ATLAS VBF+WH/ZH MC Validation Status and Wishlist for RUN-2

Carlo Pandini on behalf of the ATLAS Collaboration

Higgs (N)NLO MC and Tools workshop, 17-19.12.14

Overview and outline of this talk

ATLAS VBF+WH/ZH MC Validation Status and Wishlist for RUN-2

This talk will cover two subjects:

$\otimes$ **VH** (**1st part)** Associated Production

$\otimes$ **VBF** (**2nd part**) Vector Boson Fusion

Validation Status and Wishlist for RUN-2

- Run1 examples and lessons on signal modelling and related uncertainties
- Some MonteCarlo generators for Run2
- Signal features and modelling at $\sqrt{s} = 13$TeV

VH: associated production in ATLAS analyses

Higgs@125GeV

pb	8 TeV	13 TeV	ratio
ggF	19	44	2.3
VBF	1.6	3.7	2.3
VH	1.1	2.2	2
ttH	0.13	0.5	3.8

from CERN YR

Associated production mechanisms in ATLAS:

- ⊗ VH→ $Vb\bar{b}$
  (stand alone result: ATLAS-HIGG-2013-23)
- ⊗ VH→ VWW
  (stand alone result: ATLAS-CONF-2013-075)
- ⊗ analysis category in other channels
  ($\gamma\gamma$, WW, ZZ)

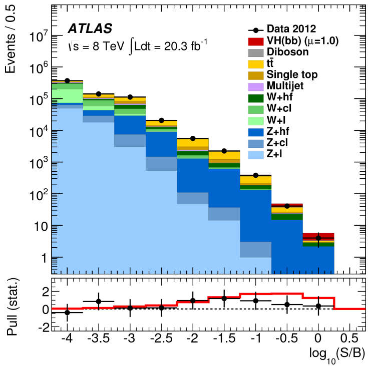
Important features of VH production in ATLAS analyses:

- **HVV** coupling measurement
- sensitive to **BSM** contribution
- **fermions** versus **bosons** couplings (via VH→ $b\bar{b}$)
- sensitive to **high-p_T Higgs boson production** (p_T^H>200GeV)
  Significant for VH→ $b\bar{b}$ (large BR(H→ $b\bar{b}$))

VH modelling: the VH→ $b\bar{b}$ example

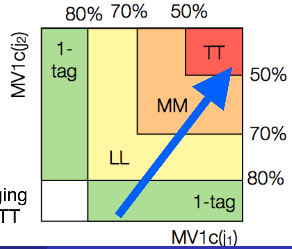
Analysis strategy:

- **Multivariate discriminator (BDT)**
- **3** Nlepton channels ($\nu\nu$,$l\nu$,ll)
- **2** Njet bins (2,3)
- **4** b-tag regions (1-tag = CR; 2-tag = SRs)
- **2** p_T^V bins ($\lessgtr$ 120GeV)



Analysis Phase Space

		1 tag (CR)
VpT > 120GeV	2 jets	2 tag LL (SR)
VpT < 120GeV	3 jets	2 tag MM (SR)
		2 tag TT (SR)



Increasing b-tagging purity from LL to TT

Results from the full Run1 statistics:

- **Significance** → **1.4(2.6)** **obs.(exp.)**

Signal strength

$\hat{\mu} = 0.51 \pm 0.31(\textit{stat.}) \pm 0.24(\textit{syst.})$

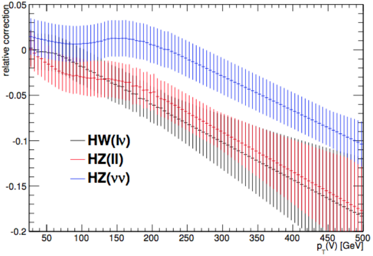
VH modelling: the VH→ $b\bar{b}$ signal

VH signal modelling:

- **qqVH** (qq→ZH, qq→WH) **Pythia8.165** (AU2-CTEQ6L1)
- **ggVH** (gg→ZH) **POWHEG+Pythia8.165**

Normalization and XS systematics:

- normalized to the CERN YR cross-sections (k^{NLO}_{ggHZ}=2)
- QCD scale / PDF uncertainties on total XS from CERN YR
- **EW** corrections from **HAWK** f(p_T^V)



Uncertainty band: max(2%, ΔEW^2)

Dominant signal systematic uncertainties:

- **PS/hadronization/UE**: POWHEG+Pythia8 vs POWHEG+Herwig
- **QCD scale**, p_T^V shape variation
- **QCD scale**, total XS(gg→ZH)

Signal	
Cross section (scale) (3:ggZH)	1% (qq), 50% (gg)
Cross section (PDF)	2.4% (qq), 17% (gg)
Branching Ratio (5)	3.3 %
Acceptance (scale)	1.5–3.3%
3-jet acceptance (scale)	3.3–4.2%
p_T^V shape (scale) (2)	S
Acceptance (PDF) (4)	2–5%
p_T^V shape (NLO EW correction)	S
Acceptance (parton shower) (1)	8–13%

Table: percentage effect on signal yield

(N): uncertainty rank according to its impact on the fitted signal strength

VH modelling: the VH$\rightarrow b\bar{b}$ example

Missing higher order corrections:

- **POWHEG+MiNLO+Pythia8** samples
- ren. (μ_R) and fact. (μ_F) scales varied by **2** and **0.5**, keeping $0.5 \leq \frac{\mu_R}{\mu_F} \leq 2$
- samples normalized to NNLO+NLO(EW) XS, to avoid double counting the inclusive uncertainties

⊗ p_T^V **differential**: (unit normalization)
Linear fit of varied p_T^V distributions, enveloping of the most discrepant variations

⇒ **non negligible change in the p_T^V shape** induced by μ_R, μ_F variations

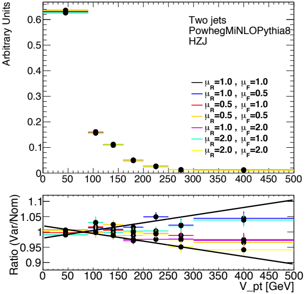
⇒ uncertainty $\sim$ 5%-10% for p_T^V > 200GeV

high-sensitivity region, even bigger role at 13TeV

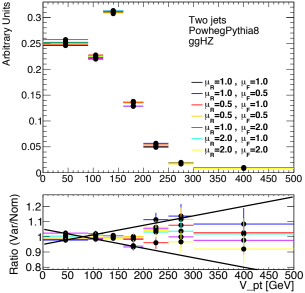
⊗ **2-3 jet acceptance**: effect of the order of 1% - 4%

⟹ applied via **Stewart-Tackmann method**

second dominant signal systematic in VH(bb)



HZJ

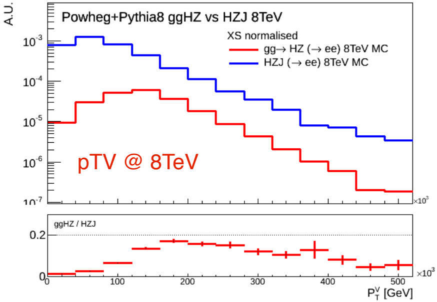


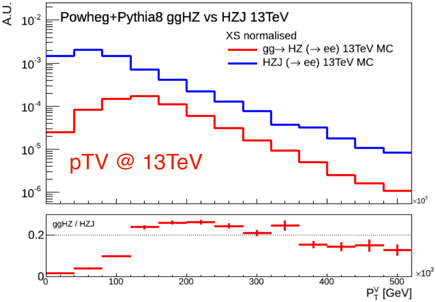
ggHZ

Towards Run2: VH modelling at 13TeV

VH Signal from 8TeV to 13TeV: the gg→ZH contribution

Plots normalized to the **POWHEG** cross section, ZH→ $e^+e^- b\bar{b}$ channel:





- **gg→ZH** was already included for the 8TeV analysis, using the POWHEG+Pythia8
- will become more important at 13TeV: increase of ∼70% of (gg/qq) from 8TeV to 13TeV for $p_T^Z \geq$150GeV
- much **larger scale uncertainties** than qq-induced processes

POWHEG+Pythia8.186:

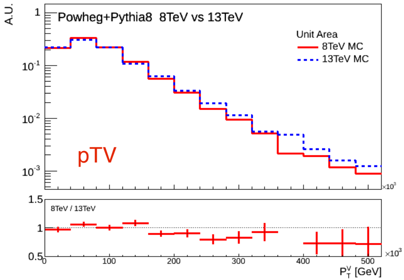
- samples generated at $\sqrt{s}$ of 8 and 13 TeV using POWHEG to produce ggHZ at LO (O(α_S^2) QCD)
- **main31 algorithm**: vetoed PS with the use of "power showers"
- simple event selection targets the 2-lepton final state (∼analysis selection)

Towards Run2: VH modelling at 13TeV

VH from 8TeV to 13TeV: ZH$\rightarrow e^+e^- b\bar{b}$ example

POWHEG+MiNLO+Pythia8.186 `HZJ`

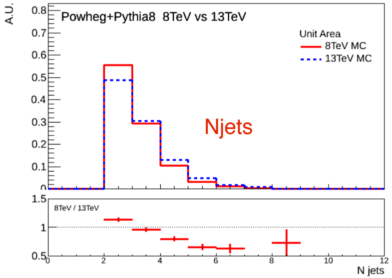
- new samples generated at $\sqrt{s}$ of **8** and **13** TeV
- **main31 algorithm**: vetoed PS with the use of "power showers"

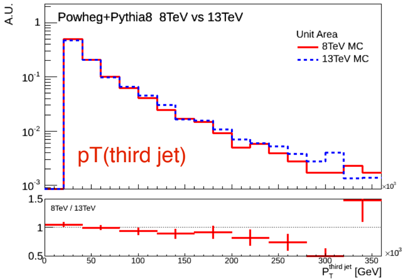


Shape comparison: samples normalized to unit area

$\Rightarrow$ harder p_T spectra and larger jet multiplicity at 13TeV

(increased contribution in high p_T^V region, most affected by systematics)





Towards Run2: VH modelling at 13TeV

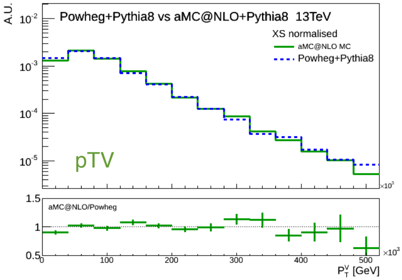
VH at 13TeV: POWHEG+MiNLO vs aMC@NLO

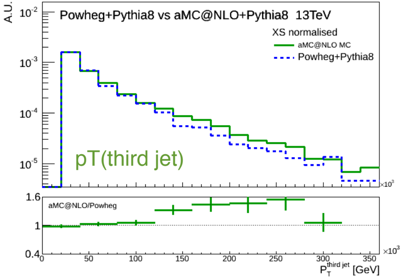
Plots normalized to their cross section, ZH$\rightarrow$ $e^+e^- b\bar{b}$ channel:

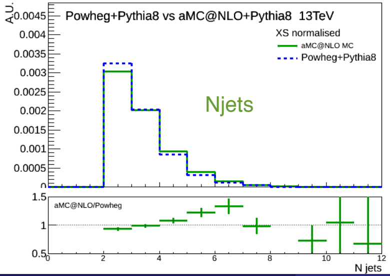
- **POWHEG+MiNLO+Pythia8.186**: HZ + 0 and 1 jet at NLO QCD [HZJ]
  main31 algorithm: vetoed PS with the use of "power showers"
- **aMC@NLO+Pythia8.186**: HZ inclusive at NLO (QCD)
  "wimpy showers" setup, global-recoil setting
  (possibility to generate NLO samples for EFT lagrangians)

Main features:

- good agreement for the total XS and for most of the differential distributions; discrepancies for the $p_T^{3^{rd}jet}$ and jet multiplicity
- compared **POWHEG** HZ vs HZJ (not shown here): differences from the 4th jet



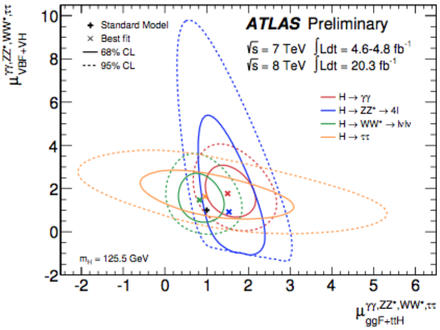




VBF: vector boson fusion in ATLAS analyses

ATLAS channels sensitive to **vector boson fusion**:

- H→ WW (ATLAS-HIGG-2013-13)
- H→ $\gamma\gamma$ (ATLAS-HIGG-2013-08)
- H→ $\tau\tau$ (ATLAS-CONF-2014-061)
- H→ZZ→4l (CERN-PH-EP-2014-170)



Evidence for VBF production at 4.1σ

ATLAS-CONF-2014-009:

not yet including updated WW, $\tau\tau$

$$\frac{\mu_{VBF}}{\mu_{ggF+ttH}} = 1.4^{+0.4}_{-0.5}(stat.)^{+0.4}_{-0.3}(syst.)$$

Signal Modelling:

- **POWHEG+Pythia8**
  normalized to full NLO QCD+EW prediction,
  approximate NNLO QCD corrections
- **EW shape corrections**
  from HAWK:
  $\Rightarrow$ $\mathbf{p_T^H}$ **reweighting**

H$\rightarrow \tau\tau$: VBF contribution

Analysis Strategy:

- 3 decay modes ($\tau_l\tau_l$, $\tau_l\tau_h$, $\tau_h\tau_h$) $\otimes$ Category (Boosted, **VBF**)
- MVA analysis trained separately against VBF signal
  VBF-topology variables: $\Delta\eta(j,j)$, m_{jj}, object centrality, ...

Channel	*VBF* category selection cuts
$\tau_{\text{lep}}\tau_{\text{lep}}$	At least two jets with $p_{\text{T}}(j_1) > 40$ GeV and $p_{\text{T}}(j_2) > 30$ GeV $\Delta\eta(j_1, j_2) > 2.2$
$\tau_{\text{lep}}\tau_{\text{had}}$	At least two jets with $p_{\text{T}}(j_1) > 50$ GeV and $p_{\text{T}}(j_2) > 30$ GeV $\Delta\eta(j_1, j_2) > 3.0$ $m_{\tau\tau}^{\text{vis}} > 40$ GeV
$\tau_{\text{had}}\tau_{\text{had}}$	At least two jets with $p_{\text{T}}(j_1) > 50$ GeV and $p_{\text{T}}(j_2) > 30$ GeV $p_{\text{T}}(j_2) > 35$ GeV for jets with $\lvert\eta\rvert > 2.4$ $\Delta\eta(j_1, j_2) > 2.0$

Analysis results:

- **Significance** $\rightarrow$ **4.5(3.5)** **obs.(exp.)**

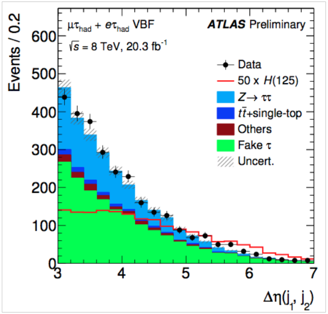
$$\hat{\mu} = 1.42^{+0.27}_{-0.26}(stat.)^{+0.32}_{-0.24}(syst.) \pm \mathbf{0.10(theory\ syst.)}$$

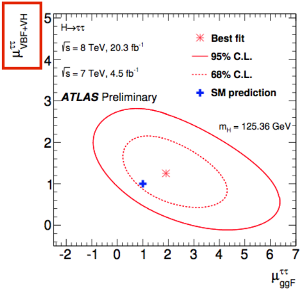
$\Rightarrow$ **The dominant contribution to H$\rightarrow \tau\tau$ evidence is from VBF category**

Production modes fit:

$$\hat{\mu}^{\tau\tau}_{\mathbf{ggF}} = 1.93^{+0.78}_{-0.77}(stat.)^{+1.19}_{-0.80}(syst.) \pm \mathbf{0.29(theory\ syst.)}$$

$$\hat{\mu}^{\tau\tau}_{\mathbf{VBF+VH}} = 1.24^{+0.48}_{-0.45}(stat.)^{+0.31}_{-0.28}(syst.) \pm \mathbf{0.08(theory\ syst.)}$$





VBF: main theoretical uncertainties (H$\to \tau\tau$)

VBF Higgs: small theory uncertainties in itself $\Rightarrow$ the H$\to \tau\tau$ example

- **UE / Parton Shower / Hadronization**:
  POWHEG+Pythia8 vs **POWHEG+Herwig**
  leads to a 5%-10% effect on the signal yield
- **ME+PS matching**:
  POWHEG+Herwig vs **aMC@NLO+Herwig**
  leads to $\sim$ 5% impact on the signal yield
- **Missing higher orders** (scale variations):
  range 1.4%-2%
  (can be larger with third jet selections (CJV, p_T^{tot}))
- **EW corrections**: differential modelling $d\sigma/dp_T^H$
  from truth p_T^H spectrum of **HAWK**, and associated uncertainty
- **PDFs**: PDF4LHC recommendation
  2%-3% effect on the total normalization,
  small shape effect up to $\sim$5% in the tails of the dijet distributions

Source of Uncertainty	Uncertainty on μ
Signal region statistics (data)	$^{+0.27}_{-0.26}$
Jet energy scale	$\pm$ 0.16
Tau energy scale	$\pm$ 0.07
Tau identification	$\pm$ 0.06
Background normalisation	$\pm$ 0.12
Background estimate stat.	$\pm$ 0.10
BR ($H \to \tau\tau$)	$\pm$ 0.08
Parton shower/Underlying event	$\pm$ 0.04
PDF	$\pm$ 0.03

Note: branching ratio uncertainty not negligible for H$\to \tau\tau$

$\Rightarrow$ VBF: **UE/PS/Hadronization** and **ME+PS matching** dominant signal systematic uncertainties

H→WW

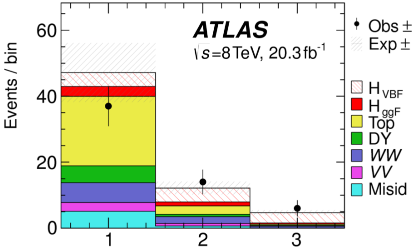
Anaysis Strategy: 3 OS dilepton channels (eμ, ee, $\mu\mu$)

- VBF category defined using **Centrality requirements**

$$C_X = |\eta_X - \frac{\eta_{j1} + \eta_{j2}}{2}| / \frac{\Delta\eta(j,j)}{2} \Rightarrow C_{j3} > 1 \text{ (CJV)}$$

$\Rightarrow C_{l(1,2)} < 1$ (central H decay products)

- **BDT discriminant** to select VBF-like events (3 SR bins)



$\Rightarrow$ Evidence of VBF H production at **3.2σ** !

$$\frac{\mu_{VBF}}{\mu_{ggF}} = 1.26^{+0.61}_{-0.45}(stat.)^{+0.50}_{-0.26}(syst.)$$

Key Features

- **Multivariate selection** of VBF-enriched region
  $\Rightarrow$ not trivial effect of scale variations in VBF-bins
- **Dominant signal systematic**:

  PS/UE/Hadronization from POWHEG+Pythia8 vs POWHEG+Herwig
  $\Rightarrow$ up to 14% in the most signal-like bin

  ME+PS matching
  $\Rightarrow$ 4.2% overall normalization

Uncertainty source	$n_j = 0$	$n_j = 1$	$n_j \geq 2$ ggF	$n_j \geq 2$ VBF
Gluon fusion				
Total cross section	10	10	10	7.2
Jet binning or veto	11	25	33	29
Acceptance				
Scale	1.4	1.9	3.6	48
PDF	3.2	2.8	2.2	-
Generator	2.5	1.4	4.5	-
UE/PS	6.4	2.1	1.7	15
Vector-boson fusion				
Total cross section	2.7	2.7	2.7	2.7
Acceptance				
Scale	-	-	-	3.0
PDF	-	-	-	3.0
Generator	-	-	-	4.2
UE/PS	-	-	-	14

$\Rightarrow$ Effect of all uncertainties on BDT is checked: **only PS/UE/Hadr. as shape**

ggF contamination in VBF regions

Contamination issue: ggF events can end up in VBF regions (especially with "loose" selections)

H$\rightarrow\tau\tau$ event yields in VBF-category:

Process/Category	*VBF*		
BDT output bin	all bins	second last bin	last bin
Fake backgrounds	1680 ± 50	8.2 ± 0.9	5.2 ± 0.7
$Z\rightarrow\tau\tau$	878 ± 29	7.6 ± 0.9	4.3 ± 0.7
Top	85 ± 16	0.1 ± 0.4	0.5 ± 0.4
$Z\rightarrow\ell\ell(\ell\rightarrow\tau_{\mathrm{had}})$	46 ± 21	1.2 ± 0.8	0.4 ± 0.3
diboson	66 ± 12	1.1 ± 0.4	0.5 ± 0.2
$ggF:\ H\rightarrow\tau\tau\ (m_H = 125\ \mathrm{GeV})$	17 ± 6	1.0 ± 0.4	1.2 ± 0.6
$VBF:\ H\rightarrow\tau\tau$	31 ± 8	4.5 ± 1.1	9.2 ± 2.3
$WH:\ H\rightarrow\tau\tau$	0.5 ± 0.4	< 0.1	< 0.1
$ZH:\ H\rightarrow\tau\tau$	0.2 ± 0.1	< 0.1	< 0.1
Total signal	48 ± 12	5.5 ± 1.3	10.4 ± 2.5
Total background	2760 ± 40	18.2 ± 2.3	10.9 ± 2.8
Data	2830	22	21

Examples from VBF-sensitive analyses

ggF contamination (%)

- H$\rightarrow\gamma\gamma$: VBF-tight(loose) = 20%(40%)
- H$\rightarrow WW$: 30%
- H$\rightarrow\tau_l\tau_h$: 35%-43%
  (of which, 11%-15% from last BDT bin)

Note on the H+(2-jets) modelling in VBF regions: **POWHEG** (H inclusive) NLO+PS prediction

- modelling in the $\geq$2-jets regions for now relied on the **parton shower**
- In addition: **POWHEG** p_T^H spectrum in 2-jet bin reweighted to match HJJ **POWHEG+MiNLO** prediction

Alternative MC tools for H+2jets modelling are available, but not yet fully used in VBF categories (under study):

- HJ, HJJ **POWHEG+MiNLO**
- MEPS@NLO
- aMC@NLO FxFx
- (new) **GoSAM ggH HJJJ** (up to three jets) at NLO can be included in **Sherpa** (under developement)

CP-violation from Higgs sector: can be probed by VBF production

$\Rightarrow$ Simple CP-odd observable: $\epsilon_{\mu\nu\rho\sigma} b_+^\mu p_+^\nu b_-^\rho p_-^\sigma = 2 p_{T,+} p_{T,-} \sin \Delta\phi(j,j)$ (Klämke, Zeppenfeld arXiv:hep-ph/0703202)

REPOLO

Generating full-detector simulation samples for different mixtures of CP-odd / CP-even states can be very resource intensive:

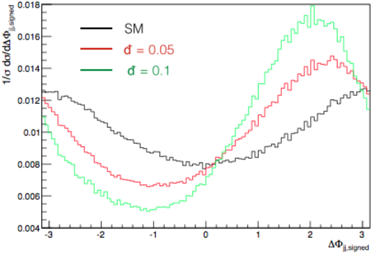
$\Rightarrow$ reweight from SM samples: $\frac{|M_{BSM}|^2}{|M_{SM}|^2}$

$\Rightarrow$ Leading Order ME for reweighting

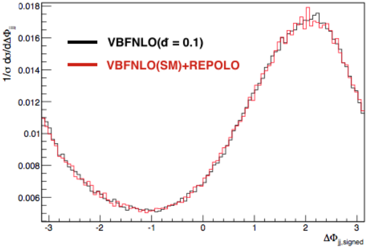
Under study:

- **REPOLO** w/ **VBFNLO**
- **REPOLO** w/ **POWHEG**
- **aMC@NLO** (EFT available at NLO)
  $\Rightarrow$ validate reweighting POWHEG samples

REPOLO Validation (w/ VBFNLO or POWHEG): $\tilde{d} \rightarrow$ additional CP-odd HVV-couplings for SM Lagrangian



SM vs 2 mixture scenarios in VBFNLO



REPOLO validation

Note: no form factors in these first studies (will be included)

Conclusions and Outlook

VH Associated Production in ATLAS

- **POWHEG** `ggHZ`
  - ⇒ large uncertainties as today, increased contribution at 13TeV
- **POWHEG+MiNLO** `HVJ`
  - ⇒ increased contribution at high-p_T^V
  - ⇒ important effect on the shape of the p_T^V from scale variations
- **aMC@NLO** `VH` (interest for `VH+1jet` with FxFx, EFT models)
  - ⇒ overall agreement with **POWHEG+MiNLO** for bosons kinematic, first differences on the additional jets

VBF Vector Boson Fusion in ATLAS

- **Modelling tools well tested** during Run1: small syst. uncertainties
- **PS/hadronization/UE** systematics dominant for VBF
- **ME+PS matching** uncertainties not negligible
- **Reweighting tools** (REPOLO) under study

Thank you for your attention!
And many thanks for any feedback from the community!

BACK-UP