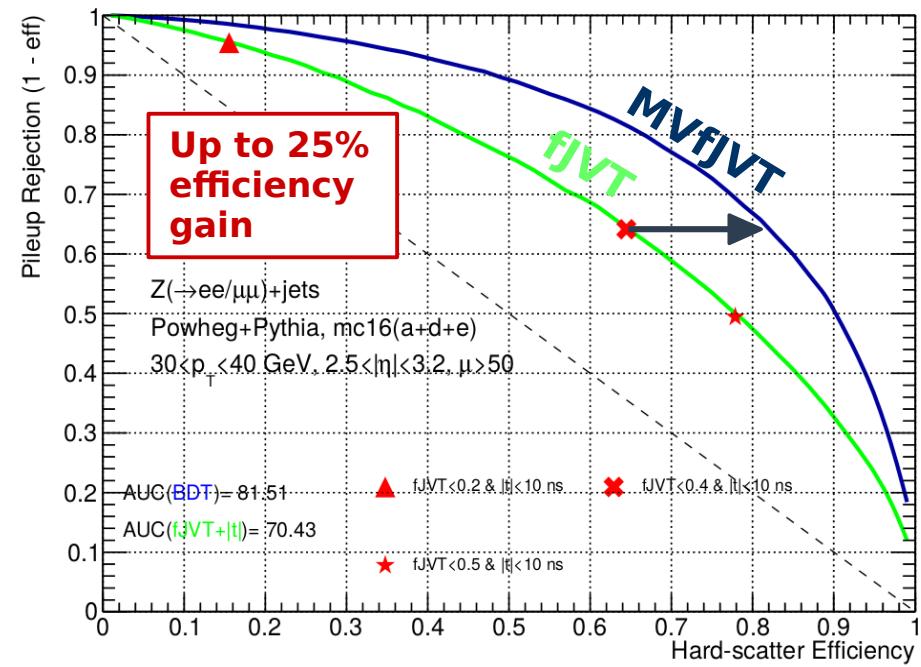
### fJVT, $|t|$ , $w$ , ISOsum

→ Provides most of the separation

$P(\text{EM})_{\text{sum}}$ ,  $E_{\text{sum}}$  (Sum over all clusters,  $\Delta R < 0.6$ )

$\langle \lambda^2 \rangle_{\text{lead}}$ ,  $\sigma_\eta^{\text{lead}}$  (Leading cluster properties)

→ Performance fine-tuning



# Efficiency correction

## > Potential data/MC performance discrepancy

- Requires MC tagging efficiency correction

## > Improved efficiency with MVfJVT

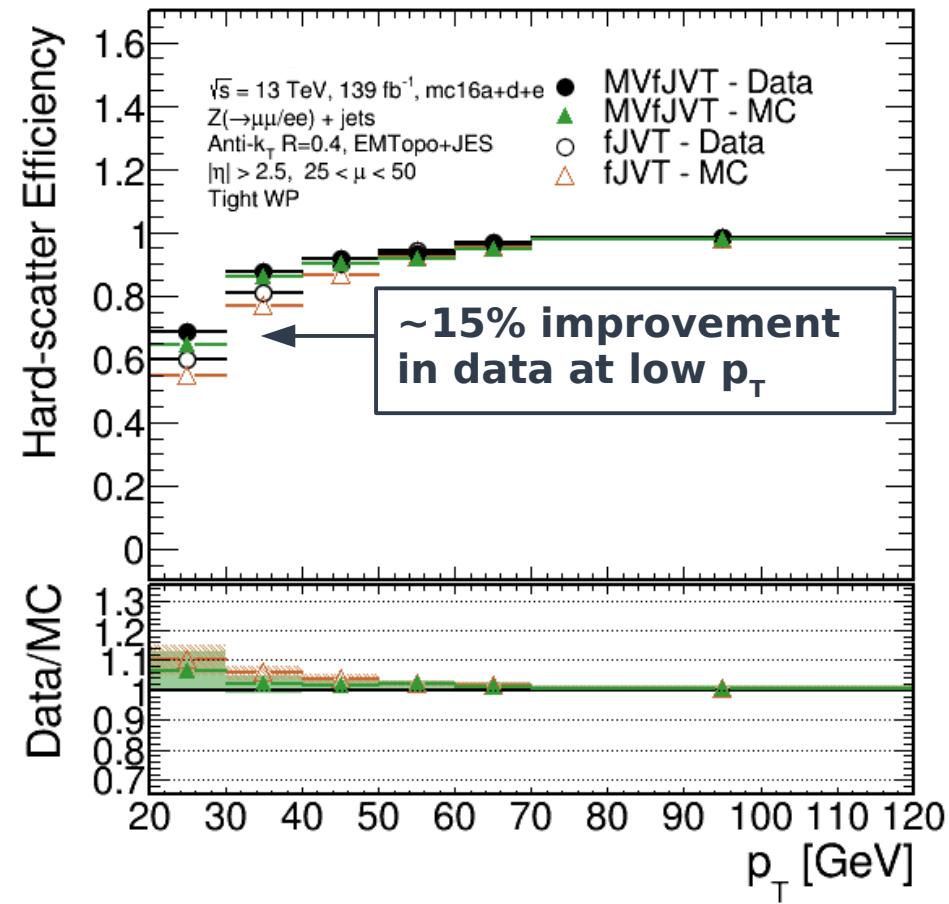
- Overall smaller corrections
- Comparable uncertainties

fJVT, Tight operating point,  $25 < \mu < 50$

$p_T$ bin [GeV]	[20,30]	[30,40]	[40,50]	[50,60]	[60,70]	[70,120]
Relative unc. [%]	4.5	2.3	1.5	0.9	1.0	0.8
Statistical [%]	0.5	0.5	0.5	0.7	0.9	0.7
PU estimate [%]	0.2	0.0	0.0	0.0	0.0	0.0
$ \eta $ -dependence [%]	1.9	0.8	0.1	0.2	0.0	0.1
Year dependence [%]	4.1	2.2	1.3	0.5	0.5	0.3

MVfJVT, Tight operating point,  $25 < \mu < 50$

$p_T$ bin [GeV]	[20,30]	[30,40]	[40,50]	[50,60]	[60,70]	[70,120]
Relative unc. [%]	6.4	2.7	1.4	1.1	1.2	0.9
Statistical [%]	0.5	0.5	0.6	0.7	0.9	0.7
PU estimate [%]	0.2	0.0	0.0	0.0	0.0	0.0
$ \eta $ -dependence [%]	2.5	1.0	0.3	0.6	0.3	0.2
Year dependence [%]	5.4	2.3	1.2	0.6	0.7	0.3
Modelling [%]	2.4	1.0	0.2	0.1	0.2	0.5



# Future improvements

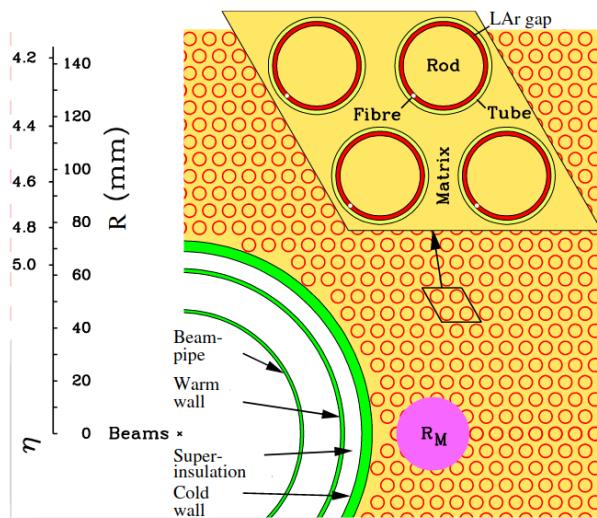
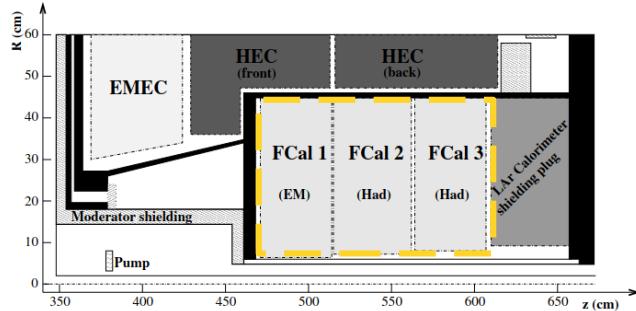
Eur.Phys.J.C 77 (2017) 9, 580

## > More pileup expected in future LHC runs

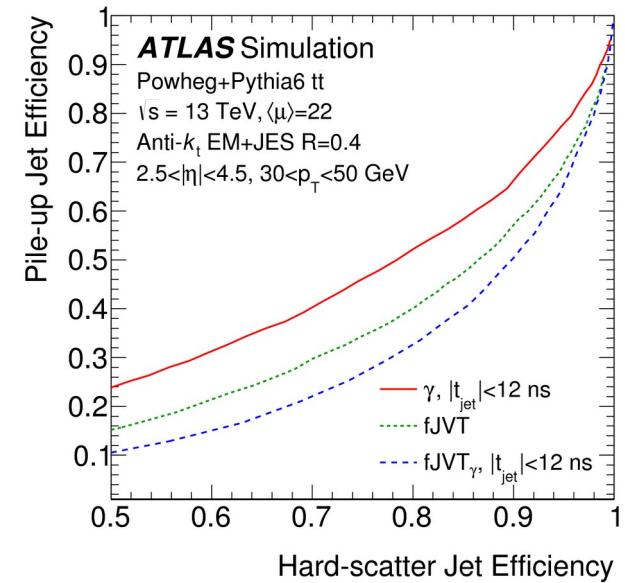
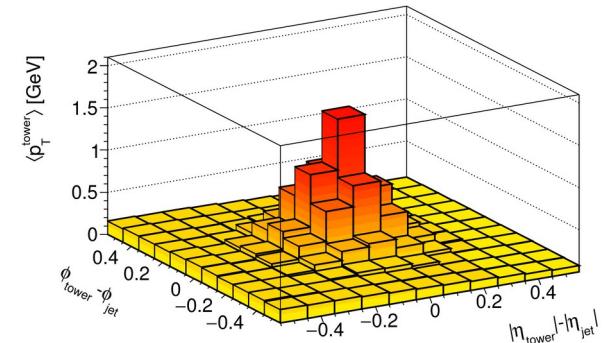
- Up to  $\langle\mu\rangle \sim 200$  at HL-LHC
- Developments ongoing to dampen the impact

## > Run 3: Calorimeter towers ?

- Fixed-grid ( $\Delta\eta \times \Delta\phi$ ) cell projection
- Improves granularity for  $|\eta| > 3.2$  (non-projective towers)
- “Proof of concept” in original fJVT paper



**ATLAS Simulation**  
Powheg+Pythia6 tt  
 $\sqrt{s} = 13 \text{ TeV}, \langle\mu\rangle=22$   
Anti- $k_t$  EM+JES R=0.4  
 $2.5 < |\eta| < 4.5, 20 < p_T < 50 \text{ GeV}$



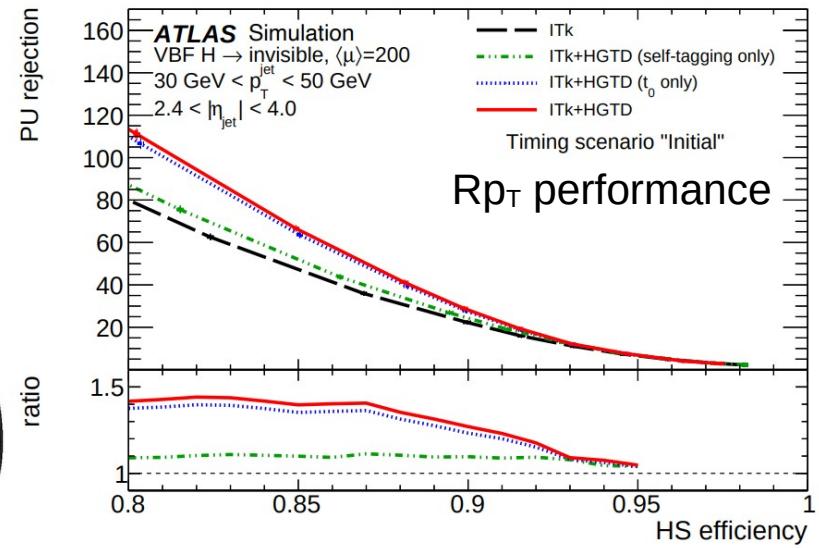
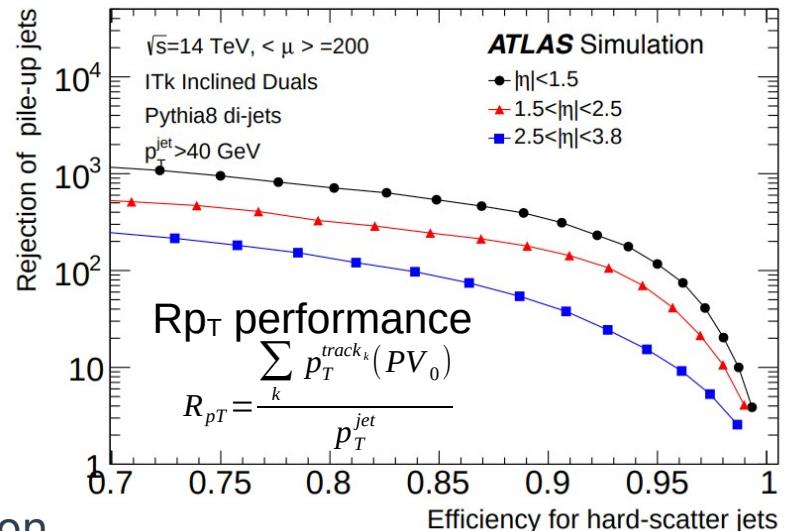
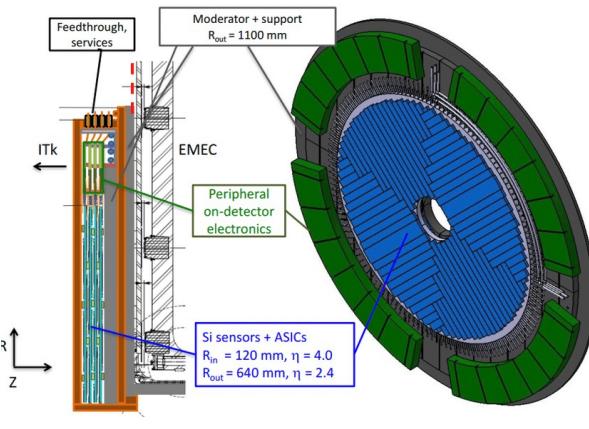
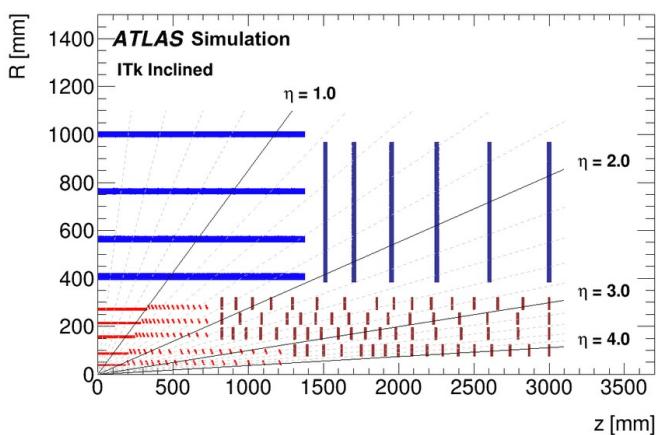
# Future improvements

## > More pileup expected in future LHC runs

- Up to  $\langle \mu \rangle \sim 200$  at HL-LHC
- Developments ongoing to dampen the impact

## > HL-LHC: New detectors

- ITk (Extended tracker), and HGTD
- Track-based identification up-to  $|\eta|=4$
- Improved out-of-time (/stochastic) pileup suppression
- Great R&D potential for further improvements



# Conclusion

## Vector Boson Scattering

- Observation of WZjj-EW production presented
- Illustrates typical challenges of VBS analyses
  - Low signal yield, large & poorly characterised backgrounds
  - Efficiently handled through use of MV method and clever fitting
- Rapidly evolving topic
  - 3 leptonic channels now observed, many more being studied
  - Active implication from theorists (NLO computations, SM-EFT)
- Bright future ahead
  - Possible evidence of  $V_L V_L$  scattering at HL-LHC
  - Tremendous QGC & SM-EFT constraints improvement expected

# Conclusion

## Forward pileup jets suppression

- Current main tools in ATLAS described
- Recent inclusion of shape-based pileup-jet taggers
  - Allows discrimination potential for stochastic pileup jets
  - Overall performance improved w.r.t. “topology-only” tagger
- Several developments to come
  - Near future: calorimeter towers for higher granularity
  - HL-LHC: extended tracking and potential timing detector
  - Active R&D about how to best use these new tools