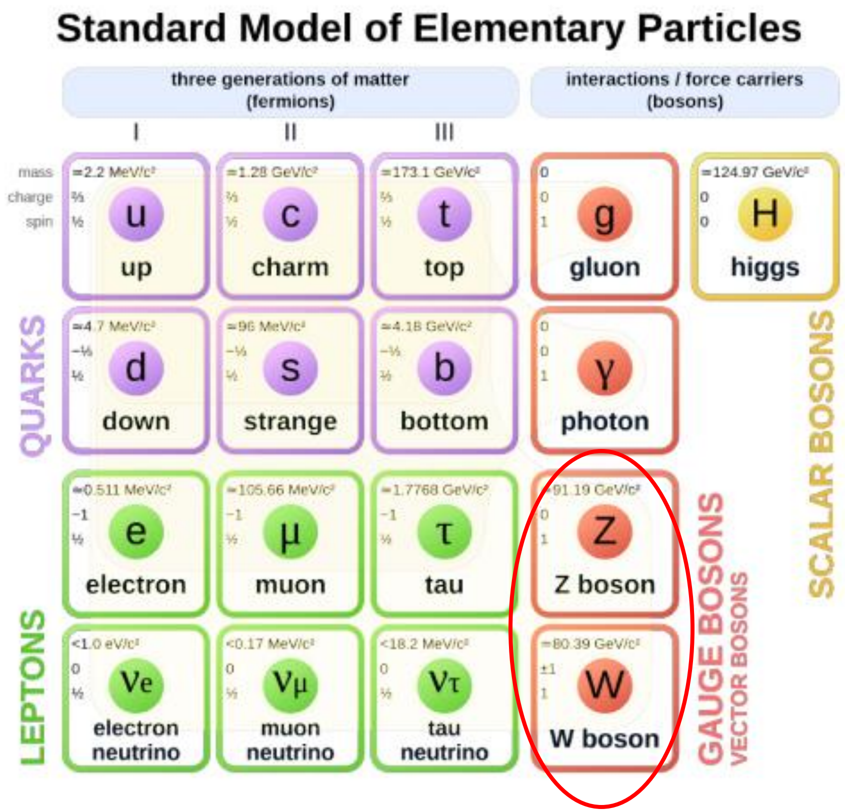
Optimization of the selection of WZjj VBS events in the fully leptonic channel using a cut-based approach

GraSPA Summer School
July 21st, 2026

Enzo Durand

First … Some theory

- One of the objective of the ATLAS experiment is to probe the **electroweak sector**



- The electroweak theory unifies the weak interaction with the electromagnetic interaction

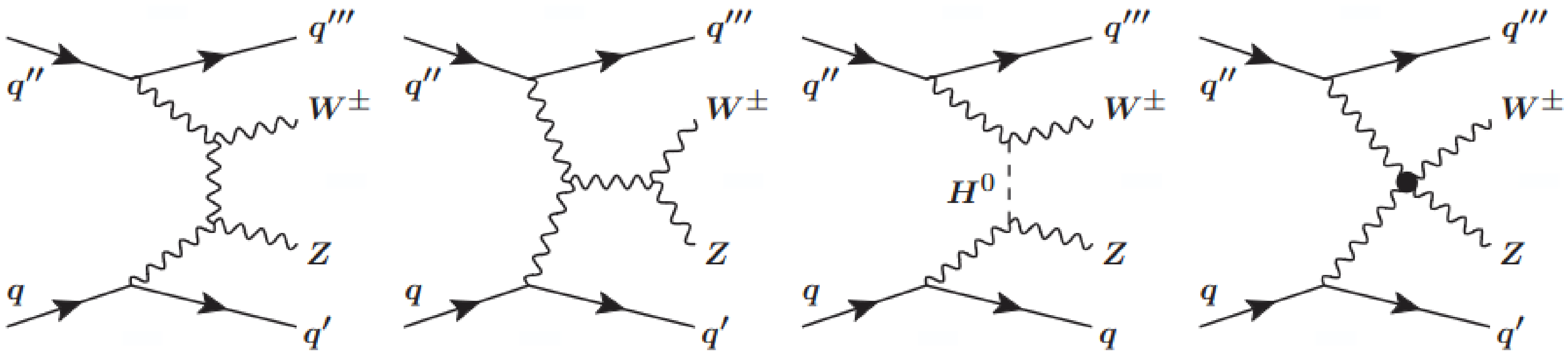
$$SU(2)_L \otimes U(1)_Y$$

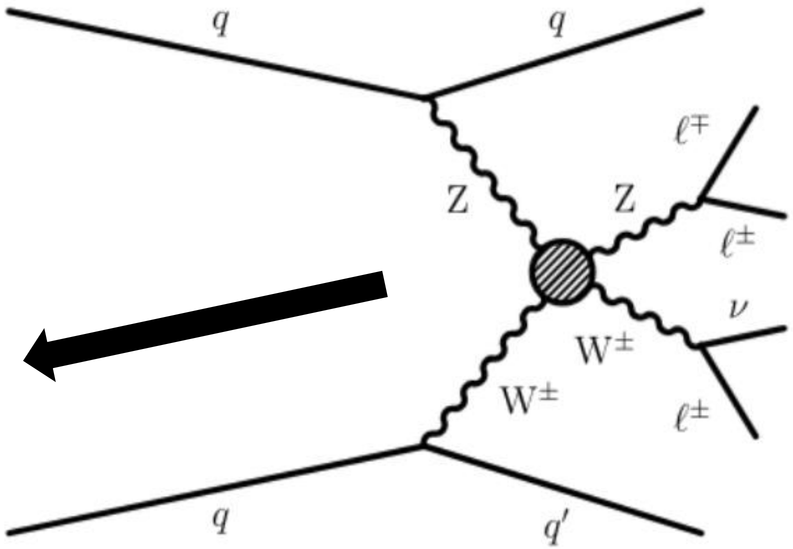
- This group has four massless gauge bosons and any mass term in the lagrangian would break the symmetry
- **Brout-Englert-Higgs mechanism**: spontaneously breaks the symmetries of the group so that a mass term appears for the three bosons of the weak interaction
- **Nambu-Goldstone theorem**: the spontaneous breakdown of symmetry creates three massless bosons
  → the Goldstone bosons disappear through a gauge transformation

- Interactions between electroweak gauge bosons arise from the Yang-Mills lagrangian
  → Every coupling with W± and Z bosons respecting charge conservation are allowed
  → **Triple** and **quartic couplings**
- New phenomena beyond the Standard Model can alter the couplings of vector bosons!

Vector Boson Scattering (VBS)

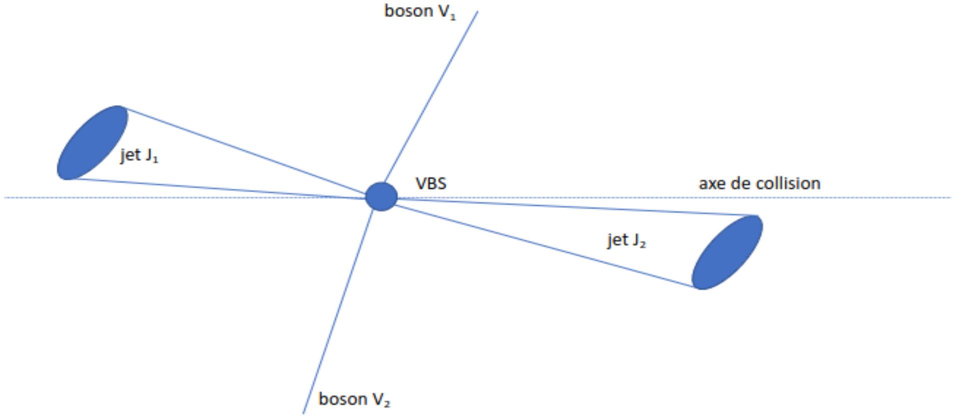
- WZ electroweak in the fully leptonic channel, denoted WZjj → lllvjj
- Clear signature: 3 leptons (electrons or muons), 2 tagging jets and 1 neutrino





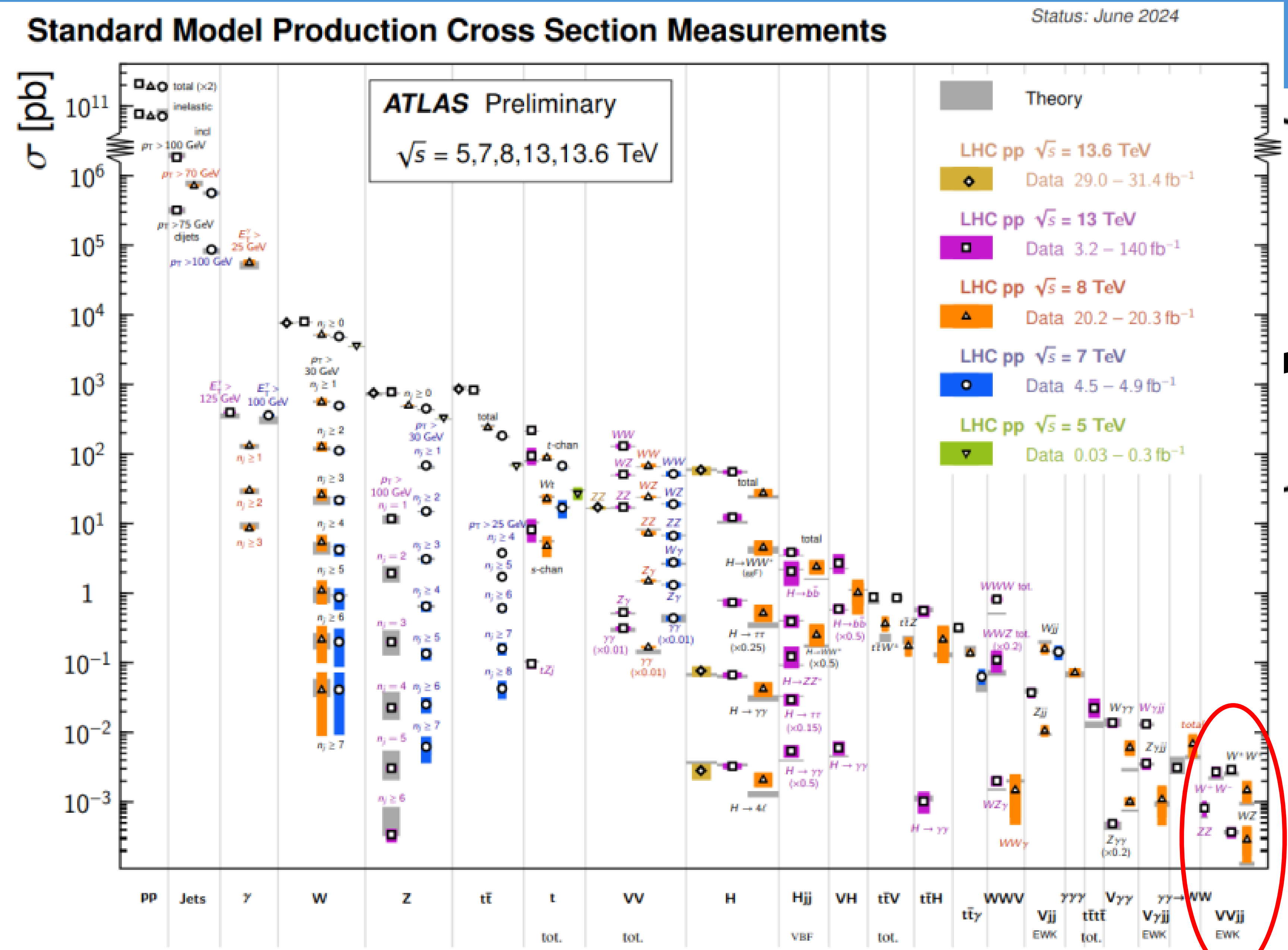
Encompasses several processes!

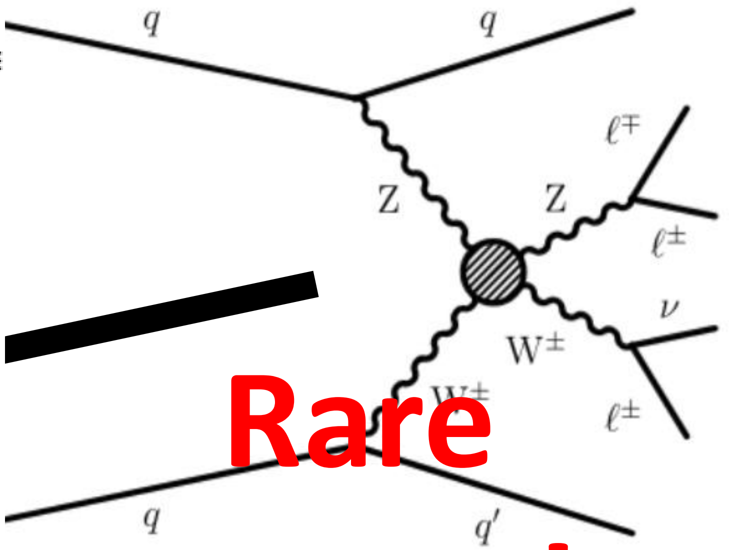
- Diboson electroweak production associated with two jets is not limited to VBS events
- VBS topology is kinematically clear



Ve

- W
- Cl

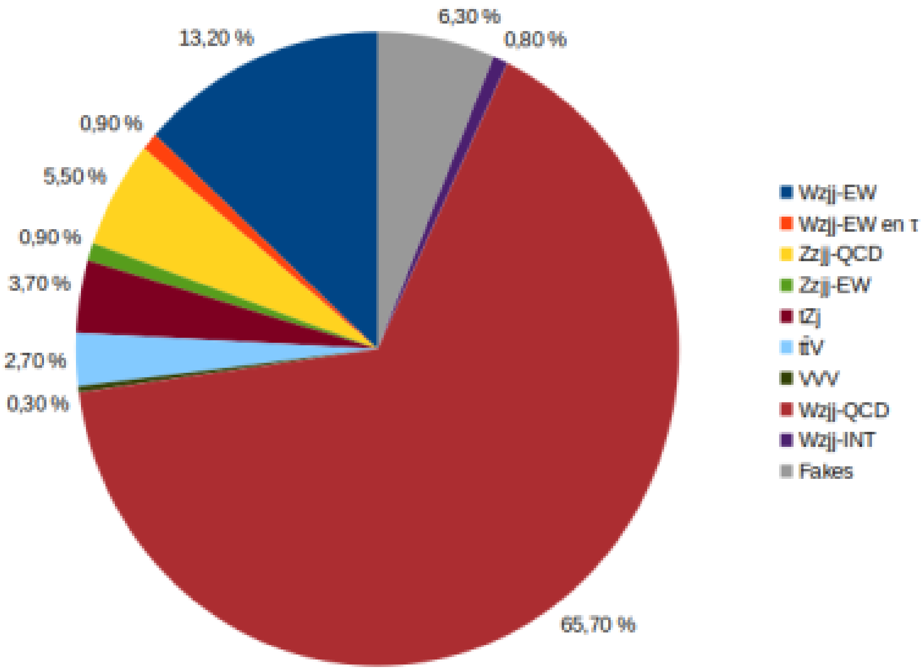


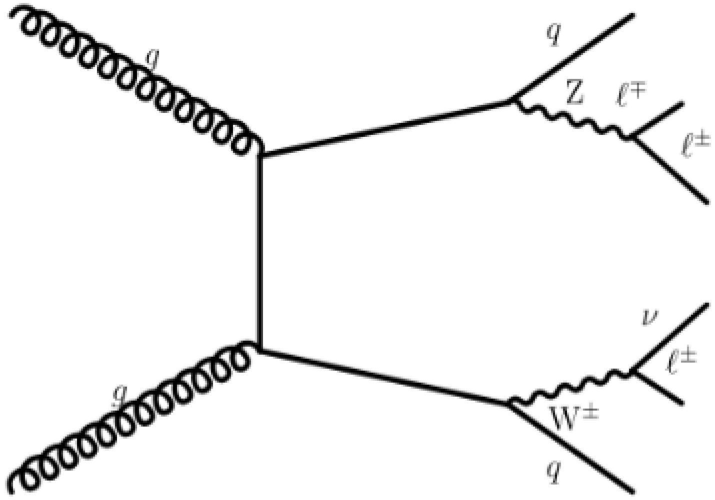


Rare events!

Experimental difficulties

- Among all diboson processes, WZjj is an experimental compromise... which comes at a price: many background contributions
  → Irreducible background (same final state) contributions: major one coming from the WZjj-QCD
  → Several reducible background (mis-reconstructed/identified objects) contributions





- <u>Objective</u>: exploit the kinematic signature and properties of the different processes to discriminate the signal versus background events

Step by step refinement of the phase space

- Different physics processes populate different regions of the detector
- Signal and background contributions exhibit different kinematic signatures

  → Use the observables (angles, momenta, … of particles) to restrict the high-dimensional phase space to a region of interest

BUT no single observable is sufficient to separate signal vs. background contributions

Observable	Criteria
Jet multiplicity	≥ 2
p_T tagging jets	> 40 GeV
$\lvert\eta_j\rvert$ tagging jets	< 4.5
$\eta_{j_1} \cdot \eta_{j_2}$	< 0
m_{jj}	> 500 GeV
$\lvert\eta\rvert$ lepton	< 2.5
$p_T^{\ell_{Z_{1,2}}}$, $p_T^{\ell_W}$	> 15, > 20 GeV
Z mass window	$\lvert m_{\ell\ell} - m_Z^{\mathrm{PDG}}\rvert < 10$ GeV
m_T^W	> 30 GeV
$N_{\mathrm{b-jet}}$	$= 0$

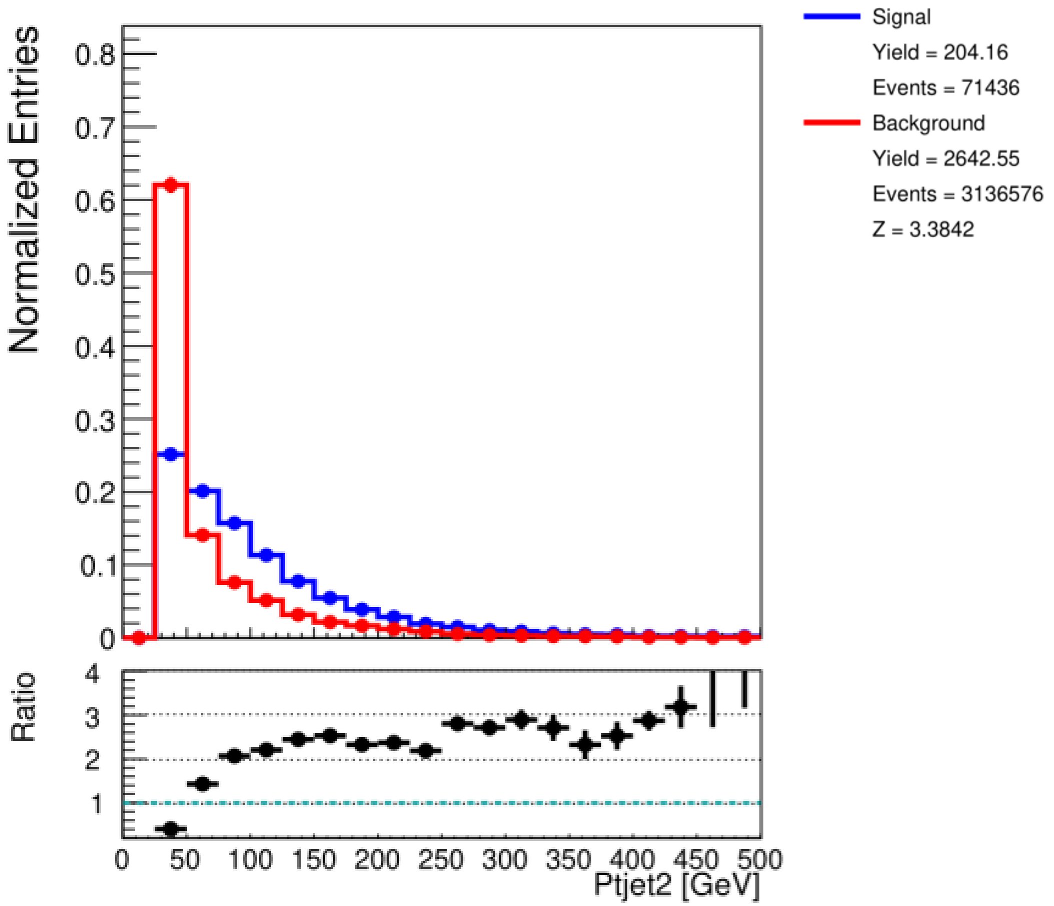
→

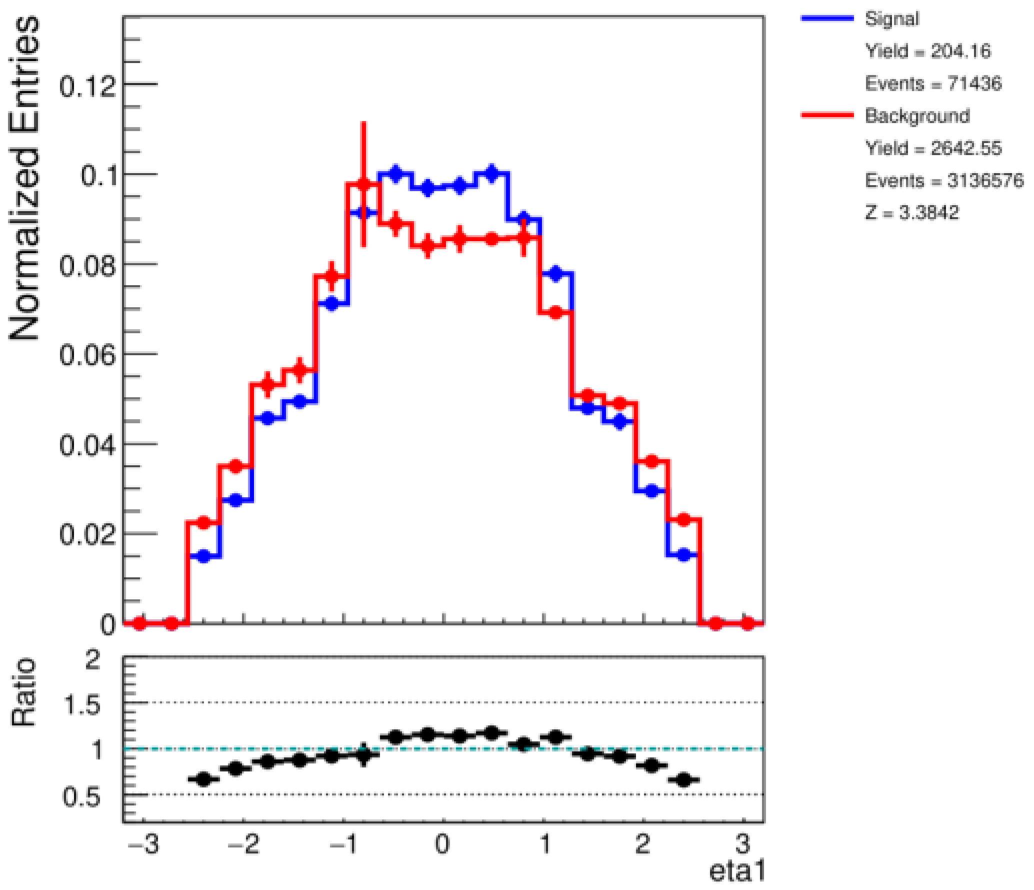
Process	Baseline selection
$WZjj$-EW	141
$WZjj$-QCD	698
$WZjj$-EW τ	8
VVV	4
ZZ_{jj}-EW	15
ZZ_{jj}-QCD	51
tZj	23
$t\bar{t}V$	21
$Z\ell\ell$	37
Fakes	25
Total MC	**1022**
Significance	**3.9**

- <u>Objective</u>: study other configurations and optimize the cut thresholds to maximize the significance

Results

- **First**: looking at the observables' distribution, some look discriminating while others do not





Results

- **Second**: optimize the cut threshold on the selected discriminating observables
  → Use a Bayesian optimization algorithm to find a near optimal configuration

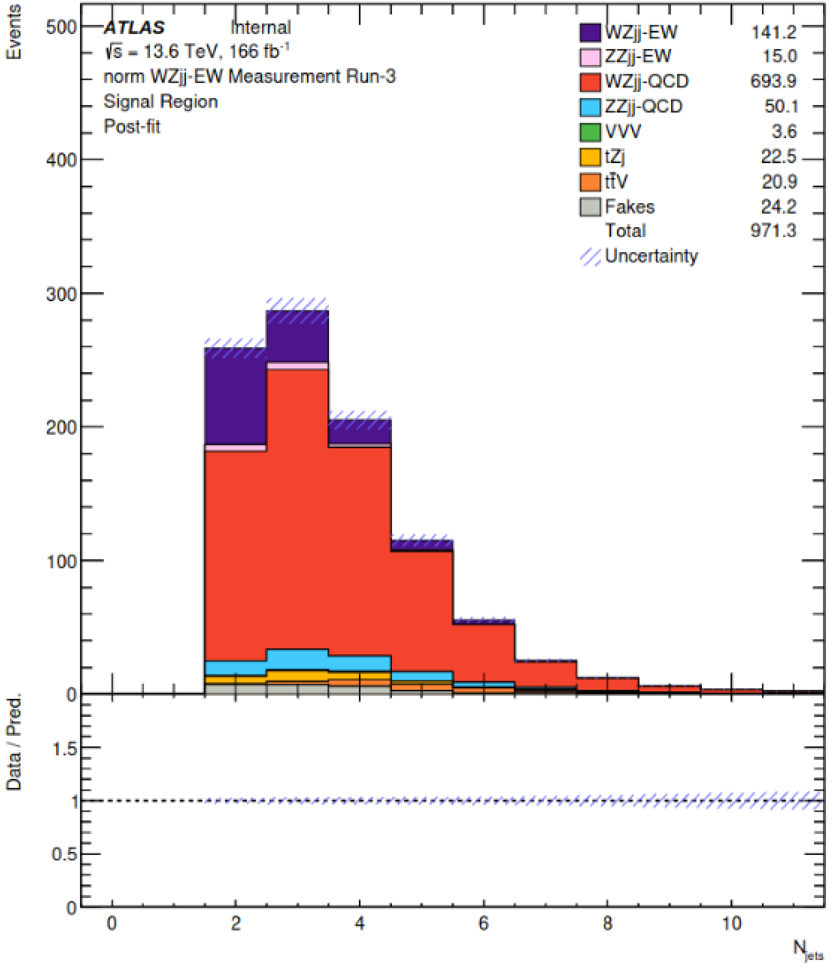
Observable	Nominal selection	Enlarged selection	Optimized selection
$p_T^{j_1}, p_T^{j_2}$	$> 40, > 40$ GeV	$> 20,\ 20$ GeV	$> 65,\ 45$ GeV
$\lvert\eta_j\rvert$ *tagging* jets	< 4.5	< 4.5	< 4.5
$\eta_{j_1} \cdot \eta_{j_2}$	< 0	< 0	< 0
m_{jj}	> 500 GeV	> 400 GeV	> 800 GeV
$p_T^{\ell_{Z_1}}, p_T^{\ell_{Z_2}}, p_T^{\ell_W}$	$> 15, > 15, > 10$ GeV	$> 10, > 10, > 10$ GeV	$> 10, > 10, > 10$ GeV
Z mass window	$\lvert m_{\ell\ell} - m_Z\rvert < 10$ GeV	$\lvert m_{\ell\ell} - m_Z\rvert < 10$ GeV	$\lvert m_{\ell\ell} - m_Z\rvert < 10$ GeV
m_T^W	> 30 GeV	> 20 GeV	> 20 GeV
$N_{\mathrm{b-jet}}$	$= 0$	$= 0$	$= 0$

Process	Nominal selection	Optimized selection
$WZjj$-EW	141	107
$WZjj$-QCD	698	315
$WZjj$-EW τ	8	6
VVV	4	2
ZZ_{jj}-EW	15	11
ZZ_{jj}-QCD	51	19
tZj	23	13
$t\bar{t}V$	21	9
$Z\ell\ell$	37	11
Fakes	25	13
Total MC	**1023**	**506**

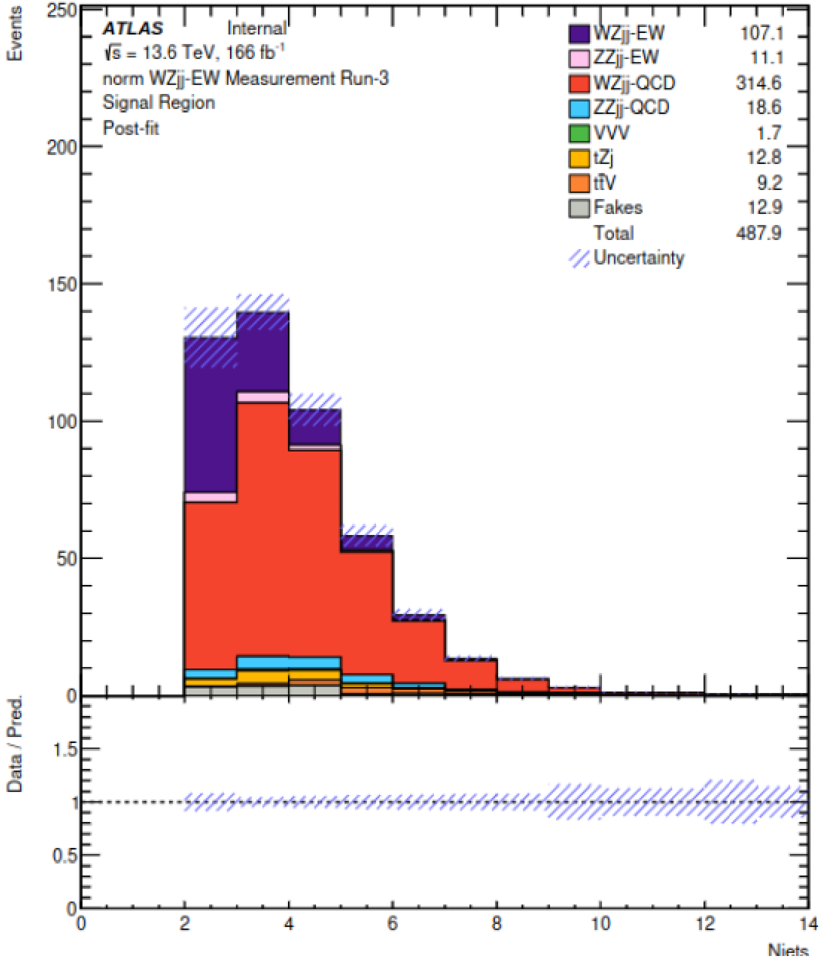
- **Third**: study if there are other observables with a discriminating power through a more rigorous approach
  → Use a beam search algorithm to find a near optimal configuration
- Answer: no other observables bring additional information, or the highlighted observables are correlated to the ones already used

Results

- **Fourth**: perform a fit to compare the expected significance obtained with the baseline and the optimized selections



Expected significance: 3.1



Expected significance: 4.1

Conclusion

- The VBS electroweak WZjj process provides a unique probe of the electroweak sector

- Final signal region with a higher purity: maximization of the number of WZjj events in the signal region
  → By fitting on simple kinematic observables, the expected significance is higher

- However, we lost a lot of statistics (nearly 50%). Is it good? … It depends on the type of analysis we want to do
  → For multivariate training: a large phase space is often preferable to maximize the available statistics and improve the learning process
  → For precision measurements: a more selective phase space can increase the sensitivity by focusing on the regions where the signal or the Wilson coefficients have the strongest impact

Thank you for your attention!

Questions?